

University of Groningen

## Structure-function relationships in type I signal peptidases of bacilli

Roosmalen, Maarten Leonardus van

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2001

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Roosmalen, M. L. V. (2001). *Structure-function relationships in type I signal peptidases of bacilli*. s.n.

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# 7

## Chapter Seven

**SUMMARY AND GENERAL CONCLUSIONS**

**SAMENVATTING EN ALGEMENE  
CONCLUSIES**

# SUMMARY AND GENERAL CONCLUSIONS

Every living cell, either being part of a complex cellular network or just a unicellular organism, produces proteins inside the cell that are destined for extracellular locations. Consequently, these proteins have to be transported across one or more membrane(s) that subdivide, or surround the cell. Secretory proteins are synthesized as pre-proteins that start with a signal peptide. This signal peptide is required to target these proteins to the pre-protein translocase in the membrane, and to initiate the translocation process. During or after their translocation across the membrane, the signal peptide is removed from the exported pre-proteins. This cleavage of the signal sequence is performed by a membrane-bound endopeptidase named signal peptidase (SPase). After cleavage of the signal peptide by SPase, the mature protein can be released from the membrane, and usually folds into a functional protein. To perform its functions in the secretion process, the signal peptide has a typical tripartite structure, consisting of a positively charged N-region, a hydrophobic H-region and a polar C-region that contains the SPase cleavage site. Studies of sequences surrounding the SPase cleavage site led to the formulation of the “-1, -3” or “Ala-X-Ala” rule, defining the preferred residues (*i.e.* Ala) at the -1 and -3 positions relative to the cleavage site as critical determinants for signal peptide recognition and cleavage.

The largest number of type I SPases has, thus far, been found in the Gram-positive eubacterium *Bacillus subtilis*, which contains five paralogous chromosomally-encoded enzymes of this type (SipS, SipT, SipU, SipV and SipW). In addition, certain strains of *B. subtilis* (*natto*) contain plasmid-encoded type I SPases, denoted SipP. Even though the presence of paralogous SPases is not unusual in eubacteria, archaea and eukaryotes, the number of paralogous SPases in *B. subtilis* is unusually high, suggesting that some of these SPases may have specialized functions. Indeed, it was shown, that SipS, SipT and SipP, unlike SipU, SipV and SipW, are essential for the secretion of degradative enzymes and cell viability. These and

other observations indicate that significant differences exist in the substrate specificities of the various SPases of *B. subtilis*.

Despite considerable similarities between the known type I SPases, only few residues are strictly conserved in all known enzymes of this family. In particular, the type I SPases of eubacteria, mitochondria and chloroplasts (P-type SPases), differ considerably from their homologues in archaea and the ER membrane (ER-type SPases). In addition to an N-terminal membrane anchor domain (A), all P-type SPases contain four well-conserved domains (B to E). These domains include residues involved in substrate recognition and catalysis. Specifically, domain B contains a strictly conserved Ser residue and domain D a strictly conserved Lys residue. Together, these residues form a Ser-Lys catalytic dyad. The domains B to E of the P-type SPases are also conserved in the ER-type SPases. Nevertheless, instead of a Lys residue, domain D of the ER-type SPases contains a strictly conserved His residue, which is required for catalysis. Apart from the residues involved in catalysis, very little is known about the factors that determine the activity of SPases from eubacteria other than *E. coli*. Considering the fact that the type I SPases of *B. subtilis* display different substrate specificities *in vivo*, these enzymes appeared to be attractive models for the identification of important determinants for SPase activity. Furthermore, these enzymes are structurally very different from the SPase of *E. coli*, as they are much smaller, and contain only one membrane anchor instead of two.

The research described in this thesis is mainly concerned with the analysis of the relationship between structure and function of the type I signal peptidases (SPases) of *Bacillus* species.

In Chapter 2, we investigated whether the unique membrane anchor of *Bacillus* SPases is required for *in vitro* activity. Even though the P-type SPases of bacilli are homologous to the unique SPase I of

*E. coli*, they are structurally different from the latter enzyme, being almost half the size and containing one membrane anchor instead of two. Soluble forms of SipS of *B. subtilis*, SipS of *Bacillus amyloliquefaciens* and SipC of the thermophile *Bacillus caldolyticus* were produced. Of these three proteins, only a hexa-histidine tagged soluble form of SipS of *B. amyloliquefaciens* could be isolated in significant quantities. This protein displayed optimal activity at pH 10, which is remarkable considering the fact that the catalytic domain of SPases is located in an acidic environment at the outer surface of the membrane of living cells. Strikingly, in contrast to what has been previously reported for the soluble form of the *E. coli* SPase, soluble SipS was shown to be active in the absence of added detergents. This observation can be explained by the fact that a highly hydrophobic surface domain of the *E. coli* SPase, implicated in detergent-binding, is absent from SipS.

