Biocatalysis and Biotransformation, 2013; 31(6): 335–342

informa
healthcare

ORIGINAL ARTICLE

Improved soluble expression of the gene encoding amylolytic enzyme Amo45 by fusion with the mobile-loop-region of co-chaperonin GroES in *Escherichia coli*

LEI WANG[1], HILDEGARD WATZLAWICK[1], OLAFUR FRIDJONSSON[2], GUDMUNDUR HREGGVIDSSON[2,3] & JOSEF ALTENBUCHNER[1]

[1]*Universität Stuttgart, Institut für Industrielle Genetik, Stuttgart, Germany,* [2]*Matis Ltd., Reykjavik, Iceland, and* [3]*Department of Biology, University of Iceland, Faculty of Life and Environmental Sciences, Reykjavik, Iceland*

Abstract
The gene encoding the amylolytic enzyme Amo45, originating from a metagenomic project, was retrieved by a consensus primer-based approach for glycoside hydrolase (GH) family 57 enzymes. Family 57 contains mainly uncharacterized proteins similar to archaeal thermoactive amylopullulanases. For characterization of these family members soluble, active enzymes have to be produced in sufficient amounts. Heterologous expression of *amo45* in *E.coli* resulted in low yields of protein, most of which was found in inclusion bodies. To improve protein production and to increase the amount of soluble protein, two different modifications of the gene were applied. The first was fusion to an N-terminal His-tag sequence which increased the yield of protein, but still resulted in high amounts of inclusion bodies. Co-expression with chaperones enhanced the amount of soluble protein 4-fold. An alternative modification was the attachment of a peptide consisting of the amino acid sequence of the mobile-loop of the co-chaperonin GroES of *E.coli*. This sequence improved the soluble protein production 5-fold compared to His_6-Amo45 and additional expression of chaperones was unnecessary.

Keywords: *Metagenome mining, glycoside hydrolase family 57, expression in* E.coli*, chaperones, Fusion with the mobile-loop of GroES*

Introduction

The starch industry depends on biocatalytic processes to modify and fractionate starch. There is a constant need for new and better enzymes for starch and carbohydrate processing. The target is highly thermostable and highly active enzymes as well as provision of new primary products such as oligosaccharides of defined size, composition and degree of branching, new types of linkages, cyclic or more complex polysaccharides, and secondary sugar derivatives such as substituted starches and polyols. Most industrial enzymes, for example, starch-modifying enzymes of the α-amylase family have been obtained so far from organisms isolated from the natural environment. Only a small fraction of the bacterial population in soil and other habitats can be propagated in the laboratory which means there is a huge and largely unexplored wealth of genes in the environment. Metagenomics is an approach to extract DNA and prospective genes instead of isolating microorganisms from habitats and screen for useful genes. The mining of DNA from a geothermal habit from Iceland resulted in novel amylolytic enzymes of the glycoside hydrolyse family 13 by applying a consensus primer-based approach for this glycoside hydrolase (GH) family (Labes et al. 2008). Besides GH13 family, amylolytic enzymes were also grouped in some other GH families, for example, family 70 and 77 (Cantarel et al. 2009). Recently new members of amylolytic enzymes, especially branching enzymes, were classified in GH family 57 (Palomo et al. 2010). As these enzymes exhibit interesting features like branching activities, a similar approach of mining environmental DNA

Correspondence: Hildegard Watzlawick, Universität Stuttgart, Institut für Industrielle Genetik, Allmandring 31, 70569 Stuttgart, Germany. E-mail: hildegard.watzlawick@iig.uni-stuttgart.de

(Received 16 May 2013; revised 21 August 2013; accepted 21 October 2013)

ISSN 1024-2422 print/ISSN 1029-2446 online © 2013 Informa UK, Ltd.
DOI: 10.3109/10242422.2013.858712

was used for α-amylases of the family GH57 and is described in this paper. Applying the consensus primer-based oligonucleotides approach on a metagenomic library derived from thermophilic microorganisms resulted, among others, in the gene *amo45*.

In order to characterize the gene product, overexpression was necessary. Here we chose *E.coli* as it is the best known and most commonly used host for expression of recombinant proteins. Despite many advantages, for example, number of vectors with strong promoters which facilitates high yields of protein production, some proteins are difficult to produce in large amounts and in an active form. This is caused by high or low GC content of the genes, rare codons in the host, instability of the mRNA or protein, and inclusion body formation.

