



Universiteit
Leiden

The Netherlands

A functional genomics study of extracellular protease production by *Aspergillus niger*

Braaksma, M.

Citation

Braaksma, M. (2010, December 15). *A functional genomics study of extracellular protease production by Aspergillus niger*. Retrieved from <https://hdl.handle.net/1887/16246>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/16246>

Note: To cite this publication please use the final published version (if applicable).

SAMENVATTING

De schimmel *Aspergillus niger* wordt veel gebruikt in de industriële biotechnologie voor de productie van verschillende metabolieten, waaronder citroenzuur, en een grote verscheidenheid aan enzymen. Van nature is *A. niger* in staat grote hoeveelheden eiwit uit te scheiden. Door deze eigenschap is bij de opkomst van genetische modificatietechnieken ook onderzocht of deze schimmel kon worden gebruikt als productieorganisme voor de commerciële productie van enzymen die van nature in andere schimmels of andere organismen voorkomen. Deze soortvreemde (heterologe) eiwitten worden vaak efficiënt tot expressie gebracht als het eiwitten van andere schimmels betreft. Wanneer het echter om heterologe eiwitten met een andere oorsprong dan schimmels gaat, is de productie tot op heden meestal onvoldoende om rendabel te zijn. Afbraak van deze heterologe eiwitten door eiwitafbrekende enzymen, zogenaamde proteases die de schimmel van nature uitscheidt, wordt als één van de oorzaken gezien van deze tegenvallende opbrengsten.

De recente opmars van systeembioïologie, waaronder ook de onderzoeksgebieden genomics, transcriptomics, proteomics en metabolomics vallen, heeft een nieuwe impuls gegeven aan de ontwikkeling en toepassing van *A. niger* als een productieorganisme. Het werk dat in dit proefschrift beschreven is, had tot doel de productie van extracellulaire proteases in *A. niger* te onderzoeken, waarbij gebruik is gemaakt van een systeembioïologie aanpak.

In **Hoofdstuk 1** wordt de rol van proteases in stam- en procesverbetering van *A. niger* en andere aspergilli beschreven. Ook wordt vooruitgeblikt op de rol die systeembioïologie zou kunnen spelen om het begrip van het proteolytische systeem van aspergilli te vergroten. Zowel klassieke mutagenese als moleculair genetische methoden worden tot op de dag van vandaag succesvol toegepast om stammen te genereren met een verlaagde protease activiteit. Het gebruik van de resulterende stammen voor de productie van heterologe eiwitten kan tot een significante toename van de geproduceerde hoeveelheden eiwit leiden. Strategieën waarbij een gen wordt uitgeschakeld hebben zich niet alleen gericht op individuele protease genen. Ook is gekeken naar het effect van het uitschakelen van specifieke regulatoren van protease genen. Een andere mogelijkheid is het uitschakelen van regulatiegenen die de expressie sturen van een breder scala aan genen, waaronder bepaalde protease genen, als reactie op externe factoren zoals pH of koolstofbron. Deze laatste aanpak lijkt echter ongeschikt voor het genereren van protease-deficiënte gastheerschimmels voor eiwitproductie. Het uitschakelen van zulke regulatoren resulteert namelijk ook in pleiotrope groeidefecten.

Ook de selectie en ontwikkeling van fermentatiecondities waarbij de productie van proteases wordt onderdrukt kan, naast stamontwikkeling, helpen de productie van heterologe eiwitten te verbeteren. Hier is echter nooit systematisch onderzoek naar verricht. De enkele studies die over dit onderwerp zijn verschenen hebben zich met name gericht op de kweekparameters pH, koolstof- en stikstofbron.

Hoofdstuk 2 beschrijft een systematische studie naar de invloed van verschillende omgevingsfactoren op de productie van extracellulaire proteases door *A. niger*. Hierbij is gebruik gemaakt van gecontroleerde batch fermentaties. In het eerste deel van de studie is door steeds één parameter per experiment te veranderen geprobeerd vast te stellen welke factoren met name protease activiteit beïnvloeden. Vervolgens zijn vier geselecteerde factoren, namelijk pH, koolstofbron, stikstofbron en concentratie van de stikstofbron, verder onderzocht. Hiervoor werd een aanpak gebruikt waarbij elke factor op twee niveaus werd getest en in alle mogelijke combinaties met de andere factoren. De resultaten toonden onder andere een duidelijk interactie-effect aan tussen stikstofbron en concentratie van de stikstofbron. Dergelijke interactie-effecten tussen omgevingsfactoren in relatie tot extracellulaire protease productie door *A. niger* zijn nog niet eerder beschreven. Het selecteren van omgevingsfactoren om protease activiteit te reduceren is door deze interactie-effecten niet eenvoudig, omdat onverwachte antagonistische of synergetische effecten kunnen optreden.

