Identification and characterization of starch and inulin modifying network of Aspergillus niger by functional genomics

Yuan, X.L.

Chapter 2

Construction of inulin and starch specific *Aspergillus niger* cDNA expression libraries

Xiao-Lian Yuan, Mark Arentshorst, Cees A.M.J.J. van den Hondel and Arthur F.J. Ram

Abstract

To screen for genes encoding enzymes with starch or inulin modifying activities from *Aspergillus niger*, cDNA expression libraries were generated using Gateway cloning technology. RNA was isolated from different time points during submerged growth in cultures containing starch or inulin as carbon sources. RNA from the different time points (16, 32, 48 and 64 h after inoculation) of inulin or starch grown cultures were isolated and equally pooled for cDNA synthesis. The synthesized cDNAs were cloned into both *Escherichia coli* expression vector (pSPORT1) and *Saccharomyces cerevisiae* expression vector (pDEST-YES52). Analysis of 24 randomly picked clones from the inulin and starch specific *E. coli* cDNA expression libraries showed that 95% of clones harbor cDNA inserts with average size of 1.4 kb for inulin libraries and 100% of clones contain cDNA inserts with average size of 1.8 kb for starch libraries, respectively. For construction of the *S. cerevisiae* expression libraries the construction procedure was more complex. First, cDNA fragments were cloned into a mammalian expression vector (pCMV-SPORT6) and the resulting inserts from the primary libraries were transferred to pDONR201 to give so-called entry libraries. The inserts from the starch or inulin entry libraries were transferred to the destination vector (pDEST-YES52), to give the *S. cerevisia*e expression libraries. Analysis of the yeast expression libraries showed that 33% of the ampicillin resistant cDNA clones from the starch library and 54% of clones from the inulin library contained a cDNA in the yeast expression vector. The remaining clones were from previous libraries, either primary vector pCMV-SPORT6 or entry vector pDONR201 or a combination of the two. To correct for the relatively low percentage of *E. coli* clones harboring the right expression vector, a relative large number of primary transformants were produced to have enough correct yeast expression clones for screening. As only cDNAs clones inserted in pYES-DEST52 will be transformed to *S. cerevisiae* the presence of other plasmids in the pool of clones will not hamper the screening procedure. Analysis of size of the inserts showed, in general, a reduction of the average length of the insert in subsequent libraries. The average length of the inserts in the yeast expression library for inulin was 1.2 kb and 1.6 kb for starch. This suggested that the construction of cDNA libraries can be improved by reducing cDNA transfer steps, e.g. by insertion of cDNAs directly into an entry vector or by insertion cDNAs directly into a *S. cerevisiae* expression vector.

Introduction

Aspergillus niger are distributed worldwide and commonly present on decaying plant debris. As a saprophytic fungus, *A. niger* produces a variety of hydrolytic enzymes that are able to break down the plant polysaccharides into smaller molecules, which can be served as their nutrient source. Many enzymes secreted by *A. niger* have already found their applications in the baking, starch, textile and food and feed industries (Schuster et al., 2002; De Vries et al., 2001; Semova et al., 2006).

Enzymes secreted by *A. niger* that are involved in starch or inulin catabolism have been previously described. The known starch degrading enzymes include glucoamylase

(*glaA*) (Nunberg et al., 1984), acid alpha-amylase (*aamA*) (Boel et al., 1990; Brady et al., 1991), alpha-amylases (*amyA* and *amyB*) (Korman et al., 1990) and alpha-glucosidase (*aglA*) (Kimura et al., 1992). Well described inulin degrading enzymes consist of invertase (*sucA*) (Bergès et al 1993, Boddy et al., 1993; L'Hocine et al., 2000), endo-inulinase (*inuA* and *inuB*) (Ohta et al., 1998; Aikimoto et al., 1999), and exo-inulinase (*inuE*) (Arand et al., 2002; Kulminskaya et al., 2003; Moriyama et al., 2003). Since *A. niger* grows well on starch and inulin as carbon and energy sources, it is predicted that a full set of enzymes converting starch into oligosaccharides, maltose, glucose or enzymes converting inulin into oligofructose, sucrose, glucose and fructose are present in this organism.