The studies described in Chapter 3 were aimed at answering the question whether the high-level production of truncated soluble *Bacillus* SPases in *E. coli* is precluded by proteolysis. The results show that this is indeed the case, and that soluble forms of the *Bacillus* SPase SipS are subject to self-cleavage. First, catalytically inactive soluble forms of this signal peptidase were not prone to degradation; in fact, these mutant proteins were produced at very high levels in *E. coli*. Second, the purified active soluble form of SipS displayed self-cleavage *in vitro*. Third, as determined by amino-terminal sequencing, at least one of the sites of self-cleavage (between Ser15 and Met16 of the truncated enzyme) strongly resembles a typical SPase cleavage site. Self-cleavage at the latter position results in complete inactivation of the enzyme, as Ser15 forms a catalytic dyad with Lys55. Ironically, self-cleavage between Ser15 and Met16 can not be prevented by mutagenesis of Gly13 and Ser15, which conform to the “-1, -3” rule for signal peptidase recognition, because these residues are critical for signal peptidase activity.

Chapter 4 describes the identification of the fifth chromosomally-encoded type I SPase of *B. subtilis*, denoted SipW. Surprisingly, this newly identified SPase belonged to the so-called ER-type SPases,

which are structurally and mechanistically distinct from the P-type SPases that are commonly found in prokaryotes and organelles of eukaryotes. Furthermore, the studies in Chapter 4 were aimed at defining the function of the complete chromosomal *sip* gene family of *B. subtilis*. Interestingly, viable mutant strains lacking as many as four of these genes could be obtained. As shown with a temperature-sensitive SipS variant, only cells lacking both SipS and SipT were not viable, which may be due to jamming of the secretion machinery with secretory pre-proteins. Thus, SipS and SipT are of major importance for protein secretion. This conclusion was underscored by the observation that only the transcription of the *sipS* and *sipT* genes is temporally controlled *via* the DegS-DegU regulatory system, in concert with the transcription of most genes for secretory pre-proteins.

The studies described in Chapter 5 were aimed at the characterization of differences between *major* and *minor Bacillus* SPases, and the identification of domains in these enzymes that are critical for their specificity. Molecular phylogeny was successfully used to predict *major* and *minor* SPases, which are defined by their importance for the viability of *B. subtilis*. Notably, the functional difference between *major* and *minor* SPases is not clearly reflected in sequence alignments. The results obtained with molecular phylogeny were verified with SPases from various bacilli. As predicted, the latter enzymes behaved as *major* or *minor* SPases when expressed in *B. subtilis*. Strikingly, molecular modeling indicated that the active site geometry is not a critical parameter for classification of *major* and *minor Bacillus* SPases. Even though the substrate binding site of the *minor* SPase SipV is smaller than that of other known SPases, SipV could be converted into a *major* SPase without changing this site. However, replacement of amino-terminal residues of SipV with corresponding residues of the *major* SPase SipS was sufficient for conversion of SipV into a *major* SPase. This suggests that differences between *major* and *minor* SPases are based on activities other than substrate cleavage site selection. Taken together, the results show that *major* and *minor Bacillus* SPases can be

distinguished by phylogenetic analyses and that critical information for their role in cell viability is provided by residues that are located amino-terminally of the catalytic Ser residue.

In Chapter 6 we investigated which of the conserved elements within the first 42 residues of SipS, including the actual transmembrane domain, are important for SPase activity. It was previously shown that the catalytic domains of SPases can be separated from their membrane anchors without the loss of activity. Consequently, few studies have addressed the importance of the membrane anchor domain for *in vivo* signal peptidase activity. The results described in Chapter 6 show that only a limited number of single mutations in the anchor domain inactivate SipS. These include mutations that significantly reduce the length of the hydrophobic core, the charge at the carboxyl-terminal end of the actual transmembrane segment, the integrity of the S3 substrate binding pocket, or the accessibility of the active site in general. Strikingly, a conserved arginine residue of SipS was not essential for activity, whereas the corresponding arginine residue of the *Escherichia coli* SPase was previously shown to be essential. A possible explanation for the fact that only few mutations in the membrane anchor domain affected the *in vivo* activity of SipS is that various residues in this domain have redundant functions.