Hiernaast is in deze studie niet alleen gekeken naar het effect van procesparameters op de maximale protease activiteit, maar ook op nog vijf andere fenotypen gerelateerd aan protease. De resultaten duiden op significante verschillen in het effect van de omgevingsfactoren op de verschillende protease-gerelateerde fenotypen. De pH beïnvloedde bijvoorbeeld duidelijk de eindwaarden van protease activiteit, maar niet die van protease productiviteit. De uitgevoerde experimenten resulteerden in een grote en evenredig verdeelde variatie van protease activiteit. Dit maakt de resultaten van deze experimenten een geschikt startpunt voor een systeembioologie aanpak (Hoofdstukken 3-5).

Met de beschikbaarheid van de genoomsequentie van *A. niger* is het mogelijk geworden *in silico* het secretoom, oftewel de subset van het proteoom dat alle uitgescheiden eiwitten omvat, te voorspellen. Door alle voorspelde eiwitten te analyseren op de aanwezigheid van een signaalpeptide (SP) aan de N-terminus kan een lijst worden opgesteld met potentieel uitgescheiden eiwitten. Echter, de genmodellen die voor deze voorspellingen worden gebruikt bevatten soms fouten, waardoor het *in silico* voorspelde secretoom niet een accurate weergave is van het daadwerkelijke secretoom. In **Hoofdstuk 3** wordt een verbeterde lijst gepresenteerd

met potentiële SP eiwitten gecodeerd door het *A. niger* genoom. Bij het opstellen van deze lijst zijn, naast de SP voorspellingen van *A. niger* CBS 513.88, ook de SP voorspellingen van de beste homologen van *A. niger* ATCC 1015 en drie naburige *Aspergillus* soorten meegenomen. Ons uitgangspunt was dat de SP voorspelling van een *A. niger* non-SP eiwit zeer waarschijnlijk onjuist is, wanneer voor de meerderheid van de homologe eiwitten wel een SP voorspelt is en vice versa.

Als aanvulling op de *in silico* data is een shotgun proteomics aanpak gebruikt om het secretoom van *A. niger* te onderzoeken bij zowel groei- als koolstoflimiterende condities. In onze studie hebben we in het kweekmedium meer dan 200 eiwitten met een SP voorspelling gevonden. Daarnaast zijn ook tenminste twee uitgescheiden eiwitten gevonden waarvoor geen SP voorspeld was. Eén van deze twee eiwitten is een aspartylprotease (An01g00370) dat sterk overeenkomt met aspergillopepsin apnS van *A. phoenicis*. Blijkbaar wordt de export van dit eiwit niet door een SP gedirigeerd. Behalve dit aspartylprotease werden negentien andere proteases in deze studie gevonden die allen wel een SP hadden.

Het secretoom verandert met de groeicondities. Bij groei op sorbitol werd een breed scala aan koolhydraatacterende enzymen gevonden, terwijl bij groei op galacturonzuur vooral een specifieke subset van pectinolytische enzymen werd geïnduceerd. Koolstoflimitatie induceerde de expressie van proteases, zoals ook al was waargenomen in Hoofdstuk 2. Echter, niet zozeer het aantal verschillende proteases veranderde bij koolstoflimitatie, maar de relatieve bijdrage van specifieke proteases nam toe. Hieronder ook het belangrijkste protease in *A. niger*, het aspartylprotease aspergillopepsin.