To identify additional genes encoding enzymes with starch or inulin modifying activities, one approach is to use the deduced amino acids sequences of known starch or inulin degrading enzymes and to perform Blast or HMM searches against the *A. niger* genome to identify related proteins. This approach has been successfully used to predict or identify additional enzymes involved in starch or inulin catabolism (chapters 3 and chapter 7). The disadvantage of this approach is that completely new starch or inulin modifying enzymes can be missed if these enzymes may have amino acid sequences which differ from the known enzymes. Consequently, those proteins are not recognized in Blast of HMM searches. In addition, site activities of enzymes form other Glycosyl Hydrolyases (GH) families might be missed by performing only the bioinformatics based genome mining approach. An alternative approach to identify starch or inulin modifying enzymes is to construct cDNA expression libraries in combination with a High Throughput Screening (HTS) using appropriate substrates. The screening of cDNA libraries for interesting enzyme activities has been proven to be a powerful and efficient method for investigation of the function of gene products (van der Vlugt-Bergmans and van Ooyen, 1999; Meeuwsen et al., 2000).

Construction of high-quality (full length) cDNA libraries in proper expression vectors is essential for successful screening and further facilitates the cDNA product characterization. Genetic manipulation of *E. coli* colonies are easy to handle in HTS and large amount of recombinant proteins can be expressed in a short time. However, the expression of eukaryotic proteins in *E. coli* can be problematic, due to aggregation, formation of insoluble inclusion bodies, and/or degradation of the expression product (Hannig and Makrides, 1998; Baneyx, 1999). Eukaryotic hosts e.g. yeast species (*Pichia pastoris* and *S. cerevisiae*) or filamentous fungi (e.g. *A. niger*) as an alternative expression system, provide the specific cellular environment for the expression of eukaryotic proteins. However, possible bottlenecks in the use of *P. pastoris* or *S. cerevisiae* as host for the expression of *Aspergillu*s cDNAs include that expression could suffer from lower yields of heterologous proteins (Buckholz and Gleeson, 1991; Punt et al., 2002; Holz et al., 2003) and that these organisms are more laborious in HTS compared to *E. coli*. .

Gateway Cloning Technology is a universal system for cloning and subcloning DNA/cDNA fragments into many expression vectors (Ohara and Temple, 2001). This technology uses the λ-recombination system transfer cDNA fragments between vectors that

contain recombination sites for the λ-recombinase machinery while maintaining reading frame and orientation. This approach is characterized by several advantages over conventional restriction-assisted cloning methods which include the facts that, (i) it eliminates restriction digestion for directional cloning, ii) it generates lower level of chimeric clones, iii) produces higher amount of full-length cDNA clones and iv) gives higher cloning efficiency (Hartley et al., 2000; Ohara and Temple, 2001). Moreover, this approach facilitates the construction of same cDNAs input into different expression vectors, especially when it is not clear which system or host background will provide sufficient expression levels to allow interesting protein purification. Thus, this approach reduces enormous amount of work and further prevents the significant barrier to progress.

In this paper, we have used the Gateway cloning technology to construct *A. niger* cDNA expression libraries to screen for starch and inulin degrading activities. Molecular characterization of the *E. coli* and *S. cerevisiae* libraries indicated that useful expression libraries have been obtained. Approaches that might further improve the construction of cDNA expression libraries as well as the screening are discussed.

Materials and Methods

Strains and culture conditions

A. niger strain N402 used in this study was derived from the wild-type strain *A. niger* van Tieghem (CBS 120.49, ATCC 9029) (Bos *et al*., 1988). Growth curves were determined by inoculating 2 x 10^6 spores/ml in 50 ml medium in 150-ml of Erlenmeyer flasks. Cultures were incubated at 30 °C in a rotary shaker at 300 rpm for the time indicated. Each flask contained 50 ml of Minimal Medium (pH 6.5) (Bennett & Lasure, 1991) supplemented with 0.1% (w/v) casamino acids and 2% (w/v) carbon source. Inulin (Sensus, Frutafit® TEX, Cosun, The Netherlands) and starch (Windmill starch, Avebe, The Netherlands) were used as carbon sources. Mycelia were harvested over a Myracloth filter followed by freezing in liquid nitrogen. The samples were stored at -80 °C prior to the isolation of total RNA or to the determination of biomass. For determination of biomass, the frozen mycelia were freeze dried. Conidiospores were obtained by harvesting spores from a complete medium-plate (minimal medium with 0.5% (w/v) yeast extract and 0.1% (w/v) casamino acids) containing 1% (w/v) glucose, after 4-6 days of growth at 30 °C, using a 0.9% (w/v) NaCl solution.