In conclusion, the observations described in this thesis provide novel insights into the relationships between the structure and function of the type I SPases of *Bacillus* species. Biochemical characterization, combined with molecular modeling and functional analyses studies indicate that the SPases found in *Bacillus* can be used as a paradigm for the study of type I SPases from other organisms.

## SAMENVATTING EN ALGEMENE CONCLUSIES

Iedere levende cel, hetzij als onderdeel van een complex cellulair netwerk, hetzij slechts een eencellig organisme vormend, produceert eiwitten die ten dele bestemd zijn voor locaties aan de buitenzijde van de cel. Deze eiwitten moeten dus getransporteerd worden door één of meerdere membranen heen die de cel onderverdelen of begrenzen. Eiwitten bestemd voor secretie worden gemaakt als precursor eiwit dat aan het begin een signaal peptide draagt. Dit signaal peptide is nodig om deze eiwitten naar het eiwit translocase systeem in het membraan te dirigeren, en om het translocatie proces te laten beginnen. Tijdens of na de translocatie wordt het signaal peptide verwijderd van het geëxporteerde precursor eiwit. Het afknippen van het signaal peptide wordt uitgevoerd door een membraan-gebonden enzym dat signaal peptidase (SPase) genoemd wordt. Nadat het signaal peptide door SPase geknipt is kan het mature eiwit van het membraan loslaten, waarna het zich meestal tot een functioneel eiwit vouwt. Om zijn functies in het secretie proces te kunnen uitoefenen, bezit het signaal peptide een driedelige structuur bestaande uit een positief geladen N-domein, een hydrofoob H-domein en een polair C-domein. In dit laatste domein bevindt zich ook de knipplaats voor SPase. Onderzoek aan sequenties in de directe omgeving van de knipplaats hebben geleid tot de zogenaamde "-1, -3" of "Ala-X-Ala" regel, waarmee de voorkeursresiduen (Ala) op de -1 en -3 posities van de knipplaats worden aangegeven. Dit zijn de kritische determinanten voor herkenning en klieving van het signaal peptide.

Het grootste aantal type I SPases wordt, tot nog toe, aangetroffen bij de Gram-positieve bacterie *Bacillus subtilis* die vijf door het chromosoom gecodeerde paraloge enzymen van dit type bevat (SipS, SipT, SipU, SipV en SipW). Daarnaast bevatten bepaalde *B. subtilis* (*natto*) stammen ook een plasmide-gecodeerde type I SPase, dat SipP wordt genoemd. Hoewel het voorkomen van paraloge SPases in bacteriën, archaea en eukaryoten niet ongewoon is, is het bij *B. subtilis*

bijzonder hoog. Dit zou kunnen betekenen dat sommige van deze enzymen een gespecialiseerde functie hebben. Dit blijkt zo te zijn: aangetoond is dat SipS, SipT en SipP, in tegenstelling tot SipU, SipV en SipW, essentieel zijn voor de secretie van degraderende enzymen en voor overleving van de cel. Deze en andere waarnemingen duiden erop dat er significante verschillen bestaan in de substraat specificiteit van de SPases van *B. subtilis*.

Ondanks aanzienlijke overeenkomsten tussen de bekende type I SPases, zijn slechts enkele residuen strikt geconserveerd in alle bekende enzymen die tot deze familie behoren. De type I SPases van bacteriën, mitochondriën en chloroplasten (P-type SPases) verschillen echter van hun homologen in archaea en het ER membraan (ER-type SPases). Naast een N-terminaal membraan anker domein (A) hebben alle P-type SPases vier goed geconserveerde domeinen (B-E). Deze domeinen bevatten residuen die betrokken zijn bij substraat herkenning en katalyse. Domein B bevat een strikt geconserveerd Ser residu en domein D een strikt geconserveerd Lys residu. Samen vormen deze residuen een Ser-Lys diade. De domeinen B tot E van de P-type SPases zijn ook geconserveerd in de ER-type SPases. In domein D van het ER-type SPase komt echter een strikt geconserveerd His residu voor dat nodig is voor katalyse. Naast de residuen die bij katalyse betrokken zijn, is met uitzondering van het SPase van *E. coli*, betrekkelijk weinig bekend over andere factoren die de activiteit van SPases bepalen. Aangezien SPases van *B. subtilis* een verschillende *in vivo* substraat specificiteit bezitten, lijken deze enzymen aantrekkelijke modellen voor de identificatie van belangrijke determinanten van SPase activiteit te zijn. Bovendien verschillen deze enzymen structureel verschillend van het SPase van *E. coli*: ze zijn veel kleiner en zijn slechts van één membraan anker voorzien, zulks in tegenstelling tot het SPase van *E. coli* die er twee heeft.