Eén van de toepassingsgebieden van systeembioïologie is het optimaliseren van microbiële productieprocessen. In **Hoofdstuk 4** hebben we onderzocht hoe de keus van een kwantitatief fenotype voor optimalisatie de uitkomst beïnvloedt van een optimalisatiestrategie gebaseerd op een metabolomics aanpak. Hiervoor hebben we de productie van de industrieel belangrijke *A. niger* producten glucoamylase en protease bestudeerd. Voor beide producten zijn verschillende kwantitatieve fenotypen verwant aan productie gedefinieerd. Daarbij is ook rekening gehouden met de verschillende tijdstippen waarop monsters genomen zijn. Met behulp van de multivariate data-analyse techniek partial least squares (PLS) is het informatiegehalte van de metabolomics dataset in relatie tot al deze verschillende kwantitatieve fenotypen geëvalueerd. De resultaten lieten zien dat het tijdstip van bemonstering een veel grotere invloed heeft op het informatiegehalte van de data dan de manier waarop een kwantitatief fenotype is gedefinieerd.

Uit analyse van de metabolieten die door de PLS analyse als belangrijk zijn aangemerkt voor de twee bestudeerde producten bleken verschillende suikers met glucoamylase activiteit te correleren. Echter, de geïdentificeerde disaccharides kunnen zowel mogelijk glucoamylase productie induceren, of producten zijn die door glycosyltransferase activiteit van glucoamylase uit glucose worden gevormd. Vooral in het eerste geval zijn de disaccharides van belang voor stamoptimalisatie. Voor optimalisatie van glucoamylase productie in *A. niger* werd ook glucose-6-fosfaat isomerase als één van de andere interessante aanknopingspunten voor verder onderzoek geïdentificeerd.

Om protease activiteit in *A. niger* te reduceren werden geen direct voor de hand liggende aanknopingspunten voor verder onderzoek gevonden. Er was van te voren al rekening gehouden met het feit dat de relatie tussen intracellulaire metabolieten en extracellulaire protease activiteit gecompliceerd zou zijn, omdat bij proteolytische afbraak verschillende enzymactiviteiten betrokken zijn, zoals ook al in Hoofdstuk 3 was gevonden.

Hoofdstuk 5 geeft een genoom-brede blik op de transcriptierespons van *A. niger* door genen op basis van co-expressie te groeperen. Co-expressie netwerken van genen werden gemaakt op basis van twee verschillende DNA microarray datasets. De ene dataset is gebaseerd op experimenten waarbij *A. niger* relatief mild verstoord werd door kleine hoeveelheden van een inducerende component toe te voegen. Voor de andere dataset werd *A. niger* veel heviger verstoord door verschillende groeiomstandigheden op te leggen, waaronder koolstoflimitatie.

In eerste instantie werden voor zowel de individuele als de gecombineerde datasets gen co-expressie netwerken gemaakt op basis van een selectie van geconserveerde genen. Bij het vergelijken van de verschillende netwerken bleken sommige clusters van genen, zogenaamde modules, aanwezig in alle drie netwerken, terwijl andere conditiespecifiek waren. Vervolgens zijn alle *A. niger* genen, waaronder ook hypothetische en slecht geconserveerde genen, geïntegreerd in de co-expressie analyse. Dit werd gedaan door voor elk oorspronkelijke module een representatief expressieprofiel op te stellen en daarmee de expressieprofielen van alle overige genen te vergelijken. De biologische relevantie van de ontdekte modules blijkt uit de oververtegenwoordiging binnen deze modules van genen die gerelateerd zijn aan specifieke biochemische routes. Voor een aantal modules wordt dit tevens ondersteund door de aanwezigheid van geconserveerde motieven in de promotorregio van veel van de genen binnen zo'n module. Enkele van de gevonden geconserveerde sequentiemotieven zijn bekende bindingsites van DNA-bindende

transcriptiefactoren. Voorbeelden hiervan zijn de bindingsites van de transcriptiefactor CpcA, die gerelateerd is aan het aminozuurmetabolisme, en die van de transcriptiefactoren FarA and FarB, die gerelateerd zijn aan het verzuurmetabolisme. Daarnaast zijn enkele nog niet eerder beschreven mogelijke transcriptiefactor bindingsites ontdekt voor de module met genen die coderende voor ribosomale eiwitten en de module met genen gerelateerd aan algemene gen expressie processen.

Tenslotte wordt in **Hoofdstuk 6** een 'top-down' systeembioologie aanpak, gebaseerd op de informatie verkregen middels functionele genomics technologieën en in combinatie met multivariate data analyse, bediscussieerd als een methode om onbevooroordeeld aanknopingspunten voor stam- en procesverbetering te selecteren en op hun belangrijkheid te rangschikken.