The bacterial strain used for transformation and amplification of cDNA libraries was *Escherichia coli* ElectroMAX DH10B (Invitrogen). Transformation of ElectroMAX DH10B cells was performed according to the electroporation protocol as described by suppliers. The electroporation voltage in the gene pulser is 2.5 kV in a 0.1 cm gap chamber at settings of 100 Ω and 25 μF.

Northern analysis

Total RNA was isolated by grinding frozen (-80 °C) mycelium in liquid nitrogen with a pestle and mortar. Powdered mycelium (200 mg) was extracted with 1 ml TRIzol Reagent (Invitrogen, U.S.A) in accordance with the supplier's instructions. Northern analysis and generation of probes for the *A. niger* invertase (*sucA*) and the exo-inulinase (*inuE*) was performed as described in Yuan et al. (2006). The probe for glucoamylase (*glaA*) was a 0.5-kb *Kpn*I fragment from pAN56-2 and the probe for glyceraldehyde-3-phosphate dehydrogenase (*gpdA*) was a1.5-kb of *Hind*III fragment from pAB5-2. Both pAB56-2 and pAB5-2 were obtained from Dr. Peter Punt (TNO, the Netherlands).

Construction of cDNA library in *E. coli* expression vector

Total RNA was isolated using Trizol reagent (Invitrogen) according to the supplier's instructions. mRNA was isolated using FastTrack kit (Invitrogen, K1593). cDNA libraries were constructed using Superscript plasmid system with Gateway technology for cDNA synthesis and plasmid cloning (Invitrogen, 18248-013). Briefly, the first strand cDNA was primed with an oligo(dT)-primer/*Not*I-linker in the presence of SuperScript II RT and dNTPs. The second strand cDNA was synthesized by the combined actions of RNase H and DNA polymerase I. The ends of the cDNA copies were first filled in with DNA polymerase I in the presence of dNTPs, ligated with *Sal*I adaptors using T4 DNA ligase, then digested with *Not*I, to obtain *Not*I-*Sal*I termini cDNA. The obtained ds cDNAs with *Not*I-*Sal*I termini were further directionally cloned into pSPORT1. Plasmid pSPORT1 contains the ampicillin resistant gene for selection of transformants. The cloned products were transformed into *E. coli* strain ElectroMAX DH10B (MAX Efficiency®, Invitrogen) by electroporation. Independent transformants were grown into colonies on LB plates with ampicillin at a density of (1-3)*10^4 colonies per plate. Library cells were washed off from the plates with 0.9% NaCl and frozen down with 40% glycerol in 1-ml aliquots and stored at -80°C. Library DNA was isolated from the *E. coli* transformants using Plasmid Maxiprep Kit (QIAGEN) and 100-µl aliquots were stored at -20°C.

Construction of cDNA libraries in the *S. cerevisiae* expression vector

Gateway Cloning Technology provides the possibility to transfer DNA inserts between vectors using λ-att site-specific recombination and the technology was used to construct *A. niger* cDNA libraries in yeast expression vector. cDNAs with *Not*I-*Sal*I termini (see above) were first directionally cloned into attB containing vector pCMV-SPORT6 which resulted in the pCMV-SPORT6-cDNAs libraries (primary libraries). Plasmid DNA from the pCMV-SPORT6-cDNAs libraries was isolated and subjected for BP reaction with pDONR201. In short, 500 ng of the plasmid library was incubated with 300 ng of attP containing entry vector pDONR201, 4 µl of BP reaction buffer and 4 µl of BP clonase enzyme mix at 25°C for 1 hour. The reaction was stopped by adding 2 µl of proteinase K solution for 10 min at 37°C. 1µl of the BP reaction were electroporated into 50 µl of library efficiency competent cells *E. coli* strain ElectroMAX DH10B and plated out on of Kanamycin (50 µg/ml)

containing LB plates. Transformants washed from the plates and plasmid DNA was isolated which resulted in the entry library (pDONR201-cDNA). To construct final expression libraries, the entry plasmid pDONR201-cDNAs were used for LR reaction. Briefly, 300ng of entry plasmid was incubated with 300 ng of attR containing destination vector pYES-DEST52, 4 µl of LR reaction buffer and 4 µl of LR clonase enzyme mix at 25°C for 1 hour. 2 µl of proteinase K solution was added and incubated at 37°C for 10 min to stop the reaction. 1 µl of LR reaction mix were then transformed into 50 µl of strain ElectroMAX DH10B competent cells and plated out on Ampicillin (100 µg/ml) containing LB plates and resulted in the destination expression library harboring plasmid pYES-DEST52-cDNAs (see Fig. 2 for the construction procedure).

