Thèse présentée pour obtenir le grade de

Docteur de l'Université Louis Pasteur Strasbourg I

Discipline: Sciences du vivant

par Morgane Ciréfice

Engineering of a monospecific protease for blood glucose diagnostics

Soutenue publiquement le 21 Septembre 2009

Membres du jury :

Directeur de Thèse: Pr. Etienne Weiss
Ecole Supérieure de Biotechnologie, Strasbourg, France
Rapporteur Interne: Dr. Christian Boudier
Faculté de Pharmacie, Strasbourg, France
Rapporteur Externe: Dr. Pierre Martineau
Centre Régional de Lutte contre le Cancer Val d'Aurelle-Paul Lamarque, Montpellier, France
Rapporteur Externe: Pr. Clemens Posten
Institute für Bio- und Lebensmitteltechnik, Karlsruhe, Germany
Examinateur: Dr. Mara Boenitz-Dulat
Roche Diagnostics GmbH, Penzberg, Germany

To my family

Acknowledgements

My very first acknowledgements are for Dr. Rainer Schmuck and Dr. Herbert von der Eltz for the funding and the carrying out of this phD project at Roche Diagnostics in the XR-R department.

My second acknowledgements are for Dr. Uwe Kobold, who gave me the opportunity to work on a very interesting phD topic.

Sincerest acknowledgements are dedicated to Dr. Mara Boenitz-Dulat and Pr. Etienne Weiss, my supervisors at Roche Diagnostics and at the University Louis Pasteur (ULP) of Strasbourg respectively, for supporting me by circumspect advices and dedicating a lot of their time during this phD.

Special thanks are for Mr. Peter Kratzsch for his judicious advices during this phD project.

All my gratitude is to Mrs Angela Puhlmann and Mrs Jo Anne Wright for the analysis by mass spectrometry.

A lot thanks to the XR-RE and XR-R departments for the completion of this phD.

I would also like to thank Dr. Christian Boudier, Dr. Pierre Martineau and Pr. Clemens Posten, for accepting to take part to the dissertation committee.

My final acknowledgements but not the least are for my family and all the friends, who support me during this phD.

Thèse de doctorat

Ingénierie d´une protéase monospécifique pour le diagnostic du glucose sanguin

Morgane Ciréfice

Unité de recherche: Roche Diagnostics GmbH, Penzberg, Germany
DXR-RE3 Enzymology/Genetics Department

Directeur de thèse: Etienne Weiss
Université Louis Pasteur, Strasbourg

1 Introduction

Le diabète concerne actuellement 246 millions d'adultes dans le monde entier d'après les données collectées en 2006 par la Fédération Internationale du Diabète (FID). Ce chiffre devrait atteindre, 380 millions d'ici 20 ans, selon l'Organisation Mondiale de la Santé Humaine (WHO). Les laboratoires cliniques disposent de plusieurs méthodes analytiques (physiques, chimiques ou immunologiques) pour contrôler la glycémie sanguine. Cependant, aucune méthode et aucun matériel de référence n´ont été réglementés de manière uniforme au niveau international. La Fédération Internationale de Chimie Clinique et de Médecine de Laboratoire (IFCC) a décidé de constituer un groupe de travail, dans lequel Roche Diagnostics GmbH est engagé, afin de développer et standardiser un système de référence international de mesure de la glycémie sanguine (Hoelzel et al., 2004).

Le dosage sanguin de la glycémie, qui est un paramètre-clé de la surveillance du diabète à long terme, est dosé par le taux d´hémoglobine glyquée (HbA1c). Celui-ci est mesuré par chromatographie liquide à haute performance (HPLC) couplée à la spectrométrie de masse en tandem à source d'ionisation électrospray (MS-ESI) (Saudek et al., 2006). La première étape préparative de ces deux techniques consiste en une digestion protéolytique de l´hémoglobine glyquée par l'endopeptidase glutamique V8 (GluC) de *Staphylococcus aureus* (EC 3.4.21.19; Houmard et Drapeau, 1972), dans des conditions particulières qui permettent de générer à partir de l'hémoglobine un peptide de 6 résidus comprenant le groupement glyqué. Ce groupement est situé sur l'amino-acide Val en position 1 de la chaîne bêta de la protéine (Fig. 1).

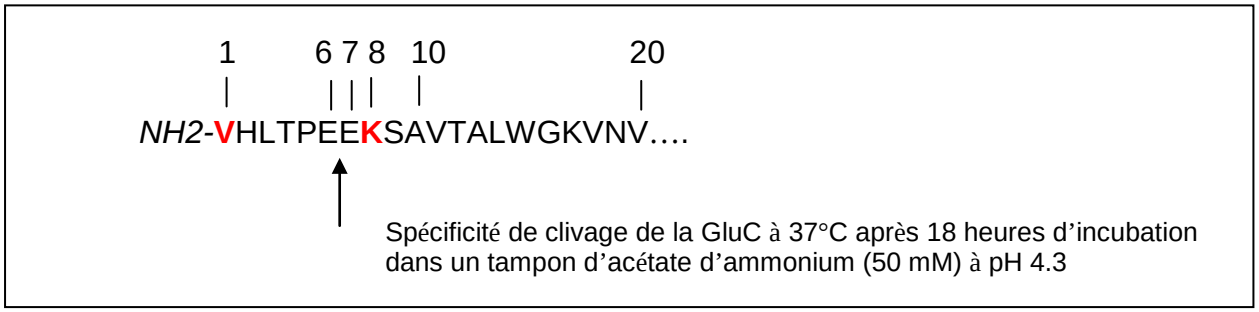


Figure 1. Spécificité de clivage de la GluC sur la chaine beta d'hémoglobine glyquée (HbA1c). L'endopeptidase glutamique de la souche V8 de *Staphylococcus aureus*, EC 3.4.21.19, hydrolyse spécifiquement après l'acide glutamique en 6ème position de la chaîne bêta d'hémoglobine. Cette hydrolyse génère dans des conditions expérimentales très spécifiques un hexapeptide glycosylé sur la valine en première position (en rouge).

Les hexapeptides ainsi obtenus après digestion sont quantifiés par HPLC/MS-ESI et le rapport entre la quantité de peptides glyqués et l'ensemble des hexapeptides générés permet de calculer le pourcentage de l'hémoglobine glyquée, par comparaison à un standard préalablement défini. Cette étape de digestion est en fait cruciale pour le dosage quantitatif du glucose sanguin et des conditions très précises de digestion par la GluC (incubation à 37°C pendant 18 heures dans une proportion de 1:100 (enzyme : hémoglobine) dans un tampon d´acétate d´ammonium à pH 4.3 (Jeppsson et al., 2002) doivent être mises en oeuvre pour effectuer le suivi de la glycémie par cette approche. Bien que très précis et fiable, ce dosage n'est pas utilisé couramment en laboratoire clinique à cause de la lenteur de la réaction enzymatique requise.

La GluC extraite de la souche V8 de *Staphylococus aureus* appartient à la famille des sérines protéases de part sa triade catalytique: histidine 51 - sérine 169 - acide aspartique 93. C'est un monomère de 268 résidus qui possède du côté amino-terminal une proséquence appelée zymogène (Drapeau, 1978). Cette proséquence rend la protéase inactive après synthèse et empêche ainsi son autodégradation. Elle est clivée lors de la sécrétion par une autre protéase, l'auréolysine (EC 3.4.24.29), également présente chez *S. aureus* (Nickerson et al., 2007; Fig. 2).

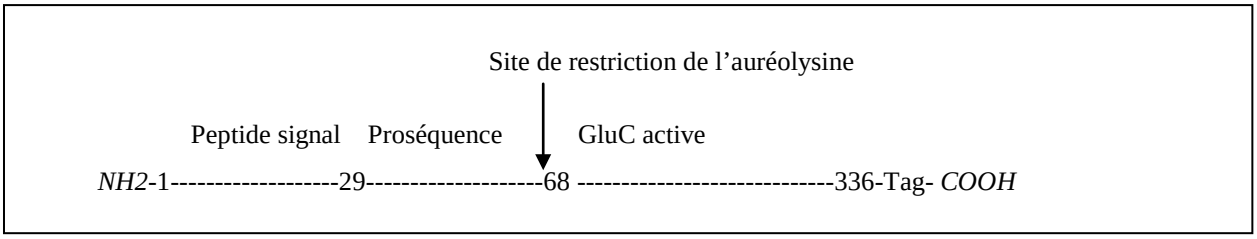


Figure 2. Structure de la GluC de la souche V8 de *Staphylococcus aureus*. La GluC possède sur sa partie amino terminale une proséquence, qui empêche son autodégradation. La GluC est activée suite au clivage de la proséquence par une autre protéase de *S. aureus*, appelée auréolysine.

Comme des travaux antérieurs à cette thèse avaient montré qu'il était possible de produire la GluC (Yabuta et al., 1995), Roche Diagnostics GmbH souhaitait obtenir une GluC dont la vitesse d'hydrolyse est accélérée par comparaison à l'enzyme parentale, de manière à réduire le temps de préparation des hexapeptides d'hémoglobine nécessaires au dosage quantitatif du glucose sanguin. L'évolution moléculaire est une technique d'ingénierie de protéines pour modifier leurs propriétés enzymatiques et est donc une méthode d'approche prometteuse tout à fait adaptée à la génération de librairies de variants pour du criblage à haut débit.

Mon projet de thèse a consisté à mettre en place une stratégie permettant par évolution moléculaire de sélectionner et de produire des mutants de GluC permettant d´appliquer en routine cette méthode de dosage par hydrolyse enzymatique. La technique d'évolution moléculaire suit généralement un procédé itératif qui se structure en 3 étapes: la GluC est tout d'abord clonée dans un organisme recombinant pour faire évoluer le gène par évolution moléculaire et ainsi produire un variant optimisé. La structure de la GluC a été étudiée pour déterminer ses positions stratégiques à muter. Une librairie de variants a pu être ainsi créée par différentes techniques de mutagénèse. Les paramètres enzymatiques de la GluC parentale ont été tout d'abord caractérisés pour mettre en place le système de criblage. Les librairies de mutants sont criblées jusqu'à cibler un variant avec de meilleures caractéristiques enzymatiques. Une fois que la reproductibilité des résultats est vérifiée, le mutant est exprimé, purifié et activé. Il peut alors être de nouveau muté pour constituer une nouvelle librairie de variants ou il peut être caractérisé avant de l'expérimenter dans le test d'hémoglobine glyquée (Fig. 3).

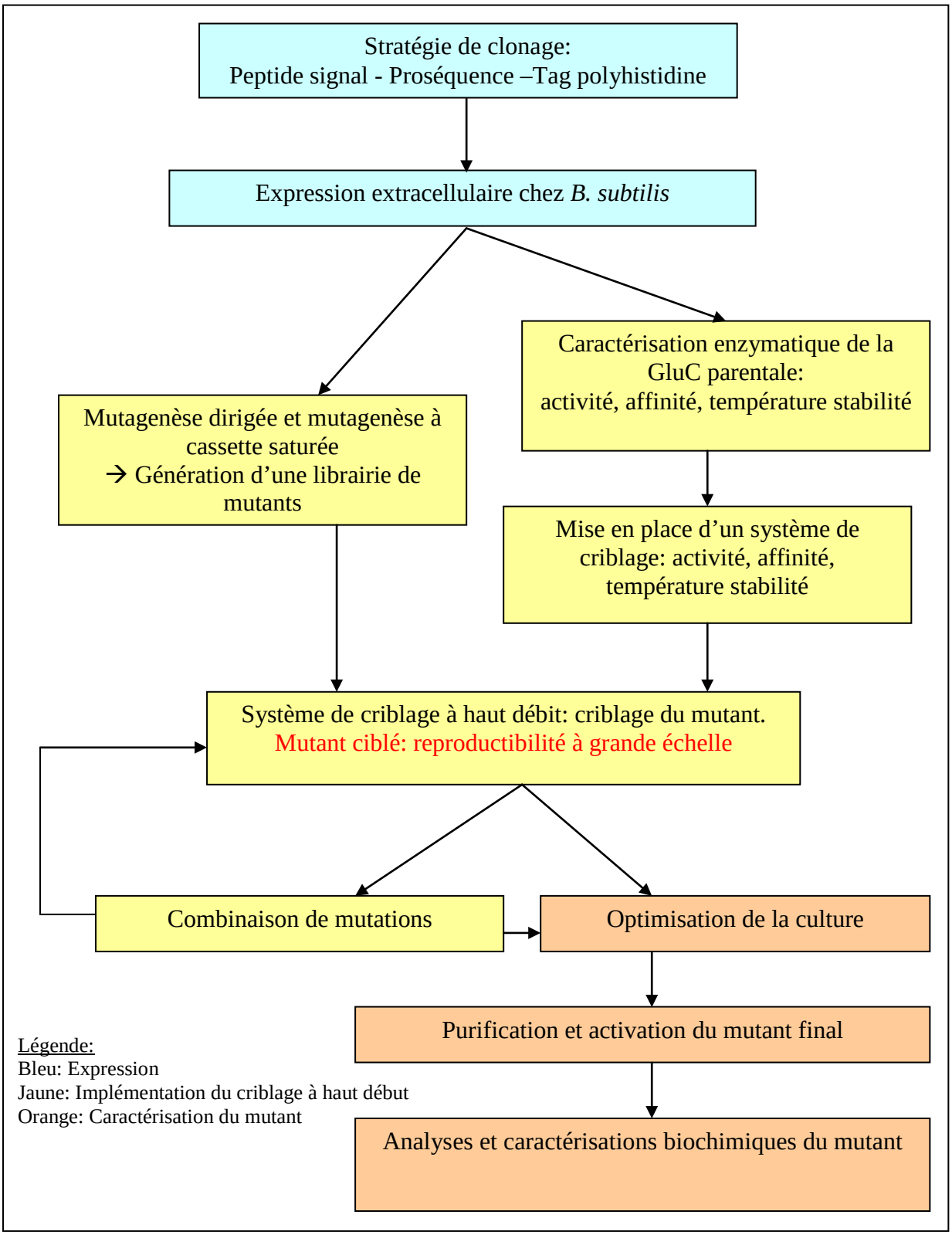


Figure 3. Représentation schématique des différentes étapes de l'évolution moléculaire de la GluC. Après clonage et expression de la V8-GluC parentale, la protéase est caractérisée pour mettre en place le système de criblage. Des librairies de mutants sont générées en parallèle. Après criblage à haut débit, le mutant est exprimé par culture de *B. subtilis* pour être finalement purifié et activé.

2 Résultats

La GluC a été ainsi exprimée chez *B. subtilis* WB600, une souche déficiente en six protéases extracellulaires, à l'aide d'un vecteur navette, le plasmide pMSE3 (Wu et al., 1991 ; Silbersack et al., 2006). Une étiquette poly-histidine a été ajoutée en position carboxy-terminale de manière à faciliter sa purification en une étape.

Les différents essais de culture de la GluC ont été donc stoppés en début d'expression pour éviter la digestion de la GluC, son activation et/ou des dégradations causées par les huit autres protéases extracellulaires de *B. subtilis.* Les activités mesurées ne correspondent pas aux activités enzymatiques réelles de la GluC parentale par l'absence de son activation complète.

Les variants de l'endopeptidase glutamique ont été créés par polymérisation en chaîne (PCR) après identification de positions stratégiques par modélisation de la structure tridimensionnelle de l'endopeptidase parentale (Prasad et al., 2004). Ces positions correspondent essentiellement aux résidus de la triade catalytique (H51, D93, S169) et aux trois acides aminés (V1, T164, N193) qui interagissent avec la triade catalytique et le substrat. Les autres positions modifiées par mutagenèse dirigée ont été déduites de l´alignement de la séquence protéique avec d´autres endopeptidases glutamiques (Barbosa et al., 1996).

Pour le criblage, un système de criblage à haut débit à multi-paramètres enzymatiques a été miniaturisé dans des plaques à 96 puits de manière à pouvoir analyser en parallèle l'affinité (Km), la thermostabilité et la spécificité de coupure de l´hexapeptide de l´hémoglobine des mutants obtenus. Les tests ont été réalisés avec des concentrations de substrat pour être à 30 % de la valeur du pseudo-Km et à une température d´incubation de 51°C pendant 30 minutes pour rester à 20 % de l´activité résiduelle de la GluC. La mise en place de l'ensemble de cette approche et le criblage de près de 12000 mutants construits a permis d'isoler un variant G166I qui a les caractéristiques recherchées.

Exprimée chez *B. subtilis*, la culture de la protéase mutée a été étudiée pour ne pas s'autoprotéolyser en position carboxy-terminale et être ainsi aisément purifiée par chromatographie d'affinité Ni-NTA grâce à l´étiquette poly-histidine. Après activation par autoprotéolyse de la proséquence à la position de la mutation N-1E, diverses analyses enzymatiques et biochimiques ont permis de montrer que ce variant, G166I avec la mutation

N-1E possède une activité spécifique 4 fois supérieure à celle de l'endopeptidase parentale en présence de substrat synthétique: Z-Phe-Leu-Glu-pNA dans un tampon de Tris-HCl à pH 7.8 (Fig. 4). Dans les mêmes conditions, la constante d'affinité (Km) est également améliorée d´un facteur de 3 (Fig. 5). De plus, ce variant est 4 fois plus actif que la protéase parentale dans un tampon d´acétate d´ammonium à pH 4.3 en présence du substrat correspondant à l'hexapeptide synthétique de l'hémoglobine non glyquée. Le pH optimal d'activité de ce variant est de 7 par rapport à la valeur de 7.8 de l'enzyme parentale et il est également plus stable à 50°C (Fig. 6 et Fig. 7). Le point isoélectrique de 4.5 n'est pas changé par la mutation. Cependant, bien que intrinsèquement plus actif que l'endopeptidase parentale en présence de substrats synthétiques, le variant isolé est moins performant que la GluC d'un point de vue activité, affinité et spécificité de coupure pour l´hexapeptide de l´hémoglobine entière dans les conditions expérimentales avec un tampon d'acétate d'ammonium à pH 4.3.

	AS (U/mg)			
Tampon	**Roche**	**NEB**	**Wt**	**Mutant**
Tris pH 7,8	44.84	30.77	42.62	196.23
KH_2PO_4 pH 7,8	47.14	40.51	54.82	206.70
KH_2PO_4 pH 5	0.1612	0.22	0.5606	5.39
Acétate d'ammonium pH 4.3	0.0140	0.01	0.0206	0.2564

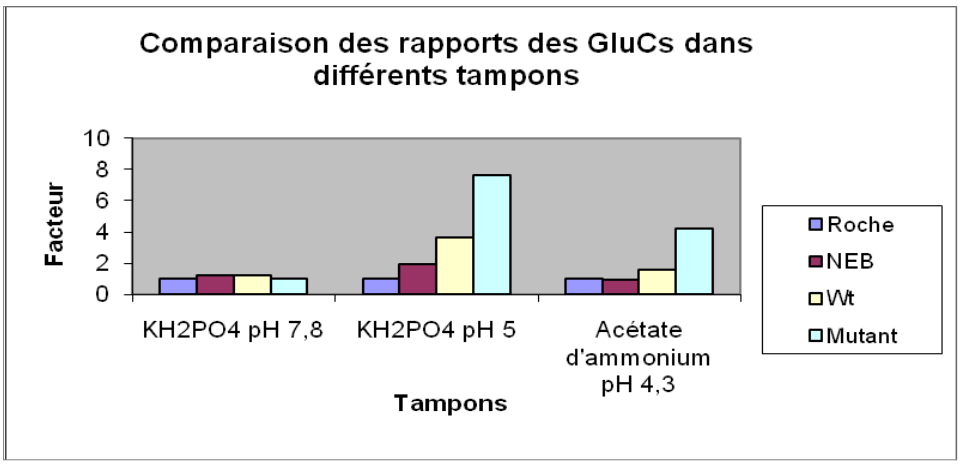


Figure 4. Activités spécifiques (AS) de la GluC parentale (Wt) et du variant (M). Les activités spécifiques du mutant (M) ont été comparées aux GluCs de Roche, de New England Biolabs (NEB) et la GluC parentale recombinante (Wt) dans des tampons Tris à pH 7.8, d'acétate d'ammonium à pH 4.3 et de phosphate de potassium à pH 5 et 7.8. Les différentes activités spécifiques ont été comparées après calcul du quotient entre l'activité dans un tampon donné et celle de référence dans le tampon Tris à pH 7.8 pour éviter les différences d'activité dues aux tampons.

	Km (mM)			
GluCs	**Km 1**	**Km 2**	**Km 3**	**Moyenne**
Roche	0,53	0,54	0,41	0,49
NEB	0,46	0,55	0,46	0,49
Wt	0,44	0,53	0,54	0,50
Mutant	0,17	0,18	0,16	0,17

Figure 5. Constante d'affinité Km de la GluC parentale (Wt) et du mutant. La constante d'affinité Km est 4 fois supérieure dans un tampon Tris à pH 7.8 avec le substrat synthétique de Roche que la GluC de Roche, de New England Biolabs (NEB) et de la GluC parentale.

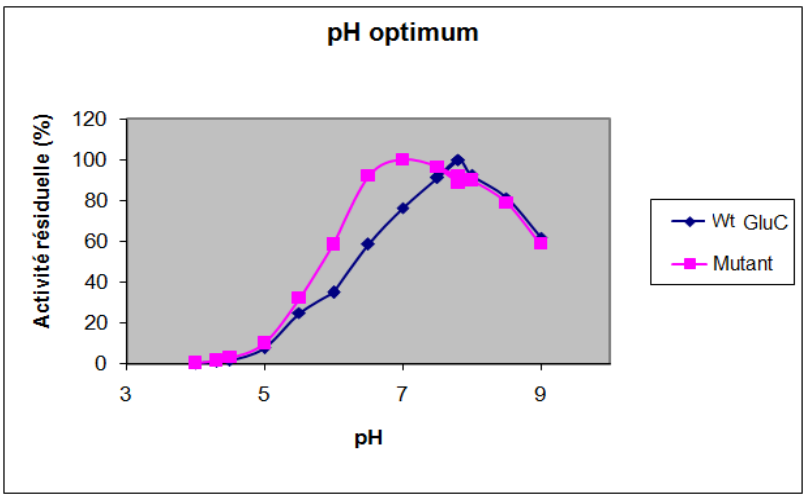


Figure 6. pH optimum de la GluC parentale (Wt) et du mutant. Le pH optimum a été déterminé dans un tampon d'acétate d'ammonium (50 mM) entre pH 4 et 6, KH_2PO_4 (0.1 M) entre pH 6 et 7.5 et Tris-HCl (0.1 M) entre pH 7.5 et 9 avec le substrat d'hexapeptide d'hémoglobine. Le pH optimum se décale de pH 7.8 pour la GluC parentale à pH 7 pour le mutant.

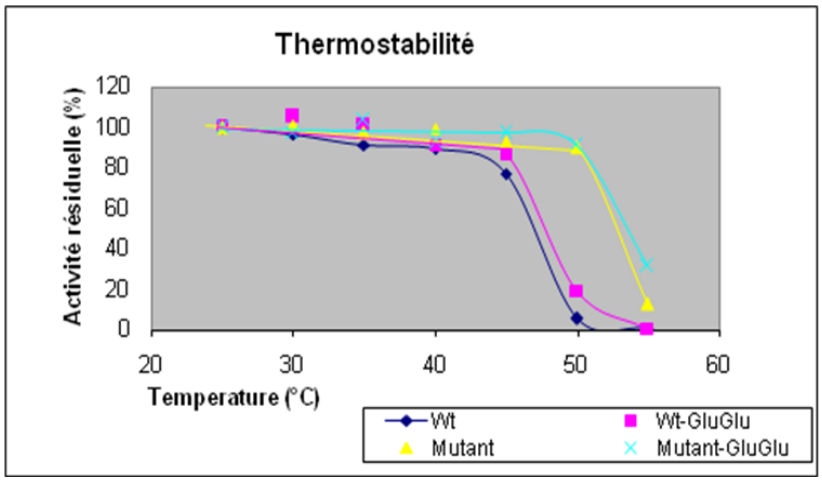


Figure 7. Thermostabilité de la GluC parentale (Wt) et du mutant avec ou sans dipeptide de Glu-Glu. L'activité résiduelle de la GluC a été ensuite calculée après 30 minutes d'incubation entre 25°C et 55°C avec ou sans dipeptide Glu-Glu (5 mM) pour tester la thermostabilité du mutant.

3 Discussion

Il existe deux isoformes de GluC de la souche V8 de *S. aureus*. La “GluC de Roche”, qui a été séquencée par Drapeau et est commercialisée comme lyophilisat de GluC par Roche GmbH (Drapeau et al., 1972). L’autre “GluC V8” a été découverte par Carmona est fournie sur le marché par New England Biolabs Ltd en tant que protéine recombinante de *B. subtilis* (Carmona and Gray, 1987). Pour ce projet, seule la GluC de Carmona était disponible. Cette GluC de Carmona a été clonée chez *E. coli* avec sa proséquence afin d’empêcher son autodégradation. La proséquence de la GluC était supposée être clivée que par une autre protéase extracellulaire de *S. aureus*. Cette affirmation a été contredite dans des travaux plus récents publiés en 2009. Il a été démontré que le clivage de la proséquence était une succession de d’hydrolyses séquentielles. La proséquence est d’abord autodigérée avant d’être heteroprotéolysée par l’auréolysine (Nickerson et al., 2008; Nemoto et al., 2009; Fig. 8). Cette découverte tardive dans le projet de thèse n’a pas eu de conséquences sur le système de criblage. Si un variant de la GluC est plus actif, il sera activé et détecté plus rapidement. Comme l’hydrolyse de la proséquence de la GluC est due à une autodigestion séquentielle, le dernier acide aminé de la proséquence, une asparagine, a été substitué par un acide glutamique (N-1E) pour favoriser la coupure de la proséquence par autodigestion lors de la sécrétion.

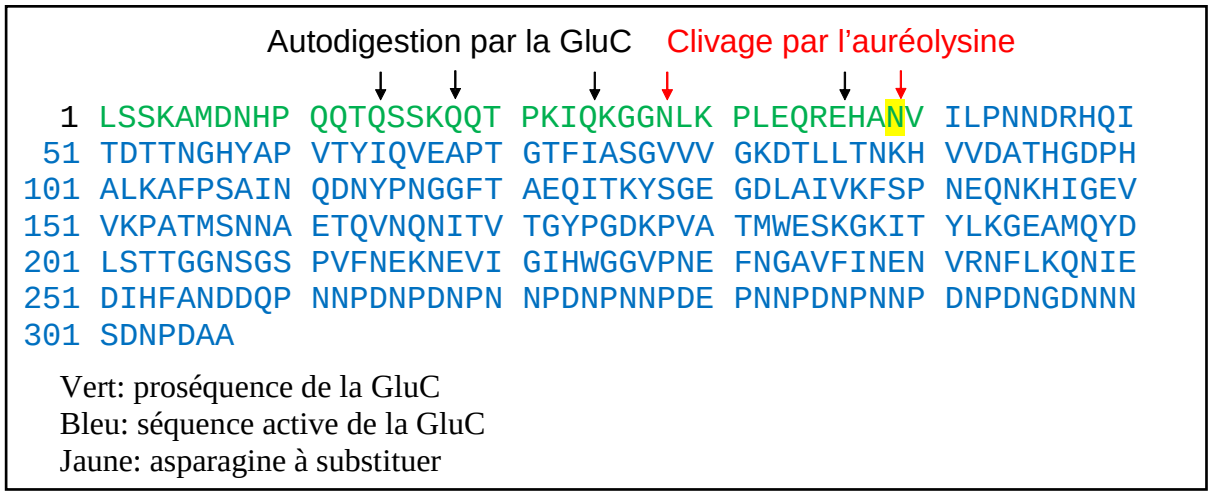


Figure 8. Maturation de la GluC par autocatalyse et hydrolyse par l’auréolysine. La GluC est protégée de son autodégradation par une proséquence du côté amino-terminal. La GluC est activée suite à une autodigestion de la proséquence suivie d’un clivage complet par une autre protéase de *S. aureus*, appelée auréolysine.

Cette étude de la GluC, menée dans un contexte industriel, a permis de montrer que les caractéristiques enzymatiques de l'endoprotéase glutamique peuvent être modulées sélectivement par le biais de mutations précises grâce à l'évolution moléculaire. L'activité spécifique du variant isolé par criblage à haut débit et caractérisé en détail après surexpression chez *B. subtilis* est nettement supérieure à celle de l'enzyme parentale en présence du substrat synthétique Z-Phe-Leu-Glu-pNA, mais ne correspond pas à ce qui était recherché : un mutant dont les constantes cinétiques de l'hydrolyse de l'hémoglobine pour l'obtention de l'hexapeptide amino-terminal de la chaîne bêta sont améliorées. Ces résultats soulignent l'importance du système de criblage mis en place lors de la sélection des mutants et confirment la première loi qui régit l'évolution moléculaire "you get what you screen for" (Schmidt-Dannert and Arnold, 1998). Il aurait été préférable d'utiliser dans les tests effectués l'hémoglobine entière en tant que substrat synthétique. L'amélioration des caractéristiques enzymatiques du variant pourraient être cependant utilisées pour d'autres tests enzymatiques telles que les mesures de routine par empreinte de masse de peptide en protéomique (Corthals et al., 1999 ; Bakthiar and Nelson, 2001).

4 Perspectives

Ce projet pourrait être continué de différentes façons. Le mutant étant 4 fois plus actif et présentant 4 fois plus d'affinité envers les acides glutamiques que la GluC parentale, d'autres cycles de direction moléculaire pourraient être également répétés pour tenter d'améliorer le mutant. Les positions autour de la mutation pourraient être mutées par mutagénèse à cassette saturée. Les acides aminés interagissant avec la position 166 pourraient être aussi testés de la même manière.

La GluC de Drapeau pourrait être mutée à la position correspondante pour vérifier l'effet de la mutation sur la protéase. L'ADN de cette GluC pourrait être extrait directement de la culture sur le site de production de Roche. La GluC est cependant seulement commercialisée sous le nom de Roche et non pas directement produite par Roche, en raison de la toxicité de *S. aureus*. Il faudrait donc modifier chaque position différente avec la GluC de Carmona par mutagénèse dirigée pour obtenir la même séquence d'ADN que celle de Drapeau et y apporter par la suite la mutation en position 166.

Cette mutation aurait pu être aussi testée sur les endopeptidases glutamiques d'autres organismes tels que *Bacillus* ou *S. epidermis*. Cela n'a pas pu être effectué en raison des brevets déposes par d'autres entreprises sur ces protéases.

Enfin, ce projet aurait pu être réalisé sans utiliser la voie de l'évolution moléculaire. La GluC a déjà été immobilisée sur des billes d'agarose dans les années 90. (Sahni et al., 1991). Cette immobilisation n'avait pas interféré avec la spécificité de la GluC envers la chaîne alpha de l'hémoglobine à pH 4 et 37°C mais son activité avait été réduite de 30 %. Dix ans plus tard, l'activité et la spécificité de la GluC ont été améliorées par adsorption de la GluC sur une phase solide en présence de Sodium Dodecyl Sulfate (Vercaigne-Marko et al., 2000). Des études sur l'immobilisation de la pepsine sur de l'oxyde d'aluminium ont démontré que les cinétiques enzymatiques de la protéase pouvaient être améliorées (Ticu et al., 2004). Cependant, la formation d'autres composés ont rendu la purification de l'hémoglobine hydrolysée plus difficile. Cette idée a cependant été développée par la compagnie Takeda, San Diego, Inc. (Lim et al., 2006). La GluC de Roche a été immobilisée sur une cartouche couplée directement au spectromètre de masse. Cette technique permet la digestion de peptides en moins de cinq minutes et serait la solution pour l'IFCC pour le transférer au test de l'hémoglobine glyquée.

Le mutant trouvé durant cette étude pourrait lui servir pour d'autres méthodes où la GluC est utilisée dans des tests analytiques de routine, comme par exemple pour des tests d'empreinte de masse de peptide en protéomique.

Ce travail de thèse a permis d'étudier les propriétés moléculaires de la GluC ainsi que ses conditions d'expression chez *Bacillus subtilis*. Ce projet donne ainsi une meilleure compréhension de la structure-fonction d'une enzyme qui a de nombreuses applications industrielles.

TABLE OF CONTENTS

Table of contents

INTRODUCTION

1. Presentation of Roche Diagnostics

Founded on the first of October 1896 by Fritz Hoffmann-La Roche in Basel, Switzerland, Roche is one of the world's leading research-focused healthcare groups in the fields of pharmaceuticals and diagnostics. Its activities consist in the discovery, development and manufacture of pharmaceuticals and diagnostic systems, which spreads out on about 150 countries and employs around 79,000 people. Roche operates world-wide under two divisions: Pharmaceuticals and Diagnostics, both belonging to the Roche Holding AG. Roche has several centers of Research & Development worldwide for each division to sell its products all over the world (Figures 1 and 2), but only one facility, in Penzberg, Germany, has a center with both divisions. In 2008, sales by the Pharmaceuticals Division totaled 36.8 billion Swiss francs and the Diagnostics Division posted sales of 9.3 billion francs. Roche has R&D agreements and strategic alliances with numerous partners, including majority ownership interests in Genentech and Chugai, and invested over 8 billion Swiss francs in R&D in 2007.

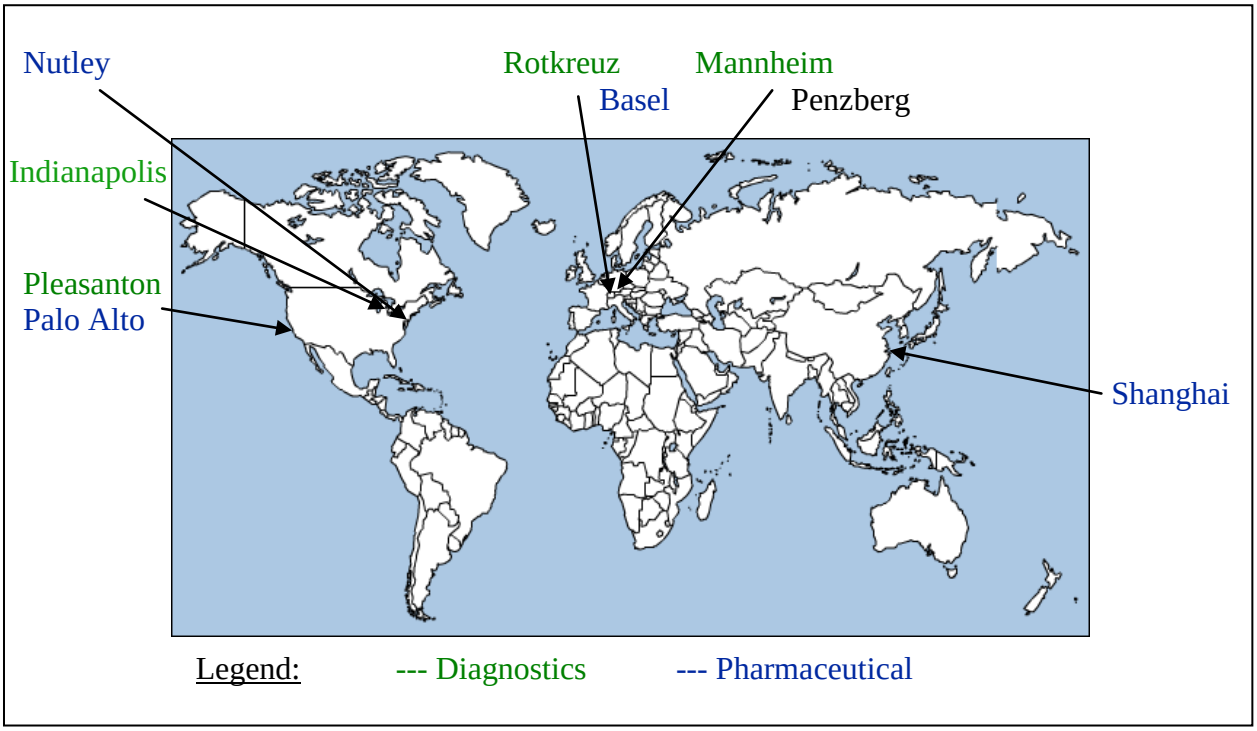


Figure 1. Roche locations of the diagnostic and pharmaceutical divisions worldwide. Roche divisions are above all located in the United States and Europe, but only one facility, in Penzberg, Germany, has a center with both divisions.

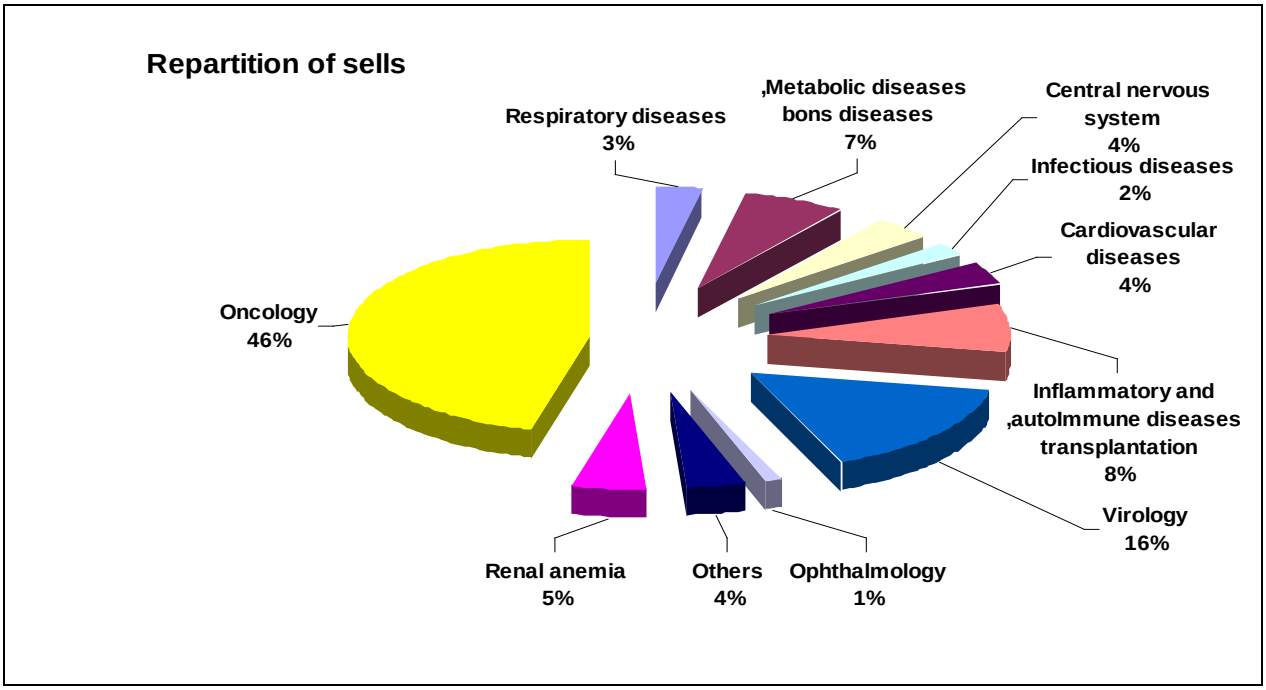


Figure 2. Repartition of sells. Sells are mainly divided into oncology (46%) and virology (16%).

Roche Pharma is engaged in combating human diseases and work in the following areas: anaemia, cardiovascular diseases, central nervous system, dermatology, infectious diseases, inflammatory and autoimmune diseases, metabolic disorders, oncology, respiratory diseases, transplantation and virology. Several products were developed by Roche Pharma: Tamiflu for the prevention and treatment of influenza, Avastin for the treatment of advanced colorectal cancer, Tarceva for the treatment of patients with certain types of lung or pancreatic cancer, Mircera, also called PEG-EPO, for the treatment of anaemia, Valium for the treatment of anxiety disorders or Bactrim for the prevention or the treatment of specific bacterial infections.

Roche Diagnostics division was established in 1968 in Switzerland. It develops new diagnostic tests and automatic analyzers for patient self-monitoring as well as services for diagnostic laboratories to perform clinical analyses for hospitals and office-based physicians.

2. Diabetes

2.1 Diabetes overview

Diabetes mellitus is defined as high blood glucose (hyperglycemia) in patients. Normally, the human body maintains blood glucose in a steady range: 70 mg/dl to 110 mg/dl (~ 3.9 to 6 mM). This optimal range can fluctuate throughout the day following diet, exercise, stress and illness.

Two main diabetes forms, types 1 and 2, were officially distinguished in 2003 by the ADA Expert Committee on the diagnosis and Classification of Diabetes Mellitus (Gavin et al., 2003). Type-1 diabetes was formerly called IDDM (Insulin Dependant *Diabetus Mellitus*) in 1979 by the National Diabetes Data Group but some other terms are used such as immune-mediated or juvenile-onset diabetes as the immune system destroys the beta cells of the pancreas in this form and it appears most commonly in childhood. Type-1 is characterized by little or no circulating insulin and can be controlled by regular injections of insulin. Type-2 diabetes was named NIDDM (Non Insulin Dependent *Diabetes Mellitus*) or adult-onset diabetes, which accounts for about 85 % to 95 % of all diagnosed cases of diabetes. In this form, the body does not respond to insulin. Other types of diabetes are defined separately besides these two diabetes categories: pre-diabetes, gestational diabetes and all other specific types of diabetes. Pre-diabetes is a high glucose level (Impaired Glucose tolerance: IGT) but not as high as normal diabetes. In most cases, people with pre-diabetes develop type-2 diabetes within 10 years. Gestational diabetes (Gestational *Diabetes Mellitus*) is developed by half of pregnant women and is similar to type-2 diabetes by a hyperglycemia diagnose. In this case, the pregnancy hormones block the action of insulin, thereby making the women less sensitive to their own insulin. It usually goes away after the child was born, but they have a 40 % to 60 % chance of developing diabetes, most of the time type-2 diabetes, within the next 5–10 years (IDF, 2007). The last class of diabetes includes all the other specific types of diabetes. In this group are included people with beta cell function defects, insulin action defects and exocrine pancreatic disease or dysfunctions (Gavin et al., 2003).

Diabetes is ranked among the leading causes of cardiovascular disease, blindness, renal failure and lower limb amputation. Type-2 diabetics are over twice as likely to have a heart attack or stroke compared to non-diabetics. The biggest concern is raised by children and adolescents, who are already developing type-2 diabetes due to obesity.

According to a survey of the World Health Organization (WHO), 171 million adults, older than 20 years old, were affected by Diabetes worldwide in 2000. At this time, the WHO estimates to reach 366 million diabetics in 2030. A recent study of the International Diabetes Federation (IDF) estimated in 2006 this number of 246 million, which represents 6 % of the world's population and has predicted that it would increase up to 380 millions of diabetics worldwide in 2025, which means 7.1 % of the adult population.

2.2 Roche and Diabetes Care

Tightly managing diabetes can delay and even prevent some of the complications associated with diabetes. This is why several studies on diabetes management and long-term complications were launched separately for type-1 and -2 of diabetes.

The Diabetes Control and Complications Trial (DCCT) was a multicenter randomized clinical trial in the US, designed to compare intensive and conventional therapies and their relative effects on the development and progression of diabetic complications in patients with type-1 diabetes over more than 6 years. For this study, insulin injections were intensively or regularly administered thanks to HbA1c assessment to patients without or with mild-renopathy.

Besides, early vascular and neurological complication development and progression, such as renal, neurological, cardiovascular and neuropsychological outcomes, were followed up by considering glycemic exposure thanks to HbA1c assessment. The intensive insulin injections were designed to achieve blood glucose as close to normal range as possible whereas regular injections are based on conventional therapy of two or three administrations a day. The DCCT concluded from this experiment that intensive therapy effectively delays the onset of diabetic retinopathy, nephropathy and neuropathy in type-1 diabetics. It confirmed by the same way the major role of the HbA1c as a blood glucose control long-term parameter.

This data study was then taken over by the Epidemiology of Diabetes Interventions and Complications (EDIC) from the DCTT to examine the longer term effects of the DCCT intervention over the next 10 years, as cardiovascular, advanced stages of retinal and renal diseases require a longer period of time to develop. Thanks to this collaboration DCCT/EDIC, a long-term follow-up of diabetes control and complications in type-1 diabetics was set up over 16 years (DCCT/EDIC, 1993; DCCT/EDIC, 1999).

In the same way, other programs were launched by the United Kingdom Prospective Diabetes Study (UKPDS) and the Steno Diabetes Center in Copenhagen (SDC) on patients with type-2 diabetes to study the cardiovascular risks and long-term complications due to type-2 diabetes (Stratton et al., 2000; Gaede et al., 2003). However, a not-yet-solved controversy was raised up after the result publications from the ACCORD (Action to Control Cardiovascular Risks in Diabetes) and the ADVANCE (Action in Diabetes and Vascular Disease) projects. (Nathan et al., 2005; Dluhy and McMahon, 2008).

In type-1 diabetes, intensive insulin therapy can delay the onset and slow progression of clinically important retinopathy by a range of 54 to 76 % according to the initial stage of the disease. Every 10 % HbA1c decrease was correlated to 39 % retinopathy development decrease. It reduced also the risk of any cardiovascular disease event by 42 % and the development of albuminuria, clinical symptom of kidney dysfunction, by 54 %. Besides, a 10 % decrease in HbA1c can lead to a 25 % decrease in risks of nephropathy development signs. Diabetic neuropathy was reduced by 60 % during this intensive treatment (DCCT/EDIC, 1999; Nakamoto, 2004).

According to studies of the National Institute of Diabetes and the Diabetes Trial Unit at Oxford University (UKPDS), in type-2 Diabetes, 1 % HbA1c decrease reduces microvascular complications by 37 % and a 21 % decrease in deaths related to diabetes. As other results are contested for the moment, further investigations are required (DCCT/EDIC, 1993).

Numerous advancements have been made in diabetes management over the past three decades. Years ago, patients checked their blood sugar level either with urine test at or blood system at hospital. In the 1970s, two new promising devices are introduced: insulin pumps and portable glucose monitors. The first portable glucose monitor was introduced by Roche Diagnostics, formerly known as Boehringer Mannheim (Dunn et al., 1997).

The 1990s exploded with tremendous advancements in new treatments, technologies, understanding of diabetes and diabetes care. On the one hand, these include rapid-acting human analogs of insulin or Thiazolidinedione such as oral medication and insulin pump (Disetronic, Switzerland). At the present time, scientists continue to look for improving current treatments and finding a cure for Diabetes. On the other hand, monitoring enables diabetics to take care of their blood glucose. The measurement by blood glucose monitor consists in applying a drop of blood to a test strip by pricking the finger. The device measures the degree of colour changes on the test strip to determine and display the sample blood glucose concentration.

A considerable choice of blood glucose monitors is commercially offered (Figure 3). As monitoring blood glucose levels always will remain a part of life for diabetics, concepts such as minimally invasive glucose sensors or non invasive systems are being studied.

A regular blood glucose level monitoring is the key-parameter of diabetes care. It can reduce or avoid life threatening complications (DCCT/EDIC, 1993; Halimi et al., 2001).

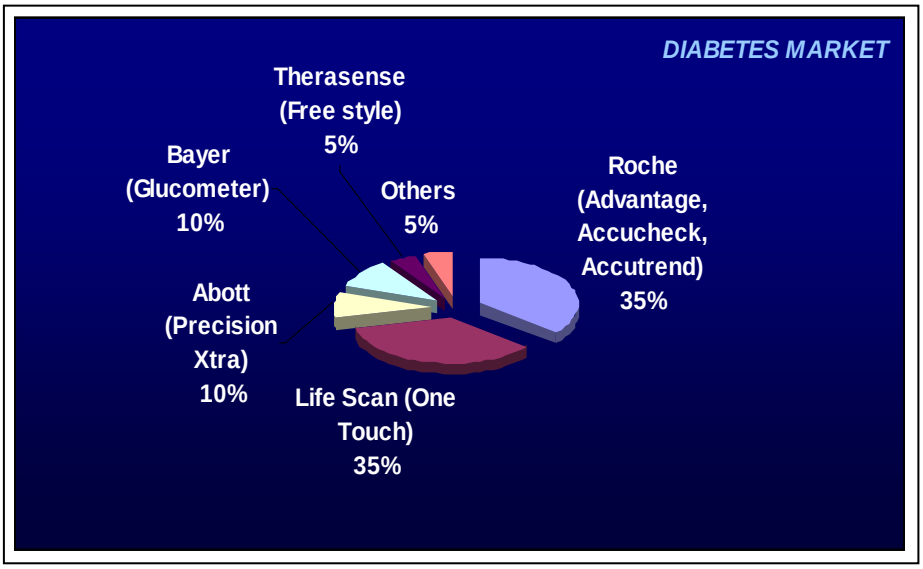


Figure 3. Diabetes market in 2008. Diabetes market represents more than 4 billion euros, which are mainly shared between Roche Diagnostics and Life Scan companies.

With a growth rate of 11 % per year, the diabetes market is more than 4 billion euros. This market is mainly shared between Roche Diagnostics and Life Scan as it is shown in figure 3. Roche diagnostics allots 1/3 of its activity to this market. About 4100 employees worked out products such as Advantage®, Accutrend® or Accu-Chek®, family of blood glucose meters, test strips, lancing and data management systems. The whole development and production of enzymes used for diabetes care platform is located in Penzberg, Germany.

3. The HbA1c measurement

3.1 HbA1c measurement and diabetes care

In diabetes management, glycemic control is correctly assessed by a combination between the self-monitoring blood glucose and the HbA1c test, which are complementary. Self-monitoring of blood glucose is the best way to manage oneself its own insulin injection according to meals or exercises. It reduces thus the hypoglycemia risks in different daily life conditions and gives more freedom and flexibility with daily activities and exercises.

The HbA1c test is a clinical blood glucose control from the previous 2-3 months (Halimi et al., 2001; Nakamoto, 2004).

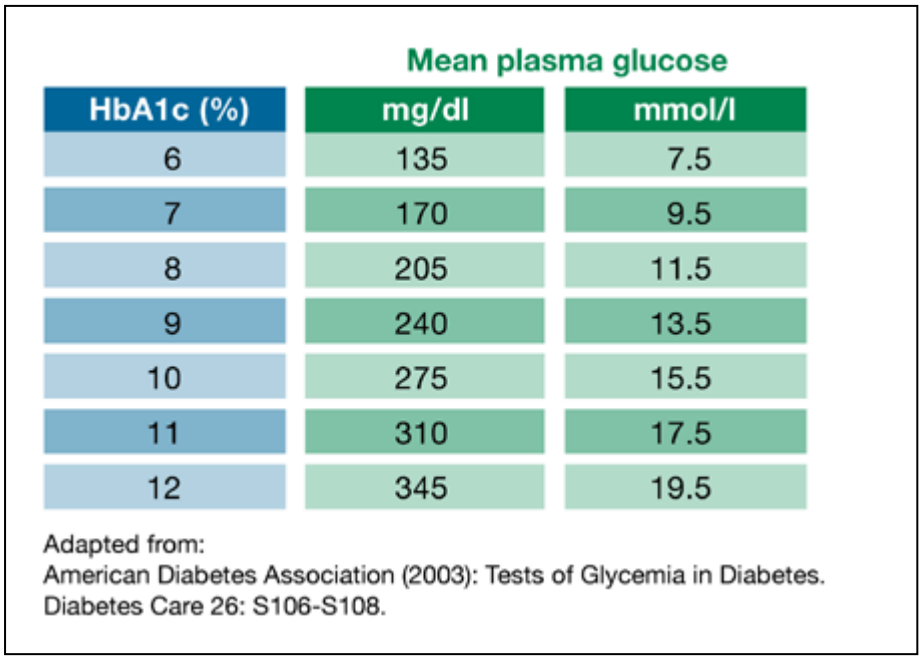
By its measurement of the glycated hemoglobin (HbA1c), it is an official long-term key parameter for diabetes management. This test is necessary for the following-up of long-term complications of the diabetes disease as well as the short-term risk of life-threatening hypoglycemia according to the DCCT (DCCT/EDIC, 1993; Mosca, 2003). This assay determines whether treatment is adequate and to guide treatment adjustments (Nathan et al., 2007).

In the blood, glucose binds reversibly to the hemoglobin in the red cells. The amount of glucose is directly proportional to the blood concentration. Blood glucose assessment is by this way directly related to Diabetes mellitus (Bunn et al., 1978; Bunn et al., 1981). The glycated hemoglobin measurement, ‘the HbA1c test’, reflects the blood glucose mean during red cell life and is therefore a completely appropriate long-term parameter of blood glucose control, which prevents from diabetes long-term complications (Rahbar et al., 1968; Rahbar et al., 1969; Koenig et al., 1976; Gabbay et al., 1977; Gonen et al., 1977) Furthermore, this test has the main advantages of not being affected by punctual glycemia variability and it can be used as a glycation model of other proteins that are involved in diabetes long-term complications (Forrest et al., 1987; Derr et al., 2003).