Het in dit proefschrift beschreven onderzoek richtte zich voornamelijk op de analyse van de relatie

tussen structuur en functie van type I signaal peptidases van *Bacillus* soorten.

In Hoofdstuk 2 werd onderzocht of het enkele membraan anker van *Bacillus* SPases nodig is voor *in vitro* activiteit. Hoewel de P-type SPases van bacillen homoloog zijn aan die van *E. coli*, verschillen ze, zoals hierboven vermeld, structureel van dit enzym. Bovendien zijn ze bijna de helft kleiner. In het hier beschreven onderzoek werden oplosbare vormen van SipS van *B. subtilis*, SipS van *B. amyloliquefaciens*, en SipC van de thermofiel *B. caldolyticus* verkregen. Van deze drie eiwitten kon alleen een van een hexa-histidine staart voorziene oplosbare vorm van SipS van *B. amyloliquefaciens* in voldoende hoeveelheid geïsoleerd worden. Dit eiwit vertoonde optimale activiteit bij pH 10. Dit is opmerkelijk, omdat het katalytische domein van SPases zich in een zure omgeving bevindt. Het buitenste membraanoppervlak van levende cellen is namelijk zuur. In tegenstelling tot verhoging van de activiteit van de oplosbare vorm van het *E. coli* SPase door detergentia, wordt het oplosbare SipS niet beïnvloed door toevoeging van detergentia. Deze waarneming zou het gevolg kunnen zijn van het voorkomen van een zeer hydrofoob oppervlakte domein in het *E. coli* SPase, betrokken bij binding van detergentia, dat in SipS ontbreekt.

Het in Hoofdstuk 3 beschreven onderzoek was gericht op het beantwoorden van de vraag of de geringe productie van getrunceerde oplosbare *Bacillus* SPases in *E. coli* veroorzaakt wordt door proteolyse. Dit bleek inderdaad het geval te zijn: de oplosbare vormen van het *Bacillus* SPase SipS zijn onderhevig aan zelf-klieving. Aan deze conclusie lagen de volgende waarnemingen ten grondslag. Ten eerste bleken katalytisch inactieve oplosbare vormen van deze SPase ongevoelig voor degradatie te zijn. In tegendeel, deze mutante eiwitten werden in grote hoeveelheden door *E. coli* geproduceerd. Ten tweede werd vastgesteld dat de gezuiverde actieve oplosbare vorm van SipS onderhevig is aan zelf-klieving *in vitro*. Ten derde werd door middel van amino-terminale sequencing, ten minste één van de zelf-klievings plaatsen (tussen Ser15 en Met16 van het getrunceerde enzym) bepaald. Deze plaats vertoont sterke overeenkomst met een

typische SPase knipplaats. Zelf-klieving op de genoemde positie resulteert in een complete inactivering van het enzym, aangezien Ser15 een katalytische diade met Lys55 vormt. Zelf-klieving tussen Ser15 en Met16 kon, ironisch genoeg, niet worden voorkomen door mutagenese van Gly13 en Ser15, die de "-1, -3" regel voor signaal peptidase herkenning volgen, omdat deze residuen kritisch zijn voor signaal peptidase activiteit.

Hoofdstuk 4 beschrijft de identificatie van het vijfde chromosomaal-gespecificeerde type I SPase van *B. subtilis*, waaraan de naam SipW is gegeven. Deze nieuw geïdentificeerde SPase bleek verrassenderwijs tot de zogenoemde ER-type SPases te behoren, die structureel en mechanistisch verschillen van de P-type SPases welke normaliter in prokaryoten en organellen van eukaryoten worden aangetroffen. Het onderzoek in Hoofdstuk 4 was bovendien gericht op het bestuderen van de functie van de complete chromosomale *sip* genfamilie van *B. subtilis*. Het kwam als een verrassing dat mutanten waarin vier van deze *sip* genen ontbraken, toch bleken te kunnen leven. Met behulp van een temperatuur-sensitieve SipS variant werd aangetoond dat, indien zowel SipS als SipT ontbraken, de cellen niet levensvatbaar zijn. Dit zou een gevolg kunnen zijn van het blokkeren van het secretie apparaat met precursor eiwitten. SipS en SipT zijn daarom van *majeur* belang voor eiwitsecretie. Deze conclusie wordt ondersteund door de observatie dat van de *sip* familie alleen de transcriptie van *sipS* en *SipT* temporeel gereguleerd wordt via het DegS-DegU regulatie systeem, gelijktijdig met de transcriptie van de meeste genen voor precursor eiwitten.