Digestion analysis

For the analysis of cDNA inserts in each library, 24 clones were randomly selected and cultured in 3 ml of LB medium containing 50 µg/ml of ampicillin (*E. coli* expression libraries, Gateway Primary libraries and Destination libraries) or 25 µg/ml of kanamycin (Gateway Entry libraries) individually. The harboring plasmids were isolated and used for digestion analysis. For each enzyme digestion reaction, 500 ng of plasmid was incubated with 10 U of the indicated enzymes, 2 µl of 10X corresponding enzyme buffer (InVitrogen), TE buffer to a total volume of 20 µl at 37°C for 2h. The digested reaction was then running in the agarose gel by electrophoresis at 80 volts for 2h.

Results and discussion

Set up growth conditions for generation of cDNA library from inulin and starch grown cultures

To construct inulin and starch specific cDNA libraries, the growth of *A. niger* in liquid medium containing inulin and starch as sole carbon source was examined. *A. niger* wild type strain N402 was cultivated in 2% (w/v) inulin or 2% (w/v) starch and the biomass during the growth period was determined at each time points (16, 32, 48, 64 and 80 h after inoculation) (Fig. 1A). In both the inulin and starch grown cultures, *A. niger* showed a typical growth curve with an exponential growth phase (between 16 and 48 h), a stationary growth phase (48-64 h) and a death or autolytic phase (64-80 h). During the stationary phase the cultures reached the highest biomass values with a density of 11.0 g/L for inulin, 9.0g/L for starch. During the autolysis phase the biomass dropped to a density of 9.0 g/L for inulin and 8.0 g/L for starch at 80 h respectively (Fig. 1A).

To examine the expression of genes encoding representative inulin or starch degrading enzymes, Northern blot analysis was performed using RNA isolated at the different time points. For the inulin culture, the genes encoding invertase (*sucA*) and exo-inulinase (*inuE*) were used as probes. *sucA* was highly expressed at very early stage (16 h) and the expression level was soon strongly reduced with the later time points. Almost no expression could be detected at the latest time point (80 h). *inuE* was expressed specifically

at 16 h and 48 h and not detectable at other time points (Fig. 1B). The lack of expression of *sucA* and *inuE* at 80 h might be due to the total consumption of inulin at this stage, possibly in combination with the degradation of mRNA during the autolytic phase of the growth.

A similar Northern blot analysis was also performed for the starch culture. The gene encoding extracellular starch degrading enzyme, glucoamylase (*glaA*) and an additional gene encoding glyceraldehyde-3-phosphate dehydrogenase (*gpdA*) were chosen to detect mRNA levels. *glaA* was expressed at exponential phase and stationary phase, 32, 48 and 64 h and no or very weak expression was detected at 16h and 80h. *gpdA* is usually a constitutively expressed gene, and often used as marker for RNA loading control. This gene was only highly expressed at early stages 16, 32 and 48h and its expression was strongly reduced and hardly detected at later stages (64 h and 80 h) (Fig. 1B). Thus, the expression of *gpdA* is only observed when the culture is actively growing. Reaching the stationary or autolysis phase of the culture, results in a strong reduction in *gpdA* mRNA levels, indicating that the constitutive expression of *gpdA* seems to be true for exponentially growing cultures.