As the red blood cell life span is 120 days, the measurement of the glycated hemoglobin should provide an assessment of average blood glucose control during these 120 days prior to the test.

Even if this measurement is heavily influenced by young red blood cells because of their greater number, the HbA1c test is an indicator of only the last 2-3 months, as blood glucose level assessment 30 days before the test determines roughly half of the HbA1c's test results (Nakamoto, 2004). In this case, a regular test of the glycated hemoglobin every three months gives a good estimation of the average blood sugar and thereby on treatment efficiency and patient´s compliance (Nathan et al., 2007).

However, the HbA1c test is not recommended for diagnostic test as it is positively correlated with the cardiovascular events. The American Diabetes Association published in 2003 the correlation between the percentage of HbA1C and the plasma glucose concentration based on the data of the DCCT (ADA, 2003; DCCT/EDIC, 1993; Figure 4). The percentage of HbA1c indicates the percentage of hemoglobin, which is linked to glucose molecules.



	Mean plasma glucose	
HbA1c (%)	mg/dl	mmol/l
6	135	7.5
7	170	9.5
8	205	11.5
9	240	13.5
10	275	15.5
11	310	17.5
12	345	19.5

Adapted from:
American Diabetes Association (2003): Tests of Glycemia in Diabetes. Diabetes Care 26: S106-S108.

Figure 4. HbA1c test. HbA1c test is a blood glucose mean indicator of the 2-3 last months. Diabetes threshold is defined around 6.5 or 7 % of HbA1c, according to the different diabetes association. Diabetics, who try to lower their HbA1c up to 7 %, have more chances to avoid long-term diabetes complications.

According to these data, the American Diabetes Association considers that a diabetic patient should reach a HbA1c percentage over 7 %, whereas other groups such as the American College of Endocrinology, the American Association of Clinical Endocrinologists, the European Association for the Study of Diabetes and the International Diabetes Federation (IDF) estimate that this value could be lowered to 6.5 % (Nakamoto, 2004). According to DCCT/EDIC findings, diabetics, which try to lower their HbA1c up to 7 %, have more chances to avoid long-term diabetes complications (DCCT/EDIC, 1999). The HbA1c threshold is not solved and is still discussed and studied today (Hadjadj et al., 2008).

The HbA1c can be influenced by anything, which affects the hemoglobin or its life-span. Anemia, blood loss, sickle cell disease, chronic kidney disease or genetic abnormality of hemoglobin can influence HbA1c results. High levels of vitamins C or E can lower HbA1c results whereas alcoholism, high aspirin quantities, chronic use of opiate-containing drugs, high levels of triglyerides and high levels of bilirubin can elevate them (Nakamoto, 2004).

3.2 Hemoglobin glycation

Diabetes is mainly characterized by blood hyperglycemia. This hyperglycemia is induced by the glycosylation of one major component of red cells, the hemoglobin. This glycosylation of the hemoglobin was revealed and studied in diabetic patients in 1968 for the first time (Rahbar, 1968). Glycosylation can be catalyzed on proteins by enzymes in the case of posttranslational protein modification. These proteins play a major role in the maintenance of membrane integrity or in protein secretion to the extracellular spaces. Proteins can also undergo glycosylation by a chemical reaction, which does not require physiologic incubation conditions but is simply dependent on the presence of high concentrations of free sugars. This non-enzymatic reaction is also called glycation (Nakamoto, 2004). This reaction between sugar and proteins is commonly called the Maillard or browning reaction. The Maillard reaction is a slow succession of three chemical steps (Figure 5). At first, the aldehyde group of the carbohydrate is condensed with the α-amino group of the N-terminal amino acid of proteins to form a labile Schiff base. The α-amino group of the N-terminal amino acid of the protein can also be substituted by the ε-amino group of lysine or hydroxyl-lysine residues.

The compound may hydrolyze back to glucose and protein or is stabilized into a slight reversible fructosamine derivative thanks to a ketoamine linkage by an Amadori rearrangement in approximately one-sixthieth of the cases. This product cyclises into a ring structure to form in the blood the final glycated protein (John and Lamb, 1993).

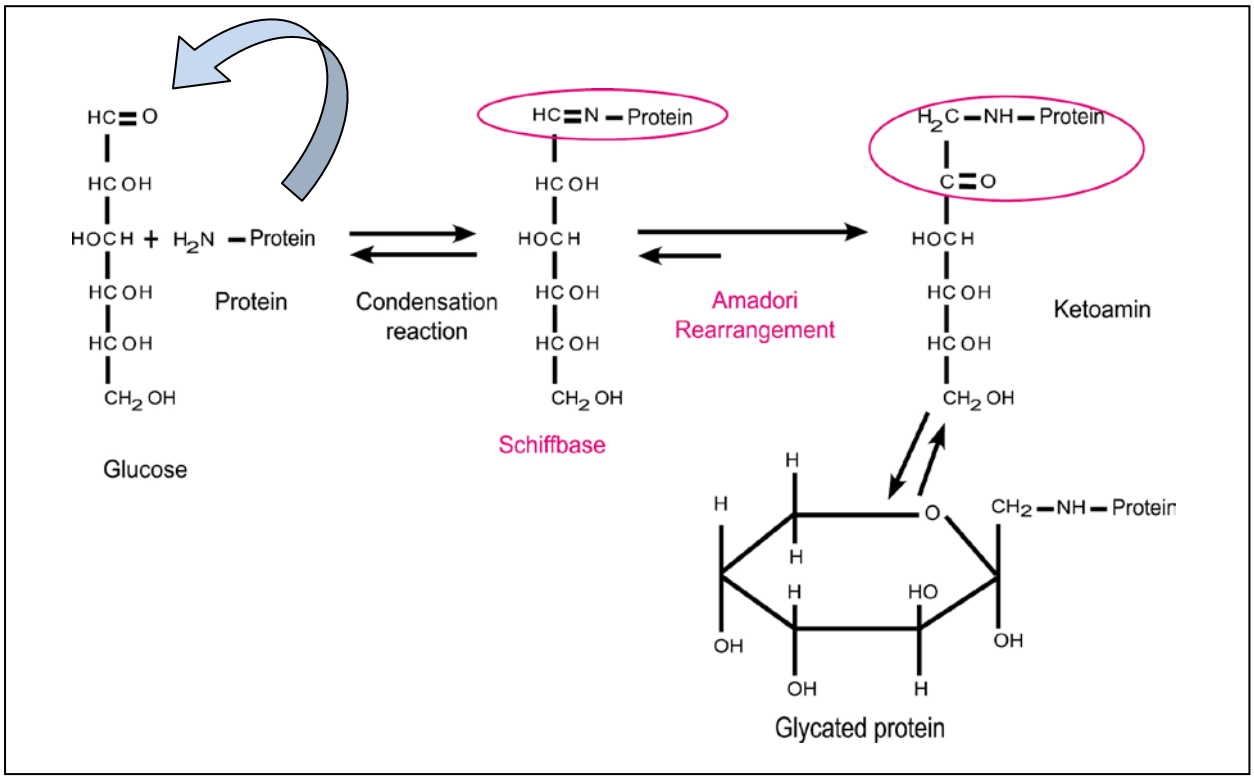


Figure 5. Maillard reaction. The Maillard reaction is a slow succession of three chemical steps. The aldehyde group of the carbohydrate is condensed with the α-amino group of the N-Terminal amino acid of proteins to form a labile Schiff base. The compound may hydrolyse back to glucose and protein or is stabilized into a slight reversible fructosamine derivative thanks to a ketoamine linkage by an Amadori rearrangement, which is cyclised into a ring structure to form the final glycated protein (John and Lamb, 1993).

In the case of the glycated hemoglobin, the Maillard reaction occurs under physiological conditions and at a specific site on the protein. This *in vivo* posttranslational non-enzymatic glycosylation takes place here between the glucose and hemoglobin (Bunn et al., 1976; Bunn et al., 1978). The glucose interacts with the N-terminal valine of the β-chain of hemoglobin. It can also take place at multiple sites at the amino-terminus of α chain and ε-amino group of several lysine residues on both α and β-chains (Rahbar, 1980).

Red blood cells travel in the human circulatory system so that oxygen and carbon dioxide can be carried between the lungs and the peripheral tissues. Oxygen and carbon dioxide bind onto the red cells thanks to their main components, hemoglobin. Adult human hemoglobin is made up of 90 % of hemoglobin A, HbA ($\alpha_2\beta_2$). The two other minor hemoglobins are the HbA2 ($\alpha_2\delta_2$) and HbF ($\alpha_2\gamma_2$), where the β-chain is substituted respectively by a delta or gamma-chain.

This tetrameric molecule associates two pairs each of alpha and beta polypeptide chains and a prosthetic heme group. The HbA1 is also divided into 4 minor components, where the HbA1c's name comes from the molecule type, which is condensed with the β-chain of hemoglobin (Bunn et al., 1976; Rahbar, 1980).

A wide range of carbohydrates (aldoses or ketoses) can participate in the glycation. Their relative interaction with the hemoglobin fluctuates to a factor 300 according to the sugar types and depends on the electronegativity of the sugar. The more electrophilic the compound is, as for example aldose, the more the reaction occurs. However, even if the glucose is one of the less reactive aldose, most glycated proteins are glucose adducts (John and Lamb, 1993).

The HbA1 nomenclature depends on the molecule's type. HbA1a1 is the glycation of the HbA1 with the fructose 1-6 diphosphate, HbA1a2 with the glucose-6-phosphate, HbA1b with the pyruvate and HbA1c with the glucose molecule, which is in majority quantity (Garrick et al., 1980; Bunn, 1981).

3.3 Techniques through the years

HbA1c was tried to be purified by different techniques over years. However, these technologies were limited by several points (Peacock, 1984).

As the HbA1c is more negatively charged as the HbA0, HbA1c was used previously to be purified by ion exchange chromatography techniques, mainly cation exchange resin such as Bio-Rex 70 (McDonald et al., 1978; Rahbar, 1980; Bunn, 1981). This method was however sensitive to experiment conditions: ionic strength, pH, the elution concentration and the temperature, which could generate very drastic variations of the HbA1c concentration.

Furthermore, it could hardly separate the HbA1c from hemoglobin variants, sickle hemoglobin or Schiff bases, which are usually formed very fast during the pre-hyperglycemia. Isoelectrofocusing electrophoresis, which was also established on the principle of the electric charge separation, could solve this problem but did not give any good accurate results (Basset et al., 1978). Other chromatography techniques, such as affinity chromatography, enabled to overcome the experiment condition dependency and can separate the HbA1c from hemoglobin variants and carbamylated hemoglobins by covalent binding of the HbA1c glucose to the boronate matrix but they could not be calibrated (Mallia et al., 1981).

High performance chromatography methods gave better separation accuracy but cost a lot of work and equipment.

A radioimmunoassay, based on the antigenic affinity of the valine glycation, was tried once but was not further improved (Javid et al., 1978). Colorimetric assays were developed according to the heme chromogenicity or chemical colorimetric reactions but they required controlled reactions, which made them fastidious and time consuming and did not enable to detect carbamylated hemoglobin, the condensation of HbA1c with urea derived cyanate (Fluckiger and Winterhalter, 1976).

All these new techniques of the early eighties gave hope that a feasible HbA1c clinical test could be implemented based on the collected data of blood glucose concentration correlation with the HbA1c assessment in diabetics (Koenig et al., 1976; Gonen et al., 1977; Bunn et al., 1978; Jovanovic and Peterson, 1981). The National Diabetes Data Group recognized for the first time the advantages of implementing a clinical HbA1c assay to measure the blood glucose concentration and recommended its standardization (Baynes et al, 1984).

3.4 Current techniques

All these techniques were improved with the science breakthrough and several methods are currently used. The DCCT implemented its study of diabetes control and long term complication by the HbA1c assessment, using an automated high-performance liquid ion exchanger chromatography based on the Bio-Rex 70 resin from Biorad (DCCT, 1987).

The Japanese and the Sweden HbA1c measurements are also performed on ion exchange HPLC, respectively the ion exchange Chromatography of Tosoh and Kyoto Daiichi and the monoS HPLC method. Today, several HbA1c tests are available on the market, based on immunoturbidimetry measurements. The first one, the Novoclone HbA1c, was introduced by Darko Diagnostics Ltd in 1991. It was followed by the Bayer DCA 2000 Point of Care HbA1c Analyser in 1992 (Goodall, 2005; John et al., 2007).

Roche Diagnostics GmbH provides several devices on the market for clinical blood glucose determination: the Cobas Integra and the Roche-Hitachi/Tina-quant families assess the HbA1c percentage by immunological detection.

The total hemoglobin is quantified by colorimetric estimation (Groche et al., 2003). Recent studies show that the Cobas-Integra family, which was previously developed by Boehringer Mannheim, is more reliable for the HbA1c measurement even it is less specific (Wang et al., 2007). Results were standardized by the National Group Standardization Program (NGSP) and the International Federation of Clinical Chemistry and laboratory medicine (IFCC). The NGSP and the IFCC are two different organizations - one American, the other one includes different laboratories worldwide - who work on the standardization of the HbA1c testing (Little et al., 2001). The NGSP is a follow-up of the American Association of Clinical Chemistry (AACC) Glycoshemogloglobin Standardization Subcommittee in collaboration with the DCCT. The NGSP is the organization, which gives the US manufacturer annual certification for the hemoglobin assay and system standardization. Both organizations standardize the HbA1C results to correlate the HbA1c percentage to the plasma glucose. The plasma glucose concentration, designated under “plasma glucose” in the equation, can be calculated in millimoles per liter (mmol/L) or in milligramms per deciliter (mg/dL) and depends on the HbA1C percentage “HbA1C (%)”. This HbA1C standardization between the DCCT and NGSP can be calculated by the following way:

- Plasma glucose (mmol/L) = 1.98 x HbA1c (%) – 4.29
- Plasma glucose (mg/dL) = 35.6 x HbA1c (%) – 77.3

HbA1c reference value = 4.8 -5.9 %

The IFCC standardization is assessed according these equations (not available in the US):
- Plasma glucose (mmol/L) = 1.73 x HbA1c (%) + 0.20
- Plasma glucose (mg/dL) = 31.2 x HbA1c (%) + 3.51
HbA1c reference value = 2.9 - 4.2 %

HbA1c reference values show hyperglycemia in the 2-3 months before analysis. HbA1c rates can increase up to 20 % in the case of a bad diabetes control. Therapeutic treatments are recommended over 8 %. Rates under references are synonyms of hypoglycaemia, which can be caused by short red blood cell life span, hemoglobin variants or short-term hypoglycaemia (Rohlfing et al., 2002; Jeffcoate et al., 2004).

Despite high technology definition, results are still limited by short red blood cell life span and hemoglobin variants, sickle and foetal hemoglobins. The advantage is given by the fact that the test is not affected by labile, acetylated or carbamylated hemoglobins.

Fortunately, due to science breakthrough, one acuter state-of-the-art technology can overcome all these measurement limits and inaccuracy: the mass spectrometry technique.

3.5 The IFCC Working Group

Today, clinical laboratories can use about 30 different analytical techniques (physical, chemical or immunological) to control blood glucose by the HbA1c measurement (Mosca et al., 2007). As there is no international uniformly agreed reference method, the International Federation of Clinical Chemistry and laboratory medicine (IFCC) decided to build a Working Group up on HbA1c standardization in 1995 to reduce interlaboratory variations. Studies show this feasibility thanks to the calibration with the same set of calibrators and/or adjusted to Designated Comparison Methods (DCM), on which are based the the NGSP, the Japanese Diabetes Society /Japanese Society of Clinical Chemistry (JDS/JSCC) and the Sweden. The main goals of the organization were:
- a mutual agreement on the definition of the HbA1c molecule
- the setting up of a Reference Laboratory Network (NRL) worldwide
- the development of a HbA1c system, in which Roche Diagnostics GmbH has been also committed, divided into three points:

- the preparation and purification of primary hemoglobin as reference material to provide a well-defined Hba0 and HbA1c mixture in large quantities
- the production of secondary reference calibrators and controls for all manufacturers
- the generation of a common Reference Method (RM) procedure for all laboratories

After first reference method developments, studies highlighted the strong existing correlation between the Reference Method of the IFCC, the Designated Comparison Methods of the NGSP, the JDS/JSCC and the Sweden (Hoelzel et al., 2004; Goodall, 2005; Miedema, 2005).

In the approved IFCC reference method measurement procedure, the HbA1c assay can be determined by Reverse Phase High-Performance Liquid Chromatography coupled to Mass Spectrometry by Electrospray Ionisation (RP-HPLC/MS-ESI) as it is performed at Roche Diagnostics (Saudek et al., 2006).

3.5.1 HbA1c reference material definition

The IFCC had to clarify the definition of the material: the HbA1c at first to standardize the HbA1c test. The HbA1c is defined as hemoglobin that is irreversibly glycated at one or both amino terminal valines of the β-chains (β-N-1-deoxyfructosyl-hemoglobin). It excludes hemoglobin, which is glycated on lysine residue of the β-chain. Furthermore, if the hemoglobin is glycated only at a lysine, it cannot be considered as HbA1c. The IFCC succeeded in preparing primary reference material: purified HbA0 and HbA1c for the reference method calibration with 98.5 and 99.5% of purity respectively (Kobold et al., 1997; Finke et al., 1998).

The Designated Comparison Methods (DCM) is the relationship between the IFCC and the NGSP ones. It can be calculated by an equation (Hoelzel et al., 2004). The HbA1c percentage defined by the NGSP is correlated to the HbA1C measurement of the IFCC by the following correcting factor:

NGSP HbA1c = 0.915 (IFCC HbA1c) + 2.15 %

The term of glycated hemoglobin was suggested to be expressed as “hemoglobin beta chain (blood)-N-(1deoxyfructos-1-yl) hemoglobin beta chain” and shorten as DOF-Hb in everyday speech (Nordin and Dybkaer, 2007). The term was officially accepted by the IFCC WG-HbA1c whereas the abbreviation was rejected. Current HbA1c refer to methods aligned to the U.S. National glycohemoglobin Standardization Program and as it is recommended by American Diabetes Association (Panteghini et al., 2007).

A new study was started in 2004 by the International Diabetes Federation, the Europeans Association for the Study of Diabetes (EASD), the American Diabetes Association, the IFCC Working Group and the NGSP to investigate and hopefully establish a relationship between HbA1c and the “average plasma glucose”: The Mean Glucose Study (Sacks, 2005).

A consensus agreement emerged in 2007 as it is described on the figure 6.

	Current HbA1c	IFCC traceable methods HbA1c
Reference interval (non-diabetics)	4 - 6 %	20 – 42 mmol/mol
Target for treatment in diabetics	< 7 %	< 53 mmol/mol
Change of therapy in diabetics	> 8 %	> 64 mmol/mol

Figure 6. Consensus agreement of a relationship between HbA1c and the “average plasma glucose between the International Diabetes Federation, the Europeans Association for the Study of Diabetes (EASD), the American Diabetes Association, the IFCC Working Group and the NGSP. Units and target values were suggested for HbA1c when measured with methods traceable to the IFCC reference system. A comparison with the current figures is given.

3.5.2 Method standardization for the HbA1c measurement

To achieve a uniform international standardization, the IFCC has established at the end of the 90ies a Working Group on HbA1c standardization. The Working Group has developed a reference system for the international standardization (http://www.ifcchba1c.net/; DCCT, 1987; Tran et al., 2003).

For the method standardization, the IFCC Working Group decided to test two available clinical techniques: the Reverse Phase High Performance Liquid Chromatography coupled with mass spectrometry by Electrospray Ionization (RP-HPLC/ESI-MS) and the Capillary Electrophoresis (CE). Both methods can be used as reference method for the HbA1c test. Simple and robust, the capillary electrophoresis has the drawback to require an off-line preparation of the sample on a reverse phase chromatography pre-column. The on-line coupling of the ESI-MS enables to reduce time in the sample preparation and even if the equipment is more expensive, it offers a more accurate specificity with low limit detection (Kobold et al., 1997). Both techniques eliminate the results deviation due to carbamylated and acetylated hemoglobins. Hemoglobin variants did not interfere in the results thanks to the enzyme specificity cleavage in the sample preparation unlike immunoassay or affinity chromatography methods (Jeppsson et al., 2002; Manley et al., 2006).

3.6 Roche Diagnostics GmbH commitment

The R&D rare reagents chemistry department of Roche Diagnostics has been committed to collaborating with the IFCC Working Group since 1999 for the HbA1c reference standardization. The HbA1c test standardization has been developed in the laboratory of Dr. Uwe Kobold by RP-HPLC/ESI-MS measurement (Kobold et al., 1997; Finke et al., 1998; Jeppsson et al., 2002).

3.6.1 Electrospray Ionization - Mass Spectrometry (ESI-MS)

The mass spectrometry, such as ESI-MS, assesses the purity, sequence or molecular weight of biomolecules with the accuracy of millidaltons. Very sensitive technique, it allows to determine molecular weight of 100 000 Da or higher, using only a few picomoles of material (Bakhtiar and Nelson, 2000). The accuracy of the molecular weight is usually better than 0.01 % of the calculated mass whereas the SDS-PAGE method has an accuracy of 5 to 10 %, which can reach 50 % by the presence of lipid or carbohydrate in the initial solution (Carr and Annan, 2001). The accuracy of the measurement enables to observe posttranslational modifications or proteolytic processing.

The principle is the generation of multiple charged or protonated ions via an electrostatic field, which transfers the ions from the soluble form into the gas phase under atmospheric pressure (Chapman, 1996). All mass spectrometers are composed of three main components: an ion source, a mass analyzer and a detector (Figure 7). The sample is injected in a thin capillary under atmospheric pressure and then droplets are created, at the tip of the capillary by formation of a Taylor cone. The sample in the droplets is then ionized by applying a strong electric field, typically around 3.5 kV, separating the positive from the negative charges. The droplets exit the capillary as a spray and travel at atmospheric pressure down an electrical gradient to the gas conducting tube (Bakhtiar et al., 2000; Carr and Annan, 2001). Two theories explain the ion formation: either the ions are expelled directly from the droplets by successive fissions, which is called the charge residue model, or the ion evaporation system, where single ionized sample molecules remains after solvent evaporation and droplet fragmentation (Dole et al., 1968; Dole et al., 1970; Iribarne and Thomson, 1975; Nguyen and Fenn, 2007). Single or multiple charged ions are then transported in the gas phase through different vacuum stages to the mass analyzer (Manisalia, 2006). Quadrupole analyzers, such as ion traps or time of flight analyzers, are commonly used for ESI-MS. Ions are there separated according their mass to charge (m/z) ratios by the application of electric fields. Separate ions are finally conducted to the detector, which is for example an electron multiplier.

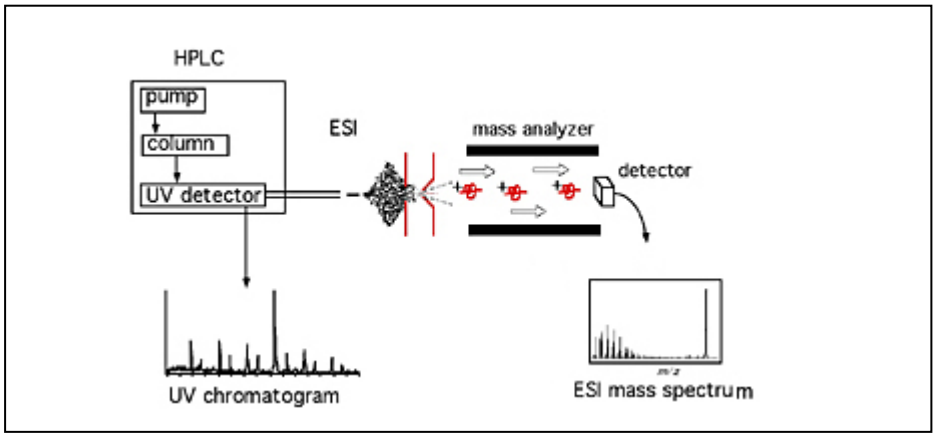


Figure 7. RP-HPLC Electrospay Ionization Mass Spectrometry. Molecules are first purified by RP-HPLC. They are eluted by increasing apolarity before arriving in a thin capillary under atmospheric pressure, where fine droplets are created at the end of the tip. By applying a strong electric field, positive and negative charges are then separated. Single or multiple charged molecules are then transported as gas phase through different vacuum stages to the mass analyzer. Ions are there separated according their mass to charge (m/z) ratios by the application of an electric or a magnetic field. Separate ions are finally conducted to the detector, where they can be more easily detected by an electron multiplier. The mass spectrum is the plot of ion abundance versus the mass to charge ratio.

The mass spectrum is the plot of ion abundance versus the mass to charge ratio. The ESI-MS technique is a very sensitive analytical method to measure molecular weights inferior than 150 kDa with an accuracy of about 0.005 % - 0.01 % to 50 kDa.

However, this technique has a low tolerance (≤ 20 mM) for non volatile buffers, alk ali metal salts and detergents (Bakhtiar et al., 2000). If necessary, samples must be purified to get free of salt and detergent as the quality of the data is very dependent on the type and the concentration of the components present in the sample. The ideal sample solution is a mixture of highly pure water and volatile organic solvents containing low level of volatile acid or base (water, acetonitrile and trifluoroacetic acid at a concentration of 0.05 % to 0.1 % by volume). The organic solvent lowers thus the liquid surface tension to facilitate the evaporation. These mixtures have the advantage to separate easily peptides and proteins when the mass spectrometer is coupled on-line to some reverse phase high performance liquid chromatography (RP-HPLC) for a necessary pre-cleaning of the sample. If some buffers are required for protein solubility or enzymatic digestion, a volatile buffer, such as ammonium bicarbonate or ammonium acetate at a concentration inferior than 30 mM, is recommended so that the formed droplets contains less solvent and ionization is more efficient (Bakhtiar et al., 2000; Manisalia, 2006).

3.6.2 The HbA1c digestion standardization for the ESI-MS

The HbA1c test is divided into three steps. The patient blood sample is first prepared for the isolation and hemolysis of erythrocytes. Human erythrocytes are separated, washed and hemolysed in water and kept in storage buffer solution. Then, the second preparative step of this HbA1c measurement by ESI-MS consists into a proteolytic digestion of the HbA1c. Different enzymatic digestion possibilities were studied for the most appropriate cleavage of the glycated hemoglobin. This is for example, the trypsin, E.C 3.4.21.4, which cleaves peptide bonds of the lysine and arginine at the carboxy-terminal side. Commonly used for peptide mapping in hemoglobinopathies, it cleaves specifically the lysine residues at the 8^{th} amino acid of the β-chain hemoglobin. However, the trypsin cannot be used to avoid the quantification of double or single glycated hemoglobins at the Lys-8 position. On the contrary, the V8-Glutamyl endopeptidase (V8-GluC) of *Staphylococcus aureus*, EC 3.4.21.19 hydrolyzes specifically after the sixth amino acid of the HbA1c β-chain, a glutamic acid.

By this way, it generates a hexapeptide, which is glycated on its first amino acid, a valine (Houmard and Drapeau, 1972; Figure 8).

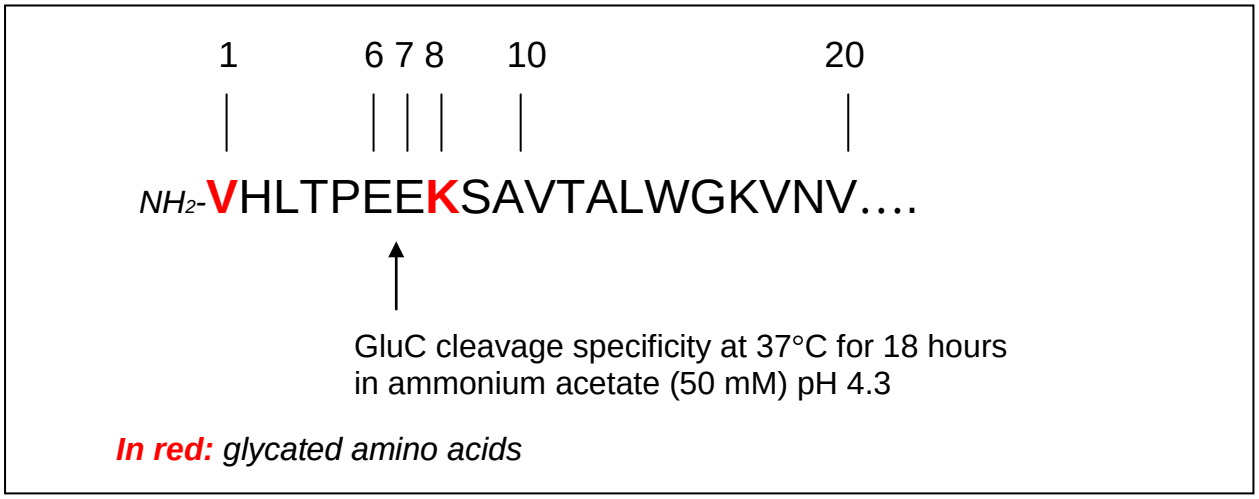


Figure 8. GluC specificity cleavage on the β-chain of the glycated hemoglobin (HbA1c). The V8-Glutamyl endopeptidase (GluC) of *Staphylococcus aureus*, EC 3.4.21.19 hydrolyzes specifically after the sixth amino acid of the HbA1c β-chain, a glutamic acid. This hydrolysis generates an hexapeptide, which is glycated on its first amino acid (valine).

As the 6th amino acid can be easily cleaved in mild denaturing conditions at pH 4.3, the V8-Glutamyl endopeptidase is the most appropriate digestion enzyme for the HbA1 test sample preparation (Kobold et al., 1997). After studying the digestion kinetic over a period of 40 hour, digestion conditions were optimized to 18 hours at 37°C.

After that the hemoglobin solution is treated with the V8-Glutamyl endopeptidase in order to split the molecule in peptides among them the glycated (HbA1c) and non-glycated N-terminal hexapeptides (HbA0) of the ß- chains, hexapeptides are then anayzed by RP-HPLC/MS-ESI. The HbA1c percentage is assessed from the glycated hexapeptides compared to the non-glycated ones and compared to a reference standard. This reference standard consists in a mixture of pure glycated and non-glycated hexapeptides of a non-diabetic person, whose HbA1c percentage was beforehand defined.

The proteolytic step is critical for the HbA1c test and must be performed in very specific conditions, which were optimized by Jeppsson et al. (2002) after a study if the digestion kinetic over a period of 40 hours. The V8-GluC must be incubated with the hemoglobin sample at 37°C for 18 hours in a proportion of 1:100 (enzyme:hemoglobin) in ammonium acetate buffer at pH 4.3.

4. Glycated hemoglobin cleavage by V8-Glutamyl endopeptidases

4.1 The V8-Glutamyl endopeptidase

4.1.1 Serine proteases

The V8-Glutamyl endopeptidase of *Staphylococus aureus* (GluC), EC 3.4.21.19, belongs to the serine protease family. According to the International Union of Biochemistry and Molecular Biology (IUBMB, 1984), proteases, peptidases or proteinases are hydrolases, which act on the peptide bond (subclass E.C 3.4.) They can be divided into two groups of enzymes: the exopeptidases and endopeptidases. Exopeptidases, E.C 3.4.11-19, cleave peptide bonds so that amino acids at the amino or carboxy terminal side of the protein are released. Endopeptidases, E.C 3.4.21-99, cleave internal bonds in the polypeptide chains.

Peptidases can be divided into five families: serine, threonine, cysteine, aspartic and metallopeptidases. Serine, threonine and cysteine peptidases differ by their catalytic action from the aspartic and metallopeptidases. The nucleophile of the catalytic site is part of amino acids in the three first groups, whereas there is the interaction of an activated water molecule in the two other peptidase groups (Rawlings et al., 2004).

Serine peptidases are characterised by the nucleophilic attack of the hydroxyl group of the serine of the catalytic triad directly on the peptide bond (Figure 9).

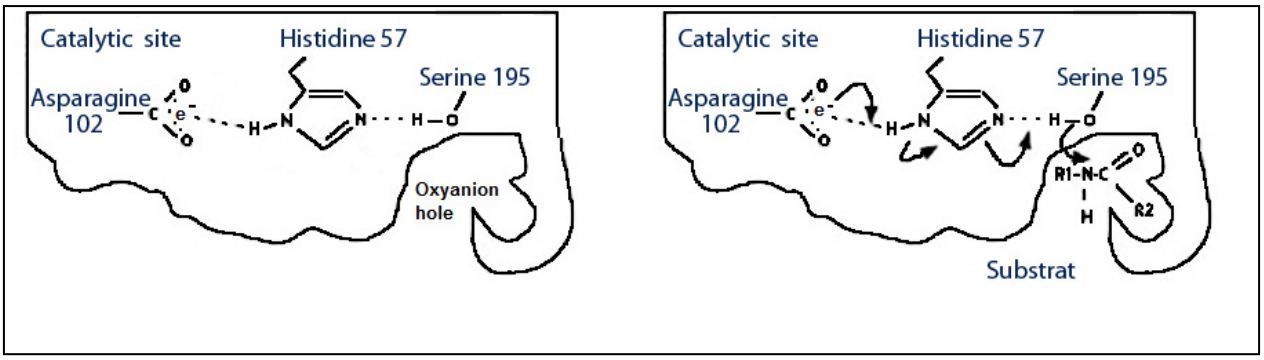


Figure 9. Catalytic mechanism of serine proteases and glutamyl endoproteases (Hedstrom, 2002). The catalytic triad (an aspartic acid, a serine and a histidine) forms an oxyanion binding site or oxyanion hole, which acts as a charge relay system for the proton transfer. Serine peptidases are characterized by the nucleophilic attack of the hydroxyl group of the serine on the peptide bond.

The serine protease group consists in around fifty families. The most well-known families are the trypsin and chymotrypsin, the elastase and the subtilisin ones. However, nine clans can be distinguished according to the tertiary structure and the order of catalytic residues in the sequences. In six of these clans, the catalytic reaction involves a proton donor, which is a histidine residue and where the third residue (usually an aspartate or exceptionally, a histidine for one of the clans) helps in the orientation of the imidazolium ring of the histidine (Ishida and Kato, 2004). In the three other clans, there is no triad but a dyad, where a lysine plays the role of the proton donor.

The catalytic triad site is the association of three amino acids: an aspartic acid, a serine and a histidine. The catalytic triad forms an oxyanion binding site or oxyanion hole. It acts as a charge relay system because of the proton transfer (Kraut, 1977).

Each amino acid has its own role. The serine acts as a nucleophile because of its hydroxyl group, which attacks the carbonyl group of the peptide bond of the substrate. The histidine behaves as a proton donor and enables the formation of the transitional tetrahedral intermediate. The carboxyl group of the aspartic acid strengthens the electrostatic effect to stabilize the protonated histidine and fixes the configuration around the active site to support the catalytic reaction (Ishida, 2006).

The serine protease mechanism is based on a symmetrical mechanism of acylation and deacylation through an oxyanionic intermediate formation (Figure 10). The polypeptide substrate binds onto the serine protease surface and the scissile amino acid bond is inserted into the catalytic site of the enzyme. The P1/S1 interaction nomenclature is described in the way that the active site of proteases is divided into subsites. According to the nomenclature of Schechter & Berger from 1967, S1, S2... are the subsites of the active site and P1, P2... the residues of the substrate occupying the corresponding subsites (Schechter and Berger, 1967).

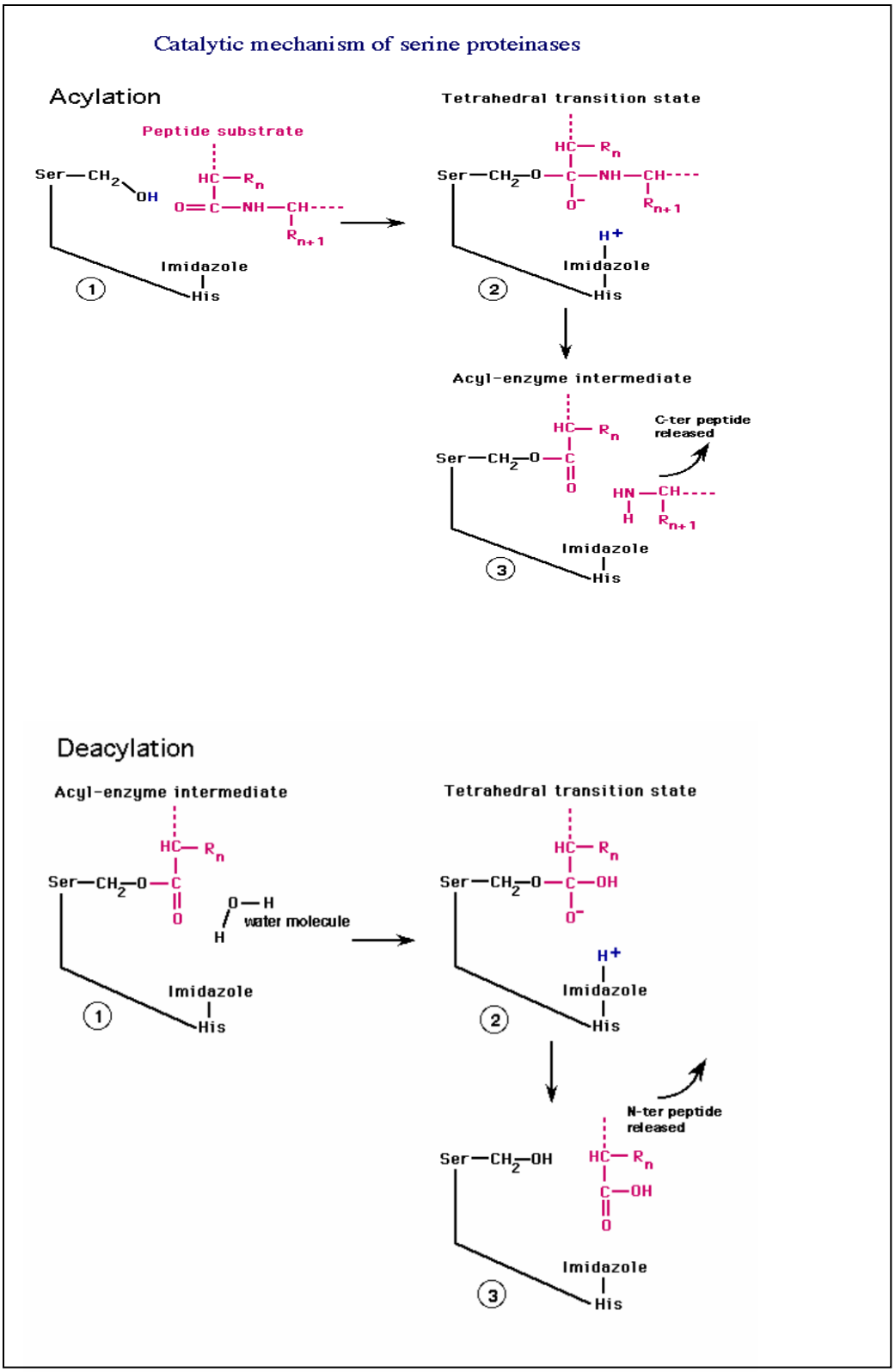


Figure 10. Serine protease reaction. The serine of the catalytic triad acts as a nucleophile because of its hydroxyl group, which attacks the carbonyl group of the peptide bond of the substrate. The histidine behaves as a proton donor and enables the formation of the transitional tetrahedral intermediate. The aspartate helps in the orientation of the imidazolium ring of the histidine.

The hydroxyl group of the serine of the catalytic triad attacks the carbonyl bond of the peptide bond. The nitrogen of the histidine accepts the hydroxyl of the serine and a pair of electron from the double bond of the carbonyl oxygen moves to the oxygen of the P1 residue. This generates a transitional oxyanionic tetrahedral intermediate. The P1 residue of the polypeptide substrate is released by break of the peptide bond and of the stationary state with the histidine. The electrons move from the negative oxygen, which reforms the carbonyl bond of the P1 residue, creating an acyl enzyme intermediate.

The deacylation of the symmetrical mechanism occurs in the presence of a water molecule. The acyl enzyme intermediate is hydrolyzed by a molecule of water to release the P1 residue peptide and to restore the hydrolyzed group of the serine. The molecule of water attacks the nucleophile (carbonyl) instead of the serine. Water replaces the amino terminal side of the cleaved peptide on the histidine residue and attacks the carbonyl carbon. By transfer of the electrons once again from the oxygen to the carbonyl of the P1 residue and stabilization by the protonated histidine, the transitional oxyanionic tetrahedral intermediate is again formed. The bond between the hydroxyl of the serine and the carbonyl carbon of the P1 residue attacks the hydrogen of the protonated histidine. The double bond of the carbonyl of the residue P1 is reformed and by this way, it releases the residue P1 of the polypeptide (Hunkapiller, 1976).

4.1.2 Different V8-Glutamyl endopeptidases

Several organisms are able to produce the glutamyl endopeptidase enzyme. The most well-known caracterized glutamyl endopeptidases come from *Staphylococcus* or *Bacillus* strains: *S. aureus*, *epidermis* or *warnerii* and *B. sutilis* or *licheniformis*, which are classified in the Merops data base (Rawlings et al., 2008). The V8 notation in the “V8-Glutamyl endopeptidase” name means that the Glutamyl endopeptidase comes from the *S. aureus* V8 strain.

Two isoforms of the V8-Glutamyl endopeptidases exist according to *S. aureus* V8 strains. A V8-GluC was first sequenced in 1978 by Drapeau, who was used to work on different *S. aureus* V8 strains (Drapeau, 1978). A second sequence was discovered by Carmona in 1987 in the wild-type *S. aureus* V8 strain (Carmona and Gray, 1987).

These two sequences differ by eleven amino acids, which are highlighted on the figure 11. The second sequence can be found in swissprot database under the "SSPA_STAU" entry name. The V8-GluC from Drapeau is bought by Roche to produce a commercial V8-GluC lyophilisate. This V8-GluC is directly purified from the *S. aureus* mutant strain and lyophilisated as it is provided in the same way on the market by other companies such as Sigma-Aldrich. The New England Biolabs company sells the V8-GluC, whose sequence was described by Carmona, as a recombinant protein expressed in *B. subtilis*.

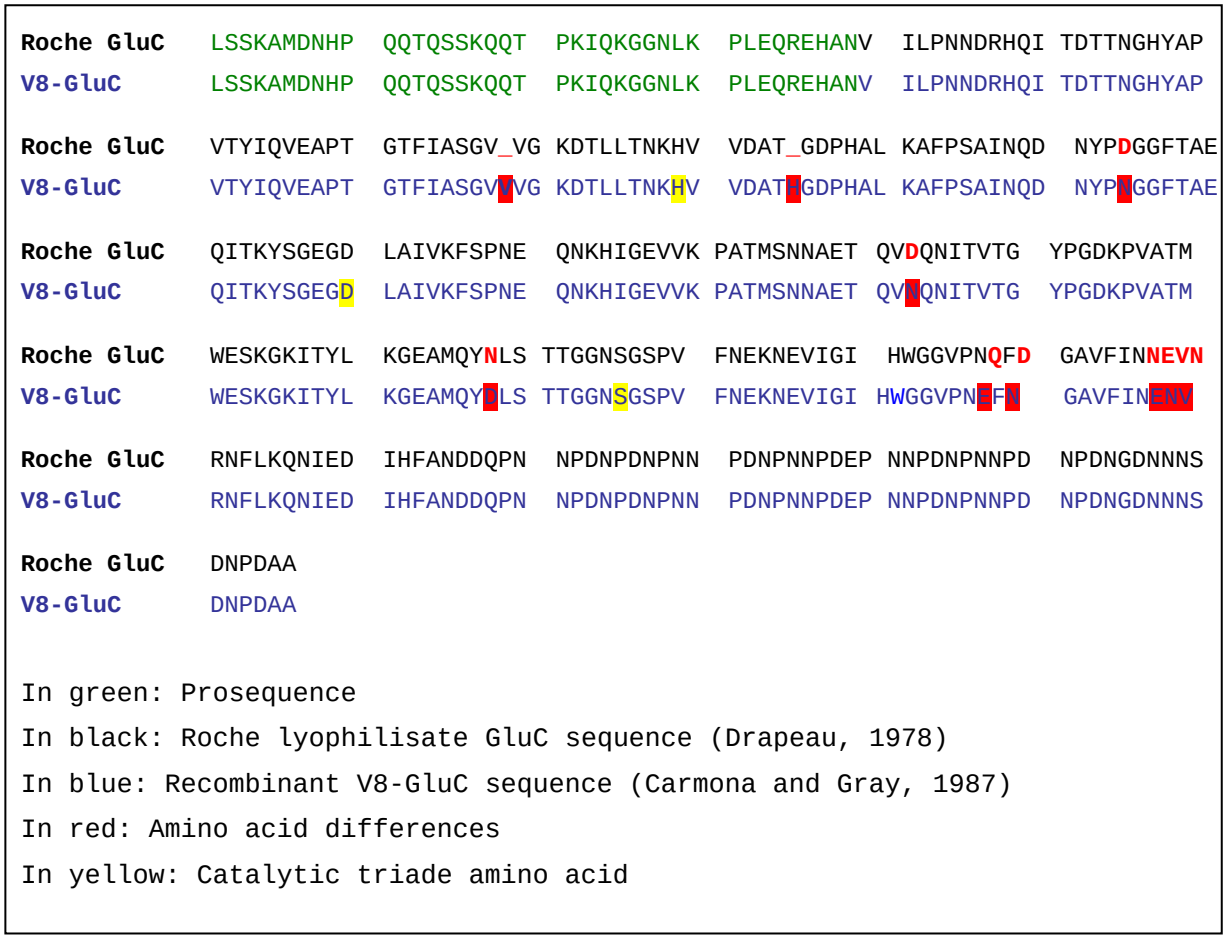


Figure 11. V8-Glutamyl endopeptidase sequences. A first V8-GluC was sequenced in 1978 by Drapeau. The second sequence, which differs by eleven amino acids, was discovered by Carmona in 1987.

The GluC from Drapeau, which is commercialized by Roche, will be abbreviated: "Roche GluC" and the V8-GluC from Carmona, which was used for cloning, will be called "V8-GluC" for the experimental part of this report. The term V8-glutamyl endopeptidase designates glutamyl endopeptidases, in general, from *S. aureus*. The GluC of New England Biolabs is abbreviated as the "NEB GluC" (Figure 12).

Sequence	Company	Production form	Production	Designation
Drapeau	Roche	Lyophilisate	Direct purification from *S. aureus*	Roche GluC
Carmona	NEB	Lyophilisate	Recombinant in *B. subtilis*	NEB GluC
Carmona	Roche (phD)	Liquid	Recombinant in *B. subtilis*	V8-GluC

Figure 12. GluC designation. The GluC at Roche from Drapeau is abbreviated: "Roche GluC" and the GluC from Carmona, which is commercialized by NEB, is abbreviated as the "NEB GluC". The term "V8-GluC" designs the recombinant GluC, which comes from Carmona's sequence and was used for this phD project.

The Roche V8-GluC is normally used for common mass spectrometry sample preparation as well as for the hemoglobin functional test of the IFCC. However, the V8-GluC from Carmona is the only available GluC sequence from the V8 *Staphylococcus aureus* strain. Glutamyl endopeptidases from other *Staphylococcus* and *Bacillus* strains could not be investigated because of patent application, which are held by companies such as Norvo Nordisk for the *Bacillus subtilis* and *licheniformis* glutamyl endopeptidases.

4.1.3 V8-Glutamyl endopeptidase characteristics

The V8-Glutamyl endopeptidase of *Staphylococus aureus* (GluC), EC 3.4.21.19, belongs to the serine protease family, which is characterized by its common catalytic triad: Histidine 51 - Serine 169 - Aspartic acid 93 (Figure 13).

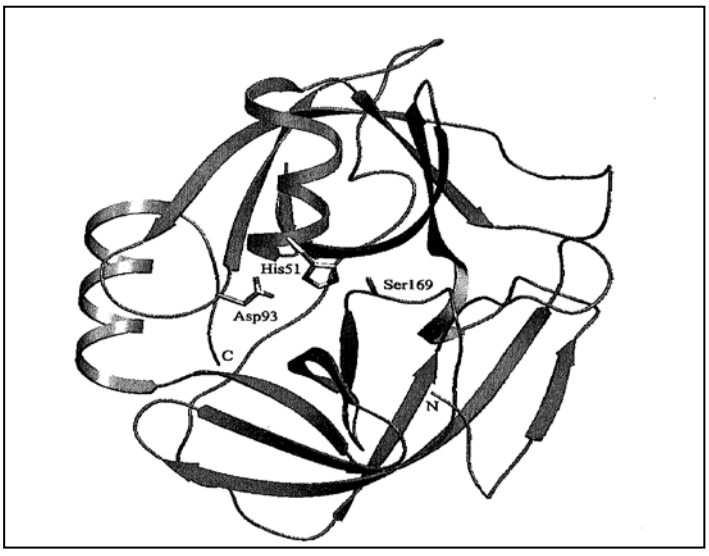
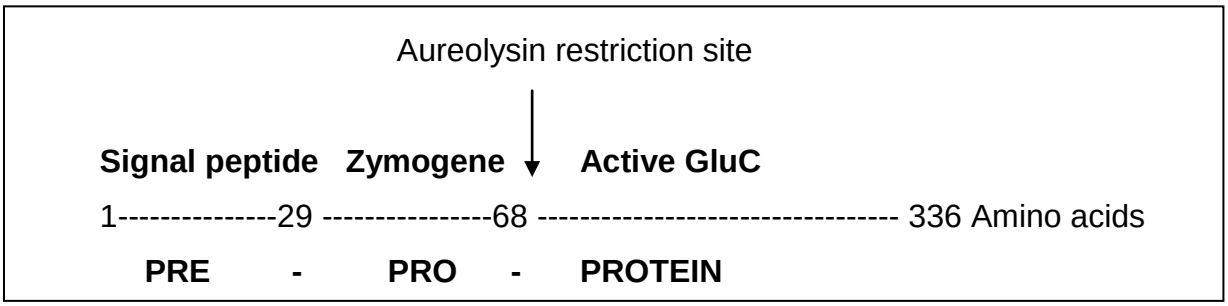


Figure 13. Catalytic triad of the V8-GluC (Prasad et al., 2004). The catalytic triad of the V8-glutamyl endopeptidase (aspartic acid 93, serine 169and histidine 51) forms an oxyanion binding site or oxyanion hole, which acts as a charge relay system for the proton transfer.

This 268 amino acid monomer preproprotein is extended on its amino terminal side by an additional sequence of 39 amino acids, which is called prosequence or zymogene. This prosequence, by its presence, inactivates the protease after its synthesis and prevents from its own autoproteolysis. This prosequence can be cleaved by another *Staphylococcus aureus* protease: Aureolysin, EC 3.4.24.29 (Nickerson, 2007; Figure 14). The V8-GluC has neither disulfide bonds nor other posttranslational modifications.



Chain sequence	Amino acids (From	To)	Length (AA)	Description	Molecular weight (Kda)
Signal peptide	1	29	29	Potential	3
Propeptide	30	68	39	Zymogene	4
Active GluC sequence	69	336	268	GluC	29
Total protein	1	336	336	Complete protein	36

Figure 14. General structure of the *Staphylococcus aureus* GluC SSPA_STAU (http://www.uniprot.org/). This 268 amino acid V8-GluC protein is extended on its amino terminal side by an additional prosequence of 39 amino acids. This prosequence, which can be cleaved by another *Staphylococcus aureus* protease, aureolysin, inactivates the protease after its synthesis in order to prevent from its own autoproteolysis.

The V8-Glutamyl endopeptidase is described in the Merops protease data base (Rawlings, 2008). This V8-GluC endopeptidase, which is encoded from the sspA gene, is known under several common synonyms: endoproteinase Glu-C, V8 Protease, V8 proteinase, staphylococcal serine proteinase, *Staphylococcus* strain V8 serine endopeptidase.

The name of the GluC comes from its specificity as the GluC cleaves peptide bonds on the carboxyl-terminal side of aspartic acid and more preferentially of glutamic acids. The GluC digestion is characterized by the following pattern specifity: P4-P3-P2-P1-P1'-P2', where P1 can be glutamic or aspartic acids but the glutamic acid is preferred. Aspartic acid is preferred at the P4 position, alanine or valine at the P3 position and the phenylalanine at the P2 position. Proline is disfavoured at P3, P1, P2 positions and aspartic acids at P1 position (Beynon and Bond, 2001).