Het onderzoek dat in Hoofdstuk 5 beschreven is, was gericht op de karakterisering van verschillen tussen *major* en *minor Bacillus* SPases en de identificatie van domeinen in deze enzymen die kritisch zijn voor hun specificiteit. De moleculair fylogenetische benadering bleek met succes te kunnen worden toegepast bij het voorspellen van de classificatie van *major* en *minor* SPases, waarvan de *major* SPases, in tegenstelling tot de *minor* SPases, voor het leven van de cel onmisbaar zijn. Het functionele verschil tussen *major* en *minor* SPases wordt, merkwaardig genoeg, niet

weerspiegeld in sequentie verschillen. De verschillende waarden van de moleculair fylogenetische benadering werden geverifieerd met SPases van verschillende bacillen. Deze enzymen van verschillende oorsprong bleken zich inderdaad als *major* en *minor* SPases te gedragen wanneer ze in *B. subtilis* tot expressie gebracht werden. Op grond van 'molecular modeling' bleek echter dat de geometrie van het actieve centrum geen kritische parameter is voor het onderscheid tussen *major* en *minor Bacillus* SPases. Hoewel de substraat-bindingsplaats van de *minor* SPase SipV kleiner is dan die van andere bekende SPases, kon SipV omgezet worden in een *major* SPase zonder wijziging van deze bindings plaats: vervanging van amino-terminale residuen van de *minor* SPase SipV met corresponderende residuen van de *major* SPase SipS bleek voldoende te zijn voor conversie van SipV naar een *major* SPase. Dit suggereert dat van de verschillen tussen *major* en *minor* SPases, andere dan substraat klievings selectie activiteiten ter grondslag liggen.

Samenvattend: onderscheid tussen *major* en *minor Bacillus* SPases kan dus gemaakt worden door fylogenetische analyses en, bovendien, het belang van deze twee categorieën voor levensvatbaarheid van de cel moet gelegen zijn in de residuen die amino-terminaal de katalytische Ser zijn gelegen.

In Hoofdstuk 6 werd onderzocht of geconserveerde elementen in de eerste 42 residuen van SipS, inclusief het eigenlijke transmembraan domein, van belang zijn voor SPase activiteit. Eerder was aangetoond dat de katalytische domeinen van SPases van het membraananker gescheiden kunnen worden, zonder verlies van activiteit, waardoor er weinig onderzoek is verricht aan het belang van het membraananker domein voor *in vivo* signaal peptidase activiteit. De resultaten beschreven in dit hoofdstuk tonen aan dat slechts een beperkt aantal aminozuur substituties in het ankerdomein tot de inactivering van SipS leiden. Evenwel, mutaties die de lengte van de hydrofobe kern reduceren, mutaties in de lading aan het carboxyl-terminale einde van het eigenlijke transmembraan segment, mutaties van de integriteit van de S3 substraat bindings plaats, of mutaties met invloed op de toegankelijkheid van het substraat tot het katalytische centrum in het algemeen, leiden alle

tot activiteitsvermindering van SipS. Een geconserveerde arginine bleek vreemd genoeg niet essentieel voor activiteit te zijn, hoewel eerder voor het corresponderende arginine residu van het *E. coli* SPase was aangetoond dat dit essentieel is. Een mogelijke verklaring voor het gegeven dat slechts weinig mutaties in het membraananker domein effect hebben op de *in vivo* activiteit van SipS, is wellicht dat verscheidene residuen in dit domein een redundante functie hebben.

Concluderend geven de observaties, die in dit proefschrift beschreven zijn, nieuwe inzichten in de relatie tussen de structuur en de functie van type I SPases van *Bacillus* soorten. Biochemische karakterisering, gecombineerd met 'molecular modeling' en functionele analyse technieken geven aan dat de SPases van *Bacillus* model kunnen staan voor onderzoek aan type I SPases van andere bacteriën.