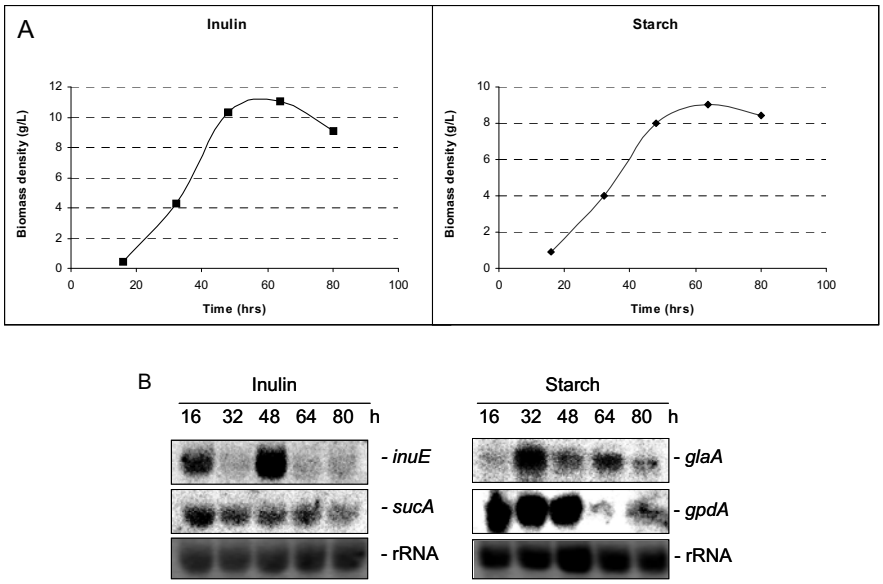

Fig. 1. Examination of growth conditions regarding to inulin or starch degrading activities. A, Growth curve of *A. niger* strain N402 on MM containing 2% (w/v) inulin or starch as sole carbon source. B, Expression analysis of *A. niger* genes on inulin or starch cultures at different time points.

Although these various genes showed different expression patterns at different time points, it was obvious that no or very weak expression of any of these genes was detected at latest time point (80 h) for both the starch and inulin cultures. The concern was that the RNA fractions of the 80 h time point contained high RNAse activity, which might cause RNA degradation of the other RNA samples upon pooling. Therefore, equal amounts of total RNA from the first four time points 16, 32, 48 and 64 h, excluding the 80 h time point, were pooled and used for generating cDNA libraries from both inulin and starch grown cultures.

Construction and characterization of inulin and starch cDNA libraries in the *E. coli* expression vector

An overview of the experimental procedure to construct the cDNA libraries is given in Fig. 2. Equal amount of total RNA from each time point (16, 32, 48, 64 h) of starch and inulin samples was pooled and mRNA was isolated using TRIzol Reagent (Invitrogen, U.S.A). The purified mRNA was used for cDNA synthesis and cloning using SuperScript plasmid system with Gateway Technology (Invitrogen). The synthesized cDNA was adapted with *Sal*I at its 5′ end and *Not*I at its 3′ end which allowed directional cloning in pSPORT1. Expression of cDNA in *E. coli* expression vector pSPORT1 is driven by the *lac*P promoter. The cDNA clones were transformed to *E. coli* by electroporation. Due to limitation of posttranslantional modification of eukaryotic proteins in *E. coli*, we have only generated a limited number of cDNA clones in *E. coli* expression vector. In total, 0.24×10^5 primary transformants for inulin cDNA library and 1.92×10^5 primary transformants for starch cDNA library were created, respectively (Table 1).

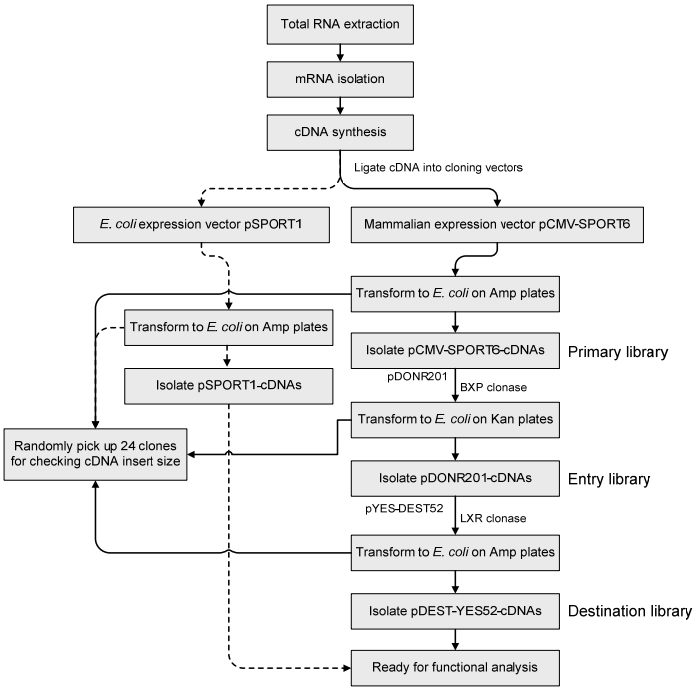

Fig. 2. Flow chart of construction of *A. niger* cDNA expression libraries using Gateway cloning technology. Solid line indicates yeast expression system; broken line indicates *E. coli* expression system.