The specificity depends on the incubation buffer. In ammonium bicarbonate buffer (50 mM) at pH 7.8 or ammonium acetate buffer (50 mM) at pH 4.0, the protease cleaves only Glu-P1' bonds. In phosphate buffer (50 mM) at pH 7.8, the protease cleaves Asp-P1' bonds as well as Glu-P1' bonds. If P1' is a bulky hydrophobic residue, cleavage is slowed down.

The V8-GluC is widely used in the mass spectrometry area, such as proteomics, because of its specificity and its detergent resistance.

4.2 Enzymatic activity test assay set up

4.2.1 Chromophore substrate: paranitroanilide bond

Spectrophotometry is based on the principle of the light absorption of a solution. The substance, the outcome or the product of a reaction, can be transparent or colored and usually absorbs the light in a defined wavelength. This wavelength is used for the spectrophotometric measurement.

Specific chromophore substrates must be prepared for measuring the V8-glutamyl endopeptidase activity and specificity (Houmard, 1976). Substrate preferences of the V8-glutamyl endopeptidase are well-delimited. The enzyme hydrolyzes all types of glutamate bonds but it can catalyze very slowly glutamate-aspartate, glutamate-asparagine and glutamate-proline bonds under specific experimental conditions (Deibler et al., 1982; Breddam et Medal, 1992).

Chromophores absorb UV lights from 200 to 400 nm and visible wavelengths of 400 up to 800 nm. The paranitroanilide, pNA or 4-NA, is used as chromophore. It is linked via an amide bond to the carboxyl function of the peptide. When this link is hydrolyzed by cysteine or serine proteases, the released group, which is called paranitroaniline, switches to a different absorption spectrum than the paranitroanilide bond. The colorless chromogenic substrate turns into a yellow compound and the paranitroaniline can be thus easily detected at a wavelength of 405-410 nm (Zimmerman et al., 1976; Gu et al., 1999; Figure 15). Paranitroaniline substrates are commonly used for the detection and measurement of serine and cysteine protease activities by spectrophotometer.

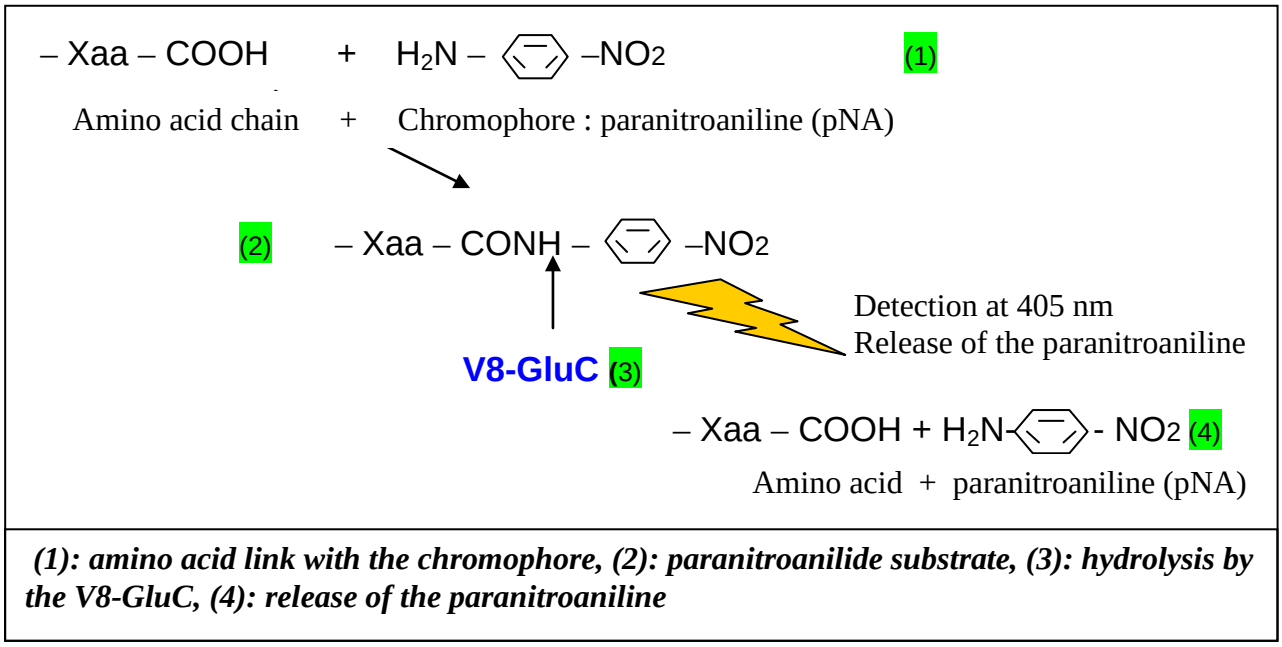


Figure 15. Reaction principle of the V8-GluC on a chromophore substrate. Paranitroaniline substrate was synthesized by addition of a chromophore on the carboxyl terminal side of the glutamic or aspartic acid. When the V8-GluC hydrolyzes the aspartate or glutamate bond, the paranitroaniline is released and can be detected at 405 nm.

4.2.2 Enzymatic characteristics

Enzyme catalysis rates are mainly affected by three important factors: temperature, pH and substrate concentration.

4.2.2.1 Temperature

The rate of any reaction increases with temperature. However, enzymes require a specific shape to work and as the temperature increases, vibrations in the enzyme molecule will interfere with its catalytic ability. At very high temperatures, the enzyme will completely denature, becoming catalytically inactive. These competing processes lead to a temperature 'optimum' for the enzyme (Figure 16).

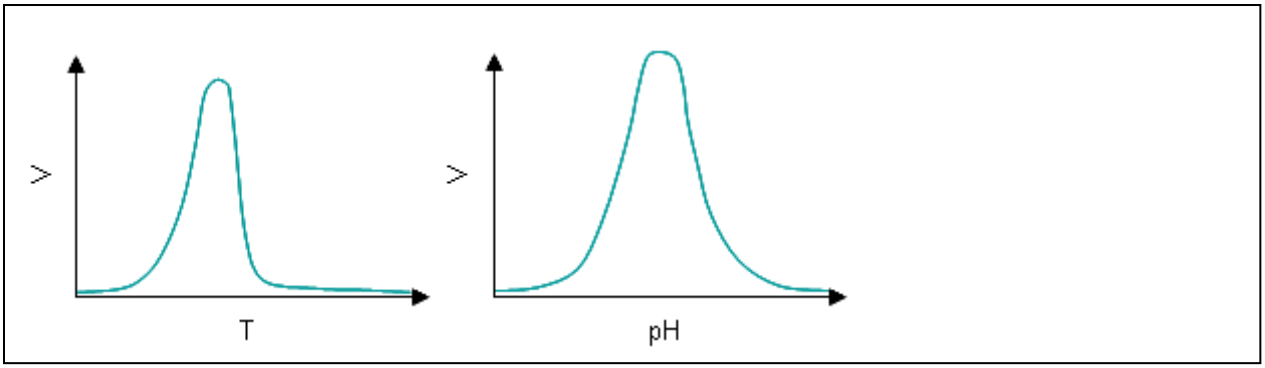


Figure 16. Influence of the temperature and the pH on the catalytic reaction of the enzyme. The speed rate of an enzyme is temperature-dependent and will be at the maximum at a defined optimum temperature. Enzyme catalysis is also affected by the pH. The pH optimum is the pH value at which an enzyme has maximum efficiency.

4.2.2.2 pH and isoelectric point

Enzymes bristle with ionisable side groups with specific values of pKa. If the active site contains a basic and an acidic amino acid that are required for catalysis, and both need to be ionised to interact with the substrate. The pH optimum is the pH value at which an enzyme has maximum potency and efficiency. The pH value can influence the enzyme conformation and it can thereby change the access of the substrate to the catalytic site. It affects by this way the activity and the stability of the enzyme.

The isoelectric point (pI), sometimes abbreviated to IEP (IsoElectric Point), is the pH value at which the molecule carries no electrical charge or the negative and positive charges are equal. Amphoteric molecules called zwitterions contain both positive and negative charges depending on the functional groups present in the molecule. The net charge on the molecule is affected by pH of their surrounding environment and can become more positively or negatively charged due to the loss or gain of protons.

4.2.2.3 Substrate concentration, affinity and Km value

Enzymes catalyze substrate by reducing the activation energy and by altering the steric constant in the Arrhenius equation. At the equilibrium, an enzyme (E) binds a substrate (S) to form an enzyme-substrate complex (E-S). The E-S complex can dissociate or irreversibly convert the substrate to a product. The Michaelis-Menten equation describes the relationship between the rate of substrate conversion by an enzyme to the concentration of the substrate (Figure 17).

In this equation, V is the rate of conversion, Vmax is the maximum rate of conversion, [S] is the substrate concentration, and Km is the Michaelis constant, the substrate concentration at which a catalized reaction proceeds at half of its maximum velocity Vmax. It enables to estimate how well the enzyme binds its substrate. Km is an apparent dissociation constant (Kd) and represents the substrate concentration at half of the maximum enzyme velocity Vmax. Therefore, a lower Km value indicates a higher affinity for the substrate. A more illustrative version of the Michaelis-Menten equation is the Lineweaver-Burk equation (Figure 17). The Lineweaver-Burk equation affords a line with a slope of Km/Vmax and y-intercept of 1/Vmax. The x-intercept, a theoretical point since 1/[S] cannot be negative, is -1/Km.

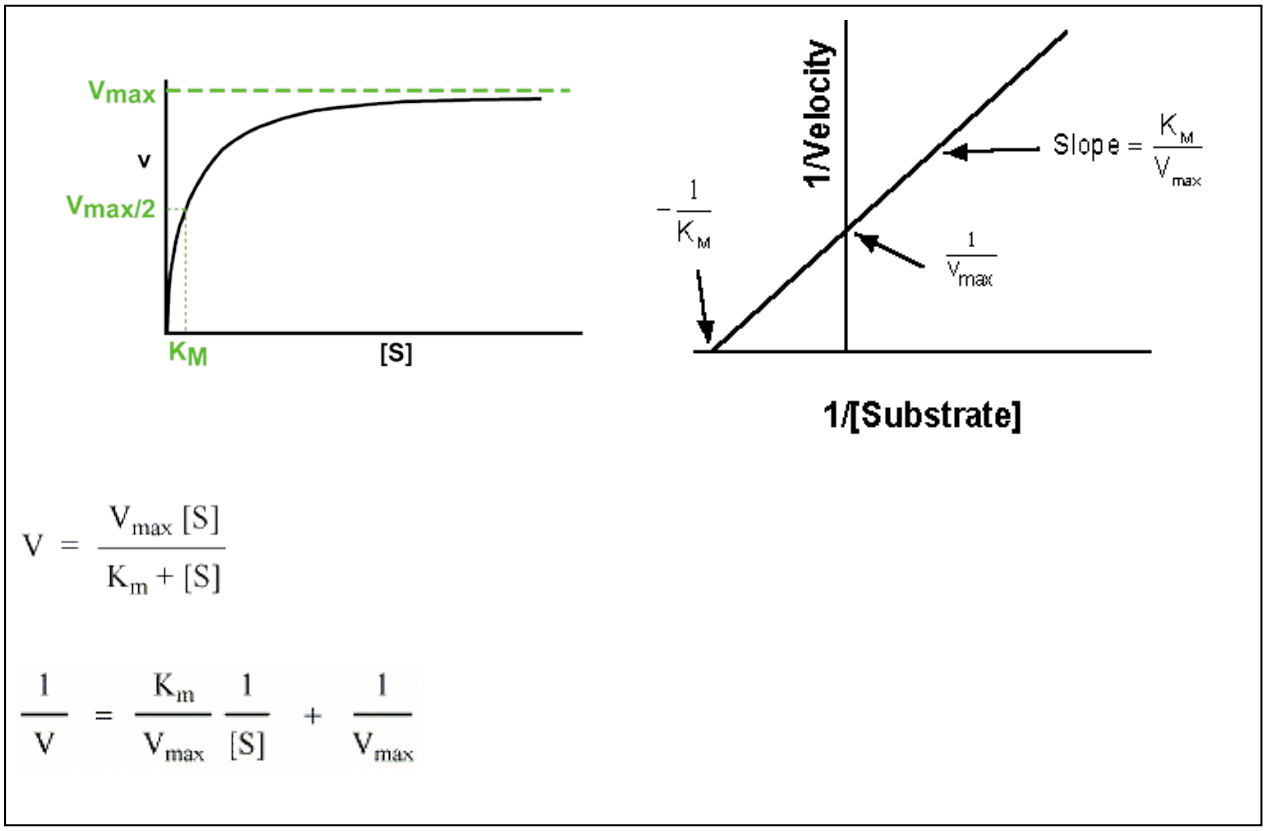


Figure 17. The Michaelis-Menten equation describes the relationship between the rate of substrate conversion, V, by an enzyme to the concentration of the substrate, S. The Km is the Michaelis constant, the substrate concentration at which a catalized reaction proceeds at half of its maximum velocity Vmax. It enables to estimate how well the enzyme binds its substrate. Km is a affinity constant of the enzyme for a particular substrate. A lower Km value indicates a higher affinity for the substrate. The Km and Vmax can be easily determined by plotting the Lineweaver-Burk equation, which is the inverse of the Michealis-menten equation. a slope of Km/Vmax and y-intercept of 1/Vmax. The x-intercept, a theoretical point since 1/[S] cannot be negative, is -1/Km.

4.3 Strategies for the cloning and expression of the V8-Glutamyl endopeptidase

4.3.1 Recombinant protein production in Bacillus subtilis

B. subtilis is one of the best well-known characterized living bacteria. Its genome has been studied by Kunst et al. since 1978 and it has become the model of the Gram positive bacteria by its metabolism and its secretion pathway systems (Fisher and Sauer, 2005; Tjalsma et al., 2004; Zweer et al., 2008). Sierro et al. even developed in 2008 a data base about *B. subtilis* transcriptional regulation. *B. subtilis* is widely used today as a cell factory for industrial enzymes and biopharmaceuticals (Westers et al., 2004).

B. subtilis is a sporulating gram positive bacterium very easy to cultivate. Furthermore, *B. subtilis* fermentation capacity has the advantage to be comparable to the *E. coli* one. A high growth rate can be reached in a simple medium. *B. subtilis* is very interesting for industrial applications as recombinant expression system thanks to its secretion pathway (Westers et al., 2004). The average yield is estimated around 0.5-1 g/L without optimization fermentation process (Terpe, 2006). *B. subtilis* is classified as S1 safe type or generally regarded as safe (GRAS) organism by the U.S Food and Drug Administration, unlike *E. coli*, which is a pathogen bacterium. *B. subtilis* is also easy to manipulate genetically. The codon usage of *B. subtilis* is less biased than the *E. coli* one (Shields and Sharp, 1987; Sharp et al., 1988). Furthermore, *B. subtilis* cells can develop genetic competence for DNA binding and uptake like *E. coli*. The optimal growth rate of wild type *B. subtilis* is 20 minutes at 37°C without plasmid. This number is decreased in the case of recombinant host. The secretion can last 3-4 days before reaching a plateau. The expression is dependent of each protein and this explains why it must be optimized for each protein, making optimization necessary.

The *B. subtilis* secretion pathway and the non-existence of outer membrane confer to *B. subtilis* a major advantage for the cloning, expression of extracellular proteins in the medium but also for the downstream process, such as purification (Young, 1980; Wong, 1995; Ling Lin et al., 2007). Overmassive production of secretory proteins can cause intracellular accumulation of protein precursors and the loss of cell viability in *E. coli* (Freudl et al., 1986).

Furthermore, *E. coli* has for drawbacks to secrete extracellular proteins only into the periplasm and proteins often aggregate into inclusion bodies, which require a special treatment for refolding (Jungbauer and Kaar, 2007; Miot and Betton, 2004; Villaverde and Carrio, 2003). The outer membrane of *E. coli* and Gram negative bacteria, in general, contain liposaccharides (LPS) or endotoxins, which need to be removed for clinical applications. In *B. subtilis*, proteins are transported extracellularly across the single cytoplasmic membrane and the thick cell wall. The translocation of proteins in *B. subtilis* is highly similar to the Sec (Secretion) machinery of *E. coli* (Freudl, 1992; van Wely et al., 2001; Yamane et al., 2004). Only few translocation steps are slightly different between Gram positive and negative bacteria by a lack of Sec factor homologues but the basic mechanistic principle remains the same as the Sec machinery in *E. coli* (Mori and Ito, 2001; Li et al., 2004; Figure 18). Other translocation machineries exist such as the TAT-pathway (Twin Arginin Translocation) or the ABC transporter (ATP-Binding Cassette) (Quentin et al., 1999; Jongbloed et al., 2000; Tjalsma et al., 2000).

1. Signal recognition particle

2. Translocation motor

3. Translocation channel

4. Signal peptidases

5. Signal peptide proteases

6. Folding factors

C. Carboxyl terminus

M. Mature protein

N. Amino terminus

R. Ribosome

SP. Signal peptide

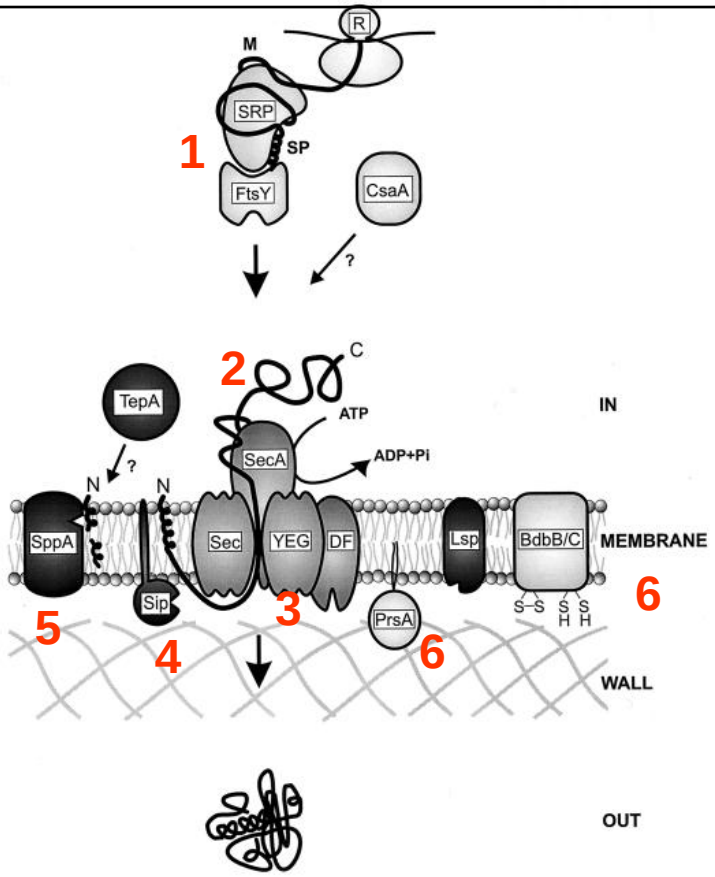


Figure 18. Sec machinery mechanisms in *B. subtlilis* (Tjalsma et al., 2000). In *B. subtilis*, the ribosome and its nascent amino acid chain forms a complex with the signal recognition particle (SRP). This complex docks onto the SRP receptor (FtsY). The SRP is then released and polypeptide translation by the ribosome is resumed. Homologues of the preprotein translocation machinery components of the *E. coli* Sec pathway have been identified in *B. subtilis*. SecA is the translocation motor and the other Sec proteins, the integral membrane proteins. First, SecA binds to acidic phospholipids and SecY is activated for recognition of SecB and the preprotein. Preprotein binding is followed by the binding and hydrolysis of ATP. The binding of ATP causes major conformational changes of SecA, leading to a release of SecB and insertion of the carboxyl terminus of SecA into the membrane. This membrane insertion, which occurs through the translocase complex promotes the translocation of a short fragment of the preprotein. Next, ATP is hydrolyzed by SecA, leading to release of the preprotein and deinsertion of SecA. Once protein translocation is initiated by SecA, further translocation is driven by both repeated cycling of SecA through ATP binding and hydrolysis and the proton motive force.

Almost all prokaryotic and eukaryotic proteins are secreted with a hydrophobic amino terminal extension, the signal peptide, which plays a crucial role in the targeting and translocation of precursor proteins across the corresponding membrane. This signal peptide is removed during or shortly after translocation by a signal peptidase. The peptidase mechanism was described by Saier et al. (1989). However, signal peptidases are specific to the organism and some of them could be identified for *B. subtilis* (Dalbey et al., 1997; Tjalsma et al., 1999). Watson compiled all known signal sequences of different organisms in 1984. This work was continued by Nielsen et al. (1997). Computational approach shows common features even if there is no common consensus sequence in signal peptides (Von Heijne, 1990). These common features were determined in *E. coli* (Izard and Kendall, 1994). The amino terminal side of signal peptides always carries positive charges. Its core is composed of a hydrophobic sequence. The signal peptide is then linked to the protein by a peptidase cleavage region.

Signal peptides from *B. subtilis* are however more hydrophobe than in *E. coli* (Zanen et al., 2005). Bolhuis (1999) defined five secretion bottlenecks for the expression of heterologue proteins in *B. subtilis*: transcription, protein folding, translocation across the plasmatic membrane, the signal peptide processing and the proteolysis. The outcome of this work showed that each raised bottleneck differs for each protein (Li et al., 2004). Even if a lot of signal peptides are available, it is thus necessary to find individually the most specific signal peptide for each protein.

The advantage of *B. subtilis* over *E. coli* is its secretion pathway, which expresses proteins extracellularly directly in the medium and the absence of endotoxins. One of its drawbacks comes from the high level of proteases, which can degrade heterologuous proteins when *B. subtilis* is used as recombinant host (Roy et al., 1986). Because of its natural life in the soil, *B. subtilis* produces natural extracellular proteases, which can interfere with recombinant protein expression (Nakamura et al., 1985; Ulmanen et al., 1986). These proteases play a major role in the “quality control” of misfolded proteins (Ling Lin et al., 2007; Harwood and Cranenburgh, 2008). Eight of them are extracellular and six other ones are bound to the cell wall (Figure 19). As these proteases can degrade recombinant proteins, several research teams try to optimize the conditions of the recombinant protein expression or to create protease deficient *B. subtilis* strains (Kawamura and Doi, 1984; Stahl and Ferrari, 1984; Wong et al., 1986). Wu et al. (1991) created a *B. subtilis* strain, WB600, deficient in 6 proteases. Any of these proteases play a vital role in the strain itself. The WB600 *B. subtilis* strain shows a residual protease activity of 0.32 % of the wild-type strain.

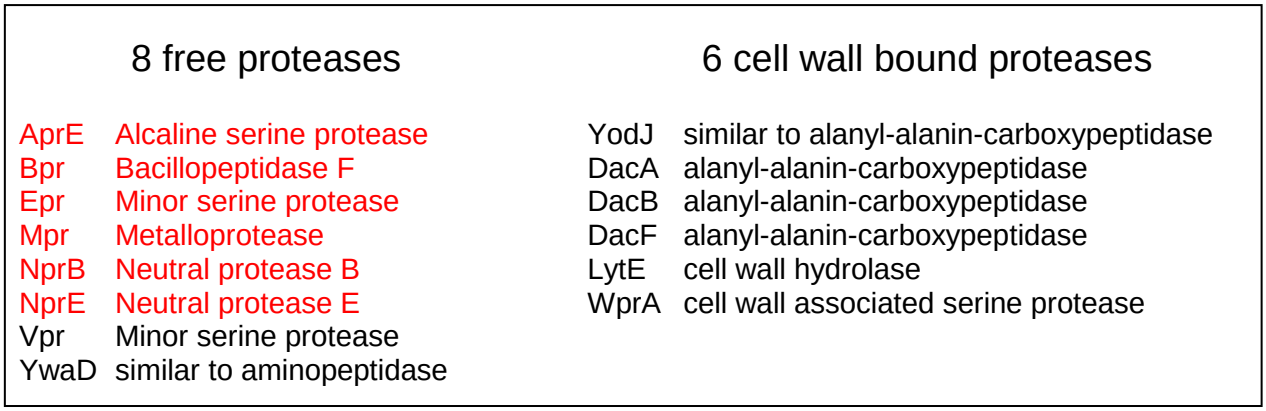


Figure 19. Extracellular proteases in *B. subtilis* WB168. *B. subtilis* produces natural extracellular, which play a major role in the "quality control" of misfolded proteins. Eight of them are extracellular and six other ones are bound to the cell wall. The knocked-out proteases in the *B. subtilis* WB600 strain are in red.

The *B. subtilis* organism was chosen as the most appropriate recombinant organism host because of its property to express proteins extracellularly. The V8-GluC could not be expressed in other "normally-used" recombinant organisms, such as *E. coli,* because of its toxicity. However, the V8-Glutamyl endopeptidase was already successfully expressed in *E. coli* thanks to a special plasmid construct (Yabuta et al., 1995). Last recent work on a glutamyl endopeptidase from *Staphylococcus epidermis* proved the feasibility of recombinant expression in *E. coli* (Nemoto et al., 2008). As other glutamic acid endopeptidase from *S. aureus* ATCC12600 or *B. intermedius* were successfully expressed in *B. subtilis*, *B.subtilis* was chosen as the most appropriate recombinant host for this project to secrete the V8-GluC extracellularly (Kakudo et al., 1992; Sharipova et al., 2007). Thanks to the secretion pathway, the recombinant host cannot be damaged by the protease activity.

Although *B. subtilis* is an efficient expression system as recombinant host, it is limited by its high secretion of extracellular proteases, which degrades secreted proteins (Wong et al., 1986). The WB168 strain of *B. subtilis* produces normally eight free proteases and six proteases, which are bound on the cell wall. Different *B. subtilis* strains were generated from the WB168 *B. subtilis* by knocking out some proteases (Kunst et al., 1997).

4.3.2 Cloning strategies

The following elements are necessary for the cloning and expression in *Bacillus subtilis* (Figure 20):

- a signal peptide for the extracellularly expression,
- a prosequence to avoid the autodegradation of the protease itself,
- the addition of a His-Tag is optional to facilitate the purification on the carboxyl-terminal side.

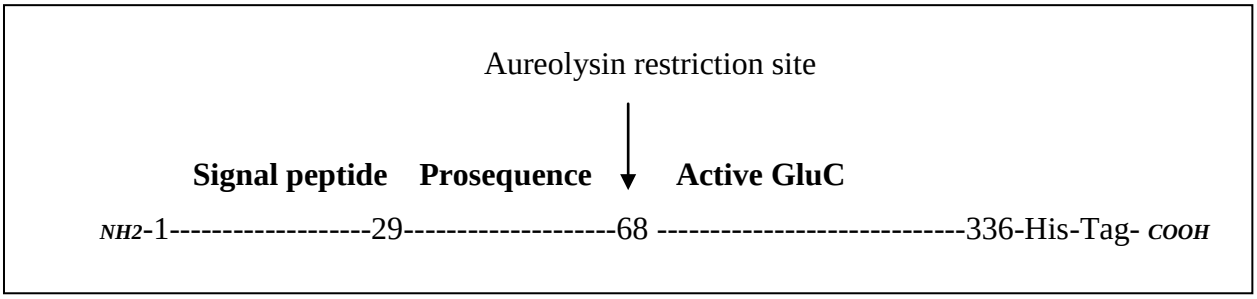


Figure 20. General structure of the *Staphylococcus aureus* GluC SSPA_STAU (http://www.uniprot.org/). The 268 amino acid V8-GluC protein is extended on its amino terminal side by an additional prosequence of 39 amino acids. The prosequence inactivates the protease after its synthesis in order to prevent autoproteolysis. For activation the prosequence can be cleaved by another Staphylococus aureus protease, aureolysin.

Prosequence cleavage occurs after secretion from *B. subtilis* as it was previously shown for the alpha-amylase of *B. subtilis* expression (Takase et al., 1988). In this work, the V8-glutamyl endopeptidase could be expressed without its initial signal peptide and prosequence. However, it was expressed in *E. coli* via a special construct and the yield of denaturation-renaturation was limited to 20 % (Yabuta et al., 1995). Otherwise, the enzyme could not be expressed without autodigestion problems up to today (Nemoto et al., 2008).

4.3.2.1 The AmyE signal peptide for the *B. subtilis* secretion

The signal peptide AmyE of the alpha-amylase of *B. subtilis* is well recognized and cleaved by a signal peptidase (Sasamoto et al., 1989). The AmyE signal peptide is one of the most appropriate signal peptide, which can be successfully used in *B. subtilis* for the expression of heterologuous proteins among several *B. subtilis* signal peptides (Brockmeier et al. 2006; Wong et al. 1986). The V8-GluC is normally secreted from *S. aureus* thanks to its own signal peptide. However, despite a similary secretion machinery, it can be expected that the signal peptide of *S. aureus* may not be recognized by the translocation machinery of *B. subtilis* as they are different organisms with their own secretion pathway (Figure 21).

```
  1 MKGKFLKVSS LFVATLTTAT LVSSPAANAL SSKAMDNHPQ QTQSSKQQTP KIQKGGNLKP
 61 LEQREHANVI LPNNDRHQIT DTTNGHYAPV TYIQVEAPTG TFIASGVVVG KDTLLTNKHV
121 VDATHGDPHA LKAFPSAINQ DNYPNGGFTA EQITKYSGEG DLAIVKFSPN EQNKHIGEVV
181 KPATMSNNAE TQVNQNITVT GYPGDKPVAT MWESKGKITY LKGEAMQYDL STTGGNSGSP
241 VFNEKNEVIG IHWGGVPNEF NGAVFINENV RNFLKQNIED IHFANDDQPN NPDNPDNPNN
301 PDNPNNPDEP NNPDNPNNPD NPDNGDNNNS DNPDAA

  1 MFAKRFKTSL LPLFAGFLLL FHLVLAGPAA ASALSSKAMD NHPQQTQSSK QQTPKIQKGG
 61 NLKPLEQREH ANVILPNNDR HQITDTTNGH YAPVTYIQVE APTGTFIASG VVVGKDTLLT
121 NKHVVDATHG DPHALKAFPS AINQDNYPNG GFTAEQITKY SGEGDLAIVK FSPNEQNKHI
181 GEVVKPATMS NNAETQVNQN ITVTGYPGDK PVATMWESKG KITYLKGEAM QYDLSTTGGN
241 SGSPVFNEKN EVIGIHWGGV PNEFNGAVFI NENVRNFLKQ NIEDIHFAND DQPNNPDNPD
301 NPNNPDNPNN PDEPNNPDNP NNPDNPDNGD NNNSDNPDAA
```

In black: AmyE signal peptide of *B. subtilis*

In red: Signal peptide of the V8-GluC *S. aureus*

In green: Prosequence of the V8-GluC

In blue: Active sequence of the V8-GluC

Figure 21. Amino acid sequence of the V8-GluC with the *S. aureus* or the AmyE from *B. subtilis* signal peptides. The V8-GluC can be cloned with its own *S. aureus* wild type signal pepdtide or with a signal peptide from an amylase of *B. subtlis*, AmyE.

4.3.2.2 The prosequence of the V8-GluC

Between the signal peptide and the active part of the GluC, there is an amino terminal extension of 20-30 amino acids, the prosequence, also called propeptide or zymogene (22). Prosequences exist in quite all exoproteases (Freudl, 1992). The propeptide does not seem to play any role for the translocation but for the correct folding after the membrane passage. The propeptide functions as an intramolecular protein chaperone by overcoming kinetic barriers in the folding pathway (Ohta et al., 1991; Harwood and Cranenburgh, 2008; Nemoto et al., 2008). The signal peptide helps folding by stabilizing intermediate complexes. The propeptide plays a very important role: it prevents the protease from being activated and thus, from autoproteolysis. It acts as competitive inhibitor of the active enzyme (Yabuta et al., 2001). This additional extension upstream from the active mature protein is removed either auto-hydrolysis or by other coexpressed proteases in the medium (Power et al., 1986; Takase et al., 1988; Chang et al., 1997).

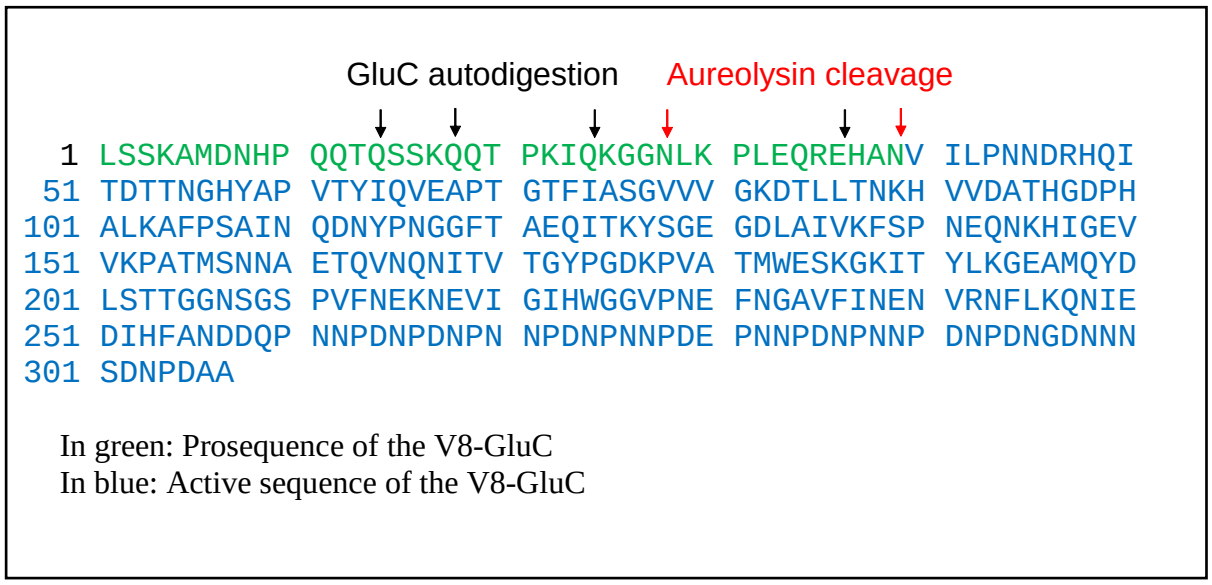


Figure 22. Maturation of the V8-GluC by autocatalysis and hydrolysis by aureolysin. The V8-GluC is extended on its amino terminal side by an additional prosequence of 39 amino acids. This prosequence, which can be cleaved by another *Staphylococus aureus* protease, aureolysin, inactivates the protease after its synthesis in order to prevent from its own proteolysis.

4.3.2.3 Aureolysin EC 3.4.24.29

The V8-Glutamyl endopeptidase propeptide is activated by prosequence hydrolysis. Unlike other proteases, the prosequence undergoes both hydrolysis mechanisms: first an autoproteolysis and then is hydrolyzed by an another coexpressed protease of *S. aureus*, a zinc metalloprotease, aureolysin EC 3.4.24.29 (Figure 22; Drapeau, 1978). This protease is also called *Staphylococcus aureus* neutral proteinase and is encoded by the Aur gene, which exists under two highly conserved allelic forms (Sabat et al., 2000; Takeuchi et al., 2001). It is registered under the AURE_STAAU code in the Swissprot data base. This protease cuts preferentially hydrophobic residues P1' like the thermolysin from *Bacillus thermoproteolyticus*. Three dimensional structure of the 301 amino acid aureolysin sequence was studied by Banbula (1998). The enzyme binds three calcium ions and one zinc ion per subunit (Potempa et al., 1989). The aureolysin and the V8-Glutamyl endopeptidase genes are encoded by the same operon, Staphylococcus Serine Protease Operonn (ssp), as well as two other cystein proteases (Massimi et al., 2002; Shaw et al., 2004).

In this way, the operon encodes for four different proteases, whose activation depends on each other (Rice et al., 2001). These four proteases are thus regulated under the same transcription control as they are located on the same operon (Dubin, 2002; Gusstafsson et Oscarsson, 2008).

The activation process of the aureolysin was up to now not known. Today the maturation of aureolysin starts being understood and autocatalysis process was suggested (Shaw et al., 2004; Nickerson et al., 2008). Aureolysin was supposed to be the only way to activate the V8-glutamyl endopeptidase by splitting the asparagin of the prosequence (Drapeau, 1978; Shaw et al., 2004). Mutant *S. aureus* strains without metalloprotease could not secrete the V8-glutamyl endopeptidase. However, after recent investigations, the activation of the V8-Glutamyl endopeptidase was discovered to occur in two different complementary ways: first an autocatalysis and then, hydrolysis by the aureolysin (Nickerson et al., 2007).

After the last alanine of the signal peptide is cleaved by the signal peptidase, the V8-GluC is matured by sequential autocatalysis and hydrolysis by aureolysin. The V8-GluC autocatalyzes itself at first by three steps in succession. The glutamine 14 is first processed, then, the glutamine 18. The autocatalytic process ends up with the glutamine 24.

Once this auto-process occurred, the aureolysin can access to the asparagine 28 and shortens the prosequence. The V8-GluC must again autocatalyze itself at the glutamic acid 36 to make the last restriction site accessible to the aureolysin. The aureolysin can then cut the last three amino acids of the prosequence, the HAN pattern for histidine 37, alanine 38 and asparagine 39, to make the V8-GluC fully active. The histidine 37 blocks the active site off via hydrogen bonds with the serine of the catalytic site so that the substrate cannot access the protease active site. The last autocatalysis processing at the glutamic acid 36 is more efficient than the aureolysin cleavage. The first valine of the active V8-glutamyl endopeptidase is an essential amino acid for the last cleavage by the aureolysin (Nemoto et al., 2008). The propeptide must be therefore sequentially process in the correct order to get an efficient maturation of the active enzyme.

4.3.2.4 Polyhistidine tag (His Tag) for the purification

Affinity tags are designed to allow one step purification, without interfering on the tertiary structure and on the protein activity. It can be used on a large range of proteins. The removal is not always required when the affinity tag, like histidine hexamer, also called polyhistidine tag or His Tag, does not cause negative effects (Terpe, 2003; Wu and Filutowicz, 1999). The most employed chromatography to purify polyhistidine tag is the Immobilized Metal Affinity Chromatography (IMAC; Porath et al., 1975; Figure 23).

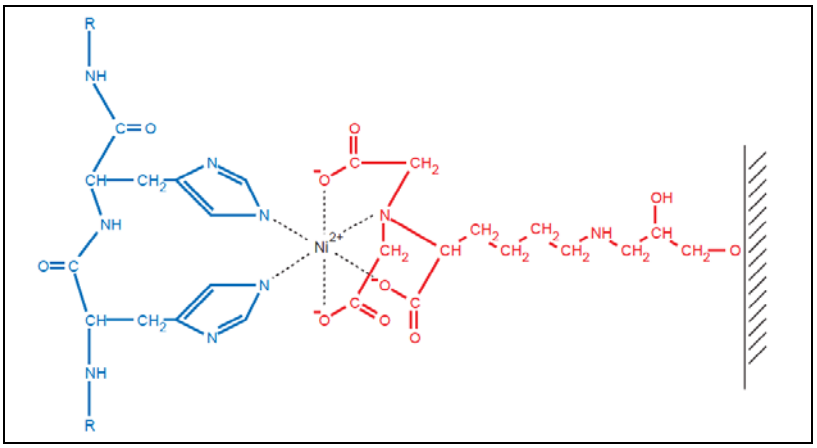


Figure 23. Interaction between neighboring residues in the polyhistidine tag and Ni-NTA matrix. Ni-NTA Agarose is composed of Ni-NTA coupled to Sepharose® CL-6B. Nitrilotriacetic acid (NTA) is a tetradentate chelating adsorbent, which occupies four of the six ligand binding sites in the coordination sphere of the nickel ion, leaving two sites free to interact with the polyhistidine tag.

Here, Nickel NTA (Ni-NTA) was used (Hochuli et al., 1987). A transition ion metal, the nickel, is immobilized on a matrix specific amino acid side chains. The nickel interacts with the imidazole ring of the histidine, which plays the role of electron donor, to form coordinate bonds. Histidine is eluted by changing the pH or adding free imidazole, which competes against the histidine. A wash with a low concentration of imidazole, as preliminary step, reduces non specific binding of host proteins.

4.4 Mutagenesis technique

Site directed mutagenesis and saturated mutagenesis techniques are based on the PCR principle in order to create a single mutation of an amino acid at a given position on DNA (Hutchison et al., 1978). In the site-directed mutagenesis method, a pair of complementary mutagenic oligonuleotides with the desired mutation at a given position is used to hybridize the parental DNA. For the saturated mutagenesis technique, a mixture of degenerate primers, which are designed with the twenty possible amino acids at the mutation point, is used for the PCR. In each generated clone, the amino acid at the position to mutate is thereby exchanged with one of the twenty different possible amino acids. In both mutagenesis techniques, primers bind to ten nucleotides upstream and ten nucleotides downstream around the given position to mutate. The oligonucleotide primers, each complementary to opposite strands of the vector, are then extended during temperature cycling by a DNA polymerase and the plasmid is thus amplified with the mutation at the determined position.

A nicked circular DNA is generated through the amplification cycle and it will be religated thanks to the bacteria machinery after transformation (Figure 24). However, the template DNA is still present and must be eliminated in order to avoid a clone mixture between parental and mutated clones. DNA-methyltransferase (Dam) recognizes the 5´-GATC-3´ target sequence and introduces a methylation on the sixth nitrogen of the adenine in *E. coli*. This enzyme directs mismatch repair during DNA synthesis in *E. coli* (Hattman et al., 1978; Geier and Modrich, 1979). Isolated DNA from almost all *E. coli* strains is dam methylated, unlike in-vitro synthetized DNA, and therefore susceptible to Dpn I digestion. The DpnI restriction enzyme is an endonuclease, which recognizes and digests the Dam restriction site in *E. coli*. The parental template DNA is digested and cannot be transformed unlike the non-methylated mutated DNA. However, the DpnI restriction enzyme cannot be used directly on the *Bacillus* DNA.

The parental DNA of *Bacillus subtilis* is not methylated as it is in *E. coli* due to the lack of the Dam gene (Dreiseikelmann and Wackernagel, 1981). Some studies, which were led by Guha (1985) showed that the cytosine can be methylated in *B. subtilis* but it cannot play the same role as the Dam endonuclease (Guha and Guschlbauer, 1992). Site directed mutagenesis cannot be performed on DNA, which is directly extracted from *B. subtilis*. A precloning is required in *E. coli* to collect the methylated DNA, which is transformed in *B. subtilis* afterwards.

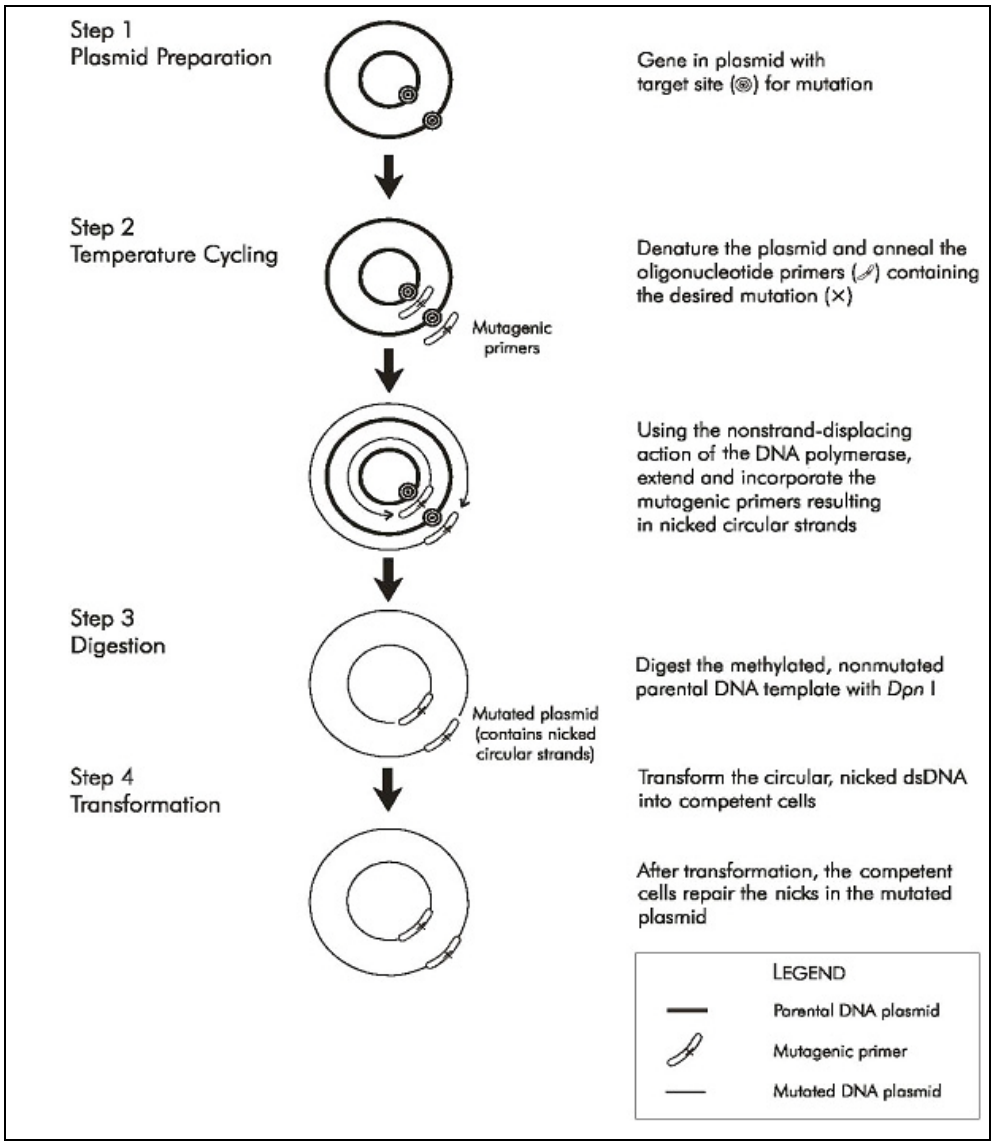


Figure 24. Site directed mutagenesis. Primers are designed with the desired mutation. After amplification by PCR, a nicked, circular DNA is generated. This mutated DNA is unmethylated unlike the DNA template. The template can be eliminated by enzymatic digestion with a restriction enzyme specific for methylated DNA, the DpnI.

5. Aim and motivation of the PhD

The HbA1c test is a long-term blood glucose key parameter for the management of diabetes. Roche Diagnostics GmbH has collaborated with the IFCC in the frame of the HbA1c reference standardization by ESI-MS measurement. The V8-Glutamyl endopeptidase (GluC) of *Staphylococcus aureus*, EC 3.4.21.19 was selected as the most appropriate endoprotease for the glycated hexapeptide digestion in the HbA1c test by ESI-MS measurement (Houmard and Drapeau, 1972). However, the slow proteolytic digestion, which was optimized up to 18 hours for the glycated hemoglobin sample preparation, prevents this test from being routinely employed in clinical laboratories. Roche Diagnostics would like to use a GluC mutant, which could hydrolyze faster the glycated hemoglobin, in order to reduce blood glucose sample time preparation (Wu et al., 1991; Yabuta et al., 1995).

The main aim of the PhD was to improve the GluC velocity (substrate affinity and activity) by keeping the same other enzymatic parameters: thermostability and above all specificity in the experimental conditions of clinical laboratory. A strategy was set up by directed evolution and irrational design to produce improved GluC mutants so that this HbA1c method could be then routinely applied by enzymatic digestion and RP-HPLC/MS-ESI assessment.

The PhD consisted in three main parts (Figure 25). The V8-GluC gene was first cloned in *B. subtilis* and characterized to set the screening system up. The appropriate signal peptide of *B. subtilis* was chosen for the secretion of the protease in the medium. Furthermore, a polyhistidine tag was added to facilitate the protease purification. The original wild type prosequence was also kept in order to avoid the autodigestion of the protease. This V8-GluC expression in *B. subtilis* was the basis to undertake directed evolution on the gene.

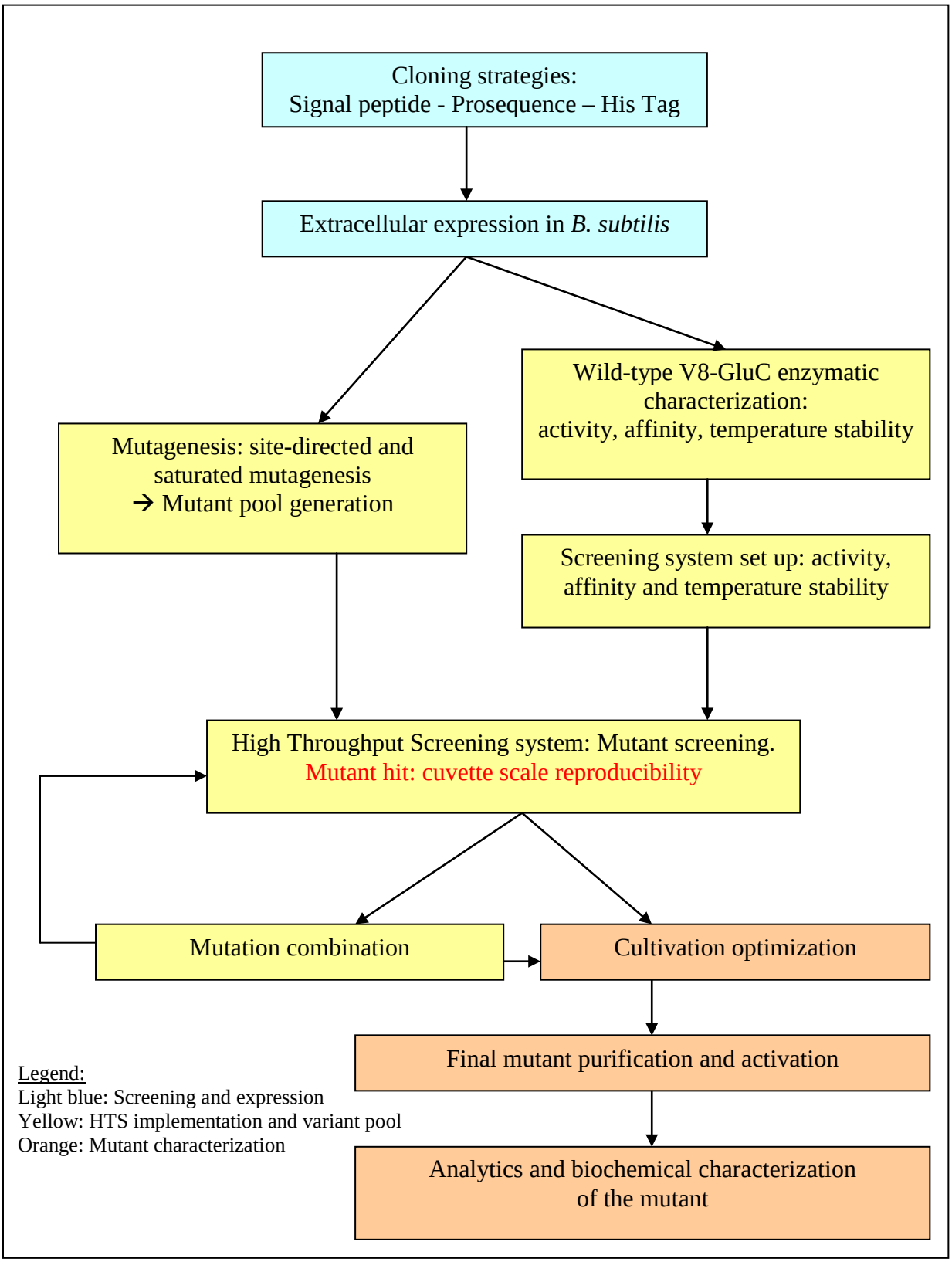


Figure 25. Project overview. After cloning and expression of the wild type V8-GluC, the protease is characterized to set the screening system up. In parallel, mutants are generated. After high throughtput screening, the mutant is expressed after *B. subtilis* cultivation to be finally purified, activated and characterized.