Table 1. Molecular characterization of *A. niger* cDNA libraries in *E. coli* expression vector.

Cloning vector	cDNA source	No. of clones	No. of clones with insert	average insert size
pSPORT-1	Inulin	0.24×10^5	95%	1.5 kb
	Starch	1.92×10^5	100%	1.7 kb

To evaluate the cDNA insert size in the libraries, 24 single clones was randomly selected for each library and the plasmid was extracted and digested with *Eco*RV (next to cDNA 5′ termini adapter *Sal*I) and *Hind*III (next to cDNA 3′ termini adapter *Not*I), thereby releasing the cDNA insert. A variety of insert sizes range from 200 bp, the size expected for a clone with no insert, to 3 kb was observed for each library (data not shown). Empty clones, based on this digestion analysis represented 5% for inulin cDNA library while 0 for starch cDNA library. When the number of empty clones in each library was subtracted, the average insert size was 1.5 kb and 1.7 kb for inulin and starch cDNA libraries respectively (Table 1). Among them, 54% of inulin clones and 61% of starch clones harbor cDNA inserts with size between 1.0-2.0 kb, while 21% of inulin clones and 26% of starch clones contain cDNA fragments with size of over 2.0 kb (data not shown). This data indicated that both inulin and starch cDNA expression libraries in pSPORT1 are of good quality.

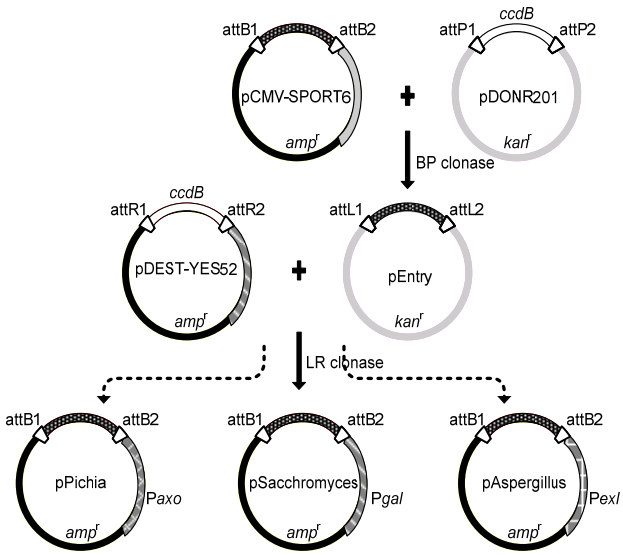

Fig. 3. Schematic representation of Gateway compatible cloning technology. BP reaction resulted in entry clones, LR reaction resulted in destination expression clones. Dot filled frame means cDNA fragment. Broken arrows mean possible destination expression clones generated from same entry clones. Possible promoters used to drive cDNA expression and antibiotic resistant genes for selection of clones are indicated.

Construction and characterization of inulin and starch cDNA libraries in the *S. cerevisiae* expression vector

To construct cDNA libraries in the *S. cerevisiae* expression vector, the same method was used as described above for the synthesis of cDNA insert flanked with *Sal*I and *Not*I restriction sites. Since no suitable entry vector was available for direct cDNA cloning in the yeast expression vector at that time of library construction (first half 2002), cDNA libraries in yeast expression vector were generated in three subsequent steps using Gateway cloning

technology (see materials and methods and Fig. 2). As shown in Fig. 3, this technology uses the λ-phage att recombination sites in combination with λ-recombinase enzyme (Clonase) to transfer cDNA segments from one vector to another. For example, cDNA fragments that are flanked by attB sequences in Gateway primary vector (e.g. pCMV-SPORT6) are efficiently transferred by recombination into entry vector (e.g. pDONR201) containing attP sequences with the use of BP Clonase. The obtained cDNA-entry clones after the BP reaction are properly oriented and now contain attL sequences. These attL flanked cDNA inserts can be efficiently transferred to other destination vectors (e.g. pDEST-YES52) containing attR sequences using LR clonase. Recombination of attL and attR sites results in the formation of attB sites, allowing subsequent transfer of the inserts to another BP reaction, if required. In this cloning system, two types of antibiotic markers are used to select for transformants. The attB-containing primary plasmid (pCMV-SPORT6) and the attR-containing destination plasmid (pDEST-YES52) contain the ampicillin resistant gene while the attP-containing donor plasmid contains the kanamycin resistant gene. Moreover, the donor vector contains the *ccdB* gene flanked by the attP sites for negative selection to prevent growth of *E. coli* harboring *ccdB* containing λ-phages, by products of BP and LR reaction, in medium, thus increasing recombination efficiency.