The second part of the PhD was based on generating a variant pool by directed evolution. The enzyme of part one was characterized in order to build a high throughput multiparameter screening system. The activity and substrate affinity, Km-value, parameters were screened .Thousands of V8-GluC variants were prepared by different mutagenesis methods: site-directed and saturated mutagenesis on strategic positions, which were defined according to the bibliographic and theoritical study of the V8-GluC structure. The screening system was based on the wild type V8-GluC enzymatic characterization by miniaturization from the cuvette scale for spectrophotometer to a 96-well Microtiterplate (MTP) scale. Mutants were screened by High Throughtput Screening (HTS) and once a positive mutant was found, its enzymatic characteristic reproducibility was checked in the 96-well microtiterplate and at the cuvette scale. If the enzymatic features of this optimized mutant could be reproduced, the mutation point is identified by DNA sequencing. A new HTS round is performed to optimize the mutant by testing all twenty amino acids at the given position by saturated mutagenesis. The outcome of the screening, the finding of a mutant with improved properties, represents normally the turning point of the project. The PhD could then follow two different directions. On the one hand, a new HTS round can be started on the identified mutant. The positive effects of several mutations can be combined to create a new mutant. On the other hand, the final goal of the PhD is considered to be achieved and the mutant can be characterized.

The final part of the PhD was finally decided to be focused on further analytic investigations. Mutant cultivation conditions were optimized so that enough material could be easily purified. After purification, the mutant was activated by cleavage of its protective prosequence and biochemically and analytically examinations were investigated.

MATERIAL

and

METHODS

1. Molecular biology

1.1 Molecular Biology - Material

- Agarose (Roche Diagnostics, Mannheim, Germany)
- Aureolysin (Biocentrum Sp. z o.o.,Krakow, Poland)
- *Bacillus subtilis*, strain WB600, ΔtrpC, ΔnprB, ΔnprE, ΔaprE, Δbpr, Δepr, Δmpr, cat, ermAM (Tryophan auxotroph, 6 inactivated proteases, chloramphenicol and erythomycin resistent) (DSZM, Berlin, Germany)
- Bromophenol Blue (Sigma-Aldrich, Seelze, Germany)
- Buffer Taq 10X (Roche Diagnostics, Mannheim, Germany)
- Calibration Kit Broad pI (pH 3-10) (Amersham Bioscience, Freiburg, Germany)
- Coomassie Blue 250 (Serva Blue G) (Boehringer Ingelheim, Heidelberg, Germany)
- DNA molecular Weight Marker X or XIV (Roche Diagnostics, Mannheim, Germany)
- dNTP (Roche Diagnostics, Mannheim, Germany)
- *Dpn* I (10 U/µL) (Roche Diagnostics, Mannheim, Germany)
- DTT (Roche Diagnostics, Mannheim, Germany)
- *E.coli* XL1-MRF´, D(mcrA)183 D(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac [F'proAB lacIqZDM15 Tn10 (Tetr)] (Stratagene, Heidelberg, Germany)
- *Eco*RI (10 U/µL) (Roche Diagnostics, Mannheim, Germany)
- DNA Buffer (Stratagene, Heidelberg, Germany)
- Glycin (Sigma-Aldrich, Seelze, Germany)
- *Hin*d III (10 U/µL) (Roche Diagnostics, Mannheim, Germany)
- PCR reaction buffer 10X (Roche Diagnostics, Mannheim, Germany)
- Pfu Turbo DNA polymerase (2.5 U/µL) (Roche Diagnostics, Mannheim, Germany)
- Primers (Thermo Electron, Dreieich, Germany)
- QIAprep™ Spin Miniprep Kit Protocol (QIA gene GmbH, Hilden, Germany)
- QIAquick™ Gel extraction Kit Protocol (QIA gene GmbH, Hilden, Germany)
- QIAquick™ PCR Purification Kit Protocol (QIA gene GmbH, Hilden, Germany)
- Taq Polymerase (5U/µL) (Roche Diagnostics GmbH, Mannheim, Germany)
- TBE 10X (Roche Diagnostics, Mannheim GmbH, Germany)
- 10X Buffer B (Roche Diagnostics, Mannheim GmbH, Germany)

- 10X Ligation buffer (Roche Diagnostics GmbH, Mannheim, Germany)
- T4 DNA ligase (1 U/µL) (Roche Diagnostics GmbH, Mannheim, Germany)
- V8-Glutamyl endopeptidase gene of *Staphyococcus aureus*, Sspa_STAU (DSMZ GmbH, Berlin, Germany)

1.2 Molecular biology: Instruments

- Electrophoresis: Amersham Biosciences, Electrophoresis power supply EPS 301 and Hoefer HE 33, Mini horizontal Submarine Unit (Freiburg, Germany)
- Electroporation: Bio-Rad *E. coli* pulser cuvette and *E.coli* Pulser (München, Germany)
- Image acquisition: Amersham Biosciences, Image master VDS (Freiburg, Germany)
  Fujifilm Thermal imaging system FTI 500 (Düsseldorf, Germany)
- Microwave: AEG Hausgeräte Micromat (Nürnberg, Germany)
- Robocycler PCR: Applied Biosystems, Gene Amp PCR system 2400 (Weiterstadt, Germany)
- Shaker: B. Braun Biotech International GmbH, Certomat® (Melsungen, Germany)
- Tabletop centrifuge: Kendro Laboratory Products, Heraeus® (Langenselbold, Germany)
- Thermomixer: Eppendorf, Thermomixer comfort (Hamburg, Germany)

2. Biochemistry

2.1 Biochemistry – Material

- Acid acetic (Sigma-Aldrich, Seelze, Germany)
- Agar Agar (Merck, Darmstadt, Germany)
- Ammonium sulphate (Merck, Darmstadt, Germany)
- Bio-Rad Protein assay (Bio-Rad, München, Germany)
- BSA (Roche Diagnostics, Mannheim, Germany)
- Endoproteinase GluC (NEB, Frankfurt am Main, Germany)
- Endoproteinase GluC (Roche Diagnostics, Mannheim, Germany)
- Ethanol (Merck, Darmstadt, Germany)

- HCl (Merck, Darmstadt, Germany)
- KH_2PO_4 (Merck, Darmstadt, Germany)
- KOH (Merck, Darmstadt, Germany)
- NaCl (Merck, Darmstadt, Germany)
- NaOH (Merck, Darmstadt, Germany)
- PEG-4000 (Fluka, Milwaukee, WI, USA)
- Synthetic substrate, HbA0(1-6)-pNA (Roche Diagnostics, Penzberg, Germany)
- Synthetic substrate, Z-Phe-Leu-Glu-pNA (Roche Diagnostics, Penberg, Germany)
- Tris (Roche Diagnostics, Mannheim, Germany)
- Water (Sartorius, Millipore Arium® 611, 0.2 µm)
- Yeast extract (Becton Dickinson and Co, Starks, USA)

2.2 Biochemistry – Instruments

- Balance: Sartorius SK-5000D balance, 3100 balance, Genius balance (Göttingen, Germany)
- Cuvette: Sarstedt cuvettes (Nümbrecht, Germany)
- MTP reader: Tecan Spectra®, Rainbow thermo (Crailsheim, Germany)
  Becton Dickinson Falcon™ 96-well Microtest (Heidelberg, Germany)
- Ovens: Salvis Thermocenter® oven (Rotkreuz, Swiss)
  Memmert Modell 2000 (Schwabach, Germany)
- pH meter: Sartorius P110 (Göttingen, Germany)
- Quartz cuvette: Brand UV-cuvettes (Wertheim, Germany)
- Robotpipettor: Beckman Coulter, Biomek® FX^P Laboratory Automation Workstation (Krefeld, Germany)
- Spectrophotometers: Varian, Cary 50 Bio (Darmstadt, Germany)
  Beckman Coulter DU 7400 (Krefeld, Germany)
- Stirrer: Janke and Kunkel, IKAMAG® Reo Basic (Köln, Germany)

- Tabletop centrifuge: Kendro Laboratory Products, Heraeus® Biofuge® fresco (Langenselbold, Germany)

- Waterbath: Boehringer Mannheim, Clinicon Mannheim, Lauda® B, Precitherm PFV, Lauda® MA6, Precitherm PFV, Lauda® M6, Precitherm PFV (Mannheim, Germany)
  Julabo U3 (Seelbach, Germany)

3. Cultivation

3.1 Cultivation - Material

- Ampicilline (Roche Diagnostics, Mannheim, Germany)
- Antifoam (Dow Corning GmbH, Wiesbaden, Germany)
- Kanamycin sulfate (aphA3) (Roche Diagnostics GmbH, Mannheim, Germany)

3.2 Cultivation - Instruments

- Incubator-shaker: B. Braun Biotech International GmbH, Certomat® (Melsungen, Germany)
  HT Infors AG (Einsbach, Germany)

4. Purification and analytics

4.1 Purification and analytics – Material

- DEAE Sepharose (Amersham Biosciences, Freiburg, Germany)
- NuPAGE® Novex Bis-Tris Gel (InVitrogen, Karlsruhe, Germany)
- NuPAGE® MES buffer (InVitrogen, Karlsruhe, Germany)
- NuPAGE ® MOPS buffer (InVitrogen, Karlsruhe, Germany)
- NuPAGE ® Sample 4x (InVitrogen, Karlsruhe, Germany)
- SeeBlue™ Pre-Stained standard (InVitrogen, Karlsruhe, Germany)
- SimplyBlue™ SafeStain (InVitrogen, Karlsruhe, Germany
- SERVAPOR® Dialysis Tubing, (MWCO 12000 - 14000; pore diameter ca. 25 Å; nominal dry flat width: 45 mm; nominal dry diameter 29 mm; approximate filling volume 6.5 mL/cm; Nominal Dry Wall Thickness 20 µm) (Boehringer Ingelheim, Heidelberg, Germany)

- Slide-A-Lyser dialysis cassettes, 3,500 MWCO with 0.5 - 3 mL capacity (Perbio Science, Bonn
  Germany)

4.2 Purification and analytics – Instrument

- Centrifuge:	Kendro Laboratory Products, Sorvall® RC-5B, refrigerated Superspeed Centrifuge, GSA Rotor (Langenselbold, Germany) Kendro Laboratory Products, Sorvall® Evolution™ MC, rotor Sorvall® SLC 6000, Super-Lite® composite (Langenselbold, Germany)
- Chromatography:	Amersham Biosciences, Peristaltic Pump P1, Uvicord™ S 2138, 2 Channel recorder 2210, Superfrac™ Fraction Collector (Freiburg, Germany) H. Hölzel, buffer mixer (Wörth / Hörlkofen, Germany)
- Electrophoresis:	Invitrogen, Xcell surelock™ (Karlsruhe, Germany)
- Filter material:	Sartorius, Cellulose nitrate filters and glass holder (Göttingen, Germany)
- French Press:	Constant systems ltd, Cell Disruption Equipment Basic Z model, One shot model (Königswinter, Germany)
- HPLC:	Amersham Biosciences HPLC Pump 2150, Biotech VWM 2141 UV Detector, Peristaltic Pump P1, Conductivity monitor (Freiburg, Germany)
- Blot:	Invitrogen iBlot™ Device with Blotting Roller (Karlsruhe, Germany)
- MS-ESI:	Q-TOF 2 (Cheshire, United-Kingdom)
- Stirrer:	Janke and Kunkel, IKA®-maxi MC (Köln, Germany)

5. Enzymatic test assay at the cuvette scale

5.1 Specific activity

5.1.1 Enzymatic test assay with the glutamate substrate at pH 7.8 in Tris-HCl buffer

6.62 mg of Z-Phe-Leu-Glu-pNA substrate (10 mM) were dissolved into 1 mL of ethanol and heated at 45°C by shaking at 1000 rpm up to complete dissolution. Substrate solutions were incubated at room temperature for 10 min before starting the enzymatic test assay. 1 mg/mL of Roche GluC or NEB GluC lyophilisates were prepared with double distilled water as positive control. V8-glutamyl endopeptidases (Roche, NEB, wild type or mutant V8-GluCs) were diluted with double distilled water up to get an absorbance of 20-30 mA/min during the spectrophotometer measurement. In this case, the Roche GluC was diluted 1000 times. 625 µL of Tris-HCl (0.1 M) at pH 7,8 buffer were pipetted into the cuvette with 75 µL of substrate Z-Phe-Leu-Glu-pNA (10 mM) and 100 µL of the V8-Glutamyl endopeptidase enzyme. After mixing, the absorption was measured at 405 nm for 5 min by a spectrophotometer at 25°C. The specific activity calculation is explained in the chapter 5.1.5 Protein concentration by Bradford.

5.1.2 Enzymatic test assay with the hemoglobin hexapeptide substrate at pH 7.8 in Tris-HCl buffer

The enzymatic test assay is analogous as described in the chapter 5.1.1. However, the synthetic hemoglobin substrate was taken instead of the glutamate substrate (Figure 26). The hexapeptide was synthetized in the Chemistry department of Roche Diagnostics GmbH. 8.148 mg of hemoglobin hexapeptide substrate were weighted and were simply mixed with 1 mL of double distilled water up to complete dissolution at room temperature. Calculations are described in figures 27 and 28.

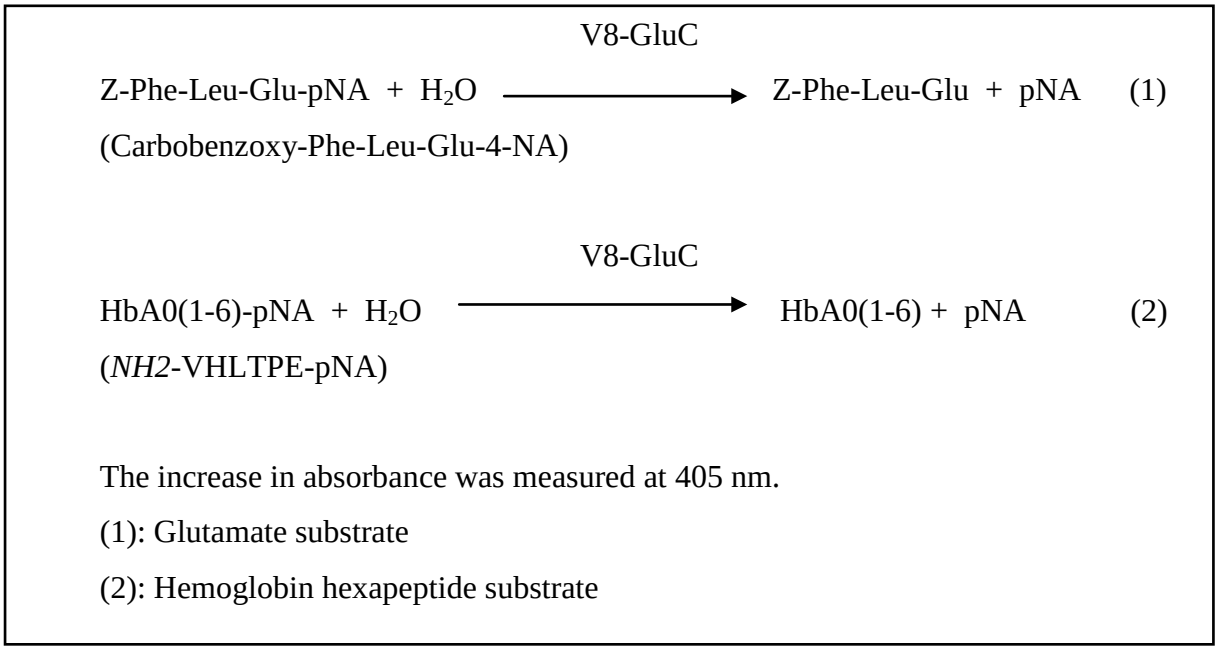


Figure 26. Experimental enzymatic test reaction. The V8-GluC hydrolyzes specifically glutamate and aspartate bonds. Two paranitroanolide substrates were synthesized for the enzymatic test assay, one as glutamate bond, the other one as hemoglobin hexapeptide.

5.1.3 Enzymatic test assay with the hemoglobin hexapeptide substrate at pH 4.3 in ammonium acetate buffer

The enzymatic test assay follows the same protocol as in the chapter 5.1.1. Enzymatic test assay with the hemoglobin hexapeptide substrate at pH 7.8 in Tris-HCl buffer. The 100 mM Tris-HCl buffer was replaced by 50 mM ammonium acetate buffer at pH 4.3. The pH was adjusted with acetic acid.

5.1.4 Enzyme activity calculations: Units (U), Units per volume (U/mL), Units per milligram of lyophilisate (U/mg lyo)

An activity unit is usually defined as the enzyme amount required for catalysing one micromole substrate amount per min under fixed conditions (temperature, buffer, pH).

The V8-glutamyl endopeptidase activity unit is the needed amount of V8-glutamyl endopeptidase (Roche, NEB, wild-type or mutant GluCs), which turns one micromole of HbA0(1-6)-pNA or Z-Phe-Leu-Glu-pNA into HbA0(1-6) + paranitroaniline or Z-Phe-Leu-Glu + paranitroaniline per min, measured at 25°C and 405 nm either with 0.1 M Tris-HCl, pH 7.8 buffer or 50 mM ammonium acetate, pH 4.3. The reaction has stoechiometric coefficients and different activities were calculated according to the Lambert-Beer law (Figures 27 and 28). Positive control was checked at each measurement with the Roche GluC lyophilisate. Measurements were repeated twice.

$A = \varepsilon \times l \times C$

A: Absorbance or optic density (dimensionless)

ε: Extinction coefficient or molar absorption coefficient at 405 nm (10,4 $mmol^{-1}.cm^{-1}$).

l: Cuvette path length (usually 1 cm)

C: Concentration (mol/L)

Figure 27. Beer-Lambert law. The absorbance is linear to the concentration and the path length. The extinction coefficient varies according to the wavelength and it depends on the temperature, pH and the solution ionic strength.

5.1.5 Protein concentration by Bradford

The protein determination assay uses the property of the Coomassie Brilliant Blue to present a shift in its maximum absorbance when it reacts with proteins (Bradford, 1976). A standard curve was set up from 0 to 10 µg/mL Bovine Serum Albumin (BSA) to determine the protein concentration of samples. 200 µL of Coomassie Brilliant Blue G-250 were mixed with 740 µL of water and 60 µL of the sample in a cuvette. The absorption was measured after 5 min at 595 nm and at room temperature (25°C). Dilution with water was required not to have an $OD_{595\ nm}$ superior than 0.6.

Samples were measured twice and the result was calculated by the measurement average. Specific activity was deduced according to the formula in figure 28.

$$\text{Vol. activity} = \frac{Vc \times (dA/min) \times d}{Ve \times \varepsilon \times l}$$

Formula, from which the following calculations were deduced:

$$\text{Lyo. activity} = \frac{\text{Vol. activity}}{\text{Lyo. concentration}}$$

$$\text{Total units} = \text{Vol. activity} \times V$$

Vol. activity: Volume activity (U/mL)
Lyo activity: Lyophilisate activity (U/mg Lyo)
Vc: Cuvette volume (µL).
dA/min: Absorbance measured on the spectrophotometer at 405 nm
d: Dilution
Ve: Enzyme volume (µL)
ε: Extinction coefficient at 405 nm
l : Cuvette path length (1 cm)
V: Sample volume (µL)
Lyo concentration: Lyophilisate concentration in mg/mL (ususally 1 mg lyo/mL).
U: Units (U)

$$SA = \frac{\text{Vol. activity}}{C}$$

SA: Specific activity (U/mg)
C: Protein concentration (mg/mL) by Bradford assessment

Figure 28. Formula for the calculation of activity and specific activity. An activity unit is usually defined as the enzyme amount required for catalyzing one micromole substrate amount per min under fixed conditions (Temperature, buffer, pH). The activity is deduced from the Beer-Lambert law. Protein concentration is calculated by Bradford measurement.

5.2 Km-value

Km-value was assessed with the glutamate substrate in Tris-HCl buffer, pH 7.8. The V8-GluC activity measurement protocol in Tris-HCl buffer at pH 7.8 at the cuvette scale was followed as it is described in the chapter 5.1.1. Instead of 10 mM glutamate substrate, different concentrations were tested. The enzyme dilution was searched for getting an absorbance with a dA/min of about 0.06 at the substrate saturation. The 10 mM glutamate substrate concentration was then diluted by ½ step dilution until about 10 mA/min were reached. The measurement was repeated three times. Km-values were assessed according to the Lineweaver-Burk linearization graph and the value average was calculated as final Km-value **(**Lineweaver and Burk, 1934**)**.

5.3 Temperature stress

The V8-glutamyl endopeptidase was incubated at different temperatures for 30 min in waterbaths. Different temperatures were tested: 25-55°C in 5°C steps. Then, the volume activity was measured at in Tris buffer, pH 7.8 at the cuvette scale according to the chapter 5.1.1. Positive control was measured at 25°C. The remaining activity average was calculated after reproducing twice the experiment.

6. V8-Glutamyl endopeptidase cloning and expression

6.1 Cloning of the V8-Glutamyl endopeptidase in the pMSE3 vector shuttle

The genomic DNA sequence of the V8-GluC was directly ordered at the German Collection of Microorganisms and Cell Cultures (DSMZ Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH). This genomic DNA is composed of the complete sequence of the V8-GluC with the wild type signal peptide, prosequence and active V8-GluC (Figure 29). The V8-GluC was cloned with another wild type signal peptide, which is recognized as efficient for the protein translocation in *Bacillus subtilis* (Brockmeier et al., 2006). The wild type signal peptide of the V8-GluC was replaced in the described way in figure 29.

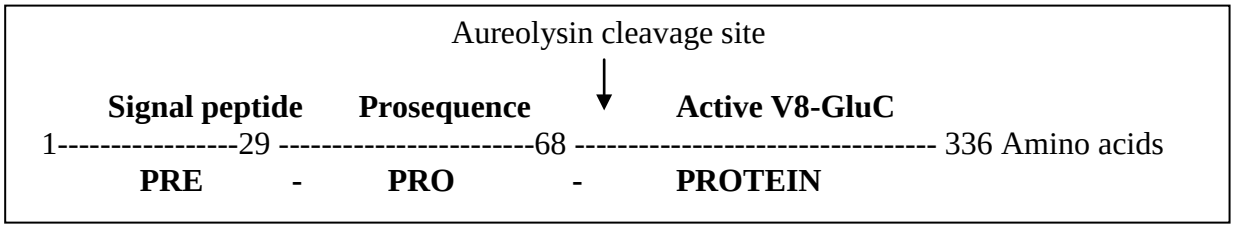


Figure 29. General structure of the V8-Glutamyl endopeptidase. This 268 amino acid V8-GluC protein is extended on its amino terminal side by an additional prosequence of 39 amino acids. This prosequence, which can be cleaved by another *Staphylococus aureus* protease, aureolysin, inactivates the protease after its synthesis in order to prevent from its own autoproteolysis.

The active V8-GluC gene was extracted with its wild type prosequence by Polymerase Chain reaction (PCR; Mullis et al., 1986; Saiki et al., 1988). Flanking primers were designed with necessary restriction sites, FseI and MluNI, in order to ligate the V8-Glutamyl endopeptidase DNA The coding sequence was added with the signal peptide between the FseI and MlunI restriction sites. As a Tag can be added only on the carboxyl-terminal side, the reverse primer was designed with a polyhistidine Tag on the 5' side. The polyhistidine tag was designed with a triplet alternation of CAT and CAC, which encodes for histidine. The CAT codon presents a higher translation frequency than the CAC codon.

The complete sequence with the promoter, signal peptide, prosequence of the V8-GluC as well as the polyhistdine tag and the terminator, was then cloned in the multisite cloning of the pMSE3 vector (Figures 30 and 31; Silbersack et al., 2006).

```
  1 MKGKFLKVSS LFVATLTTAT LVSSPAANAL SSKAMDNHPQ QTQSSKQQTP KIQKGGNLKP
 61 LEQREHANVI LPNNDRHQIT DTTNGHYAPV TYIQVEAPTG TFIASGVVVG KDTLLTNKHV
121 VDATHGDPHA LKAFPSAINQ DNYPNGGFTA EQITKYSGEG DLAIVKFSPN EQNKHIGEVV
181 KPATMSNNAE TQVNQNITVT GYPGDKPVAT MWESKGKITY LKGEAMQYDL STTGGNSGSP
241 VFNEKNEVIG IHWGGVPNEF NGAVFINENV RNFLKQNIED IHFANDDQPN NPDNPDNPNN
301 PDNPNNPDEP NNPDNPNNPD NPDNGDNNNS DNPDAA

  1 MFAKRFKTSL LPLFAGFLLL FHLVLAGPAA ASALSSKAMD NHPQQTQSSK QQTPKIQKGG
 61 NLKPLEQREH ANVILPNNDR HQITDTTNGH YAPVTYIQVE APTGTFIASG VVVGKDTLLT
121 NKHVVDATHG DPHALKAFPS AINQDNYPNG GFTAEQITKY SGEGDLAIVK FSPNEQNKHI
181 GEVVKPATMS NNAETQVNQN ITVTGYPGDK PVATMWESKG KITYLKGEAM QYDLSTTGGN
241 SGSPVFNEKN EVIGIHWGGV PNEFNGAVFI NENVRNFLKQ NIEDIHFAND DQPNNPDNPD
301 NPNNPDNPNN PDEPNNPDNP NNPDNPDNGD NNNSDNPDAA

  1 atgaaaggta aatttttaaa agttagttct ttattcgttg caactttgac aacagcgaca
 60 cttgtgagtt ctccagcagc aaacgcgtta tcttcaaagg ctatggacaa tcatccacaa
120 caaacgcagt caagcaaaca gcaaacacct aagattcaaa aaggcggtaa ccttaaacca
180 ttagaacaac gtgaacacgc aaatgttata ttaccaaata acgatcgtca ccaaatcaca
240 gatacaacga atggtcatta tgcacccgta acttatattc aagttgaagc acctactggt
300 acatttattg cttccggtgt agttgtaggt aaagatactc ttttaacaaa taaacacgtc
360 gtagatgcta cgcacggtga tcctcatgct ttaaaagcat tcccttctgc aattaaccaa
420 gacaattatc caaatggtgg tttcactgct gaacaaatca ctaaatattc aggcgaaggt
480 gatttagcaa tagttaaatt ctcccctaat gagcaaaaca aacatattgg tgaagtagtt
540 aaaccagcaa caatgagtaa taatgctgaa acacaagtta accaaaatat tactgtaaca
600 ggatatcctg gtgataaacc tgtagcaaca atgtgggaaa gtaaaggaaa aatcacttac
660 ctcaaaggcg aagctatgca atatgattta agtacaactg gtggtaattc aggttcacct
720 gtatttaatg aaaaaaatga agtgatcgga attcattggg gcggtgtacc aaatgaattt
780 aatggtgcgg tatttattaa tgaaaatgta cgcaacttct taaaacaaaa tattgaagat
840 atccattttg ccaacgatga ccaacctaat aacccagata atcctgataa ccctaacaat
900 cctgataacc ctaacaaccc agatgaacca ataaccctg acaaccctaa caaccctgat
960 aatccagaca atggcgataa caataattca gacaatccag atgcagctta a 1011//
```

Figure 30. V8-GluC amino acid sequence with the wild type or AmyE signal peptide and nucleotide sequence of the V8-GluC with the wild type V8-GluC signal peptide. In black: AmyE signal peptide of *B. subtilis*, *i*n red: signal peptide of the V8-GluC *S. aureus*, in green: prosequence of the V8-GluC, in blue: active sequence of the V8-GluC.

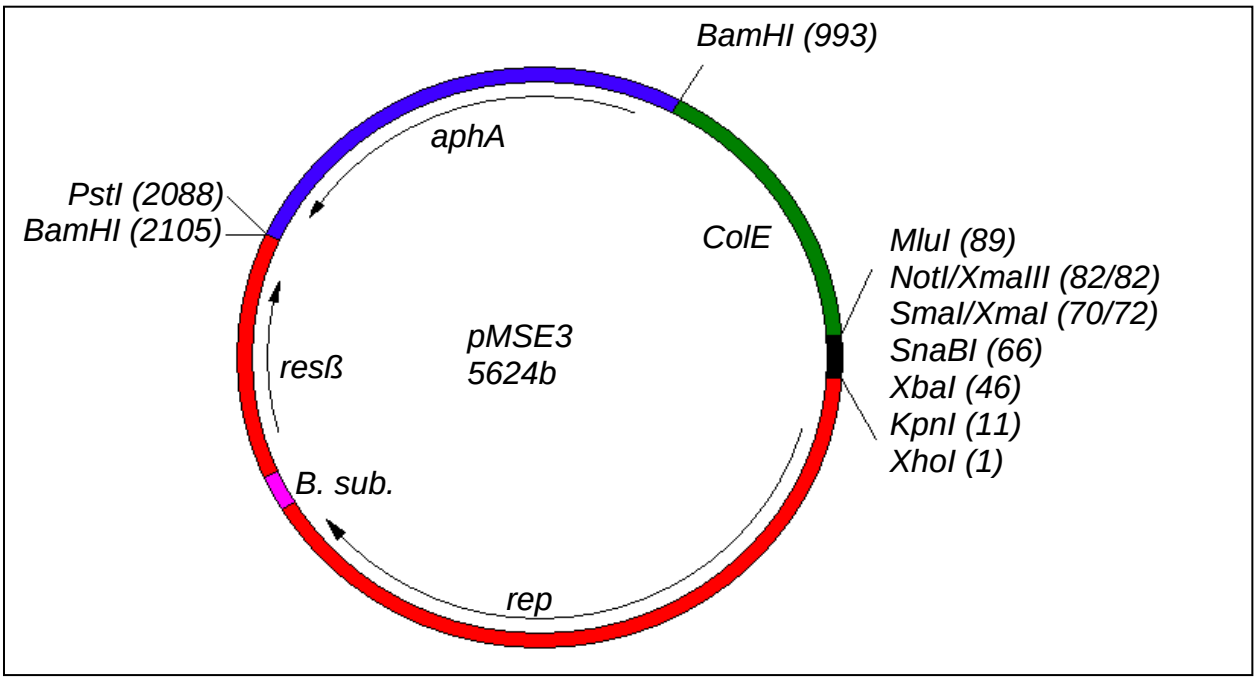


Figure 31. pMSE3 plasmid map. This vector contains two replication origins, one for *E. coli* (ColE1), one for *B. subtilis*, which was encoded in the repE gene. A multiple cloning site facilitates the insertion of sequences of interest. The plasmid harbours a kanamycin resistance gene (AphA3). The encoded resolvase gene (resβ) permits to separate plasmid dimers during replication.

6.1.1 DNA concentration and purity

The DNA concentration and purity was determined as follows: 10 µL of DNA were mixed with 490 µL of double distilled water into a quartz cuvette. Absorptions at 260 nm and 280 nm were measured and the concentration was deduced according to the following formula:
1 OD_{260} = 50 µg/mL for double stranded DNA
The DNA was pure when the ratio of OD_{260} / OD_{280} was between 1.7 and 2.

6.1.2 Polymerase Chain Reaction (PCR)

The V8-GluC gene insert was cloned by PCR by addition of the AmyE signal peptide on the amino terminal side and its prosequence (Figure 32). During the PCR, primers hybridize to the DNA template at a specific temperature of annealing, Ta (Mullis et al., 1986; Saiki et al., 1988). This temperature can be calculated in all PCR following to this formula:

Ta = Tm - 5°C

The Tm is the abbreviation of melting temperature and characterizes the stability of the DNA hybrid, which is formed between an oligonucleotide on the forward strand and its complementary on the reverse strand. It is defined as the temperature at which half of a given oligonucleotide is hybridized to its complementary on the reverse strand. It can be assessed by this way:

Tm = 69.3°C + (0.4 x % GC) - 650/length of primer

The “ % GC” term is the percentage of dGTP and dCTP in the primer.

P_{acoA} - SP_{amyE} - Staphylococcus aureus GluC

```
   1 attcagtcaa acgatgcaga ggaactagga agttaaaaag atttccaagg
  51 aaataaatac gtcgatcatt gtcaaaggcc gggtgatatc cggtcttttt
 101 tttgcatgct gtaaaacgag acaaatgaat cagtttgaga caaaacgaga AcoR bin-
 151 cacacgtctc aaactgtctc caaagtgaag atgagaagac tgattttacg ding site
 201 ggctcaaaag actggcacac ttcttgcatt tataatggtg aaccctaaat -24 -12 +1
 251 agaaggaggc gcacaaa atg tttgcaaaac gattcaaaac ctctttactg
 301 ccgttattcg ctggattttt attgctgttt catttggttc tggcaggccc FseI
 351 ggccgctgcg agcgct TTA TCTTCAAAGG CTATGGACAA TCATCCACAA
 451 CAAACGCAGT CAAGCAAACA GCAAACACCT AAGATTCAAA AAGGCGGTAA
 501 CCTTAAACCA TTAGAACAAC GTGAACACGC AAATGTTATA TTACCAAATA
 551 ACGATCGTCA CCAAATCACA GATACAACGA ATGGTCATTA TGCACCCGTA
 601 ACTTATATTC AAGTTGAAGC ACCTACTGGT ACATTTATTG CTTCCGGTGT
 701 AGTTGTAGGT AAAGATACTC TTTTAACAAA TAAACACGTC GTAGATGCTA
 751 CGCACGGTGA TCCTCATGCT TTAAAAGCAT TCCCTTCTGC AATTAACCAA
 751 GACAATTATC CAAATGGTGG TTTCACTGCT GAACAAATCA CTAAATATTC
 801 AGGCGAAGGT GATTTAGCAA TAGTTAAATT CTCCCCTAAT GAGCAAAACA
 851 AACATATTGG TGAAGTAGTT AAACCAGCAA CAATGAGTAA TAATGCTGAA
 901 ACACAAGTTA ACCAAAATAT TACTGTAACA GGATATCCTG GTGATAAACC
 951 TGTAGCAACA ATGTGGGAAA GTAAAGGAAA AATCACTTAC CTCAAAGGCG
1001 AAGCTATGCA ATATGATTTA AGTACAACTG GTGGTAATTC AGGTTCACCT
1051 GTATTTAATG AAAAAAATGA AGTGATCGGA ATTCATTGGG GCGGTGTACC
1101 AAATGAATTT AATGGTGCGG TATTTATTAA TGAAAATGTA CGCAACTTCT
1151 TAAAACAAAA TATTGAAGAT ATCCATTTTG CCAACGATGA CCAACCTAAT
1201 AACCCAGATA ATCCTGATAA CCCTAACAAT CCTGATAACC CTAACAACCC
1251 AGATGAACCA AATAACCCTG ACAACCCTAA CAACCCTGAT AATCCAGACA
1301 ATGGCGATAA CAATAATTCA GACAATCCAG ATGCAGCTTA ATAGtggcca MLuNI
1351 gcaatggaat gcatccacgt gatcgattcg gtcattttgc cacccagcaa
1401 gtaagcagtg accaaagcgg ccaaccgccg cagactgaac tcgaaaatag
1451 ctcgtcggaa acggcgagct attttagttt cttgataccc gctgacg
```

Figure 32. Nucleotide sequence with the promotor, signal peptide, V8-GluC sequence and terminator. The AmyE signal peptide was added to the amino terminal side during the PCR. The gene with the signal peptide was then ligated between the FseI and MlunI restriction sites. This sequence, with the promotor (pacoA), polyhistine tag and the terminator, was cloned in the multisite cloning of the pMSE3 plasmid.

The PCR was performed with an extension time of 1 min/kb. Following quantities were used:

- DNA template: 250 ng
- Pfu Turbo DNA polymerase: 2.5 U
- dNTPs concentrations: 200 µM each dNTP (0.8 mM total)
- Final reaction buffer 1X.
- Primers: each 150 ng/µL (= 0.3 µM). Primer design is shown in figure 33.

1.5 µL of each primer were mixed with 1 µL of dNTP mix, 1 µL of Pfu Polymerase (2.5 U/µL), 5 µL of PCR reaction buffer 10X, 250 ng of template and sterile double distilled water qs to 50 µL. The PCR was performed according to the program in figure 33.

Number of Cycles	Temperature	Duration
1	92°C	1min30s
30	92°C	45 s
	Ta = Tm-5°C	45 s
	72°C	2 min
1	72°C	1 min

Tm: Primer melting temperature

Tm = 58°C

Ta: Annealing temperature of the primers

Ta= Tm-5°C = 53°C

Tm = 69.3°C + (0.4 x % GC) - 650/length of primer

Forward primer

5´ ***GCAA*** ggccggccgctgcgagcgctTTATCT TCAAAGGCTATGGACAATC 3´

Reverse primer

5´ ***GCAA***tggccaCTA TTAAGCTGCAT CTGGATTG 3´

Reverse primer with the polyhistidine Tag

5´***GCAA***tggccaCTA**ATGATGGTGATGGTGATG** TTAAGCTGCAT CTGGATTG 3´

Figure 33. PCR program for Robocycler Temperature Cyclers.

After the PCR-product purification, the DNA concentration was measured and cut by the restriction enzymes. The DNA was again cleaned by the PCR purification kit of Qiagen. In the same way the vector was cut. 1 % agarose gel electrophoresis was run to separate the cut vector from the insert. The vector was extracted from the agarose gel by the Qiaquick gel extraction kit of Qiagen. After measurement of the DNA concentration, the vector and the gene of interest were ligated.

6.1.3 Site directed mutagenesis PCR

Site directed PCR was prepared with the following amount:

- DNA template: 250 ng
- Pfu Turbo DNA polymerase: 2.5 U
- dNTPs concentrations: 200 µM each dNTP (0.8 mM total)
- Final reaction buffer 1X.
- Primers: each 100 ng/µL (= 0.2 µM).

After that the nucleotide codon was replaced by the wanted mutation according to the *Bacillus subtilis* usage codon, primers were designed by anchoring on the upstream and downstream 12-13 amino acids around the position to mutate (Shields and Sharp, 1987). The PCR program is described in figure 34.

Number of Cycles	Temperature	Duration
1	95°C	30 sec
18	95°C	30 sec
	Ta= Tm-5°C	1 min
	68°C	2 min per kilobase
1	4°C	To infinity

Figure 34. Site directed or saturated mutagenesis PCR programs for Robocycler Temperature Cyclers. Primers were designed with the desired mutated nucleotide codon. In the case of saturated mutagenesis, wobble primers were synthesized with degenerated nucleotides at the given position.

At the end of the PCR, 1 µL of *Dpn* I (10 U/µL) were added and the PCR reaction was heated at 37°C for 1 hour. PCR product was purified according to the PCR purification protocol (See chapter 6.1.2.).

6.1.4 Saturated mutagenesis PCR

Degenerated primers at the mutagenic given position were designed likewise the site directed primers. For each mutagenic position, a mix of primers with the twenty amino acid codons at the mutagenic position was synthesized by Thermoelectron GmbH. The same protocol as the site directed mutagenesis was followed.

6.1.5 PCR purification

The PCR product was purified by a kit (Qiagen). The protocol, which is provided by the manufacturer was followed. In brief, 5 volumes of buffer PB were added to 1 volume of the PCR sample mix. A QIA quick spin column was placed onto a 2 mL Eppendorf tube. The sample was applied onto the column and it was centrifuged for 1 min at 13500 RPM. Flow-through was discarded. The QIA prep column was washed by adding 750 µL of buffer PE and the column was centrifuged for 1 min. The flow-through was discarded. The QIA prep column was centrifuged for 1 min at 13500 RPM in a tabletop centrifuge to remove residual wash buffer. The QIA prep column was placed onto a clean 1.5 mL microcentrifuge tube and 50 µL of sterile double distilled water (dd$H_2$0) were added onto the middle of the QIA prep column. After standing 1 min, it was centrifuged again at 13500 RPM in a tabletop centrifuge.

6.1.6 Restriction enzyme cutting by FseI and MluNI

1 µg DNA was mixed with 0.2 µL of restriction enzyme MluNI (1 unit) and 0.2 µL of restriction enzyme FseI (1 unit). 2.5 µL of cut buffer A, which was provided with the MlunI enzyme, were added. The solution was supplemented with 0.25 µL of BSA to get a concentration of 100 µg/mLof BSA in the final solution. The latter was adjusted by double distilled water to 25 µL. Samples were incubated for 1.5 hrs at 37°C.

Final PCR products were digested according to this protocol. Samples were then purified by following PCR purification protocol.

The vector was prepared in a different way as another gene was already cloned in *E. coli*. The insert was cut following the restriction enzyme cutting instruction and then the vector was separated from the insert by a 1 % agarose gel extraction.

6.1.7 Agarose gel electrophoresis

0.5 g of agarose were melted with 50 mL of TBE 1X (TBE 10X: 0.89 M Tris, 0.89 M Boric acid, 0.02 M EDTA disodium, pH 8.4) by microwaving at 700W. The solution was mixed with 5 µL of Ethidium Bromid and then was poured into the agarose gel cassette. 5 µL of DNA, 10 µL of water and 2 µL of DNA electrophoresis buffer (10 mL 50x TAE (121 g Tris, 28.8 g acetic acid, 9.31 g EDTA, pH 8 qs to 500 mL), 25 mL glycin, 15 mL double distilled water, Bromophenol blue and Xylene cyanol) were mixed together. Sample mixture and DNA molecular weight marker X or XIV were loaded onto the gel. The gel was run for 35 min at 95 V and 130 mA.

6.1.8 DNA gel extraction

The DNA was extracted from the agarose gel following to the QIAquick™ Gel Extraction Kit Protocol, using a microcentrifuge. The protocol was designed to extract and purify DNA of 70 base pairs to 10 Kilobases from standard low-melt gels in TAE buffer. Up to 400 mg agarose can be processed by spin column. The DNA fragment was excised from the agarose gel with a clean sharp scalpel. The gel slice was weighed in a colourless tube. 3 volumes of buffer QG were added to 1 volume of gel (100 mg = 100 µL). Tubes were incubated at 50°C for 10 min. Tubes were mixed by vortexing every 2-3 min of incubation. Once the gel slice was completely dissolved, the yellow colour was checked. If it was not yellow but orange or violet, 10 µL of 3 M sodium acetate at pH 5 were added. 1 gel volume of isopropanol was then added into the tube.

After mixing, the solution was pipetted onto a Quiaquick column and centrifuged for one min at 13000 RPM. The flow-through was discarded and 0.5 mL of buffer QG were applied onto the Quiaquick column. After centrifugation for 1 min at 13000 RPM, the column was washed by 0.75 mL of buffer and placed on a clean 1.5 mL microcentrifuge tube and 50 µL of sterile double distilled water were pipetted at the centre of the column. After standing one min, the column was centrifuged for 1 min at 13000 RPM. The recovered DNA was stored at 4°C in the fridge.

6.1.9 DNA ligation

DNA ligation was performed with the DNA rapid ligation kit of Roche. The insert was mixed with a ratio 1:3 (insert:vector) for the blunt ended DNA or a ratio, 1:5 (insert:vector) for sticky end ligation. As the DNA insert was made of both DNA end cutting, a ratio 1/5 (insert/vector) was applied to the protocol. The vector and insert DNA were dissolved in 1X concentrated DNA dilution buffer in a total volume of 10 µL in a sterile reaction vial. The maximum amount of DNA to be ligated was about 200 ng. The quantity of insert was calculated so that the applied volume was inferior than 8 µL (Figure 35).

$$\text{Quantity of insert (ng)} = \text{ratio} \frac{\text{Insert}}{\text{Vector}} \times \frac{\text{Vector quantity} \times \text{insert size}}{\text{Vector size}}$$

Figure 35. Relationship between the quantity of insert and vector for ligation. Quantities were in nanogramms (ng). The size of insert and vectors were in kilobases (Kb).

The T4 DNA ligation buffer was thoroughly mixed and 10 µL were added to the reaction vial. 1 µL of T4 DNA ligase was pipetted and after mixing thoroughly, the vial was incubated for 30 min at 25°C and for 5 min at 37°C. The ligation mixture was directly used for transformation in *E.coli* XL-MRF'.

6.2 Transformation in *Bacillus subtilis*

Competence is defined as the ability of the bacterial cell to take up free high molecular weight DNA (Fani et al., 1981). During the development of the competence, *B. subtilis* cells undergo physiological changes, whose one is the DNA and RNA synthesis decrease and appearance of new proteins (Dooley et al., 1971; Smith et al., 1981).

6.2.1 Plasmid isolation from E. coli and B. subtilis

DNA plasmids were extracted by the QIA prep Spin Miniprep Kit, using a microcentrifuge. The protocol is designed for purification up to 20 µg of high copy plasmid DNA from 1-5 mL cultures of *E. coli*. For the *Bacillus* DNA preparation, a supplementary step is added.

1.5 mL of bacteria culture were centrifuged for 5 min at 6000 RPM in an Eppendorf tube. Supernatant was discarded and pellet was kept. This step was repeated a second time in the same tube as long as there was still some bacteria culture. The bacteria cell pellet was resuspended in 250 µL of buffer P1.

For the DNA isolation from *B. subtilis*, 20 µL of lysozyme (10 mg/mL) were added and the microcentrifuge tube was vortexed for 10 sec two-three times. It was then incubated at 37°C for 10 min.

250 µL of Buffer P2 were added and the tube was gently inverted 4-6 times to mix. 350 µL of Buffer N3 were added and the tube was gently inverted 4-6 times to mix. The solution gets cloudy. The microcentrifuge tube was centrifuged for 10 min at 13500 RPM in a tabletop centrifuge. A white compact pellet was formed. Supernatants were applied onto the QIA prep column by pipetting. The QIA prep column was centrifuged for 1 min at 13500 RPM. The flow-through was discarded. The column was washed by adding 750 µL of buffer PE and the column was centrifuged for 1 min. The flow-through was discarded. The column was centrifuged for 1 min at 13500 RPM to remove residual wash buffer. The QIA prep column was placed onto a clean 1.5 mL microcentrifuge tube and 50 µL of sterile double distilled water was added onto the center of the QIA prep column. After standing 1 min, it was centrifuged again at 13500 RPM in a tabletop centrifuge.

6.2.2 Electroporation in E. coli XL-MRF'

0.5 to 1 µg of DNA amount are required for the transformation in *B. subtilis*. As the V8-Glutamyl endopeptidase cloned gene cannot be directly transformed in *B. subtilis*, it is transformed at first in *E. coli* in the pMSE3 vector shuttle. The V8-Glutamyl endopeptidase gene was transformed first in *E. coli* XL-MRF´.

E. coli competent cells were prepared as follows. 4 mL of *E. coli* medium were inoculated with 40 µL of competent cells, *E. coli* XL-MRF´, in a Falcon tube. Three Falcon tubes were prepared and incubated overnight at 37°C at 230 RPM. Each preculture was harvested in a 2L flask with 500 mL of *E. coli* medium. The flasks were incubated at 37°C at 230 RPM up to reach an optical density of OD_{600nm} = 0.5-0.7. The next steps of the experiments were performed at 4°C in the cold room. The cells were chilled in ice at 4°C and transferred to pre-cooled centrifuge tubes. Cultures were centrifuged at 4200 RPM for 20 min at 4°C. The supernatant was poured off and the pellet was resuspended and washed with 500 mL of cold sterile double distilled water by swirling. Resuspended pellet was centrifuged at 4200 RPM for 20 min at 4°C. This step was repeated a sec time. The pellet was resuspended in 10 % of cold glycerol and stock at -80°C.

Competent cells, *E. coli* XL-MRF´, were defrozen softly from - 80°C on ice. 1 µL of DNA, placed on ice, was mixed with 49 µL of the competent cells in an Eppendorf tube. Then, 50 µL were taken and put into a precooled electroporation cuvette.

The electroporation was performed with 2.5 kV. The cells were resuspended immediately with 1 mL of *E. coli* medium (50 g/L of yeast extract, 5 g/L of NaCl, pH 7) and shaken for 1 hour at 37°C and 180 RPM.

100 µL of the transformed cells were spread out directly on agar plate (20 g/L yeast extract, 5 g/L NaCl, 15 g/L agar and 30 µg/mL Kanamycine sulfate). The remaining 900 µL were centrifuged for 1 min at 6000 RPM. 800 µL of the supernatant were eliminated and the 100 µL rest were mixed with the pellet. Concentrated transformed cells were spread out agar plates. The plates were incubated at 37°C overnight.

6.2.3 Electroporation in B. subtilis

At first, new fresh competent cells were prepared every time. Three Falcon tubes, with 40 µL of WB600 *Bacillus subtilis* glycerol stock and 4 mL of *Bacillus* medium (15 g/L of Yeast extract, 10 g/L of NaCl, pH 7), were then incubated at 37°C at 230 RPM overnight. These precultures were poured into a 300-mL Erlenmeyer in 100 mL *Bacillus* medium the day after and cultivated at 37°C at 230 RPM. When the OD_{600nm} = 1.5 the culture was stopped and incubated at 4°C in ice for 1 hour. The culture was then centrifuged at 6000 RPM for 15 min at 4°C. The pellet was resuspended in 15 mL distilled and sterile cold double distilled water. It was centrifuged again at 6000 RPM for 15 min at 4°C. The pellet was resuspended in 1 mL 30 % PEG-4000 solution and put on ice.

The electroporation cuvette was precooled on ice. 10 µL of plasmid preparation (~0,5 µg) were mixed with 100 µL *Bacillus subtilis* suspension in an eppendorf. The suspension was pipetted into the electroporation cuvette. The electroporation cuvette was then dried and put between the two electrodes. 1 mL SOC Medium was prepared in a pipette and another 1 mL in a Falcon tube. Cells were electropored at 2500 V, 25 µF and 200 Ohms. As soon as possible (<1s), the 1 mL of the SOC Medium in the pipette was added into the cuvette after the electroporation. Then it was pipetted into the Falcon tube. The tube was then incubated for 1 hour at 37°C at 230 RPM.

150 µL of the transformed solution were plated onto agar plates with 10 µg/mL of kanamycin sulfate. The other 850 µL were centrifuged one min at 6000 RPM and the 600 µL of supernatant were thrown away. The 150 µL left were mixed with the pellet and plated onto agar plates with 10 µg/mL kanamycin sulfate. Agar plates were incubated overnight at 37°C.

7. Cultivation and expression of the V8-GluCs

7.1 *Bacillus subtilis* culture and expression of the V8-GluCs

3 Falcon tubes of 4 mL overnight culture of the WB600 *Bacillus* strain were incubated overnight at 37°C and 230 RPM in *Bacillus* medium with 10 µg/µL kanamycin sulfate. Negative controls were taken with the WB600 strain without the plasmid, where the V8-GluC gene was inserted. These cultures were poured into 100 mL *Bacillus* medium in a 300 mL Erlenmeyer with 10 µg/µL of kanamycin sulphate and some antifoam drops the day after and cultivated at 37°c and 230 RPM. When the OD_{600nm} reaches 1, the culture was induced with 0.5 % of acetoine (10 %). Once a mutant was found after the screening, cells were cultivated in 500 mL volume in 2 L flask.

The $OD_{600nm,}$ the enzymatic activity and the protein concentration by Bradford were measured every two hours to follow the cell grow and the V8-GuC expression. The enzymatic activity and the protein concentration by Bradford were measured on 1 mL of the supernatant after 1 min centrifugation at 6000 RPM. Pellets were kept to check by DNA extraction and observation an agarose gel that the cells did not lose the plasmid. The rest of the pellet was lyzed to see if the enzyme was completely secreted or still in the cells. SDS gel was run to follow the expression of the GluC protease. After cultivation, the culture was centrifuged at 4°C at 6000 RPM for 20 min. The pellet was discarded and the supernatant was stored at 4°C before the purification and activation steps.

7.2 SDS-PAGE analysis

SDS-PAGE was performed to separate denaturated proteins following to their apparent molecular weight (Gallagher, 2006). In this technique, proteins are reduced by β-mercaptoethanol or DTT. Pure samples and positive control were diluted with the reducing sample solution of Invitrogen to a concentration of 30 µg/mL. The culture supernatant and the flow-through of the Ni-NTA purification were diluted to a concentration of 200 µg/mL. 14 µL of the samples were mixed with 6 µL of the reducing solution (100 µL of NuPage buffer 4x, 80 µL of ddH_2O, 80 µL of NuPage sample reducing agent 10X). The mixture was heated at 95°C for 10 min. 20 µL of samples were loaded onto the NuPage gel. 20 µL SeeBlue plus 2 Marker were used as Marker. 4-12 % Bis-Tris Nu Page Gel was run with 1X MES buffer. The running conditions were 200 V constant, 35 min for the MES-SDS running buffer or 50 min for the MOPS-SDS running buffer and the expected current was 100-125 mA/gel for starting and 60-80 mA/gel for ending. The gel was placed in 100 mL ultrapure water and microwave on 700 Watts for 1 min. Then the gel was shaken on an orbital shaker for 1 min and the wash was discarded. These steps were repeated twice, to eliminate the buffer salts interference. 20-30 mL SimplyBlue^{TM} SafeStain were added and were microwaved on 700 watts for 45 sec to 1 min. The gel was shaken on an orbital shaker 5-10 min and the stain was discarded. Afterwards the gel was washed in 100 mL ultrapure water on an orbital shaker for at minimum 10 min up to 1 hour for a clear background.

8. Microtiterplate scale set up

The study of the cultivation and of the enzymatic features of the wild-type GluC permits to scale down the cuvette results to the microtiterplate scale for screening. 96-well microtiterplate scale is overviewed in figure 36. After mutant picking in a 96-well MTP, the cells were grown at 37°C overnight. Cells were induced by acetoin the day after and incubated at 37°C for 2 days more. At day+3, aureolysin was added to cut the prosequence and protease activity was measured at day+4 to determine the activity at pH 7.8 but also the specificity and stability. Microtiterplate tests were repeated three times and confirmed with cuvette scale for found mutants.

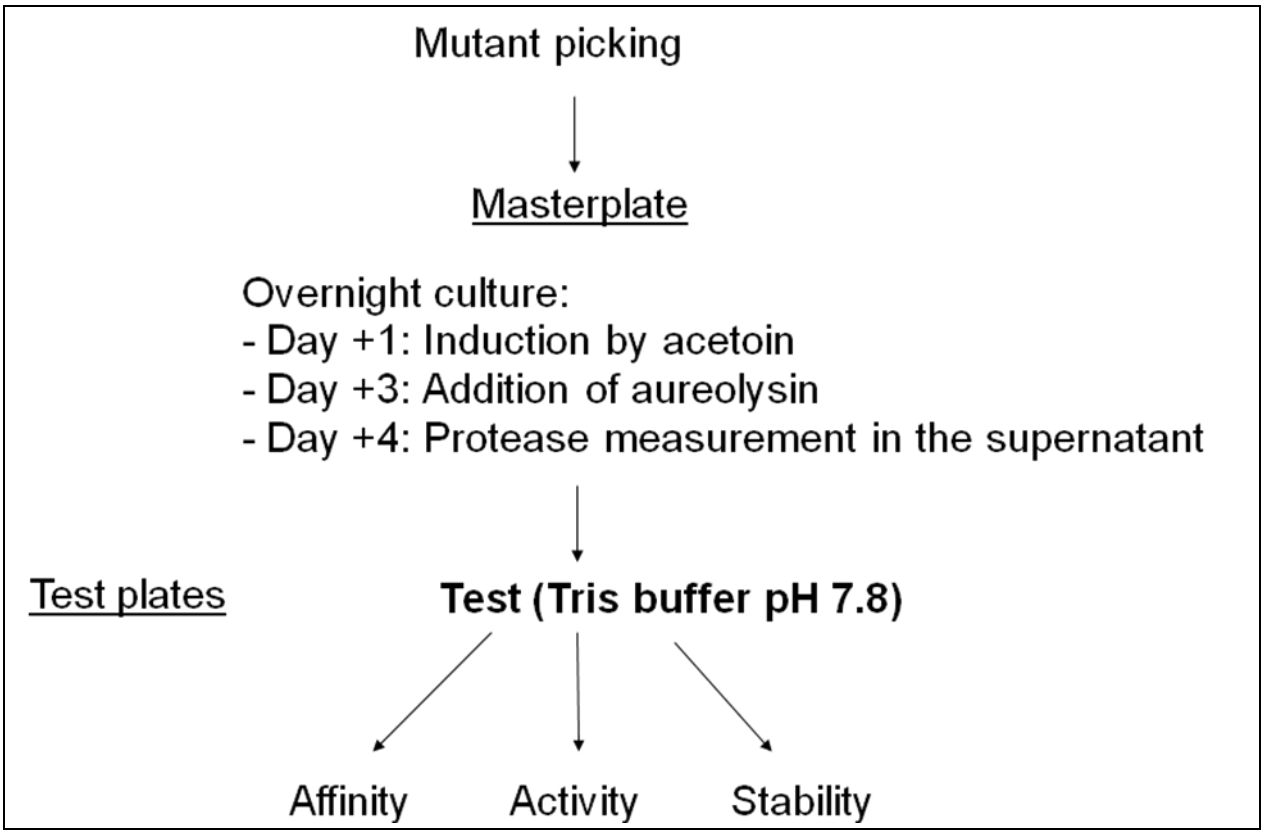


Figure 36. Overview of the High Throughput Screening system implementation.

8.1 Cultivation

Wild type or mutant clones of the V8-GluC were picked with a sterile toothpick in 96-well microtiterplates, which were filled up with 400 µL of *Bacillus* medium and 10 µg/mL of Kanamycin. After an overnight incubation at 37°C, microtiterplate cultures were induced with 50 µL of acetoin (10 %). Microtiterplates were incubated at 37°C for 2 days more. Aureolysin, which was purchased under the lyophilisate form in 20 mM of TrisHCl pH 7.8 and 5 mM of $CaCl_2$, was dissolved in *Bacillus* medium to get a 100 µg/mL final concentration. 100 µL of the Aureolysin solution were added to the microtiterplates and the microtiterplates were incubated at 37°C overnight. Supernatant was directly used for the High Throughput Screening measurement. The rest was stored for 3 days in the fridge and then frozen at -20°C with glycerol (20 % v/v). Wild type V8-GluC was always cultivated in the first row of the microtiterplate as a positive control.

8.2 Cell density measurement

100 µL of the microtiter plate culture were pipette into a transparent 96-well microtiterplate. The microtiterplace was placed into a microtiterplate reader to read the cell density at 600 nm. However, cell density could not be measured after incubation in the 96-well microtiterplates. As incubation lasts 4 days before measuring, cells fall down. Even if culture was again starkly mixed by the pipette robot, cell density cannot be evaluated because of a very strong agglomeration of the cells. Supernatant was completely clear after 4 days of incubation and it can be then directly pipetted and transferred into the test plates.

8.3 Activity measurement

The remaining activity is the percentage of calculated activity in defined experiment conditions by referring to the wild type V8-GluC activity. The screening must be set up according to the wild type V8-GluC enzymatic parameters. These parameters must be defined on a low remaining activity so that the mutant could be detected in the range of the spectrophotometer reader.

Volumes of the enzymatic assay at the cuvette scale were adjusted in order to read an absorbance of 60 mA/min at the microtiterplate scale. The 10 mM glutamate substrate stock was diluted into ethanol to 1 mM. 100 µL of the culture supernatant were diluted to 1/50. 80 µL of 0.1 M of Tris buffer pH 7.8 were added with 40 µL of glutamate substrate (1 mM). The activity of the V8-GluC was measured at an absorbance of 405 nm in the microtiterplate reader during 5 min and 25°C.

8.4 Km-value

The adequate dilution of substrate was looked for obtaining about 30 % of Km-value by using the results of the km-value measurement at the cuvette scale at pH 7.8 in 0.1 M Tris-HCl. This dilution was then used for the V8-GluC measurement at the microtiterplate scale. According to the theoretical pretests of the Roche GluC lyophilisate, a 1/8 substrate dilution was required. However, with this dilution, only 20 % of the Km value were observed after a 1/8 dilution of the recombinant wild type V8-GluC. As a consequence, a smaller dilution at 1/6 of the initial substrate concentration was used at the microtiterplate scale to screen the Km-value at the microtiterplate scale.

8.5 Temperature stress

Microtiterplates were incubated at 51°C during 30 min to get 20 % of remaining activity. The remaining activity is the percentage of calculated activity in defined experiment conditions (buffers and pH) by referring to the wild type V8-GluC activity at 25°C.

9. Purification of the V8-GluCs

9.1 Ni-NTA Spin kit

The Ni-NTA spin column was equilibrated with 600 µL of Lysis buffer (50 mM NaH_2PO_4. H_2O, 300 mM NaCl, 10 mM Imidazole, pH 8) and was centrifuged for 2 min at 1600 RPM with open lid. 600 µL of samples were loaded onto the pre-equilibrated Ni-NTA spin column and centrifuged for 2 min at 1600 RPM with closed lid to reduce the flow rate and thereby extending the binding time. These steps were repeated several times, depending on the total volume of sample to load. The flow-through was saved at -20 °C for further analysis. The Ni-NTA column was washed twice by loading twice 600 µL of wash buffer (50 mM NaH_2PO_4. H_2O, 300 mM NaCl, 20 mM Imidazole, pH 8) and centrifuging it at 1600 RPM for 2 min. The wash was frozen for further investigations at -20°C. The protein was then eluted twice with the elution buffer (50 mM NaH_2PO_4. H_2O, 300 mM NaCl, 250 mM Imidazole, pH 8) after loading 150 µL of elution buffer onto the Ni-NTA spin column and centrifuging it at 2000 RPM for 2 min.

9.2 Ni-NTA chromatography

After cultivation of the V8-GluC in *B. subtilis*, the 500-mL culture was centrifuged for 20 min at 600 RPM at 4°C. The supernatant was collected and stored at -20°C. Buffers were filtrated with 0.40 µm pore diameters and degased for 15 min by stirring. The material column (Ni-NTA agarose) was filtrated and adjusted with the equilibrating buffer and was then degased for 30 min.

The same conditions were used whatever the V8-GluC was (wild type or mutant):

- Material column: Ni-NTA agarose
- Equilibrating buffer: 30 mM NaH_2PO_4, 300 mM NaCl, 10 mM Imidazole, pH 8
- Wash buffer: 30 mM NaH_2PO_4, 300 mM NaCl, 20 mM Imidazole, pH 8
- Eluting buffer: 30 mM NaH_2PO_4, 300 mM NaCl, 250 mM Imidazole, pH 8

Buffer pHs were adjusted with 2 M of NaOH

- Binding capacity: 1:50 (protein : matrix)

After each step, a sample was collected for measuring was volume and its V8-GluC activity at pH 7.8 with the glutamate substrate at the cuvette scale. The protein concentration was also determined by the Bradford test and samples were kept for performing a SDS-PAGE.

A balance chart summarized the progress of the purification by assessing the yield and the purity. The calculations of the yield and purity are always related to the first step, which was here the cultivation. The yield is the percentage between the purification step and this initial one, by comparing the total units. The purity factor depends on the specific activity of each purification step. It corresponds to this ratio between the purification step and the cultivation step.

9.2.1 Purification of the wild type V8-GluC by Ni-NTA

The first purification of the wild type V8-GluC was run under the following conditions. The flow rate of the buffer was 7 mL/min. The chromatography was performed at a temperature of +4°C. 156.6 mg of proteins were loaded onto the column, whose volume was 15 mL. 70 mL of eluted V8-GluC were collected at the end of the process.

The purification was repeated under the same conditions to check the reproducibility. According to the specific activity and SDS-PAGE gel, all the fractions with the best purity were collected.

9.2.2 Purification of the mutant V8-GluC by Ni-NTA

The mutant V8-GluC was purified similarly as the wild type GluC. Buffers and sample were loaded onto the 20 mL column with a flow rate of 5 mL/min. The initial protein concentration was 172.5 mg and 37 mL volume of the purified mutant V8-GluC were eluted at the end. The purification of the mutant was reproducible.

The final step of purification was the dialysis of the V8-GluCS. The samples were put into a dialysis tube with a cut off of 10 kDa. The sample was dialyzed overnight at 4°C against a buffer of 50 mM Tris-HCl buffer, pH 8 in a ratio: 1:1000, which was continuously slightly stirred.

10. Biochemical characterization and analytics

10.1 Activation of the V8-GluC

The V8-GluC harbours a prosequence on its amino-terminal side to avoid its autodegradation. Once this prosequence is cleaved, the V8-GluC is active.

10.1.1 Activation of the V8-GluC by addition of GluC

The prosequence was tried to be cleaved by the addition of other GluC as the last amino acid of the prosequence was mutated to a aspartate. 0.5 U, 1 U and 3 U from a 1 mg/mL stock solution of Roche or NEB GluC were added to 50 mL of the wild type V8-GluC supernatant after cultivation. Samples were incubated for 1, 2, 3, 4 hours or 2 days at 25°C. Enzymatic activity was measured according to the chapter 5.1 in Tris-HCl buffer (0.1 M) at pH 7.8 with the Z-Phe-Leu-Glu-pNA substrate. Remaining activity was calculated by referring to the first measured activity before starting the GluC addition. A positive control was used without GluC addition. In this case, sterile double distilled water was added instead of the GluC stock solution.

10.1.2 Activation of the V8-GluC by temperature incubation

To activate the prosequence cutting, 1.5-mL dialyzed samples of the purification step were pretested and incubated for several days at different temperatures: 25°C, 30°C and 37°C. Enzymatic activity was measured in Tris-HCl buffer, pH 7.8 with the Z-Phe-Leu-Glu-pNA substrate and samples were sent to the ESI-MS department to determine the molecular weight once the enzymatic activity measurement was stable.

According to these results, the purified culture was activated at 37°C up to get a stable enzymatic activity. Time of activation depends on the enzymatic activity measurement. Once the enzymatic activity remains stable, incubation was stopped and a sample fraction was sent to the analytic department to check the molecular weight by ESI-MS.

10.2 Blot for N-Terminal sequencing

N-Terminal sequencing of the V8-GluCs was performed in an analytic department of Roche after the purification step. Blot membrane was prepared after the SDS gel step. The Blot was then stained by Coomassie Blue safe stain or Amido Black. SDS gel was run in the same way as in chapter 7.3. After washing the SDS gel with double distilled water, the gel was blotted in the iBlot^{TM} Device with blotting roller of Invitrogen GmbH. Once the lid of the device was opened, the anode stack without sealing was placed at the bottom with the tray directly at the blotting surface. The pre-run gel was placed on the PVDF membrane of the anode stack. The deionized pre-soaked filter paper was placed on the pre-run gel. The sealing of the cathode stack was removed and placed at the top over the pre-soaked filter paper with the electrode side facing up. Air bubbles were removed by the blotting roller at each step. The disposable sponge was then placed with the metal contact on the upper right corner of the lid. Blot was run for 6 min.

Membrane was stained for 5 min in Amido Black (0,1 % Amido Black, 40 % ethanol, 10 % acetic acid qs 1 L ddH_2O). The membrane was then destained for 10 min in destaining solution (40 % ethanol, 10 % acetic acid qs 1 L water). This step was repeated until the protein bands can be observed. The membrane was then dried overnight.

For Coomassie staining, the membrane was stained for 1 hour in Coomassie Blue safe stain by gentle shaking. Then it was unstained by ddH_2O by gentle shaking until the protein bands can be observed. The membrane was then dried overnight and sent to the analytic department of Roche GmbH for N-Terminal sequencing.

10.3 Storage stability after freezing at -20°C

500 µL of the V8-GluCs were frozen at -20°C and defrozen in waterbath at 25°C. This experiment was repeated 5 times. 5 mM of Glu-Glu dipeptide were added or not as stabilizer. Activity assays were performed in Tris-HCl buffer (0.1 M) at pH 7.8 with the Z-Phe-Leu-Glu-pNA substrate as described in chapter 5.1. Remaining activity was calculated compared to the first measurement before freezing samples at -20°C.

10.4 Specific activities in various buffers

Specific activity measurements in Tris-HCl buffer, pH 7.8 and ammonium acetate buffer, pH 4.3 were described in the chapter 5.1. Specific activities of the Roche, NEB, wild type and mutant GluCs were compared in different buffers. They were assessed in Tris buffer (Tris-HCl 0.1 M) and phosphate buffer (KH_2PO_4 0,1 M) at pH 7,8 with the Z-Pheu-Leu-Glu-pNA substrate (10 mM). Specific activities were also determined in phosphate buffer (KH_2PO_4 0.1 M) at pH 5 and in ammonium acetate buffer (50 mM) at pH 4.3, both with the hemoglobin peptide substrate (10 mM).

10.5 Km value assessment

Km-values of the Roche, NEB, wild type and mutant GluCs were determined in the same way as in chapter 5.2. in Tris HCl buffer (0.1 M) at pH 7.8 and with the Z-Pheu-Leu-Glu-pNA as substrate.

10.6 Temperature stress

Temperature stress was repeated in the same way as in chapter 5.3 with the wild type and the mutant V8-GluCs and with or without addition of 5 mM of Glu-Glu as stabilizer before the 30 min incubation. Activities were measured in Tris-HCl buffer (0.1 M) at pH 7.8 with the Z-Phe-Leu- Glu-pNA substrate. A temperature range between 25°C and 55°C with 5°C steps was tested. Remaining activity was determined compared to the 30 min incubation at 25°C.

10.7 Melting temperature assessment by DSC measurement

DSC (Differential Scanning calorimetry) measurement was performed in another laboratory of Roche Diagnostics GmbH. Roche, NEB, E-Wt-H and E-M-H GluCs were diluted to the same concentration of 0.3 mg/mL with 1 mL of Tris buffer 50 mM at pH 8. Lysozyme (1 mg/mL) was taken as standard.

10.8 pH optimum

The activity measurement protocol, which is described in chapter 5.1, was followed but different buffers were used to compare different pH ranges between 4 and 9 with 0.5 unit steps. The activity was measured with 50 mM of ammonium acetate for the pH range of 4 to 6. Activity measurements between pH 6 and 7.5 were performed in phosphate buffer (KH_2PO_4, 0.1 M).

The activity at pH points between pH 7.5 and 9 were assessed in Tris-HCl buffer (0.1 M) between pH 7.5 and 9. Two supplementary measurements were also performed at pH 7.8 with phosphate and Tris-HCl buffers. Activity assays were measured with the hemoglobin hexapeptide substrate because the Z-Phe-Leu-Glu-pNA cannot be dissolved correctly at acidic pHs and it forms a white precipitate in the cuvette, which interferes with the absorbance detection.

Remaining activity was assessed with the highest measured activity for the wild type and mutant V8-GluCs.

10.9 Hemoglobin functional test by ESI-MS

The functional test of the IFCC by ESI-MS was experimented with a ratio hemoglobin/enzyme of 100. Hemoglobin was not glycated (HbA0). The positive control, the Roche GluC, was freshly dissolved in double distilled water to a 0.2 mg/mL concentration.

50 µL of the Roche GluC stock solution (0.2 mg/mL) were then mixed with 2 mg/mL of hemoglobin and 50 mM of ammonium acetate buffer at pH 4.3 to get a final volume of 500 µL. Other GluC NEB, wild type or mutant V8- GluCs) were prepared in a volume to get a final concentration of 0.02 mg/mL. The V8-GluC was mixed with 2 mg/mL of hemoglobin (final concentration). Ammonium acetate (50 mM) at pH 4.3 was added to reach a total reaction volume of 500 µL. Samples were incubated at 37°C for 18 hours. The samples were then frozen at -20°C before being injected onto the ESI-MS device in the analytics department.

10.10 pH stability in 50 mM ammonium acetate pH 4.3 at 37°C for 18 hours

The Roche GluC and the mutant V8-GluCs were prepared in the same way as the Hba0 functional test by ESI-MS (Chapter 10.10). The hemoglobin volume was replaced with an equivalent volume of ammonium acetate buffer (50 mM) at pH 4.3. Samples were incubated overnight for 18 hours at 37°C and the reaction was then stopped into ice. Activity was measured in Tris-HCl buffer at pH 7.8 with the Z-Phe-Leu-Glu-pNA substrate as it was described in chapter 5.1. Remaining activity was calculated as percentage of the activity by referring to the activity before starting the incubation.

RESULTS

1. Cloning of the V8-GluC

The naked genomic DNA of the V8-Glutamyl endopeptidase was cloned with its wild type prosequence and a polyhistidine tag in the pMSE3 shuttle-vector and was then transformed in *B. subtilis*.

The V8-GluC was first cloned with its *S. aureus* wild type signal peptide. Any V8-GluC could be expressed extracellularly with this signal peptide as any enzymatic activity could be detected and any protein band could be observed on the SDS-PAGE gel. The wild type V8-GluC was finally cloned and expressed with the AmyE signal peptide of *B. subtilis* and with a polyhistidine tag in the WB600 *B. subtilis* strain.

Different forms of the enzyme were found in the culture supernatant following to different cultivation time. The complete wild type GluC, which was cloned with a mutation in the prosequence to facilitate its autoproteolysis, could be observed at a 0.50 U/mL.OD activity. The intact E-M-H GluC was detected at a 0.10 U/mL.OD activity. However, observed activities during the cultivation were not the real activities of the wild type and G166I mutant GluCs because of the prosequence activation lack.

In parallel, another WB600 *B. subtilis* strain was prepared by integrating directly the aureolysin gene in the *B. subtilis* genomic DNA. Even if the transformation of the cells with the V8-GluC gave the formation of viable clones, these clones could not be cultivated as cells started dying after acetoin induction. The aureolysin gene integration in the *Bacillus subtilis* genomic DNA had obviously a toxic effect on the cells.

2. Activity test assay set up

Two enzymatic test assays were implemented according to the reaction principle of the V8-GluC on a chromophore substrate.

At first, the six first amino acids of the hemoglobin without glycosylation were synthesized with a paranitroanilide bond at the carboxyl-terminal end. This substrate enables to check if the GluC cuts correctly and specifically after the sixth amino acid of the hemoglobin beta-chain: the glutamic acid. As the substrate could not be glycated, it is consequently called HbA0 and not HbA1c. This substrate was denominated “hemoglobin hexapeptide substrate” and is abreviated as “HbA0(1-6)-pNA” for the rest of the project. It represents the following amino acid pattern: (NH2-VHLTPE-pNA).

The hemoglobin hexapeptide substrate costs however 200-400 euros/milligram even if it was synthesized at Roche. Due to this high material cost and limited supply, another substrate, which is glutamate specific and normally produced by Roche in larger amount with a cost of 200 euros/gram, was preferentially used to check the glutamic acid cleavage specificity of the V8-GluC. It consists in a Carbobenzoxy-Phenylalanine-Leucine-Glutamate chain (Z-Phe-Leu-Glu-pNA), whose end is linked to a paranitroanilide. As this substrate is produced at Roche, the supply was guaranteed and it was also economically much cheaper compared to the hemoglobin hexapeptide substrate but its main drawback came from its insolubility in acidic pHs. In this case, it forms a white opaque cloud, which prevented the spectrophotometer ray to pass correctly through it and did not allow the activity measurement at acidic pHs. The substrate is abbreviated under the name, “glutamate substrate” or “Z-Phe-Leu-Glu-pNA”, for the project.

The V8-GluC has no calcium binding site and it is therefore not necessary to add calcium in the enzymatic test as it usually performed with other proteases like metalloproteases (Potempa et al., 1989).

The GluC activity unit was defined as the needed amount of V8-glutamyl endopeptidase (Roche, NEB, wild-type or mutant V8-GluCs), which turns one micromole of HbA0 or glutamate substrate (HbA0(1-6)-pNA or Z-Phe-Leu-Glu-pNA) into HbA0(1-6) or Z-Phe-Leu-Glu + paranitroaniline per minute at 25°C and 405 nm, either with Tris-HCl 0,1 M at pH 7,8 buffer or ammonium acetate buffer 0.050 M at pH 4.3. As the reaction had stoechiometric coefficients, the activity could be calculated from the Beer-Lambert law.

As the Roche GluC lyophilisate is usually taken for the HbA1c test, its activity was determined to use as a reference for the phD project. The Roche GluC has a lyophilisate activity of 40 U/mg Lyo and a specific activity of 40 U/mg in 0.1 M Tris-HCl buffer at pH 7.8 with the Z-Phe-Leu-Glu-pNA substrate.

As a summary, the enzymatic test assay could be correctly set up but only with defined experimental conditions, Tris-HCl buffer at pH 7.8 and with Z-Phe-Leu-Glu-pNA as a substrate. The Roche GluC lyophilisate was used as positive control with a specific activity of 40 U/mg in this enzymatic test assay for all further activity test measurements. However, the experimental conditions of the activity test differ from the HbA1C functional test.

3. High Throughput Screening system implementation

Once activity tests were defined, a screening system was established according to the wild type V8-GluC to target a faster variant among screened libraries.

The screening was implemented according to the Roche GluC lyophilisate enzymatic features at first at the cuvette scale. The enzymatic characteristics were then compared to the recombinant wild type V8-GluC to adjust them at the 96-well microtiterplate scale (Figure 37).

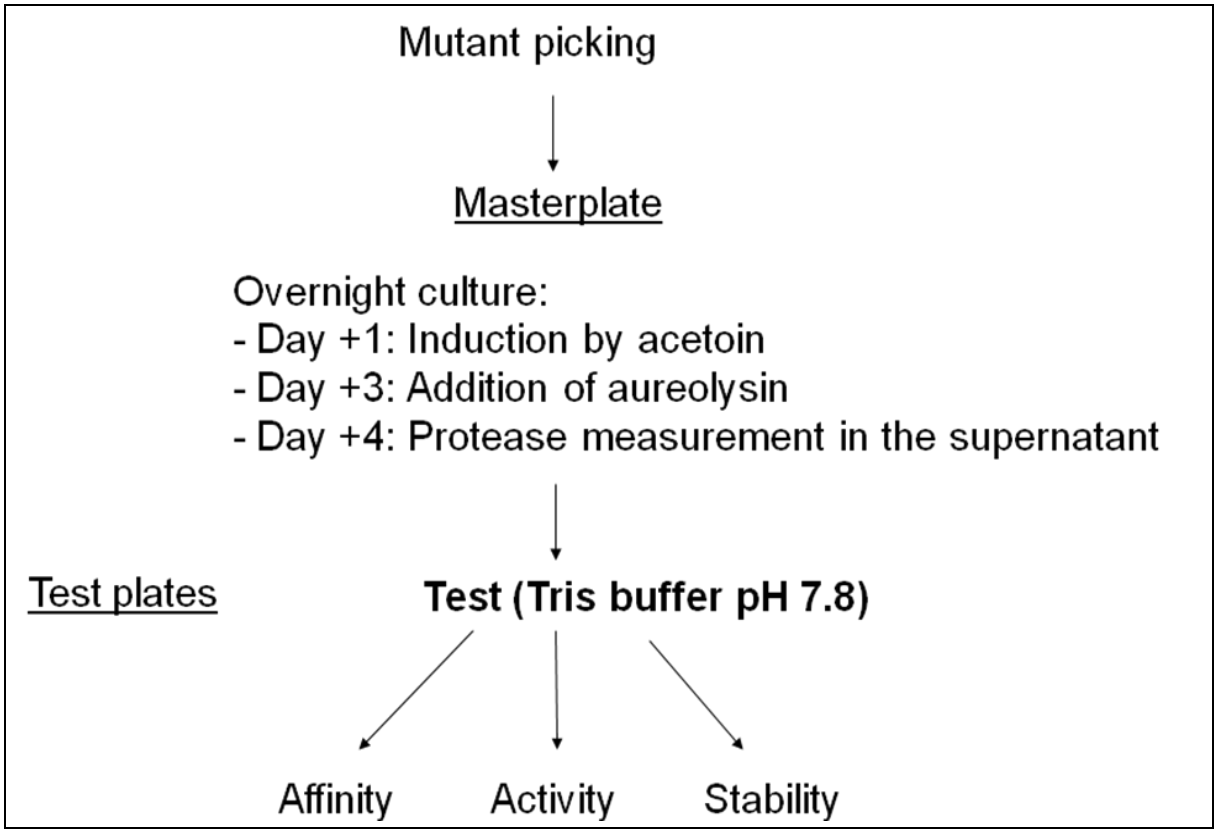


Figure 37. Schematic representation of the High Throughput Screening system implementation. After mutant picking in a 96-well MTP, the cells are grown at 37°C overnight. Cells are induced by acetoin the day after and incubated at 37°C for 2 days more. At day+3, aureolysin is added to cut the prosequence and protease activity is measured at day+4 to determine the activity at pH 7.8 but also the specificity and stability.

3.1 Cultivation and prosequence activation

Wild type V8-GluC cultures were taken as positive control for the cultivation and prosequence activation in the 96-well microtiterplate format. After acetoin induction, cell cultivation was stopped once the enzymatic activity had reached 60 mA/min. This activity was observable two days after the acetoin induction. 0.01 µg of aureolysin were then subsequently added and cells were incubated overnight at 37°C to check if a higher enzymatic activity due to the prosequence processing could be measured. This was not the case, but the aureolysin addition step was kept, as it is normally necessary for the prosequence processing (Drapeau, 1978).

As the incubation lasted 4 days before measuring, cells precipitated. Even if the culture was again starkly mixed by the pipette robot, cell density could not be evaluated because of a very strong agglomeration of the cells. The supernatant was completely clear after 4 days of incubation and it could be then directly transferred into the test plates.

3.2 Activity measurement

The main enzymatic parameter of the screening system is the activity detection. This led to the maor problem of the high substrate consumption for the screening system. As the production of the hemoglobin as substrate cost too much, it could not be used for screening. The Z-Phe-Leu-Glu-pNA substrate was taken for screening. Alternatively, the screening system was set up with the Tris buffer at pH 7.8 as the Z-Phe-Leu-Glu-pNA substrate precipitates at acidic pHs, as for example in ammonium acetate pH 4.3. It was decided that if a mutant was found with a better activity in Tris buffer at pH 7.8 with the Z-Phe-Leu-Glu-pNA substrate at the microtiterplate scale, a test with the hemoglobin substrate will be carried out afterwards at the cuvette scale in the conditions of the mass spectrometry measurement: 50 mM ammonium acetate at pH 4.3 with the hemoglobin substrate.

The enzymatic assay at the cuvette scale was adjusted in order to read an absorbance around 60 mA/min at the microtiterplate scale. This absorbance was chosen to get a flexible range of activity to detect mutants with a higher activity. An activity of 40 mA/min was measured after transfer of the activity assay with the Roche GluC from the cuvette scale to the microtiterplate scale with the wild type V8-GluC.

3.3 Km-value

Variant libraries were also screened according to their affinity, the Km-value, towards the substrate. For the Km-value screening establishment, the right dilution of substrate was chosen to get about 30 % of the Km-value at the microtiterplate scale. The Km-value was assessed with the Z-Phe-Leu-Glu-pNA as a substrate in Tris buffer at pH 7.8 with the Roche GluC at the cuvette scale (Figure 38). On the Km curve, the reaction is observed at first to be proportional to the quantity of substrate: the more substrate is added in the reaction, the faster the enzyme catalyzes the substrate. Then, it gets the saturation point where the enzyme works with a maximal speed, Vmax. The Km value was calculated around 0.40-0.50 mM for the Roche GluC in these experimental conditions according to the lineweaver Burk plot. As the initial substrate concentration of the cuvette scale protocol was around 1 mM, its dilution of 1/8 was supposed to reach the 30 % of the Km value for the microtiterplate scale-down. However, these 30 % could be observed by the microtiterplate reader after a 1/6 dilution of the cuvette scale substrate concentration. This difference is due to the difference of scale from the cuvette to the microtiterplate. The absorbance is measured through the width of the cuvette whereas it is through the well thickness at the microtiterplate scale.

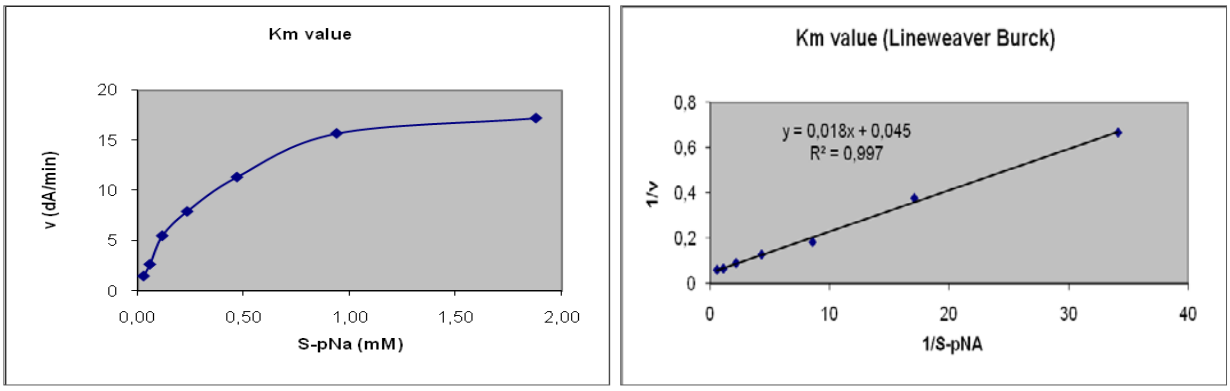


Figure 38. Km-value curve and Lineweaver Burk plot. Km-value was assessed at 0.40 mM of Z-Phe-Leu-Glu-pNA substrate (S) in Tris buffer at pH 7.8 with the Roche GluC.

3.4 Temperature stress

The temperature stability of an enzyme belongs to the enzymatic features, which can be affected severely after mutagenesis. This parameter must be therefore implemented in the screening system to control its possible fluctuations. Temperature stress was modelled by measuring the Roche GluC activity in Tris buffer at pH 7.8 with the Z-Phe-Leu-Glu-pNA substrate after 30 minutes of incubation at different temperatures in a range of 25 to 60°C with 5°C steps (Figure 39). 1°C step was tried between 50 and 55°C to get a detailed view of the remaining activity of the Roche GluC after the activity decreased to 40 %. The Roche GluC is no more stable after incubation at 40°C and a remaining activity of 20 % is observed at a temperature of 52°C. After transfer from the cuvette to the microtiterplate scale, a remaining activity around 5-10 % was observed after incubation at 51°C for 30 minutes. This difference is due to the incubation way. Temperature stress experiment was achieved in waterbath at the cuvette scale whereas it was performed in metal plates at the microtiterplate scale. As this measurement was also not uniform on the whole microtiterplate, the measurement was repeated three times to confirm the measurement repeatability. The incubation was performed at 51°C for the temperature stress screening.

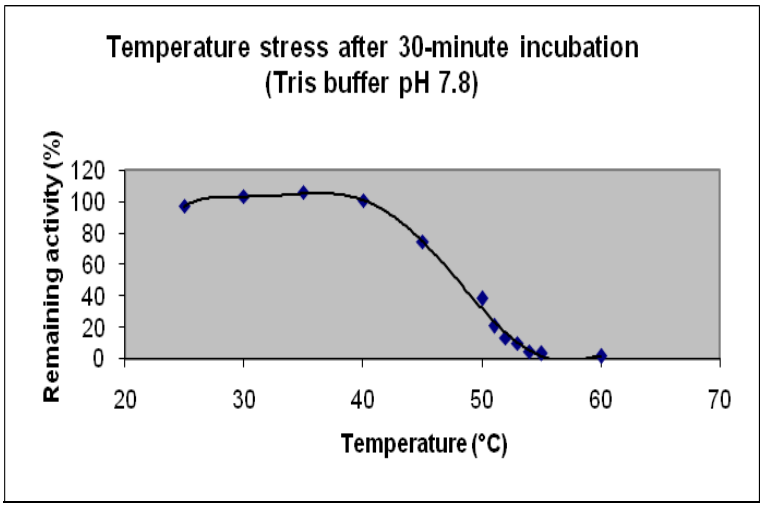


Figure 39. Temperature stress model. The Roche GluC activity was measured in Tris buffer at pH 7.8 with the Z-Phe-Leu-Glu-pNA substrate after 30 minutes of incubation at different temperatures in a range of 25 to 60°C with 5°C steps. The remaining activity was calculated by referring to the incubation at 25°C.

3.5 Specificity

As the V8-GluC can cleave aspartate and glutamate substrates, the possible specificity change of the enzyme should have been also checked during the screening. The specificity of the V8-GluC cleavage, with a aspartic acid synthetic substrate instead of a glutamic acid one, could not be also implemented in the screening assay because of substrate costs. The specificity was chosen to be tested with the hemoglobin functional test by ESI-MS afterwards.

As a conclusion of the screening system implementation, it was established on three enzymatic parameters of the wild type V8-GluC: the enzymatic activity, the Km-value and the temperature stability. The enzymatic activity was set up with an absorbance of 40 mA/min, the Km-value on 30 % of its value and the temperature stability on 5-10 % of remaining activity. They were however implemented with different buffer, pH and substrate conditions than the HbA1C functional test. Furthermore, the specificity parameter was decided to be tested afterwards on the hemglobin functional test by ESI-MS.

4. Mutagenesis strategies

Directed evolution was performed on the wild type V8-GluC in order to generate libraries of variants. Strategic amino acids to mutate were chosen at first according to the crystal structure of the already-published Roche V8-GluC (Prasad et al., 2004). Some mutagenesis positions were decided after protein modelling of the V8-GluC with the hemoglobin substrate and some amino acid positions were defined according to the homology sequence comparisons of the recombinant wild type V8-GluC with the Roche GluC and glutamyl endopeptidases from other organisms.

4.1 Mutation of the V8-GluC into the Roche GluC

First site directed mutagenesis were focused on the differences between the Roche lyophilisate GluC and the recombinant V8-GluC (Carmona and Gray, 1987; Drapeau, 1978). The Roche lyophilisate GluC differs from the recombinant V8-GluC by eleven amino acids. These differences are underlined in red in the figure 40. The eleven amino acids were exchanged by substitution, deletion or insertion. Once the complete gene was mutated, the Roche lyophilisate GluC was expressed to compare its activity to the recombinant V8-GluC. In parallel, saturated mutagenesis was tried out on each eleven different positions separately to check if one of the amino acid positions of the Roche GluC can affect the V8-GluC enzymatic parameters. Any clones of the Roche GluC showed an enzymatic activity comparable to the recombinant wild type V8-GluC. Their measured screening activities were lower than 40 mA/min.

The hemoglobin functional test showed that the specificity of both GluCs is similar. To improve the V8-GluC rather than the Roche GluC is in this case more appropriate. The hemoglobin functional test based screening did not help to explain the sequence differences between the Roche GluC and the V8-GluC.

```
Roche GluC   LSSKAMDNHP  QQTQSSKQQT  PKIQKGGNLK  PLEQREHANV  ILPNNDRHQI TDTTNGHYAP
V8-GluC      LSSKAMDNHP  QQTQSSKQQT  PKIQKGGNLK  PLEQREHANV  ILPNNDRHQI TDTTNGHYAP

Roche GluC   VTYIQVEAPT  GTFIASGV_VG KDTLLTNKHV  VDAT_GDPHAL KAFPSAINQD  NYPDGGFTAE
V8-GluC      VTYIQVEAPT  GTFIASGVTVG KDTLLTNKHV  VDATHGDPHAL KAFPSAINQD  NYPNGGFTAE

Roche GluC   QITKYSGEGD  LAIVKFSPNE  QNKHIGEVVK PATMSNNAET QVDQNITVTG  YPGDKPVATM
V8-GluC      QITKYSGEGD  LAIVKFSPNE  QNKHIGEVVK PATMSNNAET QVNQNITVTG  YPGDKPVATM

Roche GluC   WESKGKITYL  KGEAMQYNLS TTGGNSGSPV  FNEKNEVIGI  HWGGVPNQFD  GAVFINNEVN
V8-GluC      WESKGKITYL  KGEAMQYDLS TTGGNSGSPV  FNEKNEVIGI  HWGGVPNEFN  GAVFINENV

Roche GluC   RNFLKQNIED  IHFANDDQPN  NPDNPDNPNN  PDNPNNPDEP NNPDNPNNPD  NPDNGDNNNS
V8-GluC      RNFLKQNIED  IHFANDDQPN  NPDNPDNPNN  PDNPNNPDEP NNPDNPNNPD  NPDNGDNNNS

Roche GluC   DNPDAA
V8-GluC      DNPDAA
```

Figure 40. V8-Glutamyl endopeptidase sequences. In green: prosequence, in black: Roche lyophilisate GluC sequence (Drapeau, 1978), in blue: recombinant V8-GluC sequence (Carmona and Gray, 1987), in red: amino acid differences, in yellow: catalytic triad amino acid. A first a V8-GluC was sequenced in 1978 by Drapeau. This V8-GluC is sold as the Roche lyophilisate GluC. The second sequence, which differs by eleven amino acids, was discovered by Carmona in 1987 and is the recombinant wild type V8-GluC.

4.2 Mutagenesis after protein modelling and sequence alignment

However, protein modelling and sequence alignment of different glutamyl endopeptidase activity sites and specificity pockets enabled to speculate on other amino acids to exchange (Barbosa et al., 1996). Glutamyl endopeptidases from *Staphylococcus epidermis*, *Staphylococcus warnerii*, *Bacillus subtilis* and *licheniformis* and *Streptomyces griseus* present respectively 42, 68, 17, 19 and 24 % of homology with the Roche GluC. According to their homology study, the tryptophan W185, which is called 214 in the Barbosa's publication, plays an important role because of its vicinity with the catalytic triad. This residue should be important to stabilize the aspartic acid of the catalytic triad (Warshel et al., 1989). As this tryptophan is a threonine in the *Bacillus* glutamyl endopeptidases, the tryptophan was mutated into a threonine in the recombinant wild type V8-GluC to observe the effect on the enzyme parameters.

A very important tryptophan amino acid is also quoted at the position 94 in the *Bacillus* glutamyl endopeptidases, to form an aromatic lid for the protection of the catalytic site. This position is between the positions 89 and 90 of the V8-GluC and a tryptophan amino acid was inserted at this position. As this lid is also stabilized by hydrogen bonds of two other amino acids, these positions were also substituted to smaller amino acids for reasons of steric hindrance. These hydrogen bonds are formed between the amino acid 94 and the amino acids 55 and 56 in the *Bacillus* glutamyl endopeptidases. These positions correspond to the amino acids 49 and 50 in the recombinant wild type GluC sequence. The following substitutions were combined, N49A and K56G, to reproduce the same bonds as in the *Bacillus* glutamyl endopeptidases.

All these mutations were checked out but any of them gave a positive effect on the wild type V8-GluC enzyme parameters. The measured absorbance was always under 40 mA/min.

4.3 Active site and mutagenesis

A serine protease catalytic triad is always characterized by three amino acids, an aspartic acid, a serine and a histidine, to form an oxyanion binding hole, into which the substrate binds (Kraut, 1977). These amino acids are the aspartic acid 93, histidine 51 and serine 169 in the Roche GluC sequence (Prasad et al. 2004) The serine 169 of the catalytic triad was recently replaced by an alanine to show the importance of the serine 169 in the catalytic activity of the V8-glutamyl endopeptidase (Nemoto et al., 2008). However, according to the crystal structure study, three other amino acids play an additional important role in the charge relay system for the proton transfer. The valine 1, threonine 164 and asparagine 193 help in the substrate stabilization and catalysis in the active site as it is described in figures 41 and 42. The valine 1 was recently demonstrated as essential for the activity of the V8-GluC since the enzyme is deactivated without this amino acid (Nemoto et al., 2008).

Saturated mutagenesis was carried out on the catalytic triad to check the impact on the enzyme parameters: histidine 51, aspartic acid 93 and serine 169.

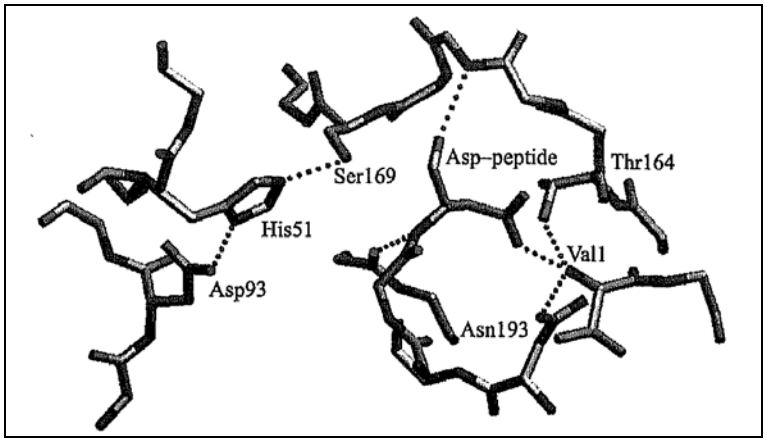


Figure 41. Catalytic triad site structure with an aspartate peptide as substrate. The valine 1, threonine 164 and asparagine 193 facilitate the aspartate substrate stabilization with the catalytic triad, the aspartic acid 93, the histidine 51 and the serine 169.

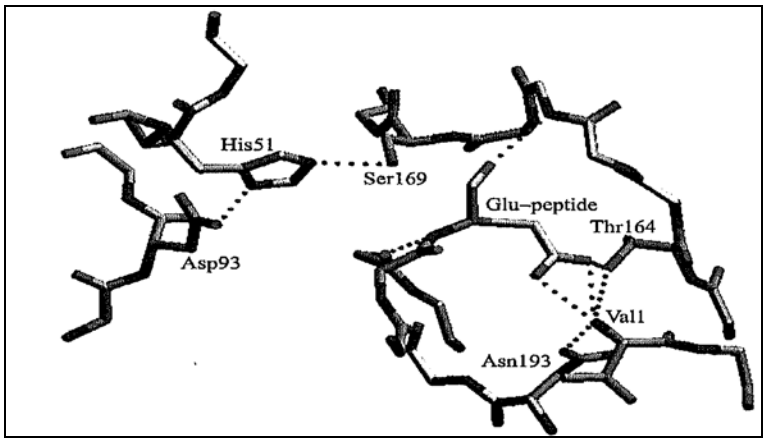


Figure 42. Catalytic triad site structure with a glutamate peptide as substrate. The valine 1, threonine 164 and asparagine 193 help to stabilize the glutamate substrate in the catalytic site, the aspartic acid 93, the histidine 51 and the serine 169.

As it was expected, no enzymatic activity could be detected anymore as the catalytic triad was directly changed in its struture itself. In the same way, the V8-GluC lost its full activity after saturated mutagenesis on the three other strategic amino acids: valine 1, threonine 164 and asparagine 193, which interact directly with the substrate. Hydrogen bonds may have been completely suppressed or not strong enough to stabilize and catalyze the substrate.

As the catalytic site amino acids and the three other amino acids: histidine 51, acid aspartic 93, serine 169 and valine 1, threonine 164 and asparagine 193, play an important role in the substrate catalysis, effects of saturated mutagenesis were also studied up to three amino acids upstream and three amino acids downstream from these strategic positions. Saturated mutagenesis was tried around the triad amino acids at the following positions 48, 49, 50, 52, 53 and 54 around the histidine 51, the positions 90, 91, 92, 94, 95 and 96 around the aspartic acid 93 and the positions 166, 167, 168, 170, 171 and 172 around the serine 169.

This was also performed on the other strategic positions, which were described according to the structure study of the Roche GluC (Prasad et al., 2004). Three amino acids, the valine 1, threonine 164 and asparagines 193, interact directly with the catalytic triad in order to stabilize the substrate in the catalytic site. The positions 2, 3 and 4 were checked out around the valine 1, the 3 amino acids upstream and downstream, 161, 162, 163, 165, 166 and 167 around the threonine 164 and as well as the positions 190, 191, 192, 194, 195 and 196 around the aparagine 193. These positions are based on the Roche GluC crystal structure as the V8-GluC crystal structure has not been studied and/or published up to now.

A protein modelling was designed with the V8-GluC and the hemoglobin substrate in a department of Roche GmbH with the Program Insight/Discover from the company Accelrys, Inc., San Diego. Previously quoted positions were confirmed as interesting positions after the protein modelling. The valine 1 was recognized as a very important amino acid because of its direct position at the entrance of the catalytic site. A substitution with a smaller acid was recommended like a glycine for example (V1G). The other amino acids, which could potentially play a role in the catalysis, are at the positions 2, 164, 165, 166, 167, 168, 184, 185, 186, 187, 188, 189, 191, 193 and 194 because of their vicinity with the catalytic site. These positions cross check the earlier hypothesis of the structure study in the literature. Only positions 164, 165, 184, 185, 186, 187, 188 and 189 were not previously listed and were also therefore mutated.

Around 12000 variants were screened according to these different mutagenesis strategies. The site directed mutagenesis at the valine 1 did not show any enzymatic improvement as well as the other screened positions.

4.4 The mutant G166I

The mutant, G166L, was found with a higher activity and a slight improvement in the Km-value towards the Z-Phe-Leu-Glu-pNA substrate at pH 7.8 during the screening of these different positions. The mutant G166I was generated from the G166L variant by site directed mutagenesis to observe the hydrophobicity effect of the leucine and isoleucine on the mutation. Screening of the G166 position was continued meanwhile by saturated mutagenesis. The G166I as another mutant, G166R, were discovered with a higher activity and better Km-value during this second screening. These three mutants G166I, G166L and G166R showed a repeatable higher activity at the microtiterplate scale.

The enzymatic features of the variants had to be determined to compare the effect of each mutation at this position. Enzymatic parameters could not be calculated at the screening scale. The three mutants G166I, G166L and G166R were cultivated in a bigger scale (50 mL culture) to check their specific activity at the cuvette scale. After cultivation, the mutant G166L was 18 times more active than the wild type GluC, the mutant G166I 32 times more and the mutant G166R 34 times more after 31.5 hours of cultivation. This factor was changed to 4.2 times more for the mutant G166L, 12.8 more for the mutant G166I and 12.4 more for the mutant G166R after 70 hours of cultivation. Km-values in Tris buffer at pH 7.5 shows a slight Km-value improvement of a factor 3.8 for the mutant G166I, compared to the Roche lyophilisate V8-GluC, whereas it is quite similar for the mutants G166I and G166R with a factor of 1.64 and 1.73 respectively. The hydrophobicity of the isoleucine has a relevant impact on the activity and affinity of the wild type V8-GluC. The mutant G166I was chosen for further characterization as it showed better activity and affinity at the cuvette scale.

4.5 Mutation of the prosequence

Screening system cultivation was set up according to previous publications, where the glutamyl endopeptidase was supposed to be activated only by digestion of the prosequence by another protease of *S. aureus*, the aureolysin. The aureolysin protease was added consequently after cultivation to cut the prosequence during the screening system. However, a new discovery on the prosequence cleavage was recently published in 2008, where the V8-GluC prosequence was demonstrated to be first autodigested before being heteroproteolyzed by the aureolysin protease (Nickerson et al., 2007; Nemoto et al., 2008; Figure 43).

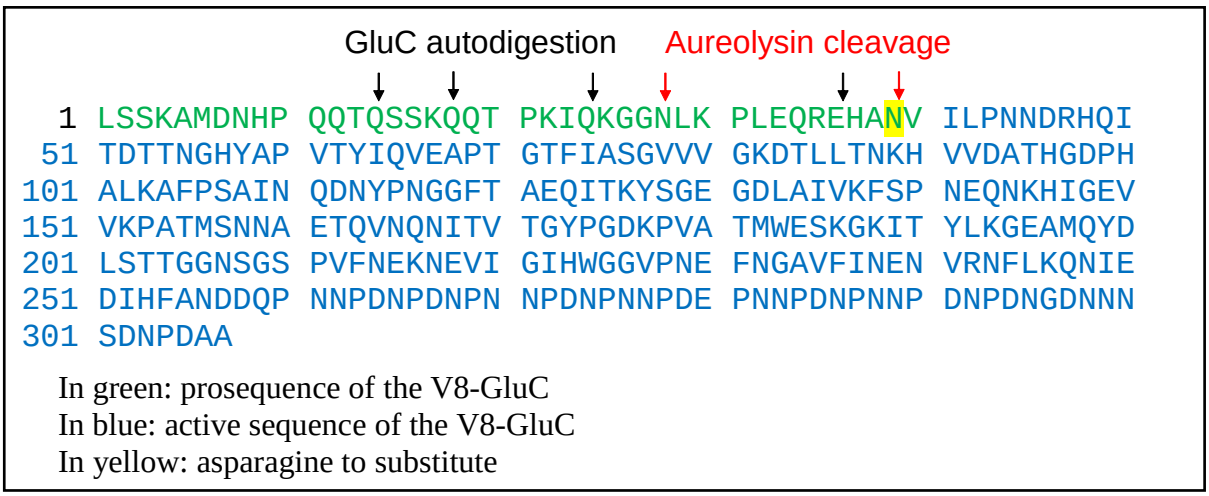


Figure 43. Maturation of the V8-GluC by autocatalysis and hydrolysis by the aureolysin. This V8-GluC is extended on its amino terminal side by an additional prosequence of 39 amino acids. This prosequence, which can be cleaved by another *Staphylococus aureus* protease, aureolysin, inactivates the protease after its synthesis in order to prevent from its own autoproteolysis.

The screening system cultivation implementation could have been more optimized consequently but this discovery does not affect the screening system. If a mutant is more active, its prosequence is faster digested and the improvement of the mutant is immediatly observed by the screening. However, hypothesis was ventured that the V8-GluC could completely autocatalyze its prosequence without addition of aureolysin, if the last amino acid of the prosequence, an asparagine, was substituted by a glutamic acid.