For the construction of the *S. cerevisiae* expression libraries, ds-cDNAs within *Sal*I-*Not*I fragments were first cloned into Gateway primary cloning vector pCMV-SPORT6. In the primary libraries, 2.80×10^5 and 1.75×10^5 independent clones were created for inulin and starch cDNA libraries, respectively. Based on the *Eco*RV-*Hind*III digestion analysis of 24 randomly selected clones, 95% of clones from inulin library and 100% of clones from starch library contain cDNA inserts. The average insert size excluding empty clones was 1.4 kb for inulin library and 2.2 kb for starch library, respectively (Table 2). For the inulin library, 71% of the clones contained an insert of 1 kb or larger and 83% of the inserts of the starch library was over 1 kb in length (data not shown). This preliminary data showed the primary cDNA libraries were satisfying and further construction of yeast expression libraries was preceded.

Table 2. Molecular characterization of *A. niger* cDNAs libraries in Gateway compatible vectors.

Gateway vector	cDNA source	No. of clones	No. of clones with insert	average insert size
pCMV-SPORT6	Inulin	2.80×10^5	95%	1.4 kb
	Starch	1.75×10^5	100%	2.2 kb
pDONOR201	Inulin	2.87×10^5	100%	1.7 kb
	Starch	1.88×10^5	100%	1.8 kb

Using the BP reaction, cDNA inserts were transferred from Gateway primary vector pCMV-SPORT6 to entry vector pDONR201 resulting in the so-called entry libraries. In the entry libraries, 2.87×10^5 and 1.88×10^5 recombinants grown on kanamycin plates were harvested for inulin and starch library, respectively. All the clones analyzed from both inulin and starch entry libraries harbor cDNA fragments. The average size was 1.7 kb for inulin library

and 1.8 kb for starch library (Table 2). The data indicated that these entry libraries were of good quality and used to produce final yeast expression libraries.

Table 3. Molecular characterization of *A. niger* cDNA libraries in yeast expression vector.

Gateway vector	cDNA source	No. of clones	% of pYES clones	No. of pYES clones	average insert size
pDEST-YES52	Inulin	2.80×10^5	54%	1.50×10^5	1.2 kb
	Starch	5.00×10^5	33%	1.50×10^5	1.6 kb

Using the LR reaction, cDNA inserts from the entry library were then transferred to destination yeast expression vector pYES-DEST52. In destination libraries, clones were selected on ampicillin LB plates. 2.80×10^5 clones for inulin library and 5.00×10^5 for starch library were generated. Surprisingly, *Eco*RV/*Hind*III digestion analysis revealed that only 54% and 33% of analyzed clones for inulin and starch libraries, respectively, consist of pYES-DEST52 backbone containing the URA3 selection marker to select for transformants in *S. cerevisiae* (Table 3). This was further confirmed by *Pst*I digestion of the clones. Proper destination plasmids should give an anticipated 500 bp fragment and only the plasmids that were identified as pYES-DEST52 vectors using the *Eco*RV/*Hind*III digestion contained the 500 bp fragment. Further digestion analysis of 11 unclear clones showed that one of them was from primary vector pCMV-SPORT6. The pattern of digestion of the remaining clones gave inconclusive results and the exact nature of these plasmids is not clear. Testing of the antibiotic resistance of the clones on ampicillin or kanamycin containing LB-plates revealed that about 30% of destination clones harbor both ampicillin and kanamycin resistance genes (data not shown). This indicates that contamination of previous vectors during recombinant reaction occurred; however the exact mechanism that is either from single cross-over during transformantion or from co-transformation with previous vectors is still unknown. Importantly, the plasmids in this library whose exact nature is unclear do not contain the URA3 selection marker. Thus, upon transformation to *S. cerevisiae* only pDEST-YES52 expression vectors will result in yeast transformants. Therefore no more effort was taken for clarification of the unexpected clones. Excluding clones with the contaminated vectors, the average cDNA insert size in pDEST-YES52 vector was 1.2 kb for inulin library and 1.6 kb for starch library (Table 3). These sizes are reasonable for production of cDNA product. Because of the presence of addition plasmids other than with the pDEST-YES52 backbone, the number of primary transformants in the libraries to reach certain amounts (150.000) of clones with a pDEST-YES52 backbone for functional screening was adjusted (Table 3).