A mutation N-1E was introduced by site directed mutagenesis on the wild type and the mutant G166I to confirm this hypothesis. Wild type and mutant G166I V8-GluCs were designated E-Wt-H and E-M-H respectively. The H designates the polyhistidine tag cloning, the E, the N-1E mutation in the prosequence and wild type and G166I mutant GluCs are abbreviated by Wt and M.

According to this mutagenesis strategy, more than 60 amino acid positions were screened after directed evolution of the wild type V8-GluC gene. One relevant mutant came out from the library screening at the position 166. Saturated mutagenesis demonstrated the importance of the hydrophobicity at this position on the protease activity and affinity towards glutamate substrate at pH 7.8. The mutant G166I was generated with an additional mutation N-1E in the prosequence in order to facilitate the prosequence auto-hydrolysis and the activation of the V8-GluC mutant.

5. Characterization of the mutant with improved activity

5.1 Expression and cultivation

The wild type V8-GluC was initially cloned in *B. subtilis* to avoid the protease toxicity by its extracellularly secretion, which facilitated the screening. After screening, the different wild type and mutant V8-GluCs were expressed in *B. subtilis* with or without the N-1E mutation to accelerate the prosequence autoproteolysis. The effect of the polyhisdine tag presence was also studied during the expression. Cell growth and enzymatic activity were measured during the cultivation (Figure 44).

B. subtilis cells were first cultivated at a lower temperature, 30°C, in order to facilitate a correct protease folding during the protease expression. As the cell growth was too slow at 30°C, the cultivation was always performed at 37°C afterwards for the expression of the V8-GluC. Acetoin concentration variations did not interfere with the *Bacillus subtilis* growth and the wild type V8-GluC expression. The same concentration, 0.5 % of acetoin (10 %), was therefore always used. The cell growth of all these V8-GluC variants reached an OD_{600nm} plateau phase of about 4 after 24 hours of cultivation. The polyhistidine tag had no negative effect on the culture growth and on the activity of the GluCs. Wild type and G166I mutant GluCs activities were detected after 48 hours of cultivation. Interestingly, the activity continued increasing even if the cell growth had already reached a plateau. All these V8-GluC variants were expressed with the prosequence to avoid their own audigestion. Observed activities of the cultivation cannot be considered as the final activities as the prosequence was not properly cleaved. However, the N-1E mutation of the prosequence caused a strong impact on the activation of the different V8-GluCs and therefore on the prosequence cleavage. The measured activity was multiplied by a factor of two for the wild type and mutant V8-GluCs, which harboured the N-1E mutation of the prosequence.

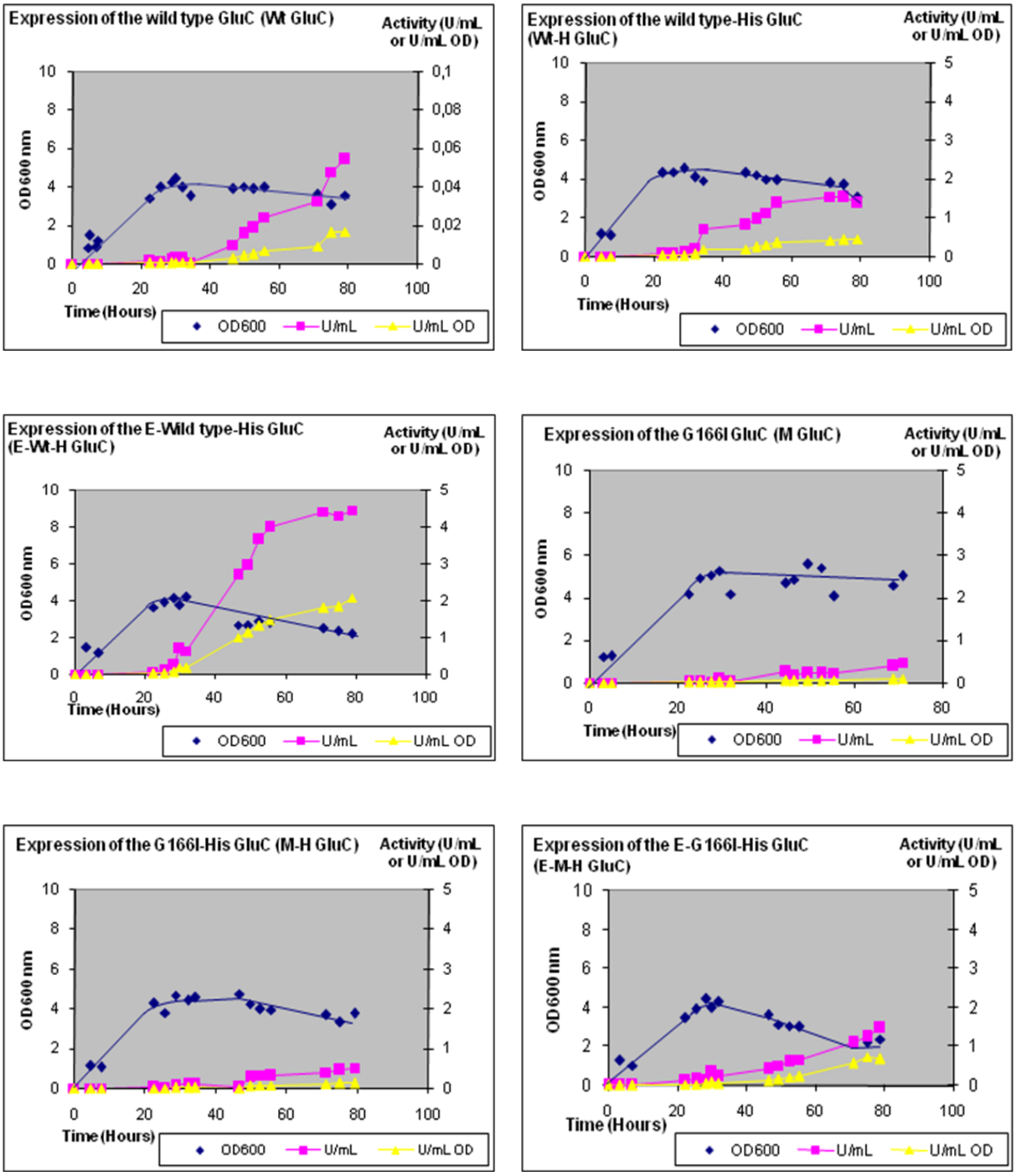


Figure 44. Cultivation, expression and activity of the wild type and mutant GluCs with or without polyhisidine tag and the N-1E mutation. The cells were cultivated and the enzymatic activity was measured over 80 hours. Wild type and G166I mutant GluCs activities can be detected after 48 hours of cultivation. The activity goes on increasing even if the cell growth has already reached a plateau at a OD_{600nm} of 4 after 24 hours of cultivation.

The protein digestion of each expressed V8-GluC protease was followed during expression. The molecular weights were investigated by ESI-MS at different time points of the cultivation. Expressed protein sequences and the prosequence cleavage were deduced by their analysis. The results showed that, for each one given V8-GluC variant, different molecular weights could be found at different cultivation times. These fluctuations of molecular weights were representative of different protease forms and emphasized the hypothesis of degradation of the protease and/or the auto-hydrolysis of its prosequence. The intact E-Wt-H GluC sequence could be observed at a 0.50 U/mL.OD activity. The intact E-M-H GluC sequence was detected at a 0.10 U/mL.OD activity. The E-Wt-H and E-M-H GluC cultivations were always stopped with these similar activities, at the early stage of the cultivation to get the intact native protease. A complete reproducibility of the protease expression was however difficult. At this early stage of the cultivation, the estimated yields of the E-Wt-H and E-M-H GluCs were 0.046 mg/mL and 0.0013 mg/mL respectively. These yields cannot be considered as maximal as the cultivation was not at its optimal expression point. Secondly, the E-Wt-H and E-M-H GluCs were also at different stages of the prosequence activation and furthermore, their own prosequence was not completely cleaved but hydrolyzed at different points. The expression yields of the E-Wt-H and E-M-H GluCs should be consequently increased.

According to the expression study of the V8-GluC variants, the presence of the polyhistidine tag did not interfere with the cell growth and protease expression. The N-1E mutation was observed to have a strong impact on the prosequence. The wild type and G166I mutant GluCs were consequently expressed extracellularly from *B. subtilis* with the N-1E mutation and the polyhisdine tag. The E-Wt-H and E-M-H GluCs were collected at the early stages of the cultivation to avoid the different digested forms, which appear after several days of expression.

5.2 Purification by Nickel-NTA affinity chromatography

After extracellular expression from *B. subtilis*, the E-Wt-H and E-M-H GluCs were purified for further analytical investigations. The E-Wt-H and E-M-H GluCs were purified on a Ni-NTA column as the polyhistidine tag was not completely digested. Purification of 10 mL culture sample was first pretested on Ni-NTA spin columns.

As the E-Wt-H and E-M-H GluCs were found to bind onto these columns, the purification was scaled up to purify the cultivation rest of the E-Wt-H and E-M-H GluCs. Purification balances are summarized in figures 45 and 46.

E-Wt-H GluC purification

Sample	Total units (U)	Yield (%)	Quantity (mg)	SA (U/mg)	Purity	Volume (mL)
Culture	1222,46	100	156,6	7,81	1	580
Flow through	1017,23	83,21	132,24	7,69	0,98	580
Wash	69,57	5,69	7,98	8,72	1,12	380
Elution	277,85	22,73	7,28	38,17	4,89	70

Figure 45. E-Wt-H GluC purification by affinity chromatography with a Ni-NTA column. The yield and the specific activity (SA) are partially representative as prosequences are not completely cleaved and GluCs are partially activated. The E-Wt-H purification did not show the same reproducibility because of the difficulty to stop the cultivation at the same activity point.

E-M-H GluC purification

Sample	Total units (U)	Yield (%)	Quantity (mg)	SA (U/mg)	Purity	Volume (mL)
Culture	141,23	100	186,3	0,76	1	540
Flow through	70,2	49,71	158,76	0,44	0,58	540
Wash	4,6	3,26	9	0,51	0,67	450
Elution	28,09	19,89	12,1	2,32	3,06	37

Figure 46: E-M-H GluC purification by affinity chromatography with a Ni-NTA column. The yield and the specific activity (SA) are partially representative as prosequences are not completely cleaved and GluCs are partially activated.

A purity effect was observed with a factor of 4 for the E-Wt-H GluC and a factor of 3 for the E-M-H GluC. However, these yields and specific activities are not completely representative after the purification step as the prosequences must be completely cleaved to activate the GluCs. If the enzyme was completely activated by the complete prosequence cleavage, yields and purity factors would be higher.

The molecular weight measurement by ESI-MS confirmed that the prosequence was still present in the expressed proteases. According to the molecular weight assessment by ESI-MS, in the purification trial of the E-Wt-H GluC, 90 % of the purified E-Wt-H GluC were activated with the right sequence, [V40-D305] and a molecular weight of 29850 Da. 5 % had still a prosequence piece, [G26-H310], with a molecular weight of 30979 Da and the last 5 % of the purified E-Wt-H GluC were an autodigestion of the E-Wt-H GluC, [N44-H309] with a molecular weight of 28880 Da. The E-Wt-H purification did not show the same reproducibility because of the difficulty to stop the cultivation at the same activity point. Furthermore, the E-Wt-H was quite completely activated after this purification and this could not be reproduced.

After the E-M-H GluC purification, 100 % of the mutant GluC were found not to be completely activated and still contained the prosequence with [R35-H311] as sequence and a molecular weight of 30240 Da.

Some GluCs were lost during the flow through and wash steps. Some carboxyl-terminal digested side of the E-Wt-H or E-M-H GluCs were found in the flow through and the wash. The whole amount of E-WT-H and E-M-H GluCs, which were expressed, could not bind onto the Ni-NTA because of the carboxyl-terminal side hydrolysis.

During this purification step, both E-Wt-H and E-M-H GluCs were purified by Ni-NTA chromatography but they were obviously degraded during the purification. The activities and the purification balance were not representative as the E-Wt-H and E-M-H GluCs were not activated. The E-Wt-H and E-M-H GluC samples were dialyzed against a 50 mM Tris-HCl buffer pH 8 for long-term storage before further analysis such as the protease activation by prosequence cleavage.

5.3 Prosequence cleavage and glutamyl endopeptidase activation

The purified and dialyzed E-Wt-H and E-M-H GluCs were expressed as inactive with their prosequence. The aim was to avoid its toxic effect towards *B. subtilis* during its expression and to avoid the auto-hydrolysis of the protease itself. As the E-Wt-H and E-M-H GluCs had the N-1E mutation to facilitate the prosequence hydrolysis by autodigestion, it was also tried to activate it by incubation with other GluCs such as the Roche and NEB commercial lyophilisate GluCs. They were added at different concentrations of 0.5 U, 1 U and 3 U after incubation at 25°C for several days. Any activation could be observed after two days of incubation with these commercial V8-GluC lyophilisates.

As some of the expressed GluCs were found under different protein forms during the ESI-MS analysis, it was assumed that the expressed GluCs were partially activated during the cultivation at 37°C. The protease activation could not be observed without the N-1E mutation during the cultivation. At the opposite, the E-Wt-H and E-M-H GluCs continued to auto-hydrolyze themselves on a long-term storage. The purified and dialyzed E-Wt-H and E-M-H GluC samples were consequently incubated for several days at different temperatures: 25°C, 30°C and 37°C to try to activate the prosequence auto-hydrolysis. Samples were incubated at 25°C in case of too rapid activation at the other temperatures. Results are shown in figure 47.

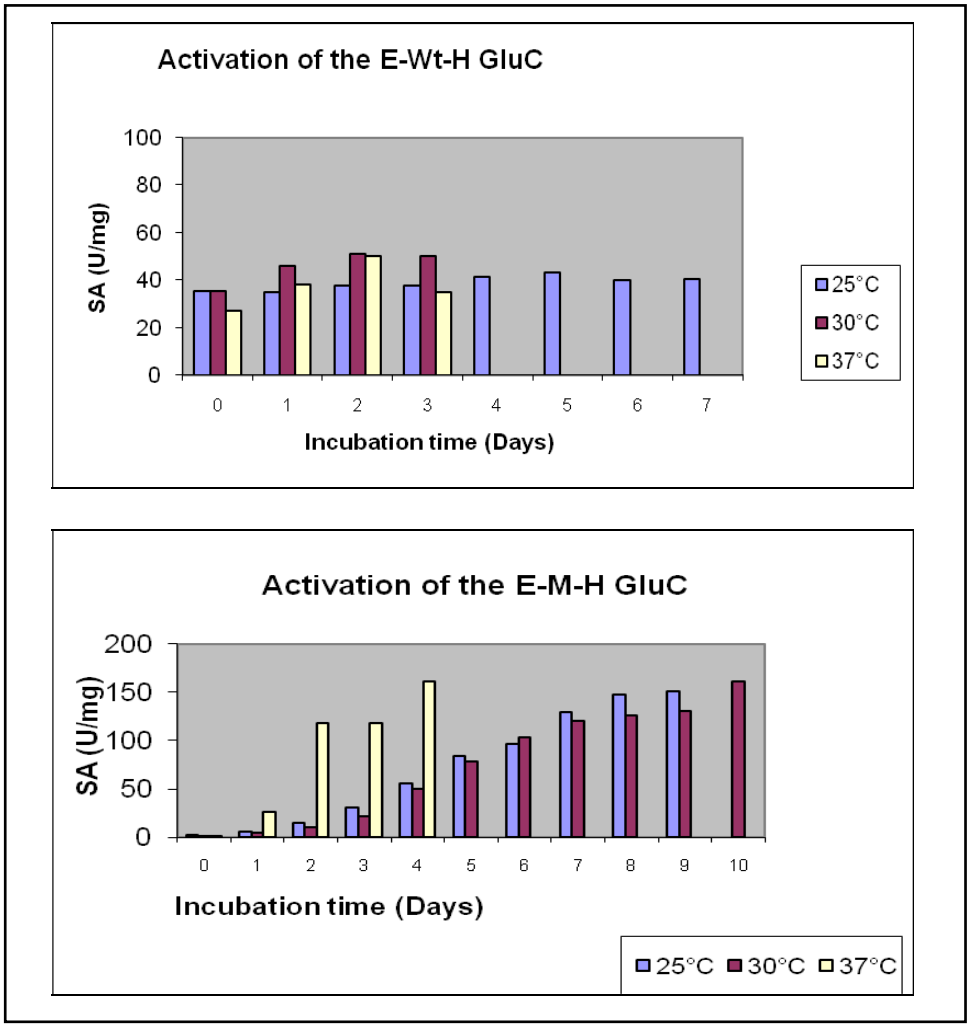


Figure 47. Prosequence cleavage and activation of the E-Wt-H and E-M-H GluCs. Purified and dialyzed E-Wt-H and E-M-H GluC samples were then simply incubated for several days at different temperatures: 25°C, 30°C and 37°C to try to activate the prosequence auto-hydrolysis. The specific activity (SA) was then measured.

The E-Wt-H GluC had a specific activity of about 45 U/mg after incubation for 2 days at 30°C or 37°C in the 1.5 mL pre-test sample. This specific activity was reproduced at 37°C at a larger scale with the rest of E-Wt-H GluC volume after the purification step. This purified and dialyzed E-Wt-H GluC sample contained 80 % of the following sequence, [V40-D305], with a molecular weight of 28880 Da and 20 % of the sample were digested under the [V40-D294] form with a molecular weight of 27730 Da after incubation at 37°C for 2 days. The major form of the E-Wt-H GluC was already autodigested at the carboxyl-terminal side, where the polyhisidine tag and the two last alanines were missing.

After activation in 1.5 mL pre-test sample, the E-M-H GluC reached a specific activity of 160 U/mg after incubation for 4 days at 37°C. The total volume of the purified and dialyzed E-M-H GluC was incubated according to this pre-test and after activation, the E-M-H GluC showed a specific activity of 190 U/mg after incubation for 2 days and 3.5 hours at 37°C. The specific activity of the E-M-H GluC changed from 160 to 190 U/mg. This fluctuation could be explained by the larger volume and therefore the amount of GluC, which was incubated. The greater amount of GluC was incubated, the more the GluC was activated. The E-M-H GluC contained 57 % of the [V40-D305] sequence with a molecular weight of 28940 Da, 27 % of the [V40-H313] sequence with a molecular weight of 29936 Da and 16 % of the [V40-H309] as sequence with a molecular weight of 29310 Da. The two last alanines of the carboxyl-terminal side were also digested as in the E-Wt-H GluC.

A SDS-PAGE was also performed as an analysis of the purification and activation of the E-Wt-H and E-M-H GluCs (Figure 48).

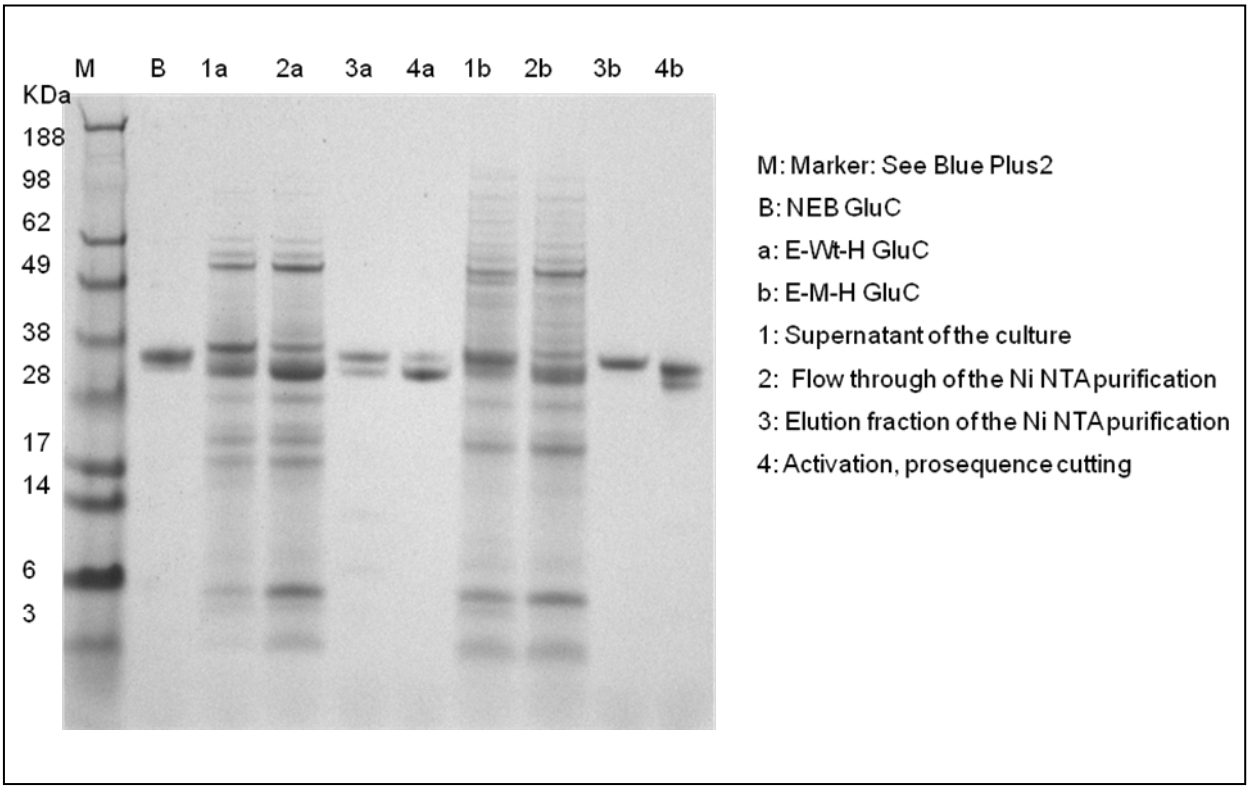


Figure 48. SDS gel of the E-Wt-H and E-M-H GluCs after culture, purification and activation. The E-Wt-H and E-M-H GluCs should have a molecular weight of 29846 and 29902 Da or respectively 29023 and 29079 Da without polyhistidine tag. Autodigested forms of both GluCs can be also observed in the eluted sample.

The E-Wt-H and E-M-H GluCs were observed to be pure as much as the standards on the SDS-gel after purification and activation.

The E-Wt-H and E-M-H GluCs should have a molecular weight of 29846 and 29902 Da respectively with the intact sequence, [V40-A307], and the polyhisdine tag. This molecular weight should decrease to 29023 and 29079 Da for the complete sequence without polyhisidine tag. However, the majority of the E-Wt-H and E-M-H GluCs were activated under the [V40-D305] form with a molecular weight of 28880 Da and 28940 Da respectively. The SDS gel confirmed the presence of the autodigested E-Wt-H and E-M-H GluCs in the flow through of the Ni-NTA purification. The E-Wt-H and E-M-H GluCs were completely purified. The presence of autodigested forms were also observed in the eluted sample. The E-Wt-H GluC is strongly autodigested on this SDS gel. It could come from the continuous E-Wt-H GluC autocatalysis. This autodigestion was also observed for the E-M-H GluC. The first six amino acids from the amino-terminal side were sequenced and the amino-terminal sequencing confirmed that the prosequence was correctly cleaved. The first six amino acids of the active GluC sequence, VILPNN, were distinctly sequenced without additional amino acids of the prosequence.

As a conclusion, the E-Wt-H and E-M-H GluCs were activated by auto-hydrolysis of the prosequence. Even if they were present under degraded forms, only some histidines of the polyhistine tag and the two last carboxyl-terminal alanines were mainly missing. The specific activities of the E-Wt-H and E-M-H GluCs were assessed at 45 U/mg and 190 U/mg respectively in Tris buffer at pH 7.8 with the Z-Phe-Leu-Glu-pNA substrate.

Once the E-Wt-H and E-M-H GluCs were expressed, purified and activated, further biochemical and analytical characterizations were investigated. The Roche Lyo GluC was taken as a positive control. The NEB GluC was used as a direct reference for the recombinant wild type GluC, E-Wt-H.

5.4 Storage stability at -20°C

The purified and activated E-Wt-H and E-M-H GluCs were stored before further investigations. Their stability was controlled for storage at -20°C. Storage stability at -20°C was tested by freezing the E-Wt-H and E-M-H GluCs at -20°C overnight and defreezing them the day after. This experiment was repeated 5 times over one week, as samples were once frozen over the week-end. Their activity was measured after defreezing. 5 mM of Glu-Glu dipeptide were also added before freezing to check a potential stabilization effect.

The E-Wt-H and E-M-H GluCs were stable through the freezing and defreezing over 1 week. The addition of the Glu-Glu dipeptide did not bring any effect. The liquid E-Wt-H and E-M-H GluCs could be stored at -20°C for further characterization.

5.5 Specific activity in different buffers

After activation, the final specific activity of the E-Wt-H and E-M-H GluCs were calculated and compared to the other glutamyl endopeptidases. Specific activities of the E-Wt-H and E-M-H GluCs were assessed in different buffers and with different substrates. Tris buffer (Tris 0.1 M pH 7.8) and potassium phosphate buffer (KH_2PO_4 0.1 M pH 7.8) were first tried out with the Roche substrate, Z-Pheu-Leu-Glu-pNA. Activities were also tested in potassium phosphate buffer (KH_2PO_4 0.1 M pH 5) and ammonium acetate buffer (NH_4-Ac. 50 mM pH 4.3) with the HbA0(1-6) pNA substrate. Outcomes are summed up in figure 49. The specific activities of the E-M-H GluC were compared to the Roche Lyo V8-GluC. The activities of the E-Wt-H and NEB GluCs were also calculated as references. Specific activities of the E-Wt-H, NEB and Roche GluCs could be directly compared to the E-M-H GluC in Tris buffer at pH 7.8. The specific activities in other buffers, ammonium acetate at pH 4.3 or potassium phosphate at pH 5 and 7.8, were analyzed by comparing the ratio between the specific activity in the desired buffer to the respective specific activity in Tris buffer at pH 7.8 in order to avoid the buffer effect on the GluC activity and specificity.

Buffers	SA (U/mg)			
	Roche	**NEB**	**E-Wt-H**	**E-M-H**
Tris buffer pH 7.8	44.84	30.77	42.62	196.23
KH_2PO_4 buffer pH 7.8	47.14	40.51	54.82	206.70
KH_2PO_4 buffer pH 5	0.1612	0.22	0.5606	5.39
Ammonium acetate buffer pH 4.3	0.0140	0.01	0.0206	0.2564

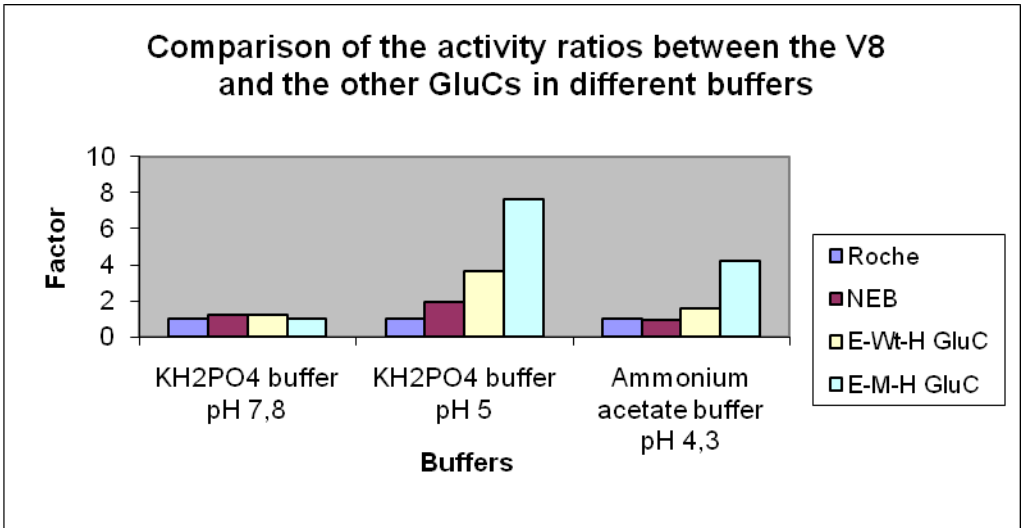


Figure 49. Specific activities (SA) of the E-Wt-H and E-M-H GluCs in different buffers. The E-Wt-H and E-M-H were compared to the Roche Lyo V8-GluC and NEB GluCs in Tris buffer pH 7.8 or potassium phosphate buffers at pH 7.8 or 5 with the Z-Phe-Leu-Glu-pNA as a substrate. Their activities were also assessed in ammonium acetate buffer at pH 4.3 with the HbA0(1-6)-pNA substrate. The ratio between the specific activity in ammonium acetate or potassium phosphate buffer and the respective specific activity in Tris buffer at pH 7.8 were calculated in order to avoid the buffer effect on the GluC activity and specificity.

The E-M-H GluC is 4 times more active in Tris buffer pH 7.8 with Z-Phe-Leu-Glu-pNA as substrate than the Roche, NEB, E-Wt-H wild type GluCs. The E-M-H GluC is 4 times more active in ammonium acetate buffer at pH 4.3 with HbA0(1-6)-pNA as substrate than the Roche, NEB and E-Wt-H wild type GluCs. The E-M-H GluC is also around 2 times more active than the Roche, NEB and E-Wt-H wild type GluCs in potassium phosphate buffer at pH 5 but there is no activity difference in KH_2PO_4 buffer at pH 7.8.

The E-M-H GluC has a higher specific activity than the Roche and NEB GluCs whatever the buffer, pH or glutamate substrate is tested. This improvement is due to only one amino exchange of a glycine to a isoleucine at the position G166I.

5.6 Km-value

One other screening characteristic after the activity was the V8-GluC affinity towards the Z-Phe-Leu-Glu-pNA substrate. Once the E-Wt-H and E-M-H GluCs were purified and activated, their affinity could be defined by their Km-Value. The Km-values of the different GluCs, Roche, NEB, E-Wt-H and E-M-H, were assessed by the Lineweaver Burk plotting. The average was calculated after repeating the Km-value experiment 3 times as it is shown in figure 50.

	Km Value (mM)			
V8-GluC	**Km 1**	**Km 2**	**Km 3**	**Average**
Roche	0,53	0,54	0,41	0,49
NEB	0,46	0,55	0,46	0,49
E-Wt-H	0,44	0,53	0,54	0,50
E-M-H	0,17	0,18	0,16	0,17

Figure 50. Km values of the E-Wt-H and E-M-H GluCs with the Z-Phe-Leu-Glu-pNA substrate in Tris buffer 0.1 M pH 7.8. The E-M-H mutant Km value was compared to the Roche, NEB and the E-Wt-H GluCs.

Roche, NEB and the E-Wt-H GluCs have the same Km-value of 0.50 mM towards the Z-Phe-Leu-Glu-pNA substrate in Tris buffer at at pH 7.8. The E-M-H GluC Km-value is 0.17. The E-M-H GluC has 3 times more affinity towards the Z-Phe-Leu-Glu-pNA substrate in Tris buffer at pH 7.8. than the E-Wt-H GluC, NEB and Roche lyophilisate GluCs.

The Km-values of the E-M-H mutant GluC could not be determined with the hemoglobin hexapeptide substrate because of its cost. However, the Roche GluC was not at the maximal speed, Vmax, when the hemoglobin hexapeptide was applied with the same amount as the protocol with the Z-Phe-Leu-Glu-pNA substrate. The hemoglobin hexapeptide was not in this case at the saturation concentration. The Roche V8-GluC is therefore less affine towards the hemoglobin hexapeptide than the Z-Phe-Leu-Glu-pNA substrate. The Roche V8-GluC Km-value should be higher towards the hemoglobin hexapeptide substrate than the Z-Phe-Leu-Glu-pNA one.

By the G166I mutation, the E-M-H GluC mutant presents more affinity towards the glutamate substrate in Tris buffer at pH 7.8. Its affinity could not be investigated with the hemoglobin hexapeptide substrate or at other pHs.

5.7 Temperature stability

As mutagenesis could influence other enzymatic parameters of the E-Wt-H wild type GluC, the thermostability of the E-M-H GluC, was investigated. The stabilities of the E-Wt-H and E-M-H GluCs were tested stepwise from 25 to 55°C by GluC incubation for 30 minutes. 5 mM of Glu-Glu dipeptide were also added to control a potential stabilization effect. Results are plotted in figure 51.

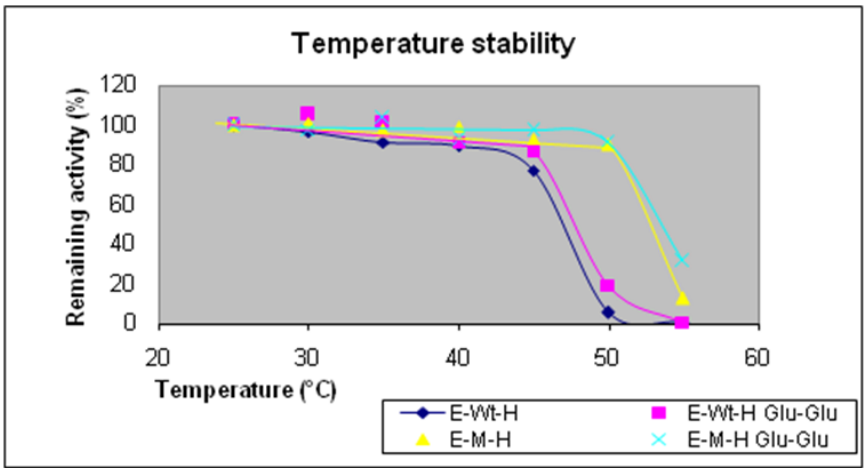


Figure 51. Temperature stability of the E-Wt-H and E-M-H GluCs. The stability of the E-Wt-H and E-M-H GluCs were tested after 30 minutes of incubation with or without 5 mM of Glu-Glu dipeptide. The remaining activity is the percentage of calculated activity in defined experiment conditions (buffers and pH) by referring to the V8-GluC activity at 25°C.

The E-M-H GluC is more stable than the E-Wt-H GluC over 50°C but any activity can be detected over 55°C for both GluCs. The stabilization effect of the Glu-Glu dipeptide cannot be observed neither on the E-Wt-H GluC nor on the E-M-H GluC. Both E-Wt-H and E-M-H GluCs are not stable anymore after 50°C and 55°C respectively.

5.8 Melting temperature assessment by DSC

The thermostability could be also checked by measuring the melting temperature by DSC (Differential Scanning Calorimetry). Differential Scanning Calorimetry (DSC) measures the amount of required energy to raise the temperature of a protein in solution.

The protein unfolding due to heat denaturation, is correlated to this energy variation by referring to a standard. The unfolding transition is recognized as a sharp endothermic peak centered at the transition midpoint, Tm, also called melting temperature. This measured Tm indicates the protein thermostability. After the thermostability study of the E-Wt-H and E-M-H GluCs by incubation and measurement of the activity, the melting temperature of the E-Wt-H, E-M-H, Roche and NEB GluCs were also measured by DSC (Figure 52).

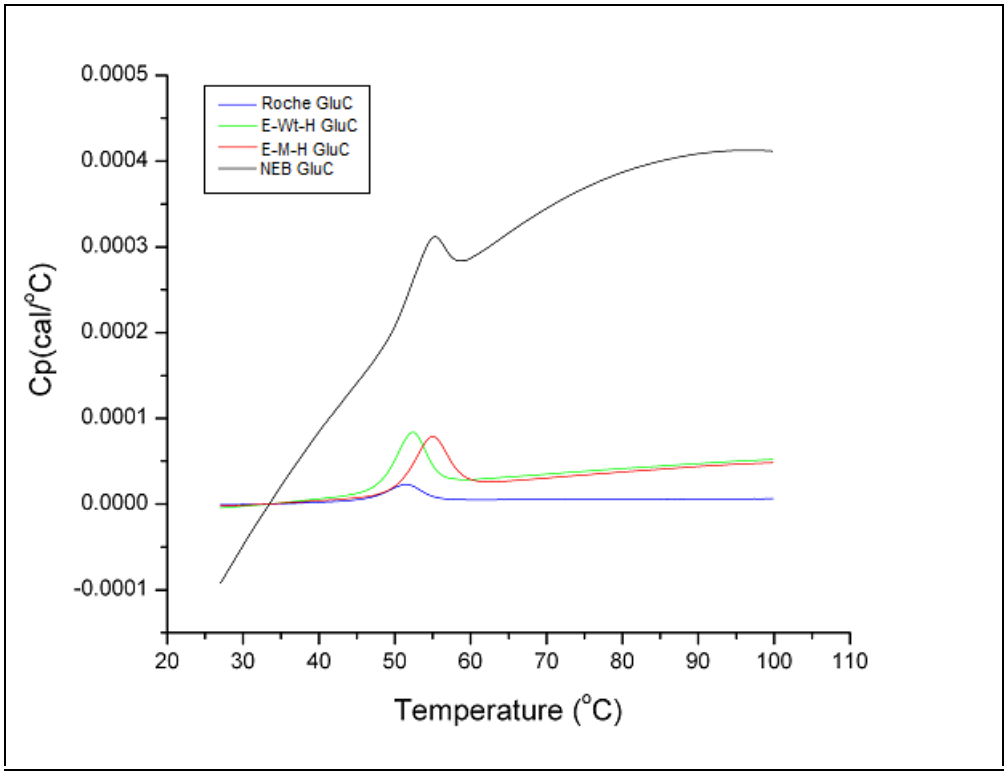


Figure 52. Melting temperature measurement by DSC of the Roche, NEB, E-Wt-H and E-M-H GluCs. The unfolding transition is recognized as a sharp endothermic peak centered at the transition midpoint, Tm. This melting temperature indicates the protein thermostability.

The melting temperatures of Roche GluC was assessed at 51.4°C, the NEB GluC at 52.33°C, the E-Wt-H at 52.4°C and the E-M-H GluC at 55.15°C. The Roche GluC is more sensitive to the temperature than the other GluCs due to the eleven amino acid position difference. The NEB and E-Wt-H GluCs are the same in the scope of the DSC. The E-M-H mutant is more thermostable than the other GluCs.

The melting temperature assessment confirmed the higher thermostability of the E-M-H GluC, which was observed during the temperature stability experiment. The mutant E-M-H GluC is slightly less susceptible to unfolding and denaturation.

5.9 pH optimum

The pH is an important enzymatic factor, which is strongly correlated to the enzyme activity. Very acidic or basic pHs can even affect the enzyme up to a complete loss of activity. The pH optimum is defined as the pH value, for which the enzyme activity is the most efficient. As the mutagenesis of the wild type V8-GluC at the position 166 improves its activity, the pH optimum may be also influenced. The pH optimum was determined in ammonium acetate (50 mM) between pH 4 and 6, KH_2PO_4 (0.1 M) between pH 6 and 7.5 and Tris-HCl buffer (0.1 M) between pH 7.5 and 9 with the hemoglobin hexapeptide substrate. This substrate is the only one, which can be diluted in all buffers and which gives a direct correlation to the HbA0 functional test by ESI-MS measurement. The pH optimum curves are drawn in figure 53.

The pH optima of the E-Wt-H and E-M-H GluCs are respectively at 7.8 and 7 with the hemoglobin hexapeptide substrate. The E-M-H GluC remains at its highest activity on a wide pH range from pH 6.5 to 8 whereas the E-Wt-H GluC is mainly at its highest activity only at one pH peak at 7.8. An activity loss of 20 % can be already observed at pH 7 for the E-Wt-H wild type GluC compared to the E-M-H mutant GluC.

Only one amino exchange in the E-Wt-H GluC sequence, a glycine into isoleucine, can create a shift of the pH optimum from pH 7.8 for the E-Wt-H GluC to pH 7 for the E-M-H GluC. The E-M-H GluC pH optimum is extended to a wider pH range. Only one pH optimum peak for both E-Wt-H and E-M-H GluCs is observed whereas two optimum pH peaks are usually present at pH 4 and 7.8 with the whole hemoglobin as substrate (Drapeau et al., 1972). The pH optimum of the E-M-H GluC could not be determined with the whole hemoglobin substrate to compare it with the Roche GluC. The pH optimum of the V8-GluC depends on the substrate.

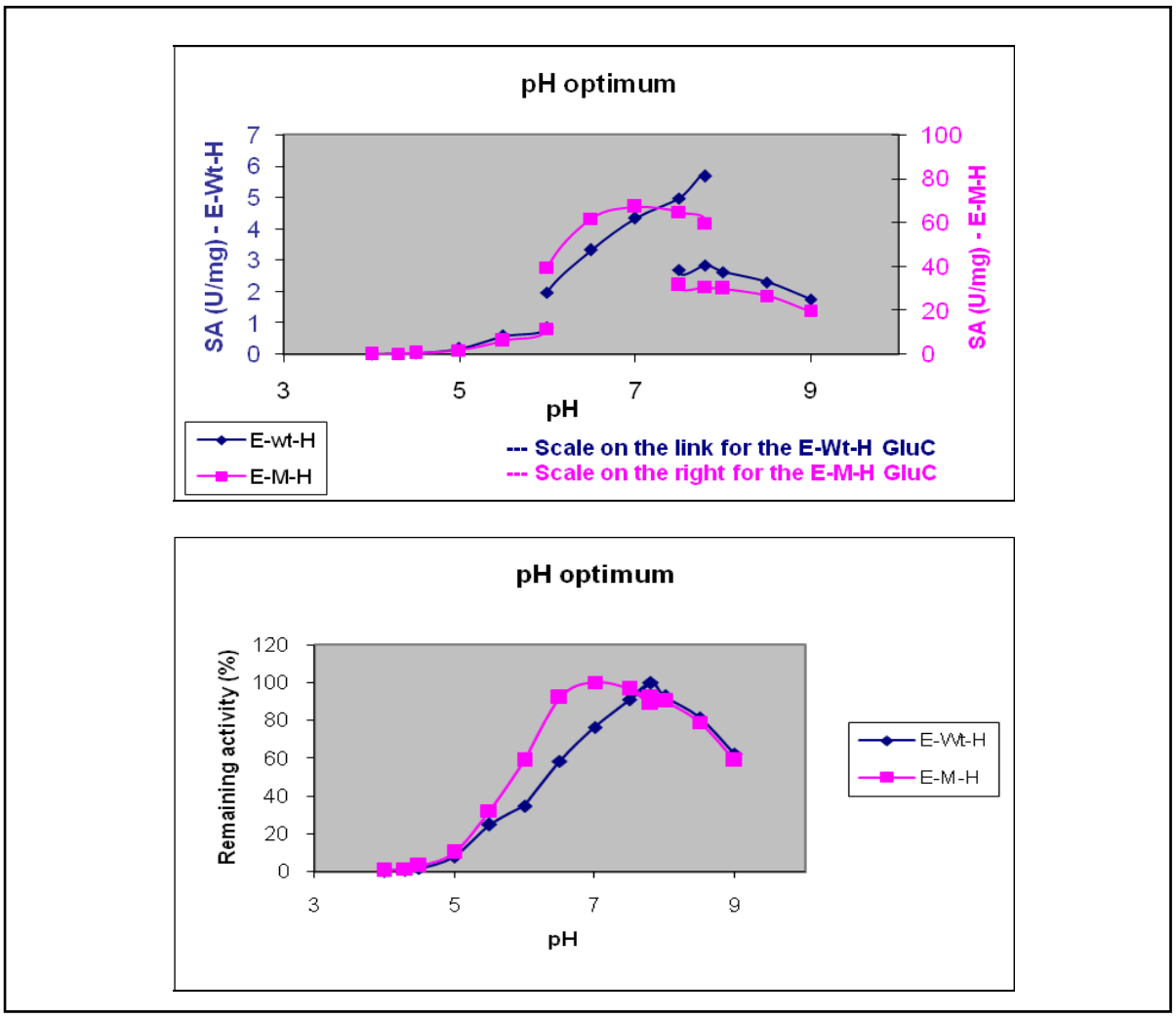


Figure 53. The pH optimum of the E-Wt-H and E-M-H GluCs. The pH optimum was determined in ammonium acetate buffer between pH 4 and 6, in KH_2PO_4 buffer between pH 6 and 7.5 and in Tris-HCl buffer between pH 7.5 and 9 with the hemoglobin hexapeptide substrate.

5.10 Isoelectric point

Enzymes are characterized by their isoelectric point (pI). The isoelectric point, also abbreviated to IEP, is defined as the pH value at which the molecule carries no electrical charge or the negative and positive charges are equal. As the G166I mutation created the amino exchange of a glycine into an isoleucine, the isoelectric point of the E-M-H GluC may have been affected. The isoelectric points of the different possible GluC constructs were calculated theoretically by the bioinformatics program of Roche and are summarized in the figure 54.

V8-GluC	Pro-sequence	N-1E mutation	Wild type	G166I mutant	2 last Alanines missing	His Tag	pI
P-Wt	x		x				4,5753
PE-Wt	x	x	x				4,52,82
P-Wt-H	x		x			x	5,1325
PE-Wt-H	x		x			x	5,0164
Wt			x				4,2361
Wt-H			x			x	4,5955
Wt-D			x		x		4,2361
P-M	x			x			4,5783
PE-M	x	x		x			4,5282
P-M-H	x			x		x	5,1325
PE-M-H	x			x		x	5,064
M			x				4,2361
M-H			x			x	4,5955
M-D			x		x		4,2361
Roche			x				4,1477
NEB			x			x	4,5955

Figure 54. Isoelectric points of different GluCs by bioinformatics. *Legend: P: prosequence, E: N-1E mutation in the prosequence, Wt: Wild type, M: G166I mutant, H: polyhisidine tag, D: 2 last missing alanine at the carboxyl-terminal side.* The G166I and N-1E mutations do not interfere on the isoelectric point unlike the prosequence. The eleven different positions in the Roche GluC sequence give a more acidic isoelectric point of 4.1477 than the the NEB, E-Wt-H and E-M-H GluCs with a pI of 4.5955.

The wild type and mutant G166I with or without the N-1E mutation and/or polyhisdine tag have the same pI. The whole prosequence causes a strong impact of 0.5 pH units on the isoelectric point. The N-1E mutation brings a slight effect on the GluCs of 0.1 pH unit. The polyhisidine tag adds 0.3 pH units more in the GluCs sequence in general. The NEB GluC has a similar isoelectric point as the recombinant E-Wt-H and E-M-H GluCs after suppression of the prosequence even if there are two missing alanines in the recombinants, as they do not play any role on the isoelectric point calculation. The eleven different positions in the Roche GluC sequence give a more acidic isoelectric point of 4.1477 than the the NEB, E-Wt-H and E-M-H GluCs with a pI of 4.5955.

The mutation G166I, unlike the prosequence autodigestion, does not influence the isoelectric point of the wild type V8-GluC. By contrast, the Roche GluC pI differs from the NEB, E-Wt-H and E-M-H GluCs because of its 11 amino acid exchange. This pI differences must be considered for experiments, which are based on this analytical feature, such as ion exchange chromatography for example.

5.12 Hemoglobin functional test

The functional test of the IFCC is the measurement of the glycated hemoglobin (HbA1c) by ESI-MS. The glycated hemoglobin is digested by the Roche GluC in specific experiment conditions in order to assess by ESI-MS the ratio between glycated and non-glycated hemoglobin hexapeptides. The hemoglobin was not glycated (HbA0) for the ESI-MS analysis.

5.12.1 Wild type V8-GluC

The specificity pattern of the hemoglobin cleavage by the Roche V8-GluC was defined in order to compare it to the E-M-H GluC mutant. They were assessed after ESI-MS analysis (Figure 55).

Chain alpha-hemoglobin		
Retention Time [min]	Masse[Da]	Residue
14,5	2306.2	Res 1-23 α
16,3	3039.4	Res 1-30 α
28,7	12103,0	Res 31-141 α

Chain beta-hemoglobin		
Retention Time [min]	Masse[Da]	Residue
11,6	**694.4**	**Res 1-6 β**
13,1	1304.6	Res 91-101 β
17,7	1744.9	Res 7-22 β
18,9	2422.2	Res 1-22 β
21	4842.6	Res 44-90 β
23	2680.4	Res 27-47 β
23,5	2095.1	Res 27-43 β
24,3	2437.2	Res 23-43 β
25	2265,6	Res 102-121 β
26,5	4928.9	Res 102-146 β

Figure 55. Measurement by ESI-MS of the Roche GluC specificity on hemoglobin after 18 hours incubation at 37°C (1:100, enzyme: substrate) in 50 mM ammonium acetate pH 4.3.

The alpha-chain of hemoglobin was digested into three residues by the Roche V8-GluC. The beta-chain was cleaved into ten residues The hexapeptide of the beta-chain was eluted with a molecular weight of 694.4 Da at a retention time of 11.6 minutes. The same pattern specificity was obtained with the E-M-H GluC.

The G166I mutation of the E-M-H GluC does not affect the cleavage specificity of the hemoglobin in the functional test conditions.

The HbA0 test is usually performed under strict conditions after 18 hours incubation at 37°C (1:100, enzyme: substrate) in 50 mM ammonium acetate pH 4.3. As the substrate specificity is buffer dependent and pH and buffers have a strong impact on the activity of the glutamyl endopeptidase, the hemoglobin was incubated with the Roche GluC in different buffers, Tris buffer at pH 7.8, ammonium acetate at pH 4.3 and pH 7. The formation of the hexapeptide was compared to the initial amount of total hemoglobin by ESI-MS analysis.

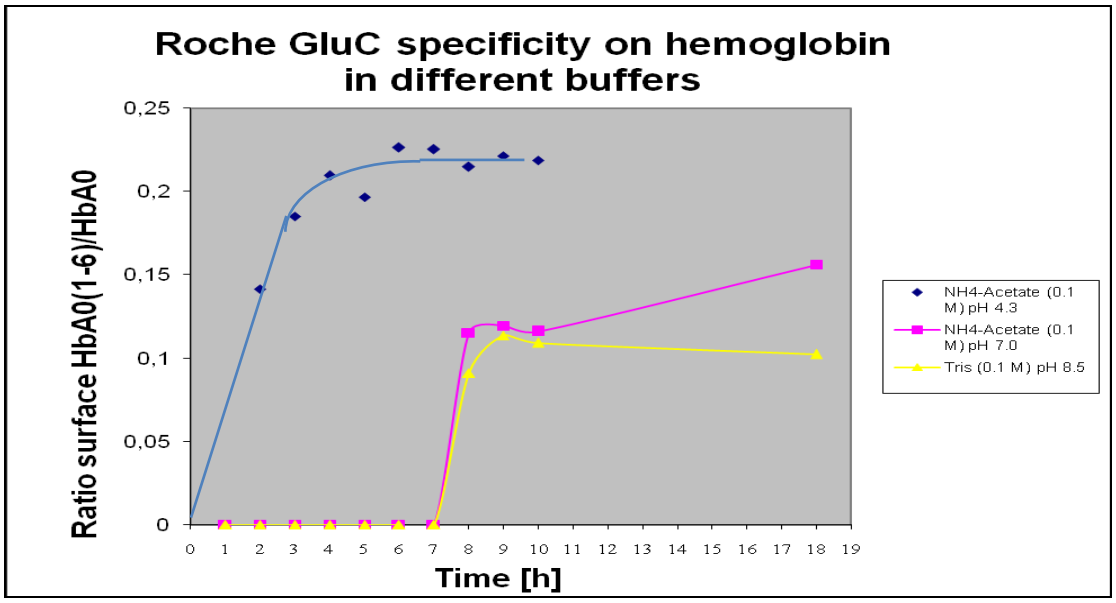


Figure 56. Roche GluC specificity towards the hemoglobin in different buffers. The Roche GluC was incubated in ammonium acetate and Tris buffer after 18 hours incubation at 37°C and a ratio 1:100 (enzyme:substrate).

According to the results in figure 56, the Roche GluC is around two times more active and specific towards the hexapeptide in ammonium acetate buffer at pH 4.3. The hexapeptide is already completely cleaved after 5 hours of incubation whereas it did not start before 7 hours in Tris buffer at pH 7.8. The cleavage is however better in ammonium acetate buffer than Tris buffer.

The conditions of the ESI-MS standard IFCC test could not be changed. Directed evolution emerged as the solution to generate a faster V8-GluC variant by keeping its other enzymatic parameters.

5.12.2 Mutant V8-GluC

First, the pH stability of the E-M-H GluC was compared to the Roche GluC in the IFCC conditions.

HbA0 functional test was then performed to test the specificity of the E-M-H GluC mutant compared to the Roche GluC. The E-M-H GluC mutant was incubated in the exact conditions of the IFCC protocols, for 18 hours at 37°C of the glutamyl endopeptidase with the haemoglobin in a ratio 1:100. The specificity towards the non-glycated hemoglobin was checked by ESI-MS measurement (Figure 57).

The parameters of the IFCC protocols were also tested. Different buffers were also tried during the HbA0 functional test. According to the previous specific activity study of the mutant, the ammonium acetate buffer was replaced by a potassium phosphate buffer at pH 5.5 or 7 and a Tris buffer at pH 7, all buffers with the same concentrations of 50 mM.

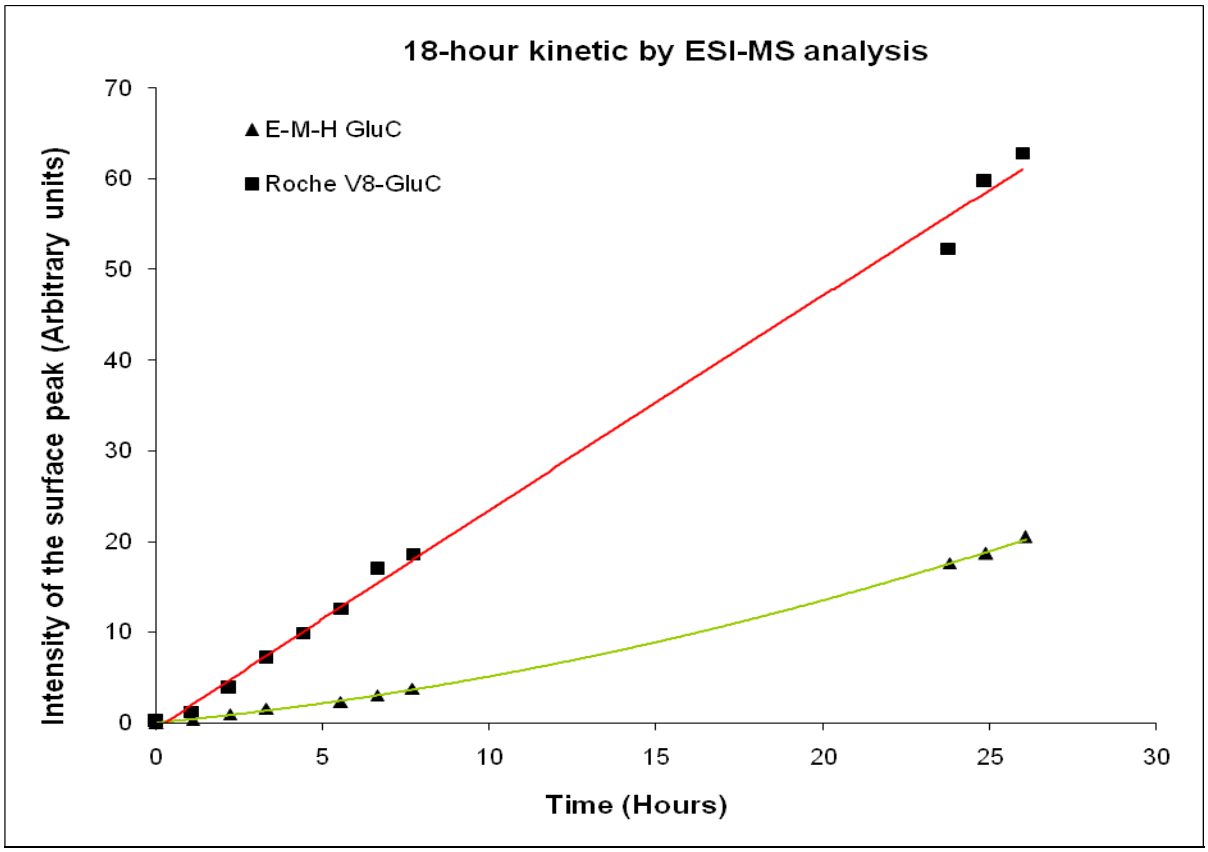


Figure 57. Mutant G166I GluC kinetics. The 18-hour kinetic shows that the specificity towards the hemoglobin hexapeptide is 3 times greater with the Roche GluC than the E-M-H GluC

The Roche and the E-M-H GluCs did not show any relevant activity loss after incubation at 37°C for 18 hours in 50 mM ammonium acetate pH 4.3. They were both stable under these conditions. The E-M-H GluC seems to have consequently the same pH stability in ammonium acetate buffer pH 4.3 as the Roche GluC.

The cleavage specificity of the E-M-H GluC after the 18-hour incubation of the IFCC test remains the same as the Roche GluC, as described in the chapter 5.12.1. However, the 18-hour kinetic showed that the velocity of the hemoglobin hexapeptide cleavage was 3 times greater with the Roche GluC than the E-M-H GluC (Figure 57). The Roche GluC cleaves still faster the hemoglobin hexapeptide than the E-M-H GluC.

The buffer exchange did not improve this result. The Roche GluC activity was already tested in Tris acetate buffers with 100 mM and 50 mM concentrations in previous experiments and any activity improvement could have been also observed.

The E-M-H GluC mutant is as stable as the Roche GluC in the IFCC test. Its hemoglobin cleavage specificity is similar to the Roche GluC. The cleavage speed of the hexapeptide is however still lower even if different buffer and pHs conditions were tested.

In summary, once the E-M-H mutant GluC was targeted by screening, it was expressed extracellularly from *B. subtilis*, purified and activated without major degradation. Further characterizations were then investigated. Its activity is 4 times better than the other GluCs in Tris buffer at pH 7.8 with Z-Phe-Leu-Glu-pNA substrate and ammonium acetate at pH 4.3 with the hemoglobin hexapeptide. Its affinity is improved with a factor 3 towards the Z-Phe-Leu-Glu-pNA substrate. Its pH optimum is extended up to pH 6. It is also more thermostable and its isoelectric point is not affected by the mutation. Its stability and specificity in the hemoglobin test does not differ compared to the Roche GluC. However, any velocity improvement could be observed in the hemoglobin hexapeptide cleavage with the use of the E-M-H GluC mutant.

DISCUSSION AND CONCLUSION

Diabetes is mainly characterized by blood hyperglycemia. This hyperglycemia is induced by the glycosylation of one major component of red cells, the hemoglobin. In diabetes management, glycemic control is correctly assessed by a complementary combination between the self-monitoring blood glucose and the HbA1c test. Self-monitoring of blood glucose is the best way to manage oneself its own insulin injection according to meals or exercises. It reduces thus hypoglycemia risks in different daily life conditions and gives more freedom and flexibility with daily activities and exercises. The HbA1c test is a clinical blood glucose control from the previous 2-3 months. Blood glucose determination via the rate of glycated hemoglobin (HbA1c) is a key-parameter for long-term diabetes care. A regular test of the glycated hemoglobin every three months gives a good estimation of the average blood sugar and thereby on treatment efficiency and patient´s compliance. Roche Diagnostics GmbH committed itself with the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) to develop an international reference standard system of the HbA1c test.

The first preparative step of the HbA1c measurement by ESI-MS consists in a proteolytic digestion of the HbA1c. The V8-Glutamyl endopeptidase (GluC) of *Staphylococcus aureus*, EC 3.4.21.19 was selected as the most appropriate endoprotease for the glycated hexapeptide digestion in the HbA1c test by ESI-MS measurement (Houmard and Drapeau, 1972). The V8-Glutamyl endopeptidase can hydrolyze specifically after the 6th amino acid of the HbA1c β-chain, a glutamic acid, to generate a hexapeptide, which is glycated on its first amino acid, a valine (Houmard and Drapeau, 1972).

However, the slow proteolytic digestion, which was optimized up to 18 hours for the glycated hemoglobin sample preparation, prevents this test from being routinely employed in clinical laboratories.

Furthermore, the V8-GluC is well-known for its buffer dependency specificity (Drapeau, 1977; Houmard and Drapeau, 1972). The HbA1c measurement must occur under strict experimental conditions, in a buffer of ammonium acetate at acidic pH, which confers to the GluC the cleavage specificity toward the glutamate bond of the glycated hemoglobin hexapeptide. The experiment conditions must be kept in order to keep this cleavage specificity of the V8-GluC. As former publications demonstrated the GluC production feasibility, Roche Diagnostics GmbH would have liked to get a GluC mutant, which could hydrolyze faster the glycated hemoglobin, in order to reduce blood glucose sample time preparation (Wu et al., 1991; Yabuta et al., 1995).

The aim of this phD consisted in setting a strategy up to generate GluC mutants by directed evolution and irrational design, to select among them the GluC variant with the desired characteristics by High Throughput Screening. The desired mutant should be produced after complete characterization so that this HbA1c method could be then routinely applied by enzymatic digestion and RP-HPLC/MS-ESI assessment.

Directed evolution is a protein engineering technique to evolve proteins with the desired enzymatic characteristics and is therefore here a very promising approach to generate variant libraries for high throughput screening system. Directed evolution follows an iterative process, which can be divided into three main parts in order to achieve this purpose. At first, the wild type V8-glutamyl endopeptidase is cloned in a recombinant organism to apply directed evolution on the gene and produce the optimized variant afterwards. A pool of variants can be then generated by different mutagenesis techniques according to the structure study. The screening system is implemented according to the enzymatic characterization of the wild-type GluC on the different enzymatic criteria to maintain or improve. High Throughput Screening system is performed up to get a mutant hit. After that results reproducibility are checked out, the mutant can be cultivated, purified and activated or can undergo new variation generation. Once mutant characteristics are assessed, new mutant features can be tested on the functional HbA1c test (Figure 58).

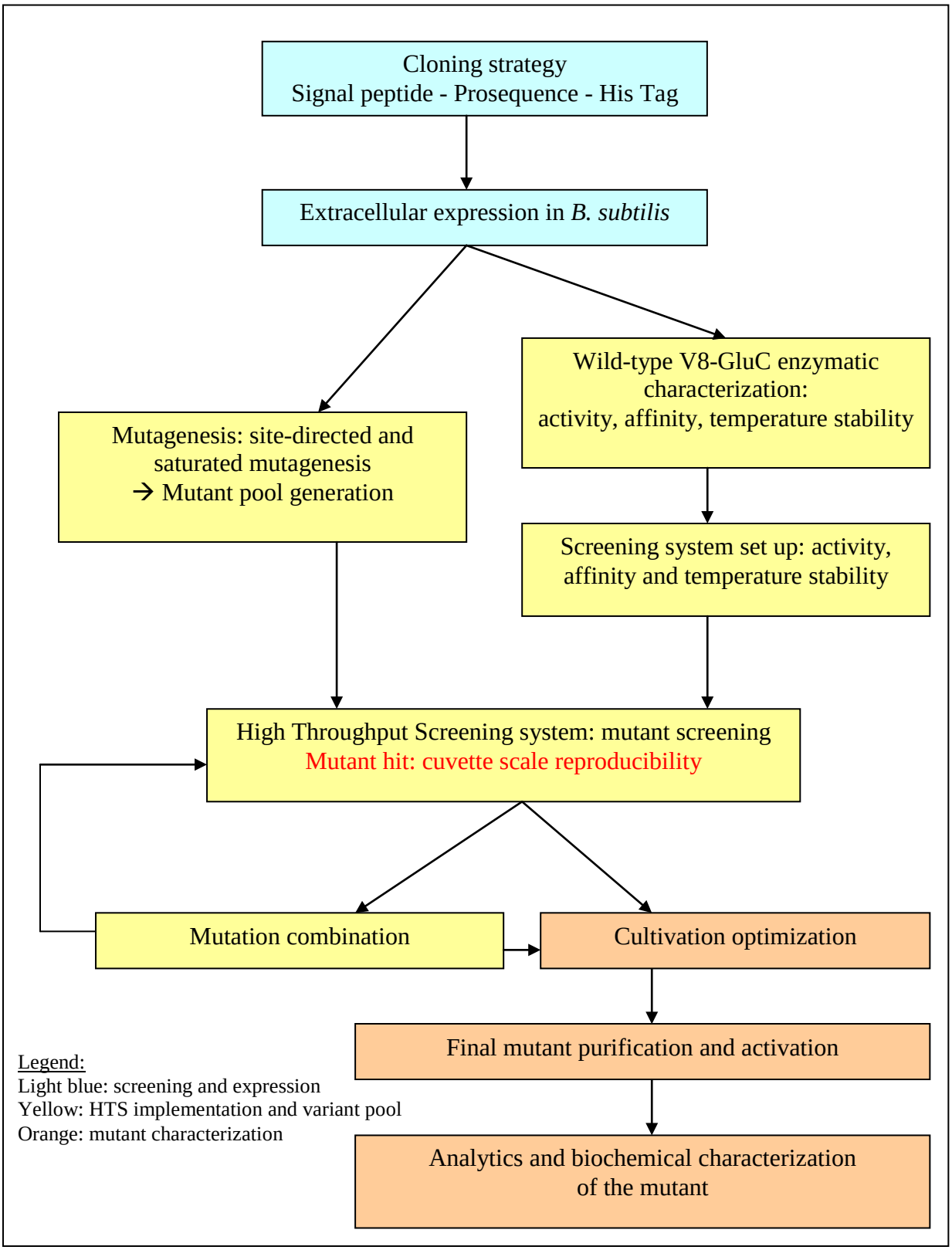


Figure 58. Project overview. Followed strategy for the directed evolution, screening system implementation on the wild type V8-GluC and the final characterization of a V8-GluC mutant.

1. Cloning and expression

The first step of the phD project consisted of the cloning and the expression of the V8-GluC.

1.1 Different V8-Glutamyl endopeptidases

Two different glutamyl endopeptidase isoforms are synthetized by the V8 *Staphylococcus aureus* strain. The “Roche GluC”, which was sequenced by Drapeau, is commercialized as lyophilisate GluC by Roche GmbH (Drapeau et al., 1972). The other “V8-GluC” was discovered by Carmona and is provided on the market by the New England NEB Ltd as a recombinant protein from *B. subtilis* (Carmona and Gray, 1987). The V8-GluC from Carmona was the only available GluC sequence from the V8 *Staphylococcus aureus* strain at the German Collection of Microorganisms and Cell Cultures (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH). The V8-GluC sequence from Carmona was used thus for the cloning. The recombinant form of the Roche GluC was prepared by site directed mutagenesis of the V8-GluC from Carmona.

1.2 *B. subtilis* as a recombinant host

The V8-GluC protease could not be expressed in the conventionnal recombinant organism, *E. coli*, because of its toxicity. Indeed, the V8-Glutamyl endopeptidase was already expressed in *E.coli* with a yield of 0.1g/L and after inclusion body solubilization (Yabuta et al., 1995). Last recent work on a glutamyl endopeptidase from *Staphylococcus epidermis* proved the feasibilty of recombinant expression in *E. coli* but with a yield of 3 to 6 mg/L (Nemoto et al., 2008). Other glutamic acid endopeptidase from *S. aureus* ATCC12600 or *Bacillus intermedius* were more successfully expressed in *Bacillus subtilis* (Kakudo et al., 1992; Sharipova et al., 2007). The NEB GluC is claimed to be expressed from *B. subtilis*. The *B. subtilis* secretion pathway and the non-existence of outer membrane confer to *B. subtilis* a major advantage above all for the cloning, expression of extracellular proteins in the medium but also for the downstream process, such as purification (Young, 1980; Wong, 1995; Lin et al., 2007).

1.2.1 The WB600 strain of B. subtilis

One of *B. subtilis* major drawbacks comes from the high level of proteases, which can degrade heterologuous proteins (Roy et al., 1986; Won et al., 1986). The WB168 strain of *B. subtilis* produces normally eight free proteases and six proteases, which are bound on the cell wall. As any of these proteases play a vital role in the strain, Wu et al. (1991) worked on a *B. subtilis* strain, the WB600, which is deficient in six of these proteases. The WB600 strain of *B. subtilis* was therefore chosen for the V8-GluC expression. The WB600 *B. subtilis* strain has a residual protease activity of 0.32 % of the wild-type strain but the remaining proteases were demonstrated not to interfere with the recombinant V8-GluC in the enzymatic test assay of this phD project.

1.2.2 The pMSE3 vector shuttle

Besides, recombinant proteins cannot be directly transformed in *B. subtilis*. A pre-transformation in *E. coli* is required in order to collect a sufficient DNA amount for the transformation in *B. subtilis*. This drawback has nevertheless the advantage that the *B. subtilis* protein expression system is based on the same principles as the *E. coli* protein expression system. Very slight differences can be observed during the transcription, protein folding translocation through the plasmatic membrane and the signal peptide processing (Li et al., 2004).

Recombinant proteins need to be cloned in an expression vector shuttle, which can be transferred into *E. coli* and *B. subtilis*. A first vector was developed by Ehrlich (1978). Then, series of vector shuttles for gene cloning, expression and regulation were constructed in order to transform *E. coli* and *B. subtilis* (Brückner, 1992; Kakudo et al., 1992; Ohashi et al., 2003; Brockmeier et al., 2006; Nguyen et al., 2007). For this project, the pMSE3 vector shuttle was offered by Silbersack et al. (2006). The pMSE3 plasmid harbours two replication origins, one for *E. coli*, one for *B. subtilis*. These replication origins stabilize the plasmid in both organisms in order to facilitate its transfer between both recombinant hosts.

1.2.3 The AcoA promoter

The AcoA promoter is commonly used among a lot of appropriate promoters for *B. subtilis* (Geissendorfer and Hillen, 1990; Henner, 1990; Le Grice, 1990; Kim et al., 1996; Lam et al., 1998). This promoter was identified during research on the proteome and transcriptome of *B. subtilis*. Transcription is repressed by glucose, but can be induced by acetoin, an alternative carbon source, as soon as glucose is exhausted. During growth on glucose, *B. subtilis* produces overflow metabolites, which are secreted in the cultivation medium. Besides weak organic acids, large quantities of the neutral metabolite acetoin (3-hydroxy-2-butanone) are produced, so that after exhaustion of the preferred carbon sources, acetoin and organic acids are reused as alternative energy source (Nakano et al., 1997; Huang et al., 1999). The acetoin is remetabolized by an acetoin dehydrogenase complex in the early stationary phase of *B. subtilis* (Bernhardt et al., 2003). Thanks to acetoin, the low basic level can be multiplied to a strong transcription and translation capacity (Ali et al., 2001). The transcription yield depends on the acetoin concentration but could not be studied during this phD project. The basic concentration was therefore used for the wild type and mutant V8-GluC expression.

1.3 The V8-GluC extracellular expression in *B. subtilis*

The WB600 *Bacillus subtilis* strain was chosen as the most appropriate recombinant host for this project to secrete the V8-GluC extracellularly. Thanks to the secretion pathway, the recombinant host could not be damaged by the protease activity of the V8-GluC. The V8-GluC was cloned with the AmyE signal peptide, its wild type prosequence and a polyhistidine tag to be expressed in the WB600 *Bacillus subtilis* strain after cultivation. *Bacillus subtilis* cell growth reaches a plateau phase after 24 hours of cultivation, which can be explained by the longer generation time of 40 minutes of *B. subtilis* compared to the generation time of 20 minutes of *E. coli*.

1.3.1 The AmyE signal peptide

Any V8-GluCs could be expressed extracellularly with their wild type signal peptide from *S. aureus* as expected. The *S. aureus* signal peptide could not be recognized by the Sec machinery mechanism or the other translocation systems of *B. subtilis* such as the TAT-pathway (Twin Arginin Translocation) or the ABC transporter (ATP-Binding Cassette). At the opposite, *Bacillus subtilis* signal peptides, such as the AmyE signal peptide, are routinely used for the extracellular expression of heterologous proteins in *B. subtilis* (Quentin et al., 1999; Jongbloed et al., 2000; Mori and Ito, 2001; Tjalsma et al., 2000; Li et al., 2004).

The AmyE signal peptide is one of the most appropriate signal peptide, which can be successfully used in *B. subtilis* for recombinant protein expression among several *B. subtilis* signal peptides (Brockmeier et al. 2006; Wong et al. 1986).The signal peptide AmyE of the alpha-amylase of *B. subtilis* is well recognized and it cleaves completely the signal peptidase (Sasamoto et al., 1989). The *S. aureus* signal peptide of the V8-GluC was replaced as a consequence by the AmyE signal peptide during the cloning step.

1.3.2 Degradation of the V8-GluCS during their expression

The carboxyl-terminal tail of the V8-glutamyl endopeptidase presents usually a particular 12-fold repeat pattern of the tripeptide proline-aspartate-asparagine. This pattern may protect the V8-glutamyl endopeptidase from its own digestion and it was also postulated that it may be involved in cytoplasmic membrane transport (Drapeau, 1978).

In the first trials, the cultivation was stopped once the V8-GluC activity could be detected and measured. According to ESI-MS analysis of the molecular weight and sequence of the protease, the V8-GluC was expressed under several digested forms at the carboxyl-terminal side. As all these digested V8-GluC forms had also different enzymatic activities besides, it confirmed the protector role of the V8-GluC tail against its autodigestion and it raised also the importance and the role of this pattern on the protease activity.

Fusion tags facilitate protein detection and purification. As a polyhistidine tag could be added on the NEB GluC as well as on the glutamyl endopeptidase of *Staphylococcus epidermis* for its expression in *E. coli* (Namoto et al., 2008), a polyhistidine tag was added by PCR during the V8-GluC cloning step. It did not interfere with the enzyme activity of the wild type and the mutant GluCs. Consequently, its removal was not required (Terpe, 2003; Wu and Filutowicz, 1999). However, the polyhistidine tag and the two alanines at the carboxyl-terminal end were not found in the recovered protein after the first cultivation trials. This may be due to autodigestion as the two alanines are located after an aspartic acid. The other possibility comes from a digestion by the other present proteases of *B. subtilis* as the WB600 *B. subtilis* strain has normally still a residual protease activity, which was measured distinctly after 144 hours by Kakudo (1992).

1.3.3 The prosequence cleavage of the V8-GluC

Researchers have been always confronted to these autodigestion problems for the glutamyl endopeptidase expression when the protease was not cloned with its prosequence (Nemoto et al., 2008). The V8-glutamyl endopeptidase was expressed once without its initial prosequence but it was via a tedious cloning work in *E. coli* and the expression yield after denaturation and renaturation was limited to 20 % (Yabuta et al., 1995).

The prosequence is present in quite all proteases at the amino-terminal side between the signal peptide and the active protease sequence (Freudl, 1992). It plays a role in the correct protease folding after the membrane passage (Ohta et al., 1991; Harwood and Cranenburgh, 2008; Nemoto et al., 2008). The wild type prosequence of the V8-GluC has no influence on the refolding of the V8-GluC unlike other proteases and there was any obligation to clone it in this phD project (Yabuta et al., 1995). The V8-GluC was however cloned with its prosequence for its second very important role, a competitive inhibitor of the active enzyme, which prevents the protease from its activation and thus, from its own potential autoproteolysis (Yabuta et al., 2001).

1.3.3.1 The activation process of the V8-GluC

Signal peptide and prosequence processes occur after secretion from *B. subtilis* as it was previously shown for the alpha-amylase expression in *B. subtilis* (Takase et al., 1988). In the first cultivations, enzymatic activity was detected after 48-hour cultivation even if the last three amino acids of the prosequence, histidine-alanine-asparagine (HAN), were still present. Two hypothetic conclusions came out from this observation. The V8-GluC was already active after the removal of the major part of its prosequence. However, this activity remained partial as the prosequence was not completely cleaved. Secondly, the V8-GluC prosequence could be hydrolyzed by other proteases of *B. subtilis* or the V8-GluC could autodigest itself at the upstream position of the HAN pattern, as it is a glutamic acid. This was already observed in previous studies, where a glutamyl endopeptidase from another *S. aureus* strain, ATCC 12600, was expressed in a non-protease deficient *Bacillus subtilis* strain, ISW1214 (Kakudo et al., 1992). In this publication, transformed *Bacillus* cells reached a OD_{600nm} maximum at 72 hours and produced the maximum amount of glutamyl endopeptidase after 96 hours. This statement confirmed the expression results of this phD project. The V8-GluC production was delayed compared to the cell growth, mainly due to the activation by the prosequence digestion.

Protease prosequences are usually removed either by autodigestion or by other coexpressed proteases in the medium (Power et al., 1986; Takase et al., 1988; Chang et al., 1997). Both mechanisms occur in the case of the V8-GluC. The V8-Glutamyl endopeptidase propeptide is activated by prosequence cleavage at first by sequential autoproteolysis and then, by a coexpressed zinc metalloprotease of *S. aureus*, the aureolysin EC 3.4.24.29 (Drapeau, 1978; Nickerson et al., 2007). Aureolysin was initially supposed to be the only protease to activate the V8-glutamyl endopeptidase by splitting the last amino acid of the prosequence, an asparagine (Drapeau, 1978; Shaw et al., 2004). Mutant *S. aureus* strains without metalloprotease could not secrete the V8-glutamyl endopeptidase. However, after recent investigations by Nickerson et al. (2007), it was discovered that activity could not be always measured, even if the V8-GluC was expressed. The activation of the V8-Glutamyl endopeptidase occurs indeed in two different complementary ways: first a sequential autocatalysis and then aureolysin hydrolysis (Ohara-Nemoto et al., 2008).

1.3.3.2 The N-1E mutation in the prosequence

To strenghten this previous finding in the phD project, a mutation N-1E was introduced in the prosequence of the wild type and the mutant G166I GluCs to accelerate its cleavage. This activation process could thus happen by auto-hydrolysis or through the intervention of other commercial glutamyl endopeptidases. Activation tests by incubation of the purified V8-GluC demonstrated the activation of the V8-GluC with the N-1E mutation by auto-hydrolysis under defined conditions whereas the cleavage by other V8-GluCs did not work. This underlines the importance of appropriate conditions to hydrolyze the prosequence: glutamyl endopeptidase amount, incubation temperature and buffer as experiment conditions could not be optimized in the activation test with other V8-GluCs. It would raise however the problem of distinction between the commercial V8-GluCs and the recombinant one during the purification step.

These investigations explained why the prosequence of the V8-GluC was hydrolyzed by its own autodigestion in the first cultivation experiments. Consequently, cultivation trials were stopped at the early stage of the expression in order to avoid the carboxyl hydrolysis of the protease by other proteases from *B. subtilis* or by itself. As the cultivation was stopped at its early stage, the prosequence hydrolysis process could not start. The V8-GluC was kept in its inactive form until its purification to prevent from degrading itself. Further experiments could be tried out to solve these degradation problems by adding a stabilizator. For example, the NEB company uses currently a Glu-Glu dipeptide, as stabilizator, which has however no effect on stability concerning the E-M-H GluCs.

1.3.3.3 Impact of the prosequence cleavage on the screening system establishment

The sequential prosequence hydrolysis process was published only at the end of this phD project. This raises now the problem that the screening system could have been established in another way. The aureolysin was added during the screening without knowing if the autocatalysis had already started or not. The aureolysin had no effect on the V8-GluC as long as the autocatalysis did not start. This explained the screening system pretest results where the V8-GluC did not show any relevant activity increase after the aureolysin incubation. Sequential hydrolysis could not have time to occur so that the aureolysin could access its cleavage site.

Even if the screening system was not set up on the right bases for the activation of the V8-GluC, mutants with the desired enzymatic features could be however correctly screened. The prosequence activation is based on the V8-GluC autodigestion and therefore the V8-GluC activity, one of the screened enzymatic parameter. If a V8-GluC mutant is characterized by a higher activity, the prosequence is faster hydrolyzed and the mutant is more rapidly activated and targeted in the screening system. As the activation step with the aureolysin incubation was kept in the screening system, the G166I mutant is anyway more active than the wild type V8-GluC. The wild type V8-GluC, E-Wt-H, with the N-1E mutation would be however interesting to test on the screening system to improve it.

As a conclusion of the cloning and expression step of this phD project, a N-1E mutation was introduced in the prosequence to facilitate the own V8-GluC activation by auto-hydrolysis. Enzymatic activity could not be considered as real activity in the expression experiments as long as the V8-GluC was not activated by prosequence cleavage. As the activity could not be measured and the V8-GluC protein was not visible on a SDS-PAGE gel after the short expression time, the cultivation could not be optimized under these conditions and could be hardly compared to other protease expression. The activation must be performed in defined conditions but the prosequence and carboxy-terminal side degradations may be avoided by addition of a stabilizator. The late comprehension of the prosequence cleavage process did not affect the screening system implementation.

2. Mutagenesis of the V8-GluC

As the V8-Glutamyl endopeptidase could not be directly transformed in *B. subtilis*, a pre-transformation in *E. coli* was required to collect a sufficient DNA amount for the transformation in *B. subtilis*. The V8-Glutamyl was thus cloned in an expression vector shuttle, which can be easily transferred into both organisms. Even if this cloning complicated the directed evolution on the gene, it gave the advantage that all directed evolution techniques, suitable in *E. coli*, could be transferred to *B. subtilis*.

2.1 The recombinant wild type V8-GluC and Roche GluC

As two V8-GluCs exist, the first priority step was to clone and express the Roche GluC of Drapeau (1972) by site directed mutagenesis of the V8-GluC sequence of Carmona (1987). As the recombinant Roche GluC expression was tested at the screening scale, its expression could be proved only by its activity measurement. Any enzymatic activity could be however measured after screening and further analysis should be investigated.

Two assumptions could be however ventured. In the first one, the glutamyl endopeptidase might be incorrectly folded and it might be expressed inactive in form of inclusion bodies. In the second one, the recombinant Roche GluC activity could not be detected because of the non-cleavage of the prosequence. As the screening system was set up on the wild type V8-GluC, if the recombinant Roche GluC was at least as much active as the recombinant wild type V8-GluC, the prosequence would have been cleaved by autoproteolysis and some activity could have been detected. The Roche GluC could not be therefore expressed or was not correctly folded. Only further investigations could optimize the Roche GluC expression.

Furthermore, the screening of the 11 amino acid positions, which differ between the Drapeau and Carmona sequences, could not help to explain the sequence differences as any significant enzymatic parameter result were observed. The hemoglobin functional test showed that the specificity of both GluCs is similar. As their specificity was similar in the functional test and the recombinant Roche was not correctly expressed, to improve the V8-GluC rather than the Roche GluC was in this case more appropriate.

It would be however interesting to test the N-1E mutation on the prosequence hydrolysis as well as the G166I mutation as it may interfere differently between the exchanged amino acids of Drapeau sequence and the Carmona one. A new protein modelling or computer analysis could help to study the amino acid interactions between the amino acid exchange mutation and the Drapeau GluC sequence.

2.2 The G166I mutation

Residue contacts of the G166I position were determined with the LPC/CSU software. This software enables to analyze protein structure and complex by providing a common definition of inter-atomic contacts and by complementing contacting surfaces (Sobolev et al., 2005). The program analysis shows that the valine 1 is directly in contact with the 166 position. The valine 1 plays a fundamental role in the catalysis because of its spatial position directly at the entrance of the catalytic site path so that the substrate can bind. The amino acid 166 is also directly in contact with the amino acids upstream and downstream, which were also tested by saturated mutagenesis but the screening at the positions 165, 167, 168 and 169 did not give any improvement on the enzymatic parameters of the screening. This underlines however that the amino acid 166 is also in the surrounding of the serine 169 of the catalytic triad by the side chain. It would be interested to start new screening rounds after mutagenesis of these contact positions on the mutant G166I and to observe the effect of the mutation on the Roche GluC enzyme.

In summary of the mutagenesis part of this project, the recombinant Roche GluC could not be expressed correctly. As its specificity in the hemoglobin test is the same as the recombinant wild type V8-GluC, the wild type V8-GluC was chosen for the screening system establishment.

As the mutant G166I was found, it would be interested to mutate this position in the Roche GluC sequence and to perform saturated mutagenesis with the amino acids, which interact with this position. The wild type and mutant G166I V8-GluCs were designated E-Wt-H and E-M-H respectively, where the H designates the polyhistidine tag cloning, the E, the N-1E mutation in the prosequence and wild type and G166I mutant GluCs are abbreviated by Wt and M.

3. Characterization of the mutant with improved activity

3.1 Purification and activation of the wild type and mutant GluCs

After their extracellularly expression from *B. subtilis*, the E-Wt-H and the E-M-H GluCs were purified. The Roche V8-Glutamyl endopeptidase is usually purified from *S. aureus* by ammonium sulphate and acetone precipitations and a chromatography step on a weak anion exchanger, a DEAE cellulose. A preparative electrophoresis was added as last polishing step (Drapeau et al., 1972). The DEAE sepharose is the most common purification step to purify glutamyl endopeptidases from *S. aureus* (Yoshikawa et al., 1992). Proteases, which are directly purified from *B. subtilis*, follow the same purification scheme. Most of the time, proteases are purified by an ammonium sulphate precipitation step and an anion exchanger chromatography (Gerze et al., 2005; Orhan et al., 2005). For the glutamyl endopeptidase of *B. intermedius*, a gel filtration step was added at the end for the purification polishing (Rebrikov et al., 1999). Glutamyl endopeptidase of *B. subtilis* was isolated by following the same purification flow chart but the anion exchanger step was substituted by an affinity chromatography (Okamoto et al., 1997). This purification process was followed for the glutamyl endopeptidase of *B. licheniformis* but in this case, a prepurification was run with an anion exchanger (Kakudo et al., 1992; Svendsen and Breddam, 1992). Recombinant proteins, which are purified today from the WB600 *B. subtilis* strain, undergo an ammonium sulphate fractionation, anion exchange chromatography and gel filtration (Liu et al., 2008).

At the beginning, the polyhistidine tag of the E-Wt-H and E-M-H GluCs were degraded or completey digested and the E-Wt-H and E-M-H GluCs could not be purified by affinity of the polyhisidine tag on the Nickel-NTA matrix. A DEAE sepharose was thus performed but the recovered GluCs were autodigested and could not be purified with the expected final activities compared to the Roche GluC. Analysis of the isoelectric point explains these fickleness and irrelevance of the results by the change of the isoelectric point due to the autodegradation of the E-Wt-H. Separation in ion exchange chromatography depends on the reversible adsorption of charged solute molecules to immobilized ion exchange groups of opposite charge. The slightest charge change, due to hydrolysis or degradation of the protease, can cause a change of the isoelectric point of the molecule and therefore cause a big impact on ion exchange chromatography.

The cultivation conditions were studied to get the non-degraded prosequence and carboxyl terminal side. The expression was therefore stopped at the early cultivation stage. As the E-Wt-H and E-M-H GluCs could be expressed with their polyhistidine tag, as well as the NEB GluC, they were both finally purified by affinity chromatography on a Nickel-NTA column. The specific activity was not the final one, as the enzyme was not completely activated and the yield and purity factor of the balance purification were thereby not completely representative. Partial prosequence hydrolysis had time to occur even if the purification step was performed at +4°C.

A great loss of the E-Wt-H and E-M-H GluCs could be observed in the flow through due to the polyhistidine tag degradation. Furthermore, only 1/6 of the yield was reached by this purification step as the prosequence was only partially hydrolyzed and the enzyme was consequently not completely activated. Despite these results, the Nickel-NTA purification step was conserved for the purification because of the previously encountered purification problems as mentioned above. The purification diagram should be optimized if it would be used at the production scale for example.

The E-Wt-H and E-M-H GluCs could be however purified here by Nickel-NTA chromatography and activated as shown by the SDS-PAGE.

3.2 Characterization of the wild type and mutant GluCs

3.2.1 Activity and affinity

The E-M-H GluC was expressed to check the screened enzymatic parameters and determine other biochemical and analytical parameters after purification and activation. The activity and affinity of the V8-GluC were the targeted characteristics to improve for the hemoglobin functional test. The E-M-H GluC is four times more active than other commercial GluCs (Roche, NEB) in Tris buffer at pH 7.8 and in ammonium acetate buffer at pH 4.3. Its affinity is also three times superior in Tris buffer at pH 7.8 towards glutamate substrate.

As the NEB and the E-Wt-H GluCs were cloned with the same GluC sequence, they should have the same specific activity of about 40-45 U/mg towards glutamate substrate at pH 7.8 but the NEB GluC has a specific activity of only 30 U/mg. This slight difference of specific activity towards glutamate substrate at pH 7.8 between the NEB and the E-Wt-H GluC could come from the lyophilisate composition of the NEB. The lyophilisate is supposed to contain the Glu-Glu dipeptide as a stabilizer to avoid the autodegradation of the NEB GluC. The Glu-Glu dipeptide cuts by this way the activity of the NEB GluC, which cannot reach in this case its maximal activity. The Glu-Glu dipeptide may also increase slightly the protein concentration. All the decrease of activity and the increase of the protein concentration because of the Glu-Glu dipeptide explain the specific activity difference between the NEB and E-Wt-H GluC.

Specific activity and affinity cannot be compared to previous research publications as they focused on the Roche GluC and not on the NEB GluC (Carmona and Gray, 1987; Drapeau et al., 1972). Furthermore, the enzyme specificity and thus activity and affinity are highly substrate and buffer dependant (Sorensen et al., 1991). Any previous work was published with the same conditions, substrate and buffers, as this phD project. For its characterization, the E-M-H GluC could be compared only to commercial Roche or NEB lyophilisate V8-GluCs or to the E-Wt-E wild type recombinant V8-GluC.

According to the first enzymatic characterization, the E-M-H GluC is four times more active towards glutamate substrate at pH 7.8 or pH 4.3 and three times more affine towards glutamate substrate at pH 7.8 than the other wild type GluCs, Roche, NEB GluCs.

3.2.2 pH optimum and temperature stability

After the activity and affinity assessment, the E-M-H GluC was characterized on two main intrinsic enzymatic parameters: the temperature stability and the pH optimum.

As mutagenesis can affect the thermostability, the E-M-H GluC thermostability was checked. The E-M-H GluC is more stable than the E-Wt-H GluC over 50°C but any activity can be detected over 55°C for both GluCs. The melting temperature assessment by the DSC experiments confirms the higher thermostability of the E-M-H GluC with a melting point of 55°C. This temperature unstability is mainly due to some autodigestion as deduced in previous work on the Roche GluC (Drapeau et al., 1972).

The activity is strongly correlated with the pH optimum of an enzyme. As it belongs to the intrinsic feature of an enzyme, the pH optimum was assessed for the E-M-H GluC. The pH optimum is shifted from pH 7.8 for the E-Wt-H GluC to pH 7 for the E-M-H GluC. The pH optimum of the Roche GluC is substrate dependent and varies between pH 7 and pH 8 (Nagata et al., 1991). The pH optimum of the E-M-H GluC is extended on a broader pH range from pH 6.5 to pH 8 in this phD project. The E-M-H GluC catalyzes at its maximum activity between pH 6.5 to pH 8. The V8-GluC was already claimed with an optimum pH of 7.2 with an aspartate substrate (Sorensen et al., 1991) but this substrate could not be tested on the E-M-H GluC.

Furthermore, only one optimum pH peak was observed whereas two optimum pH peaks are usually present at pH 4 and 7.8 with the whole hemoglobin as substrate (Drapeau et al., 1972). The Roche GluC catalyzes the whole hemoglobin at its maximum efficiency at two pHs, pH 4 and pH 7.8. These two pH optima may be caused by hemoglobin conformation variations due to the pH change. The Roche GluC has normally only one pH optimum with other synthetic substrates The E-M-H GluC could not be determined however with the whole hemoglobin as a substrate to compare it to the Roche GluC.

By this enzymatic characterization, the E-M-H GluC is shown to be more thermostable than the other commercial GluCs (Roche and NEB GluCs). Furthermore, its pH optimum is also extended to a broader range of pH.

3.2.3 Specificity and hemoglobin functional test

The E-M-H GluC activity, affinity towards glutamate substrates at pH 7.8 as well as it thermostability were improved significantly thanks to one mutation point, G166I, on the E-Wt-H wild type GluC. Another main enzymatic characteristic to assess is its specificity. The V8-GluC is well-known for its two cleavage specificity towards glutamate and aspartate substrates. However, the enzymatic parameters could not be tested towards aspartate substrates. The specificity of the E-M-H GluC was therefore tested on the hemoglobin functional test.

The V8-GluC was already commonly applied on blood samples as second digestion step to determine human hemoglobin variants (Vasseur et al., 1991; Wajcman et al., 1994). Specificity of the V8-GluC was also studied on the alpha hemoglobin by using its condensation feature (Acharya et al., 1987). Two kinetic phases are classically observed during hydrolysis by the Roche V8-GluC. Two steps appeared in the hydrolysis kinetics because some bonds are split faster than the others. This is due to the substrate competition between the original substrate and the formed peptides during the protease digestion (Adler-Nissen, 1985). According to computer investigations, two peptide sequences should be theoretically hydrolyzed from the alpha and beta chains. Glutamic acids at positions 82 on the alpha chain and 100 on the beta chain should be split by the V8-GluC, but this was not observed experimentally (Ivanov et al., 1997). Alpha and beta chains of hemoglobin under their native form are indeed very resistant to the V8-GluC digestion. Partial hydrolysis can occur under denaturing conditions only from isolated chains or after trypsic digestion (Iyer et al., 1987; Vasseur et al., 1991; Wajcman et al., 1994; Apostol et al., 1997). Aspartic bonds are generally not hydrolyzed in native hemoglobin.

The specificity of the E-M-H GluC was checked by the hemoglobin functional test. Even if the G166I mutation on the wild type GluC improves its activity and affinity, this mutation does not improve its cleavage velocity compared to the Roche V8-GluC in the HbA0 functional test at pH 4.3. The 18-hour kinetic at pH 4.3 shows indeed that, even if the hemoglobin specificity remains the same and the E-M-H GluC has the same pH stability than Roche V8-GluC, the velocity towards the hemoglobin hexapeptide is 3 times greater with the Roche GluC than with the E-M-H GluC.

After characterization, the E-M-H GluC was shown to be four times more active towards glutamate substrate at pH 7.8 or pH 4.3 and three times more affine towards glutamate substrate at pH 7.8 than the other wild type GluCs, Roche, NEB GluCs. The E-M-H GluC is also more thermostable than the other commercial GluCs (Roche and NEB GluCs) and its pH optimum is also extended to a broader range of pH. The hemoglobin specificity and the pH stability in the hemoglobin test are not affected by the mutation but the Roche V8-GluC till cleaves 3 times faster the hemoglobin hexapeptide than with the E-M-H GluC.

This result difference comes from the screening system establishment. The screening system in 96-well microtiterplate should be usually the miniaturized duplicate of the hemoglobin functional test. The scale-up transfer is indeed the crucial step of the screening system implementation, as it was previously quoted by F.H. Arnold as the first law of directed evolution “You get what you screen for” (Schmidt-Dannert and Arnold, 1998). However, two factors were changed in the screening system implementation.

At first, the screening system was set up at pH 7.8 instead of a pH 4.3 because of the substrate. Normally, the V8-GluC is well-known to cleave glutamic and aspartic acids at neutral pH and its specificity was previously shown buffer-dependent (Drapeau, 1977; Houmard and Drapeau, 1972). Beta chain of oxidized insulin is split 10 times faster in phosphate than in bicarbonate buffer (Sorensen et al., 1991). However, even if the screening was not set up at a pH 4.3, the mutant E-M-H GluC is three times more active at pH 4.3 with the hemoglobin hexapeptide substrate than the wild type E-Wt-H GluC and the cleavage specificity of the hemoglobin was not affected in the functional test. The transfer of screening from pH 7.8 to pH 4.3 was not the limiting factor for identifying a better mutant.

The second altered factor of the screening system establishment was the substrate. The enzymatic test assay was initially implemented on the hemoglobin hexapeptide instead of the whole hemoglobin in order to detect the V8-GluC activity by spectrophometer. The substrate was in this project the limiting screening factor. This limiting factor is a direct application of the first law of the directed evolution “You get what you screen for”, where the importance of the right screening implementation is emphasized to get a mutant with the desired properties (Schmidt-Dannert and Arnold, 1998).

During this phD project, other possibilities could have been studied and tested besides the directed evolution technology in order to improve the functional hemglobin test.

Monovalent anions, such as borate, bicarbonate, chloride, acetate and nitrate, inhibit indeed the V8-GluC by ionic strength increase, which disrupts ionic interactions (Austen and Smith, 1971; Houmard, 1976). Ammonium bicarbonate and acetate generate thus only inhibitory effects. This inhibition of the Roche GluC is caused by the ionic strength increase, which disrupt ionic interactions. In the opposite, phosphate and sulphate activate the enzyme activity due to their tetrahedral configuration unlike the planar configuration of the mentioned anions (Sorensen et al., 1991). After this discovery, a lower enzyme to substrate ratio was recommended in a phosphate buffer than a high enzyme to substrate ratio in ammonium bicarbonate buffer.

Even if this recommendation can be applied to the HbA1c test in ammonium acetate buffer, it cannot be used because of the specificity loss of the V8-GluC. The glycated hexapeptide cannot be therefore digested in these buffer conditions. Furthermore, the buffer exchange in the assays of the hemoglobin test did not bring any improvement to this result. The experimental conditions of the hemoglobin test to accelerate the cleavage can be hardly changed because of the cleavage specificity.

Another possibility could be to use other glutamyl endopeptidases for the hemoglobin functional test, glutamyl endopeptidases from *Bacillus subtilis* and *Staphylococcus epidermis*. The ones from *Bacillus* are patented by Novo Nordisk and the one from *Staphylococcus epidermis* is the main research theme of a japanese research team (Nemoto et al., 2008; Ono et al., 2008). Hence, these glutamyl endopetidases could not be used for the phD project. Furthermore, the specificity of these glutamyl endopeptidases are not well established towards the hemoglobin as substrate and their specificity are also dependent on the other amino acid residues around the cleavage site (Breddam et al., 1992).

As a conclusion after characterization of the mutant GluC, the E-M-H GluC is four times more active in Tris buffer at pH 7.8 and in ammonium acetate buffer at pH 4.3. Its affinity is also three times superior in Tris buffer at pH 7.8 towards glutamate substrate. The E-M-H GluC is also more thermostable and its pH optimum is extended to a broader pH range. Its specificity is the same as the Roche GluC in the hemoglobin functional test. Despite all promising results of the specific activity and affinity improvements of the E-M-H GluC on the glutamate synthetic substrates in Tris buffer at pH 7.8, at the cuvette scale, the E-M-H GluC does not cleave faster the hemoglobin hexapeptide in the hemoglobin functional test at pH 4.3 because of the wrong implementation of the screening system on the hemoglobin hexapeptide instead of the whole hemoglobin molecule.

The V8-GluC specificity was consequently altered by two factors in the screening system. In this phD, the substrate choice plays above all a major role in the screening system. The screening system should have been indeed established on the whole hemoglobin intead of the hemoglobin hexapeptide only. The E-M-H mutant GluC demonstrated during this phD that it would have been also necessary but not obligatory to use the pH 4.3 for the right cleavage specificity in order to be in the same conditions as the hemoglobin functional test and to apply the most correctly as possible the first law of directed evolution “You get what you screen for”.

4 Conclusion and outlook

The HbA1c test is a long-term blood glucose key parameter for the management of diabetes. Roche Diagnostics GmbH has collaborated with the IFCC in the frame of the HbA1c reference standardization by ESI-MS measurement. The V8-Glutamyl endopeptidase of *Staphylococcus aureus*, EC 3.4.21.19, was selected as the most appropriate endoprotease for the glycated hexapeptide digestion in the HbA1c test by ESI-MS measurement (Houmard and Drapeau, 1972). However, the slow proteolytic digestion, which was optimized up to 18 hours for the glycated hemoglobin sample preparation, prevents this test from being routinely employed in clinical laboratories.

The aim of this phD was focused on the production of a tailor-made V8-GluC, improved on its velocity and affinity towards the glycated hemoglobin by directed evolution and irrational design, so that the HbA1c functional test could be then routinely applied by enzymatic digestion and RP-HPLC/MS-ESI assessment. The tricky part of this project was to keep the same other enzymatic parameters: thermostability and above all specificity of the optimized V8-GluC by respecting the experimental conditions of clinical laboratory. The V8-GluC must be indeed incubated with the hemoglobin sample at 37°C in a proportion of 1:100 (enzyme:hemoglobin) in ammonium acetate buffer at pH 4.3.

Firstly, the V8-GluC was cloned and expressed. After that the high throughput multiparameter screening system was implemented, the wild type V8-GluC was mutated by site-directed and random mutagenesis strategies according to its structure knowledge. A mutation could be identified as efficient during the screening, G166I. A variant was expressed with this mutation in *B. subtilis* by extracellular secretion. Another mutation, N-1E, was added in the prosequence to facilitate the prosequence cleavage. The E-M-H mutant was purified by Ni-NTA affinity chromatography thanks to the polyhistidine tag addition. It was then activated by autoproteolysis to be further characterized. The specific activity of the variant is four times enhanced and the affinity constant (Km) is three times improved in a Tris-HCl buffer at pH 7.8, compared to the wild type glutamyl endopeptidase. The specific activity is also four times better in ammonium acetate buffer at pH 4.3 with the hemoglobin hexapeptide as substrate and it is more thermostable at 50°C.

However, although the variant is more active with the synthetic substrate, it is less competitive than the wild type V8-Glutamyl endopeptidase regarding to the activity, affinity and cleavage specificity of the hexapeptide of the whole hemoglobin. This comes from the screening system, which was established on other conditions than the functional test. The screening system should have been implemented on the whole hemoglobin as a substrate in order to remain as close as possible to the hemoglobin functional test and to respect the first law of the directed evolution “you get what you screen for” (Schmidt-Dannert and Arnold, 1998).

The project could be continued according to different aspects. At first, the screening system could be implemented differently for the prosequence activation with the breakthrough on its sequential hydrolysis and new directed evolution round could be started. New protein engineering could be performed on the E-M-H mutant GluC as a consequence and the optimization of the V8-GluC could be continued. Site directed mutagenesis in the surrounding of the glycine 166 would be also interesting to mutate as some of them interact directly with this amino acid. The other possibility is to substitute the position G166I in the sequence of the Roche GluC to observe the effect on its activity, affinity and specificity towards the glycated hemoglobin. However, it was not possible to get the protein expressed with the Drapeau sequence after mutagenesis of the Carmona’s sequence. One approach in this case would be to extract the DNA from the fermentation broth. It was however not possible as Roche does not produce the GluC by itself but buys the purified lyophilized GluC enzyme.

The G166I substitution could be also tested on glutamyl endopeptidases from other organisms even if the *Bacillus subtilis* and *Bacillus licheniformis* glutamyl endopeptidases are patented by Novo Nordisk.

The long HbA1c digestion, which lasts normally 18 hours, prevents fast measurement turnover. Different research teams tried to improve it by alternative methods, such as a microwave assisted enzymatic digestion (Vesper et al., 2005). However, even if it works with the trypsin protease enzyme, which is commonly used in mass spectrometry measurements, the method cannot be used with the V8-GluC in the HbA1c test because of high temperature.

Other alternative methods than the directed evolution technique may be possible to implement for solving the IFCC request about the HbA1c test. The V8-GluC has been already coupled with cross linked agarose beads in the early 90's (Sahni et al., 1991). This immobilization does not seem to interfere with the specificity of the V8-GluC on the alpha chain of hemoglobin at pH 4 and 37°C but the activity was decreased by 30 %. Ten years later, specificity and activity are both improved by adsorption of the V8-GluC onto a solid phase in presence of SDS (Vercaigne-Marko et al., 2000). In the same way, pepsin was also immobilized to digest bovine hemoglobin (Ticu et al., 2004). This idea was also developped by Takeda, San Diego, Inc. (Lim et al., 2006). The Roche GluC was immobilized onto a cartridge and directly coupled online to the mass spectrometer. By this method, the enzymatic digestion occurs in one to 5 minutes, depending on the buffer flow rate. The IFCC could use this method as alternative to a tailor-made V8-GluC and transfer it to the HbA1c functional test.

The GluC mutant is however more active and its specificity towards glutamic acid is improved. It could be therefore used for other enzymatic tests, where the V8-GluC is routinely involved. The V8 Glutamyl endopeptidase is indeed commonly used for peptide mapping in proteomics or in amino acid sequencing by its specific digestion in mass spectrometry. It is the third most frequently used protease in protein sequence studies (Keil and Tong, 1988). In these circumstances, the mutant GluC could be used in mass spectrometry fields such as some mass peptide fingerprint in proteomics tests for example (Bakthiar and Nelson, 2001; Corthals et al., 1999).

Thanks to this phD, the molecular features of the V8-GluC and its expression conditions in *Bacillus subtilis* could be studied and analyzed. The phD project gave also by this way a better understanding of the relationship between the structure and the function of the V8-GluC, an enzyme, which is commonly used for industrial applications. The aim of this phD could not be achieved to find a faster V8-GluC mutant for the HbA1c test. However, one position, playing a major role in the activity and affinity increase, could be located in the wild type V8-GluC sequence. Furthermore, the correct screening system establishment was demonstrated to be the crucial step for a successful screening, which confirms the first law of directed evolution: "we get what we screen for" of F. H. Arnold. During this phD project, I could learn by this way a wide range of engineering techniques and to be confronted to their limits.