Future considerations for construction and screening of cDNA expression libraries

cDNA expression cloning is a powerful and efficient tool for investigation of secreted hydrolytic enzymes of fungal origin (Dalboge and Heldt-Hansen, 1994; Christgau et al., 1995; van der Vlugt-Bergmans and van Ooyen, 1999). However, selection of optimal protein expression system is very important for screening and functional analysis of the interesting cDNA clones.

Both bacteria expression system and eukaryotic system have advantages and disadvantages for expressing eukaryotic genes. For example, bacteria strain *E. coli* is a valuable host for the efficient, cost-effective and high level production of heterologous proteins, however, the problems of protein accumulation and protein processing, protein folding and posttranslational modification for eukaryotic genes is still an issue (Hannig and Makrides, 1998). On the other hand, yeast strains *Pichia pastoris* and *Saccharomyces cereviseae* as well as *Aspergillus* strains have benefit for protein processing and protein posttranslation modification, but have limitation for lower yield and host strain instability (Buckholz and Gleeson, 1991; Da Silva and Bailey, 1991; Punt et al., 2002; Holz et al., 2003). Nevertheless, modification on both eukaryotic system and prokaryotic system has been improved for expression of eukaryotic genes (http://www.invitrogen.com). In the original experimental set-up, bacteria strain *E. coli*, yeast strains *P. pastoris*, *S.s cerevisiae* and *A. niger* were all considered as good candidates for expressing *A. niger* cDNA clones. Therefore we have used Gateway cloning technology which allows efficient constructing the cDNA libraries in variant expression vectors (Ohara and Temple, 2001) to construct the cDNA expression libraries (see Fig. 2 and Fig. 3).

Construction of *E. coli* expression libraries was easy by one single ligation step. The synthesized cDNA was cloned into *E. coli* expression vector pSPORT1 by single ligation step with directional cloning (see Fig. 2). Both the percentage and sizes of cDNA inserts matched to the requirements for cDNA screening (see Table 1). Due to limitation of posttranslantional modification of eukaryotic proteins in *E. coli*, we have only generated a limited number of cDNA clones in *E. coli* expression vector (Table 1).

Construction of yeast expression libraries was complex using Gateway cloning technology (Fig. 2 and Fig. 3). The results from the yeast expression libraries may not show very efficient due to three steps cDNA transfer (Table 3). However, it is worth noting that these libraries cloned in Gateway system are easy to modify. One option is to modify the destination expression vector by replacing attR recombinant sequence with the attP sequence, by one step BP reaction, cDNA inserts could be transferred into destination expression vector. On the other hand, Gateway system has been improved in such way that the synthesized cDNA fragments are flanked by attB recombinant sequences on each end and can be directly transferred to attP-containing Gateway entry vector by BP reaction, skipping the step to the primary vector pCMV-SPORT6. By one step LR reaction, the cDNAs can be cloned into different expression vectors (Karnaoukhova et al., 2003; M. Arentshorst, personal communications). These modifications would reduce the recombination step, and thus lower the cDNA size reduction and minimize the contamination efficiency.

Our cDNA libraries have been used in preliminary transformation and screening experiments in both *E. coli* and *S. cerevisiae* to identify inulinolytic and amylolytic activities (C. Goosen, R. van der Kaaij, personal communication), but have not yielded clones with clear hydrolytic activities. PCR analysis, however, revealed the presence of known inulinolytic (*inuA, inuE* and *sucA*) and amylolytic (*glaA*) clones in these cDNA libraries,

indicating the problem occurs in protein expression and screening. Targeted overexpression of *sucB* using the same yeast expression system resulted in the expression of SucB protein in cell free lysates (Goosen et al., 2007, Chapter 6) indicating that the expression system is functional. Firm conclusions about the libraries can not be drown at the moment since only a limited number of transformants (about 3,000 clones) have been analyzed. In addition, problems were encountered when expressing cDNA from the *gal1* promoter. The expression level of induction from the *gal1* promoter in *S. cerevisiae* was lower as expected in the strain used (BY4347 *Δsuc2*). A possible improvement of the screening system is the use of a yeast strain in which the expression from the *gal1* promoter is stronger.