ABBREVIATIONS

AACC: American Association of Clinical Chemistry

ABC transporter: ATP-Binding Cassette

ACCORD: Action to Control Cardiovascular Risks in Diabetes

ADA: American Association for Diabetes

ADVANCE: Action in Diabetes and Vascular Disease

AEBSF: 4-(2-Aminoethyl)-benzenesulfonyl fluoreide hydrochloride

AG: Aktiengesellschaft

α: Alpha

Amp: Ampicilline

ATP: Adenosine TriPhosphate

B. subtilis: Bacillus subtilis

β: Beta

BSA: Bovine Serum Albumin

Cal: Calory

CE: Capillary Electrophoresis

Cm: Centimeter

C-NPU: Committee on Nomenclature, Properties and Units

Cp: Heat capacity

C-Terminal: carboxyl-terminal

CV: Column volume

Da: Daltons

Dam: DNA-methyltransferase

DCCT: Diabetes Control and Complications Trial

DCM: Designated Comparison Methods

DEAE: Diethylaminoethyl

DFP: Diisopropylfluorophosphate

DNA: Desoxyribonucleic Acid

dNTP: Deoxyribonucleotide Triphosphate

DOF-Hb: Hemoglobin beta chain (blood)-N-(1deoxyfructos-1-yl) hemoglobin beta chain

DSC: Differential Scanning Calorimetry

DSMZ: Deutsche Sammlung von Mikroorganismen und Zellkulturen

DTT: Dithiothreitol

E: Enyzme

EASD: Europeans Association for the Study of Diabetes

E.C: Enzyme classification

EDIC: Epidemiology of Diabetes Interventions and Complications

E. coli: Escherichia coli

E-M GluC: Mutant GluC with the G166I and N-1E mutations and without the His Tag

E-M-H GluC: Mutant GluC with the G166I and N-1E mutations and the His Tag

E-Wt GluC: Wild type GluC with the N-1E mutation and without the His Tag

E-Wt-H GluC: Wild type GluC with the N-1E mutation and the His Tag

ESI-MS: Electrospray Ionisation Mass Spectrometry

GRAS: Generally regarded as safe

Gluc: Glutamyl endopeptidase

GmbH: Gesellschaft mit Beschrankter Haftung

H_2O: Water

HAN: Histidine, alanine, asparagine

HbA0: Non-glycated hemoglobin

HbA1c: Glycated hemoglobin

HCl: Hydrogen Chloride

His Tag: Polyhistidine Tag

HPLC: High Performance Liquid Chromatography

HTS: High Throughput Screening

IDDM: Insulin Dependant *Diabetus Mellitus*

IDF: International Diabetes Federation

IEP: IsoElectric Point

IFCC: International Federation of Clinical Chemistry and laboratory medicine

IGT: Impaired Glucose tolerance

IUPAC: International Union of Pure and Applied Chemistry

JDS: Japanese Diabetes Society

JSCC: Japanese Society of Clinical Chemistry

Kb: Kilobase

kDa: Kilodaltons

KH_2PO_4: Potassium dihydrogen phosphate

KOH: Potassium hydroxide

Km: Michaelis Menten

kV: Kilovolts

LB: Luria Bertani

Lyo: Lyophilisate

M GluC: Mutant GluC with the G166I mutation

M-H GluC: Mutant GluC with the G166I mutation and the His Tag

MS: Mass spectrometry

MTP: Microtiterplate

NaCl: Sodium Chloride

NaH2PO4: Sodium phosphate

NaOH: Sodium hydroxide

NEB: New England Biolabs

NGSP: National Group Standardization Program

NIDDM: Non Insulin Dependent *Diabetes Mellitus*

NI-NTA: Nickel NTA

Nm: Nanometer

NRL: Reference Laboratory Network

N-terminal: Amino-terminal

OD_{600}: Optical Density at 600 nm

PCR: Polymerase Chain Reaction

PEG: PolyEthylene Glycol

PEG-EPO: Pegylated erythropoietine

pI: Isoelectric point

PMSF: Phenylmethylsulfonyl fluoride

pNA: Paranitroanilide

Qs: Quantity sufficient

R&D: Research and development

RM: Reference Method

RP-HPLC: Reverse Phase High-Performance Liquid Chromatography

RPM: Rotation Per Minute

S: Substrat

S. aureus: Staphylococcus aureus

SDC: Steno Diabetes Centre in Copenhagen

SDS: Sodium Dodecyl Sulfate

SDS-PAGE: Sodium Dodecyl Sulfate- Persulfate Ammonium Gel electrophoresis

S.I: Standard Internatiomal Unit

SRP: Signal Recognition Particule

TA: Temperature annealing

TAE: 2-Amino-2-hydroxymethyl-1,3-propanediol

TAT-pathway: Twin Arginin Translocation pathway

TBE: Tris Borate EDTA disodium

Tm: Melting temperature

U: Units

UKPDS: United Kingdom Prospective Diabetes Study

U.S: United States

UV: Ultra Violet

Vmax : Maximale velocity

WG: Working Group

WHO: World Health Organization

Wt GluC: Wild type GluC

Wt-H GluC: Wild type GluC with the His Tag

BIBLIOGRAPHY

A

Acharya, A.S., Cho, Y.J., and Iyer, K.S. 1987. *Staphylococcus aureus* V8-protease catalyzed segment exchange reaction of alpha-chain of hemoglobin S: a semisynthetic approach for the preparation of variants of alpha-chain. *Prog Clin Biol Res*. 240:3-19.

ADA. 2003. American Diabetes Association: standards of medical care for patients with Diabetes mellitus. *Diabetes Care* 26: S106-S108.

ADA. 2008. Diagnosis and classification of diabetes mellitus. *Diabetes Care*. 31 Suppl 1:S55-60.

Adler-Nissen, J. 1985. Enzymic hydrolysis of food proteins. Elsevier Applied Science Publishers, London and New York. 170-186.

Ali, N.O., Bignon, J., Rapoport, G., et al. 2001. Regulation of the acetoin catabolic pathway is controlled by sigma L in Bacillus subtilis. *J Bacteriol*. 183:2497-504.

Apostol, I., Levine, J., Lippincott, J., et al. 1997. Incorporation of norvaline at leucine positions in recombinant human hemoglobin expressed in *Escherichia coli*. *J Biol Chem*. 272:28980-8.

Austen, B.M., and Smith, E.L. 1976. Action of staphylococcal proteinase on peptides of varying chain length and composition. *Biochem Biophys Res Commun*. 72:411-7.

B

Bakhtiar, R., and Nelson, R.W. 2000. Electrospray ionization and matrix-assisted laser desorption ionization mass spectrometry. Emerging technologies in biomedical sciences. *Biochem Pharmacol* 59: 891-905.

Bakhtiar, R., and Nelson, R.W. 2001. Mass spectrometry of the proteome. *Mol Pharmacol*. 60:405-15.

Banbula, A., Potempa, J., Travis, J., et al. 1998. Amino-acid sequence and three-dimensional structure of the *Staphylococcus aureus* metalloproteinase at 1.72 A resolution. Structure. 6:1185-93.

Banting, F.G., Best, C.H., Collip, J.B., et al. 2007. Pancreatic extracts in the treatment of diabetes mellitus. 1922. *Indian J Med Res*. 125:141-6.

Barbosa, J.A., Saldanha, J.W., and Garratt, R.C. 1996. Novel features of serine protease active sites and specificity pockets: sequence analysis and modelling studies of glutamate-specific endopeptidases and epidermolytic toxins. *Protein Eng*. 9:591-601.

Basset, P., Beuzard, Y., Garel, M.C., et al. 1978. Isoelectric focusing of human hemoglobin: its application to screening, to the characterization of 70 variants, and to the study of modified fractions of normal hemoglobins. Blood. 51:971-82.

Baynes, J.W., Bunn, H.F., Goldstein, D., et al. 1984. National Diabetes Data Group: report of the expert committee on glycosylated hemoglobin. *Diabetes Care*. 7:602-6.

Bernhardt, J., Weibezahn, J., Scharf, C., et al. 2003. *Bacillus subtilis* during feast and famine: visualization of the overall regulation of protein synthesis during glucose starvation by proteome analysis. *Genome Res*. 13:224-37.
Bolhuis, A., Tjalsma, H., Smith, H.E., et al. 1999. Evaluation of bottlenecks in the late stages of protein secretion in *Bacillus subtilis*. *Appl Environ Microbiol*. 65:2934-41.
Bordusa, F. 2002. Substrate mimetics in protease catalysis: characteristics, kinetics, and synthetic utility. *Curr Protein Pept Sci*. 3:159-80.
Bradford, M. M. 1976. A rapid and sensitive method for the quantation of microgram quantities of protein utilizing the principle of protein-dye-binding. *Anal. Biochem*. 72, 248-254.
Breddam, K., and Meldal, M. 1992. Substrate preferences of glutamic-acid-specific endopeptidases assessed by synthetic peptide substrates based on intramolecular fluorescence quenching. Eur J Biochem. 206:103-7.
Brockmeier, U., Wendorff, M., and Eggert, T. 2006. Versatile expression and secretion vectors for *Bacillus subtilis. Curr Microbiol*. 52:143-8.
Brockmeier, U., Caspers, M., Freudl, R., et al. 2006. Systematic screening of all signal peptides from *Bacillus subtilis*: a powerful strategy in optimizing heterologous protein secretion in Gram-positive bacteria. *J Mol Biol*. 362:393-402.
Bruckner, R. 1992. A series of shuttle vectors for *Bacillus subtilis* and *Escherichia coli*. *Gene*. 122:187-92.
Bunn, H.F., Haney, D.N., Kamin, S., et al. 1976. The biosynthesis of human hemoglobin A1c. Slow glycosylation of hemoglobin in vivo. *J Clin Invest*. 57:1652-9.
Bunn, H.F., Gabbay, K.H., and Gallop, P.M. 1978. The glycosylation of hemoglobin: relevance to diabetes mellitus. *Science*. 200:21-7.
Bunn, H.F. 1981. Nonenzymatic glycosylation of protein: relevance to diabetes. *Am J Med*. 70:325-30.
Bunn, H.F. 1981. Modification of hemoglobin and other proteins by nonenzymatic glycosylation. *Prog Clin Biol Res*. 51:223-39.

C

Carmona, C., and Gray, G.L. 1987. Nucleotide sequence of the serine protease gene of *Staphylococcus aureus*, strain V8. *Nucleic Acids Res*. 15:6757.
Carr, S.A. and Annan, R.S. 2001. Overview of peptide and protein analysis by mass spectrometry. *Curr Protoc Mol Biol*. Chapter 10:Unit 10 21.
Chang, S.C., Su, M.H., and Lee, Y.H. 1997. Roles of the signal peptide and mature domains in the secretion and maturation of the neutral metalloprotease from Streptomyces cacaoi. *Biochem J*. 321 (Pt 1):29-37.
Chapman, J.R. 1996. Proteins and peptide analysis by Mass Spectrometry. *Humana Press*.

Corthals G.L., Gygi S.P., Aebersold R., and Patterson S.D. 1999. Identification of proteins by mass spectrometry. In *Proteome research: 2D gel electrophoresis and detection methods* (ed. T. Rabilloud), pp. 197–231. Springer, New York.

D

Dalbey, R.E., Lively, M.O., Bron, S., et al. 1997. The chemistry and enzymology of the type I signal peptidases. *Protein Sci*. 6:1129-38.

Dall, T., Mann, S.E., Zhang, Y., et al. 2008. ADA statements: Economic Costs of Diabetes in the U.S. in 2007. *Diabetes Care*. 31:1-20.

50. 68. DCCT. 1987. Feasibility of centralized measurements of glycated hemoglobin in the Diabetes Control and Complications Trial: a multicenter study. The DCCT Research Group. *Clin Chem*. 33:2267-71.

DCCT. 1993. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group. *N Engl J Med*. 329:977-86.

DCCT/EDIC. 1993. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group. *N Engl J Med*. 329:977-86.

DCCT/EDIC. 1999. Epidemiology of Diabetes Interventions and Complications EDIC). Design, implementation, and preliminary results of a long-term follow-up of the Diabetes Control and Complications Trial cohort. *Diabetes Care*. 22:99-111.

Derr, R., Garrett, E., Stacy, G.A., et al. 2003. Is HbA1c affected by glycemic instability? *Diabetes Care*. 26:2728-33.

Diebler, G.E., Nomura, K., and Kies, M.W. 1982. Limited digestion of guinea pig myelin basic protein and its carboxy-terminal fragment (residues 89-169) with Staphylococcus aureus V8 protease. *J Neurochem*. 39:1090-100.

Dluhy, R.G. and McMahon, G.T. 2008. Intensive glycemic control in the ACCORD and ADVANCE trials. *N Engl J Med*. 358:2630-3.

Dole, M., Mack, L.L., Hines, R. and al. 1968. Molecular beams of macroions I. *J Chem Phys* 49:2240-2249

Dole, M, Rheude, A, Mack, L.L. 1970. Molecular beams of macroions II. *J Chem Phys* 52:4977-4986

Dooley, D.C., Hadden, C.T., and Nester, E.W. 1971. Macromolecular synthesis in *Bacillus subtilis* during development of the competent state. *J Bacteriol*. 108:668-79.

Drapeau, G.R. 1977. Cleavage at glutamic acid with staphylococcal protease. *Methods Enzymol*. 47:189-91.

Drapeau, G.R. 1978. Role of metalloprotease in activation of the precursor of staphylococcal protease. *J Bacteriol* 1362. 607-13.

Drapeau, G.R. 1978. Unusual COOH-terminal structure of staphylococcal protease. *J Biol Chem*. 253:5899-901.
Drapeau, G.R., Boily, Y., and Houmard, J. 1972. Purification and properties of an extracellular protease of *Staphylococcus aureus*. *J Biol Chem*. 247:6720-6.
Dreiseikelmann, B., and Wackernagel, W. 1981. Absence in *Bacillus subtilis* and *Staphylococcus aureus* of the sequence-specific deoxyribonucleic acid methylation that is conferred in *Escherichia coli* K-12 by the dam and dcm enzymes. *J Bacteriol*. 147:259-61.
Dubin, G. 2002. Extracellular proteases of *Staphylococcus spp*. *Biol Chem*. 383:1075-86.
Dunn, F.L., Nathan, D.M., Scavini, M., et al. 1997. Long-term therapy of IDDM with an implantable insulin pump. The Implantable Insulin Pump Trial Study Group. *Diabetes Care*. 20:59-63.

E

Ehrlich, S.D. 1978. DNA cloning in *Bacillus subtilis*. *Proc Natl Acad Sci U S A*. 75:1433-6.

F

Fani, R., Mastromei, G., Polsinelli, M., et al. 1984. Isolation and characterization of *Bacillus subtilis* mutants altered in competence. *J Bacteriol*. 157:152-7.
Finke, A., Kobold, U., Hoelzel, W., et al. 1998. Preparation of a candidate primary reference material for the international standardisation of HbA1c determinations. *Clin Chem Lab Med*. 36:299-308.
Fischer, E., and Sauer, U. 2005. Large-scale in vivo flux analysis shows rigidity and suboptimal performance of *Bacillus subtilis* metabolism. *Nat Genet*. 37:636-40.
Fluckiger, R. and Winterhalter, K.H. 1976. In vitro synthesis of hemoglobin AIc. FEBS Lett. 71:356-60.
Forrest, R.D., Jackson, C.A., Gould, B.J., et al. 1987. Short-term effect of a 75 g oral glucose load on glycohaemoglobin levels. *Ann Clin Biochem*. 24 Pt 1.53-7.
Freudl, R. 1992. Protein secretion in gram-positive bacteria. *J Biotechnol*. 23:231-40.
Freudl, R., Schwarz, H., Stierhof, Y.D., et al. 1986. An outer membrane protein (OmpA) of *Escherichia coli* K-12 undergoes a conformational change during export. *J Biol Chem*. 261:11355-61.

G

Gabbay, K.H., Hasty, K., Breslow, J.L., et al. 1977. Glycosylated hemoglobins and long-term blood glucose control in diabetes mellitus. *J Clin Endocrinol Metab*. 44:859-64.

Gaede, P., Vedel, P., Larsen, N., et al. 2003. Multifactorial intervention and cardiovascular disease in patients with type 2 diabetes. *N Engl J Med*. 348:383-93.

Gallagher, S.R. 2006. One-dimensional SDS gel electrophoresis of proteins. *Curr Protoc Mol Biol*. Chapter 10:Unit 10 2A.

Garrick, L.M., McDonald, M.J., Shapiro, R., et al. 1980. Structural analysis of the minor human hemoglobin components: Hb AIa1, Hb AIa2 and Hb AIb. *Eur J Biochem*. 106:353-9.

Gavin, J.R., Alberti, K.G.M.M., Davidson, M.B., and al. 2003. Report of the expert committee on the diagnosis and classification of diabetes mellitus. *Diabetes Care*. 26 Suppl 1:S5-20.

Geier, G.E., and Modrich, P. 1979. Recognition sequence of the dam methylase of *Escherichia coli* K12 and mode of cleavage of Dpn I endonuclease. *J Biol Chem*. 254:1408-13.

Geissendorfer, M., and Hillen, W. 1990. Regulated expression of heterologous genes in *Bacillus subtilis* using the Tn10 encoded tet regulatory elements. *Appl Microbiol Biotechnol*. 33:657-63.

Gerze, A., Omay, D., and Guvenilir, Y. 2005. Partial purification and characterization of protease enzyme from *Bacillus subtilis megatherium*. *Appl Biochem Biotechnol*. 121-124:335-45.

Genuth, S., Alberti, K.G., Bennett, P., et al. 2003. Follow-up report on the diagnosis of *diabetes mellitus*. *Diabetes Care*. 26:3160-7.

Gonen, B., Rubenstein, A., Rochman, H., et al. 1977. Haemoglobin A1: An indicator of the metabolic control of diabetic patients. *Lancet*. 2:734-7.

Goodall, I. 2005. HbA1c standardisation destination--global IFCC Standardisation. How, why, where and when--a tortuous pathway from kit manufacturers, via inter-laboratory lyophilized and whole blood comparisons to designated national comparison schemes. *Clin Biochem Rev*. 26:5-19.

Groche, D., Hoeno, W., Hoss, G., et al. 2003. Standardization of two immunological HbA1c routine assays according to the new IFCC reference method. *Clin Lab*. 49:657-61.

Gu, Y.Q., Holzer, F.M., and Walling, L.L. 1999. Overexpression, purification and biochemical characterization of the wound-induced leucine aminopeptidase of tomato. *Eur J Biochem*. 263:726-35.

Guha, S. 1985. Determination of DNA sequences containing methylcytosine in *Bacillus subtilis* Marburg. *J Bacteriol*. 163:573-9.

Guha, S., and Guschlbauer, W. 1992. Expression *of Escherichia coli* dam gene in *Bacillus subtilis* provokes DNA damage response: N6-methyladenine is removed by two repair pathways. *Nucleic Acids Res*. 20:3607-15.

Gustafsson, E., and Oscarsson, J. 2008. Maximal transcription of aur (aureolysin) and sspA (serine protease) in *Staphylococcus aureus* requires staphylococcal accessory regulator R (sarR) activity. *FEMS Microbiol Lett*. 284:158-64.

H

Hadjadj, S., Saulnier, P.J., Torremocha, F., et al. 2008. To what extent should we lower HbA1c in diabetic subjects? *Diabetes Metab*. 34 Suppl 1:S37-40.

Halimi, S., Charpentier, G., Grimaldi, A., et al. 2001. Effect on compliance, acceptability of blood glucose self-monitoring and HbA1c) of a self-monitoring system developed according to patient's wishes. The ACCORD study. *Diabetes Metab*. 27:681-7.

Harwood, C.R., and Cranenburgh, R. 2008. *Bacillus* protein secretion: an unfolding story. *Trends Microbiol*. 16:73-9.

Hattman, S., Brooks, J.E., and Masurekar, M. 1978. Sequence specificity of the P1 modification methylase (M.Eco P1) and the DNA methylase (M.Eco dam) controlled by the Escherichia coli dam gene. *J Mol Biol*. 126:367-80.

Hedstrom, L. 2002. Serine protease mechanism and specificity. *Chem Rev*. 102:4501-24.

Henner, D.J. 1990. Expression of heterologous genes in *Bacillus subtilis*. *Methods Enzymol*. 185:199-201.

Hochuli, E., Dobeli, H., and Schacher, A. 1987. New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *J Chromatogr*. 411:177-84.

Hoelzel, W., Weykamp, C., Jeppsson, J.O., et al. 2004. IFCC reference system for measurement of hemoglobin A1c in human blood and the national standardization schemes in the United States, Japan, and Sweden: a method-comparison study. *Clin Chem* 50: 166-74.

Houmard, J. 1976. Kinetic investigation of the staphylococcal protease-catalyzed hydrolysis of synthetic substrates. *Eur J Biochem*. 68:621-7.

Houmard, J. 1976. Preparation of chromophoric substrates for the glutamoyl specific staphylococcal protease. *Int J Pept Protein Res*. 8:199-204.

Houmard, J.D. and Drapeau, G.R. 1972. Staphylococcal protease: a proteolytic enzyme specific for glutamoyl bonds. *Proc Natl Acad Sci US* 69: 3506-3509.

Huang, M., Oppermann-Sanio, F.B., and Steinbuchel, A. 1999. Biochemical and molecular characterization of the *Bacillus subtilis* acetoin catabolic pathway. *J Bacteriol*. 181:3837-41.

Hunkapiller, M.W., M.D. Forgac, and J.H. Richards. 1976. Mechanism of action of serine proteases: tetrahedral intermediate and concerted proton transfer. *Biochemistry*. 15:5581-8.

Hutchison, C.A., 3rd, Phillips, S., Edgell, M.H., et al. 1978. Mutagenesis at a specific position in a DNA sequence. *J Biol Chem*. 253:6551-60.

I

IDF. 2007. *In* Diabetes Atlas, 3rd edition. International Diabetes Federation.

Iribarne J.V. and Thomson B.A. 1975. Evaporation of small ions from. charged droplets. *J Chem Phys* 64:2287–2294

Ishida, T. 2006. Low-barrier hydrogen bond hypothesis in the catalytic triad residue of serine proteases: correlation between structural rearrangement and chemical shifts in the acylation process. *Biochemistry*. 45:5413-20.

Ishida, T. and Kato, S. 2004. Role of Asp102 in the catalytic relay system of serine proteases: a theoretical study. *J Am Chem Soc*. 126:7111-8.

Ivanov, V.T., Karelin, A.A., Philippova, M.M., et al. 1997. Hemoglobin as a source of endogenous bioactive peptides: the concept of tissue-specific peptide pool. *Biopolymers*. 43:171-88.

Iyer, K.S., and Acharya, A.S. 1987. Conformational studies of alpha-globin in 1-propanol: propensity of the alcohol to limit the sites of proteolytic cleavage. *Proc Natl Acad Sci U S A*. 84:7014-8.

Izard, J.W., and Kendall, D.A. 1994. Signal peptides: exquisitely designed transport promoters. *Mol Microbiol*. 13:765-73.

J

Javid, J., Pettis, P.K., Koenig, R.J., et al. 1978. Immunologic characterization and quantification of haemoglobin A1c. Br J Haematol. 38:329-37.

Jeffcoate, S.L. 2004. Diabetes control and complications: the role of glycated haemoglobin, 25 years on. *Diabet Med*. 21:657-65.

Jeppsson, J.O., Kobold, U., Barr, J., et al. 2002. Approved IFCC reference method for the measurement of HbA1c in human blood. *Clin Chem Lab Med* 40: 78-89.

John, W.G. and Lamb, E.J. 1993. The Maillard or browning reaction in diabetes. *Eye*. 7 Pt 2.230-7.

John, W.G., Mosca, A., Weykamp, C., et al. 2007. HbA1c Standardisation: History, Science and Politics. *Clin Biochem Rev*. 28:163-8.

Jongbloed, J.D., Martin, U., Antelmann, H., et al. 2000. TatC is a specificity determinant for protein secretion via the twin-arginine translocation pathway. *J Biol Chem*. 275:41350-7.

Jovanovic, L. and Peterson, C.M. 1981. The clinical utility of glycosylated hemoglobin. *Am J Med*. 70:331-8.

Jungbauer, A., and Kaar, W. 2007. Current status of technical protein refolding. *J Biotechnol*. 128:587-96.

K

Kakudo, S., Kikuchi, N., Kitadokoro, K., et al. 1992. Purification, characterization, cloning, and expression of a glutamic acid-specific protease from *Bacillus licheniformis* ATCC 14580. *J Biol Chem*. 267:23782-8.

Kakudo, S., Yoshikawa, K., Tamaki, M., et al. 1992. Secretory expression of a glutamic-acid-specific endopeptidase (SPase) from *Staphylococcus aureus* ATCC12600 in *Bacillus subtilis*. *Appl Microbiol Biotechnol*. 38:226-33.

Karalliedde, J., and Gnudi, L. 2008. ACCORD and ADVANCE: a tale of two studies on the merits of glycaemic control in type 2 diabetic patients. *Nephrol Dial Transplant*. 23:1796-8.

Kawamura, F., and Doi, R.H. 1984. Construction of a *Bacillus subtilis* double mutant deficient in extracellular alkaline and neutral proteases. *J Bacteriol*. 160:442-4.

Keil, B., and Tong, N.T. 1988. Databases lysis: computer-assisted ivestigation of cleavage sites in protein. *In* Methods in protein sequence analysis. B.W.-L. (Ed.), editor. Springer-Verlag, Berlin. 365-368.

Kim, L., Mogk, A., and Schumann, W. 1996. A xylose-inducible *Bacillus subtilis* integration vector and its application. *Gene*. 181:71-6.

Kobold, U., Jeppsson, J.O., Dulffer, T., et al. 1997. Candidate reference methods for hemoglobin A1c based on peptide mapping. *Clin Chem*. 43:1944-51.

Koenig, R.J., Peterson, C.M., Jones, R.L., et al. 1976. Correlation of glucose regulation and hemoglobin AIc in diabetes mellitus. *N Engl J Med*. 295:417-20.

Kraut, J. 1977. Serine proteases: structure and mechanism of catalysis. *Annu Rev Biochem*. 46:331-58.

Kunst, F., Ogasawara, N., Moszer, I., et al. 1997. The complete genome sequence of the gram-positive bacterium Bacillus subtilis. *Nature*. 390:249-56.

L

Lam, K.H., Chow, K.C., and Wong, W.K. 1998. Construction of an efficient *Bacillus subtilis* system for extracellular production of heterologous proteins. *J Biotechnol*. 63:167-77.

Le Grice, S.F. 1990. Regulated promoter for high-level expression of heterologous genes in *Bacillus subtilis*. *Methods Enzymol*. 185:201-14.

Li, W., Zhou, X., and Lu, P. 2004. Bottlenecks in the expression and secretion of heterologous proteins in *Bacillus subtilis*. *Res Microbiol*. 155:605-10.

Lim, K.B., Pengson, J., and Kassel, D.B. Takeda San Diego, Inc., San Diego. 2006. Online GluC digestion and measuremnt of hemoglobin A1C and A0 peptides using LC/MS and LC/MS/Ms.

Lineweaver, H and Burk, D. 1934. The Determination of Enzyme Dissociation Constants. *Journal of the American Chemical Society*.56: 658–666

Ling Lin, F., Zi Rong, X., Wei Fen, L., et al. 2007. Protein secretion pathways in *Bacillus subtilis*: implication for optimization of heterologous protein secretion. *Biotechnol Adv*. 25:1-12.
Little, R.R., Rohlfing, C.L., Wiedmeyer, H.M., et al. 2001. The national glycohemoglobin standardization program: a five-year progress report. *Clin Chem*. 47:1985-92.
Liu, Y.H., Lu, F.P., Li, Y., et al. 2008. Characterisation of mutagenised acid-resistant alpha-amylase expressed in *Bacillus subtilis* WB600. *Appl Microbiol Biotechnol*. 78:85-94.

M

Mallia, A K., Hermanson, G T., Krohn, R.I., et al. K. 1981. Preparation and use of a boronic acid affinity support for separation and quantitation of glycosylated hemoglobins. *Anal. Lett.*, 14, 649–661.
Manisalia, I.C., Chan, D.D.Y., and Schneider, B.B. 2006. Electrospray ionization source geometry for mass spectrometry: past, present, and future. *Trends in Analytical Chemistry*. Volume 25:243-256.
Manley, S.E., Round, R.A., and Smith, J.M. 2006. Calibration of HbA1c and its measurement in the presence of variant haemoglobins: report on questionnaire to manufacturers. *Ann Clin Biochem*. 43:135-45.
Massimi, I., Park, E., Rice, K., et al. 2002. Identification of a novel maturation mechanism and restricted substrate specificity for the SspB cysteine protease of *Staphylococcus aureus*. *J Biol Chem*. 277:41770-7.
Mayer, T.K., and Freedman, Z.R. 1983. Protein glycosylation in diabetes mellitus: a review of laboratory measurements and of their clinical utility. *Clin Chim Acta*. 127:147-84.
McDonald, M.J., Shapiro, R., Bleichman, M., et al. 1978. Glycosylated minor components of human adult hemoglobin. Purification, identification, and partial structural analysis. *J Biol Chem*. 253:2327-32.
Miedema, K. 2005. Standardization of HbA1c and Optimal Range of Monitoring. *Scand J Clin Lab Invest Suppl*. 240:61-72.
Miot, M., and Betton, J.M. 2004. Protein quality control in the bacterial periplasm. *Microb Cell Fact*. 3:4.
Mori, H., and Ito, K. 2001. The Sec protein-translocation pathway. *Trends Microbiol*. 9:494-500.
Mosca, A., Goodall, I., Hoshino, T., et al. 2007. Global standardization of glycated hemoglobin measurement: the position of the IFCC Working Group. *Clin Chem Lab Med*. 45:1077-80.
Mullis K., Faloona F., Scharf S., et al. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol*. 51:263-73.

N

Nagata, K., Yoshida, N., Ogata, F., et al. 1991. Subsite mapping of an acidic amino acid-specific endopeptidase from *Streptomyces griseus*, GluSGP, and protease V8. *J Biochem*. 110:859-62.

Nakamoto, M. 2004. H-B-A-1-C what it is and why it matters. *Diabetes Self Manag*. 21:84, 86-9.

Nakamura, K., Furusato, T., Shiroza, T., et al. 1985. Stable hyper-production of *Escherichia coli* beta-lactamase by *Bacillus subtilis* grown on a 0.5 M succinate-medium using a B. subtilis alpha-amylase secretion vector. *Biochem Biophys Res Commun*. 128:601-6.

Nakano, M.M., Dailly, Y.P., Zuber, P., et al. 1997. Characterization of anaerobic fermentative growth of *Bacillus subtilis*: identification of fermentation end products and genes required for growth. *J Bacteriol*. 179:6749-55.

Nathan, D.M. 1993. Long-term complications of *diabetes mellitus*. *N Engl J Med*. 328:1676-85.

Nathan, D.M., Cleary, P.A., Backlund, J.Y., et al. 2005. Intensive diabetes treatment and cardiovascular disease in patients with type 1 diabetes. *N Engl J Med*. 353:2643-53.

National Diabetes Data Group. 1979. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. National Diabetes Data Group. *Diabetes*. 28:1039-57.

Nathan, D.M., Turgeon, H., and Regan, S. 2007. Relationship between glycated haemoglobin levels and mean glucose levels over time. *Diabetologia*. 50:2239-44.

Nemoto, T.K., Ohara-Nemoto, Y., Ono, T., et al. 2008. Characterization of the glutamyl endopeptidase from *Staphylococcus aureus* expressed in *Escherichia coli*. *FEBS J*. 275:573-87.

Nemoto, T.K., Ono, T., Shimoyama, Y., et al. 2008. Determination of three amino acids causing alteration of proteolytic activities of staphylococcal glutamyl endopeptidases. *Biol Chem*.

Nguyen, H.D., Phan, T.T., and Schumann, W. 2007. Expression vectors for the rapid purification of recombinant proteins in *Bacillus subtilis*. *Curr Microbiol*. 55:89-93.

Nguyen, S., and Fenn J.B. 2007. Gas-phase ions of solute species from charged droplets of solutions. *Proc Natl Acad Sci U S A*. 104:1111-7.

Nickerson, N.N., Joag, V., and McGavin, M.J. 2008. Rapid autocatalytic activation of the M4 metalloprotease aureolysin is controlled by a conserved N-terminal fungalysin-thermolysin-propeptide domain. *Mol Microbiol*. 69:1530-43.

Nickerson, N.N., Prasad, L., Jacob, L., et al. 2007. Activation of the SspA serine protease zymogen of Staphylococcus aureus proceeds through unique variations of a trypsinogen-like mechanism and is dependent on both autocatalytic and metalloprotease-specific processing. *J Biol Chem* 282: 34129-38.

Nielsen, H., Engelbrecht, J., Brunak, S., et al. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng*. 10:1-6.

Nordin, G. and Dybkaer, R. 2007. Recommendation for term and measurement unit for "HbA1c". *Clin Chem Lab Med*. 45:1081-2.

O

Ohashi, Y., Ohshima, H., Tsuge, K., et al. 2003. Far different levels of gene expression provided by an oriented cloning system in *Bacillus subtilis* and *Escherichia coli*. *FEMS Microbiol Lett*. 221:125-30.

Orhan, E., Omay, D., and Guvenilir, Y. 2005. Partial purification and characterization of protease enzyme from *Bacillus subtilis* and *Bacillus cereus*. *Appl Biochem Biotechnol*. 121-124:183-94.

Ohta, Y., Hojo, H., Aimoto, S., et al. 1991. Pro-peptide as an intramolecular chaperone: renaturation of denatured subtilisin E with a synthetic pro-peptide [corrected]. *Mol Microbiol*. 5:1507-10.

Okamoto, H., Fujiwara, T., Nakamura, E., et al. 1997. Purification and characterization of a glutamic-acid-specific endopeptidase from *Bacillus subtilis* ATCC 6051; application to the recovery of bioactive peptides from fusion proteins by sequence-specific digestion. *Appl Microbiol Biotechnol*. 48:27-33.

Ono, T., Nemoto, T.K., Shimoyama, Y., et al. 2008. An *Escherichia coli* expression system for glutamyl endopeptidases optimized by complete suppression of autodegradation. *Anal Biochem*. 381:74-80.

P

Panteghini, M. and John, W.G. 2007. Implementation of haemoglobin A1c results traceable to the IFCC reference system: the way forward. *Clin Chem Lab Med*. 45:942-4.

Peacock, I. 1984. Glycosylated haemoglobin: measurement and clinical use. *J Clin Pathol*. 37:841-51.

Porath, J., Carlsson, J., Olsson, I., et al. 1975. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature*. 258:598-9.

Potempa, J., Porwit-Bobr, Z., and Travis, J. 1989. Stabilization vs. degradation of *Staphylococcus aureus* metalloproteinase. *Biochim Biophys Acta*. 993:301-4.

Power, S.D., Adams, R.M., and Wells, J.A. 1986. Secretion and autoproteolytic maturation of subtilisin. *Proc Natl Acad Sci U S A*. 83:3096-100.

Prasad, L., Leduc, Y., Hayakawa, K., et al. 2004. The structure of a universally employed enzyme: V8 protease from *Staphylococcus aureus*. *Acta Crystallogr D Biol Crystallogr*. 60:256-9.

Q

Quentin, Y., Fichant, G., and Denizot, F. 1999. Inventory, assembly and analysis of *Bacillus subtilis* ABC transport systems. *J Mol Biol*. 287:467-84.

R

Rahbar, S. 1968. An abnormal hemoglobin in red cells of diabetics. *Clin Chim Acta*. 22:296-8.

Rahbar, S., Blumenfeld, O., and Ranney, H.M. 1969. Studies of an unusual hemoglobin in patients with diabetes mellitus. *Biochem Biophys Res Commun*. 36:838-43.

Rahbar, S. 1980. Glycosylated hemoglobins. *Tex Rep Biol Med*. 40:373-85.

Rawlings, N.D., Barrett, A.J. and Woesner J.F. 2004. Handbook of Proteolytic Enzymes. Vol. 2. Elsevier Ltd.

Rawlings, N.D., F.R. Morton, C.Y. Kok, et al. 2008. MEROPS: the peptidase database. *Nucleic Acids Res*. 36:D320-5.

Rebrikov, D.V., Akimkina, T.V., Shevelev, A.B., et al. 1999. *Bacillus intermedius* glutamyl endopeptidase. Molecular cloning and nucleotide sequence of the structural gene. *J Protein Chem*. 18:21-7.

Rice, K., Peralta, R., Bast, D., et al. 2001. Description of staphylococcus serine protease (ssp) operon in *Staphylococcus aureus* and nonpolar inactivation of sspA-encoded serine protease. *Infect Immun*. 69:159-69.

Rohlfing, C.L., Wiedmeyer, H.M., Little, R.R., et al. 2002. Defining the relationship between plasma glucose and HbA1c. analysis of glucose profiles and HbA1c) in the Diabetes Control and Complications Trial. *Diabetes Care*. 25:275-8.

Roy H. Doi, S.-L.W., Fuji Kawamura. 1986. Potential use of *Bacillus subtilis* for secretion and production of foreign proteins. *Tibtech*:232-235.

S

Sabat, A., Kosowska, K., Poulsen, K., et al. 2000. Two allelic forms of the aureolysin gene (aur) within *Staphylococcus aureus*. *Infect Immun*. 68:973-6.

Sacks, D.B. 2005. Global harmonization of hemoglobin A1c. *Clin Chem*. 51:681-3.

Sahni, G., Mallia, A.K., and Acharya, A.S. 1991. Proteosynthetic activity of immobilized *Staphylococcus aureus* V8 protease: application in the semisynthesis of molecular variants of alpha-globin. *Anal Biochem*. 193:178-85.

Saier, M.H., Jr., Werner, P.K., and Muller, M. 1989. Insertion of proteins into bacterial membranes: mechanism, characteristics, and comparisons with the eucaryotic process. *Microbiol Rev*. 53:333-66.

Saiki R.K., Gelfand D.H., Stoffel S., et al. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*. 239:487-91.

Sasamoto, H., Nakazawa, K., Tsutsumi, K., et al. 1989. Signal peptide of *Bacillus subtilis* alpha-amylase. *J Biochem*. 106:376-82.

Saudek, C.D., Derr, R.L., and Kalyani, R.R. 2006. Assessing glycemia in diabetes using self-monitoring blood glucose and hemoglobin A1c. *Jama* 295: 1688-97.

Schechter, I. and Berger, A. 1967. On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun*, 27(2):157-162.

Schmidt-Dannert, C., and Arnold, F.H. 1999. Directed evolution of industrial enzymes. *Trends Biotechnol*. 17:135–136.

Sharipova, M.R., Shagimardanova, E.I., Chastukhina, I.B., et al. 2007. The expression of *Bacillus intermedius* glutamyl endopeptidase gene in *Bacillus subtilis* recombinant strains. *Mol Biol Rep*. 34:79-87.

Sharp, P.M., Cowe, E., Higgins, D.G., et al. 1988. Codon usage patterns in *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila melanogaster* and *Homo sapiens*; a review of the considerable within-species diversity. *Nucleic Acids Res*. 16:8207-11.

Shaw, L., Golonka, E., Potempa, J., et al. 2004. The role and regulation of the extracellular proteases of *Staphylococcus aureus*. *Microbiology*. 150:217-28.

Shields, D.C., and Sharp, P.M. 1987. Synonymous codon usage in *Bacillus subtilis* reflects both translational selection and mutational biases. *Nucleic Acids Res*. 15:8023-40.

Sierro, N., Makita, Y., de Hoon, M.J.L., and Nakai K. DBTBS: a database of transcriptional regulation in *Bacillus subtilis* containing upstream intergenic conservation information. *Nucleic Acids Res.* 2008, 36 (Database issue):D93-D96; doi:10.1093/nar/gkm910. http://dbtbs.hgc.jp/

Silbersack, J., Jurgen, B., Hecker, M., et al. 2006. An acetoin-regulated expression system of *Bacillus subtilis*. *Appl Microbiol Biotechnol*. 73:895-903.

Smith, H.O., Danner, D.B., and Deich, R.A. 1981. Genetic transformation. *Annu Rev Biochem*. 50:41-68.

Sobolev, V., Eyal, E., Gerzon, S., et al. 2005. SPACE: a suite of tools for protein structure prediction and analysis based on complementarity and environment. *Nucleic Acids Res*. 33:W39-43. http://ligin.weizmann.ac.il/space.

Sorensen, S.B., Sorensen, T.L., and Breddam, K. 1991. Fragmentation of proteins by *S. aureus* strain V8 protease. Ammonium bicarbonate strongly inhibits the enzyme but does not improve the selectivity for glutamic acid. *FEBS Lett*. 294:195-7.

Stahl, M.L., and Ferrari, E. 1984. Replacement of the *Bacillus subtilis* subtilisin structural gene with an In vitro-derived deletion mutation. *J Bacteriol*. 158:411-8.

Stratton, I.M., Adler, A.I., Neil, H.A., et al. 2000. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes UKPDS 35. Prospective observational study. *BMJ*. 321:405-12.

Svendsen, I., and Breddam, K. 1992. Isolation and amino acid sequence of a glutamic acid specific endopeptidase from *Bacillus licheniformis*. *Eur J Biochem*. 204:165-71.

T

Takase, K., Mizuno, H., and Yamane, K. 1988. NH2-terminal processing of *Bacillus subtilis* alpha-amylase. *J Biol Chem*. 263:11548-53.

Takeuchi, S., Saito, M., Imaizumi, K., et al. 2002. Genetic and enzymatic analyses of metalloprotease (aureolysin) from *Staphylococcus aureus* isolated from domestic animals. *Vet Microbiol*. 84:135-42.

Terpe, K. 2003. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol*. 60:523-33.

Tjalsma, H., Bolhuis, A., Jongbloed, J.D., et al. 2000. Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol Mol Biol Rev*. 64:515-47.

Tjalsma, H., van den Dolder, J., Meijer, W.J., et al. 1999. The plasmid-encoded signal peptidase SipP can functionally replace the major signal peptidases SipS and SipT of *Bacillus subtilis*. *J Bacteriol*. 181:2448-54.

Tran, D.V., Hofer, T.L., Lee, T., et al. 2003. Unique approach to derivation of random error in laboratory assays: application to glycohemoglobin testing demonstrates poor clinical performance for immunochemistry assay. *Diabetes Technol Ther*. 5:975-8.

Terpe, K. 2006. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol*. 72:211-22.

Ticu, E.L., Vercaigne-Marko, D., Huma, A., et al. 2004. A kinetic study of bovine haemoglobin hydrolysis by pepsin immobilized on a functionalized alumina to prepare hydrolysates containing bioactive peptides. *Biotechnol Appl Biochem*. 39:199-208.

Tjalsma, H., Antelmann, H., Jongbloed, J.D., et al. 2004. Proteomics of protein secretion by *Bacillus subtilis*: separating the "secrets" of the secretome. *Microbiol Mol Biol Rev*. 68:207-33.

Tjalsma, H., Bolhuis, A., Jongbloed, J.D., et al. 2000. Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol Mol Biol Rev*. 64:515-47.

U

Ulmanen, I., Lundstrom, K., Lehtovaara, P., et al. 1985. Transcription and translation of foreign genes in *Bacillus subtilis* by the aid of a secretion vector. *J Bacteriol*. 162:176-82.
Unwin, N., Martin, A. 2004. Diabetes action now: WHO and IDF working together to raise awareness worldwide. *Diabetes'voices*. 49.

V

Van Wely, K.H., Swaving, J., Freudl, R., et al. 2001. Translocation of proteins across the cell envelope of Gram-positive bacteria. *FEMS Microbiol Rev*. 25:437-54.
Vasseur, C., Galacteros, F., Groff, P., et al. 1991. Cleavage by protease from *Staphylococcus aureus* V8: an improvement in the sequence analysis of human hemoglobin variants. *J Biochem Biophys Methods*. 22:195-205.
Vercaigne-Marko, D., Kosciarz, E., Nedjar-Arroume, N., et al. 2000. Improvement of *Staphylococcus aureus*-V8-protease hydrolysis of bovine haemoglobin by its adsorption on to a solid phase in the presence of SDS: peptide mapping and obtention of two haemopoietic peptides. *Biotechnol Appl Biochem*. 31 (Pt 2):127-34.
Vesper, H.W., Mi, L., Enada, A., et al. 2005. Assessment of microwave-assisted enzymatic digestion by measuring glycated hemoglobin A1c by mass spectrometry. *Rapid Commun Mass Spectrom*. 19:2865-70.
Villaverde, A., and Carrio, M.M. 2003. Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnol Lett*. 25:1385-95.
Von Heijne, G. 1990. The signal peptide. *J Membr Biol*. 115:195-201.

W

Wajcman, H., de Montalembert, M., Girot, R., et al. 1994. The use of *Staphylococcus* V8 protease in the structural determination of human hemoglobin variants: HB Valparaiso [alpha 88(F9)Ala-->Gly] example. *Hemoglobin*. 18:275-84.
Wang, Y., Beckwith, B., Smith, C., et al. 2007. Misleading glycated hemoglobin results in a patient with hemoglobin SC disease. *Clin Chem*. 53:1394-5.
Warshel, A., Naray-Szabo, G., Sussman, F., et al. 1989. How do serine proteases really work? *Biochemistry*. 28:3629-37.
Watson, M.E. 1984. Compilation of published signal sequences. *Nucleic Acids Res*. 12:5145-64.

Wehofsky, N., and Bordusa, F. 1999. Programming of enzyme specificity by substrate mimetics: investigations on the Glu-specific V8 protease reveals a novel general principle of biocatalysis. *FEBS Lett*. 443:220-4.
Westers, L., Westers, H., and Quax, W.J. 2004. *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochim Biophys Acta*. 1694:299-310.
WHO. 1985. Diabetes mellitus. Report of a WHO Study Group. *World Health Organ Tech Rep Ser*. 727:1-113.
WHO. 2004. *In* Diabetes Action Now. WHO Library Cataloguing-in-publication data, Geneva, Switzerland.
Wild, S., Roglic, G., Green, A., et al. 2004. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care*. 27:1047-53.
Wong, S.L. 1995. Advances in the use of *Bacillus subtilis* for the expression and secretion of heterologous proteins. *Curr Opin Biotechnol*. 6:517-22.
Wong, S.L., Kawamura, F., and Doi, R.H. 1986. Use of the *Bacillus subtilis* subtilisin signal peptide for efficient secretion of TEM beta-lactamase during growth. *J Bacteriol*. 168:1005-9.
Wu, X.C., Lee, W., Tran, L., et al. 1991. Engineering a *Bacillus subtilis* expression-secretion system with a strain deficient in six extracellular proteases. *J Bacteriol* 173: 4952-8.
Wu, J., and Filutowicz, M. 1999. Hexahistidine (His6)-tag dependent protein dimerization: a cautionary tale. *Acta Biochim Pol*. 46:591-9.

Y

Yabuta, M., Ochi, N., and Ohsuye, K. 1995. Hyperproduction of a recombinant fusion protein of *Staphylococcus aureus* V8 protease in *Escherichia coli* and its processing by OmpT protease to release an active V8 protease derivative. *Appl Microbiol Biotechnol* 44: 118-25.
Yabuta, Y., Takagi, H., Inouye, M., et al. 2001. Folding pathway mediated by an intramolecular chaperone: propeptide release modulates activation precision of pro-subtilisin. *J Biol Chem*. 276:44427-34.
Yamane, K., Bunai, K., and Kakeshita, H. 2004. Protein traffic for secretion and related machinery of *Bacillus subtilis*. *Biosci Biotechnol Biochem*. 68:2007-23.
Yoshikawa, K., Tsuzuki, H., Fujiwara, T., et al. 1992. Purification, characterization and gene cloning of a novel glutamic acid-specific endopeptidase from *Staphylococcus aureus* ATCC 12600. *Biochim Biophys Acta*. 1121:221-8.
Young, F.E. 1980. Impact of cloning in *Bacillus subtilis* on fundamental and industrial microbiology. The Eighth Griffith Memorial Lecture. *J Gen Microbiol*. 119:1-15.

Z

Zanen, G., Houben, E.N., Meima, R., et al. 2005. Signal peptide hydrophobicity is critical for early stages in protein export by *Bacillus subtilis*. *FEBS J*. 272:4617-30.

Zimmerman, M., Yurewicz, E., and Patel, G. 1976. A new fluorogenic substrate for chymotrypsin. *Anal Biochem*. 70:258-62.

Zweers, J.C., Barak, I., Becher, D., et al. 2008. Towards the development of *Bacillus subtilis* as a cell factory for membrane proteins and protein complexes. *Microb Cell Fact*. 7:10.

Keywords: Diabetes, long-term blood glucose care, V8-glutamyl endopeptidase (EC 3.4.21.19), glycated hemoglobin (HbA1c), directed evolution, *B. subtilis* expression, multiparameter high-throughput screening system, enzymatic activity, Km value, temperature stability, cleavage specificity.

Summary

The blood glucose determination via the rate of glycated hemoglobin (HbA1c) is a key-parameter for long-term diabetes care. Roche Diagnostics GmbH committed itself with the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) to develop an international reference standard system for blood sugar control. The glutamyl endopeptidase (GluC) from the V8 strain of *Staphylococus aureus* (EC 3.4.21.19) is used for the HbA1c digestion. The aim of this phD project is to increase the hydrolysis speed of the GluC by directed evolution by keeping its original enzymatic characteristics (specificity and stability) in the standard conditions of the functional hemoglobin test.

A variant with the improved enzymatic features was isolated and identified after the construction and the multiparameter screening of libraries of thousands of mutants. This variant of the glutamyl endopeptidase was expressed in *B. subtilis* by extracellular secretion, and then purified by Ni-NTA affinity chromatography. It was then characterized after activation by autoproteolysis. The specific activity of the variant is three times enhanced and the affinity constant (Km) is 3 times improved in a Tris-HCl buffer at pH 7.8, compared to the wild type glutamyl endopeptidase. The specific activity is four times better in ammonium acetate buffer at pH 4.3 with the hemoglobin hexapeptide as substrate. It is also more stable at 50°C. However, although the variant is more active with synthetic substrate, it is less competitive than the wild type V8-Glutamyl endopeptidase regarding the activity, affinity and cutting specificity for the hexapeptide of the whole hemoglobin. Experiments are still discussed to optimize the conditions of the HbA1c test.

Résumé

Le dosage sanguin de la glycémie via le taux d´hémoglobine glyquée (HbA1c) est un paramètre-clé de la surveillance du diabète à long terme. Roche Diagnostics GmbH s´est engagé avec la Fédération Internationale de Chimie Clinique et de Médecine de Laboratoire (IFCC) à développer un système standard de référence international de contrôle de la glycémie. L´endopeptidase glutamique (GluC) de la souche V8 de *Staphylococus aureus* (EC 3.4.21.19) est utilisée pour la digestion de l´HbA1c. L´objectif de cette thèse de doctorat est d´augmenter la vitesse d´hydrolyse de la GluC par évolution moléculaire tout en préservant ses caractéristiques enzymatiques originelles (spécificité et stabilité) dans les conditions standardisées du test fonctionnel de l´HbA1c.

La construction de librairies de plusieurs milliers de variants, couplée à un criblage multiparamètre, a permis d'isoler un mutant de la GluC qui comporte les propriétés recherchées. Ce variant a été exprimé chez *B. subtilis* par sécrétion extracellulaire et purifié par chromatographie d'affinité Ni-NTA. Il a été ensuite caractérisé après activation par autoprotéolyse. Le variant présente une activité spécifique 3 fois supérieure et une constante d'affinité (Km) trois fois meilleure à celle de la GluC sauvage pour le substrat synthétique: Z-Phe-Leu-Glu-pNa dans un tampon de Tris-HCl à pH 7.8. Le variant est 4 fois plus actif dans un tampon d´acétate d´ammonium à pH 4.3 avec l´hexapeptide de l´hémoglobine pour substrat. Il est également plus stable à 50°C. Bien que plus actif que la GluC sauvage en présence de substrats synthétiques, le variant isolé est moins performant d'un point de vue activité, affinité et spécificité de coupure pour l´hexapeptide de l´hémoglobine entière.