To overcome such problems, codon optimization according to codon usage (Grantham et al. 1980; Ikemura 1981; Altenbuchner & Mattes 2005), co-expression of additional rare tRNA genes like *argU, ileY, ilX, leuW, proL*, and *metT* by using additional plasmid (Stratagene) or genetic-engineered strains (Novagen), fusion to maltose-binding protein (MBP) or to the N-utilization substance protein A (NusA) (Waugh 2005), and chaperone co-expression (Thomas et al. 1997; Nishihara et al. 1998) have mainly been utilized.

The chaperone family GroEL and its co-chaperone GroES play an essential role in mediating protein folding and preventing protein aggregation. The GroEL/ES chaperonin system of *E.coli* has been studied most extensively (Hartl 1996; Xu et al. 1997; Richardson et al. 1998; Sigler et al. 1998; Hartl & Hayer-Hartl 2002; Horwich et al. 2007). While GroEL contains 14 identical 57kDa-subunits which together assemble to a double-heptameric-ring, GroES contains seven identical 10 kDa-subunits forming a heptameric-ring. The GroEL double-ring consists of an equatorial ATPase domain, an intermediate hinge-domain, and an apical domain exposing hydrophobic surface for substrate binding. The GroES ring complex covers the ends of the GroEL cylinder through its seven mobile loops which fold into a β-hairpin conformation upon binding to GroEL. This has been demonstrated by crystallographic and biochemical studies (Landry et al. 1993; Xu et al. 1997; Richardson et al. 1999). The binding of GroES changes GroEL into a new conformation, in which most of the previously exposed hydrophobic residues of the apical domain are involved in inter-subunit contacts, and have been replaced on the interior wall of GroEL by mostly hydrophilic residues. The consequence is the movement of the substrate into the GroES–GroEL cavity (Xu et al. 1997; Horwich et al. 2007).

Since the mobile loop of GroES mediates GroES–GroEL interaction, a peptide containing the amino acid sequence of the mobile loop of GroES was used to improve the expression of a highly toxic protein Bax under overexpression of GroEL in *E.coli* when attached to the N-terminus of the gene (Donnelly et al. 2001).

The *amo45* gene encodes an α-amylase of the GH 57 family. When it was expressed in *E.coli*, only low amounts of protein were produced and most of it was found precipitated in inclusion bodies. To improve the protein solubility, we fused the mobile-loop encoding sequence of the co-chaperonin *groES* with the *amo45* N-terminal end and constructed an expression vector based on the L-rhamnose inducible regulatory system of *E.coli* (Wegerer et al. 2008). Here, we show that fusion with the mobile loop enhances the yield of soluble Amo45 protein.

Materials and methods

Strain and plasmids

The strain and plasmids used in this study are described in Table I.

Isolation of the amo45 gene

The *amo45* gene was isolated from a DNA sample obtained from a hot spring in Grensdalur located in the southern part of Iceland after enrichment by addition of starch (Labes et al. 2008). A PCR-based screening using degenerate primers targeted against sequences encoding conserved amino acid sequences for enzymes of GH family 57 was used to retrieve a fragment of the *amo45* gene. The primers (GH57-ECY-f, GH57-GWFF-r, Table II) were designed according to the CODEHOP strategy (Rose et al. 1998) as previously described for enzymes of the GH family13 CODEHOP primers (Labes et al. 2008; Turner et al. 2005). The PCR was carried out with TEG DNA polymerase (Matís production) with a PTC-0225 MJ Research thermal cycler as previously described (Labes et al. 2008). The resulting amplification products were separated on agarose gels and purified using GFX spin columns (GE Healthcare Life Sciences). The fragments with sizes of 800–1100 bps were selected and cloned into TOPO TA-sequencing vectors by the TA-cloning method (Invitrogen). Twenty-four clones from each band were sequenced with M13 forward and reverse primers on ABI 3700 DNA sequencers, using a BigDye Terminator cycle sequencing ready reaction kit (PE Applied Biosystems, Foster City, CA). Similarity searches using BLAST were performed on the NCBI server (http://www.ncbi.nlm.nih.gov) and

Table I. Strains and plasmids used in this study.

Strain or plasmid	Relevant features	Reference
Plasmids		
pGro7	*ara*P_{BAD} expression vector containing genes for GroES and GroEL, Cm^r	TaKaRa BIO INC.
pJOE5751.1	L-rhamnose-inducible expression vector containing *His*$_6$*-eGFP*, Amp^r	Wegerer et al. (2008)
pJOE3075	L-rhamnose-inducible expression vector, Amp^r	Stumpp et al. (2000)
pLEI45.1	expression vector containing *His*$_6$*-amo45*, Amp^r	This study
pLEI88.1	expression vector containing *mobile-loop-eGFP*, Amp^r	This study
pLEI90.1	expression vector containing *mobile-loop-amo45*, Amp^r	This study
Strain		
E.coli K12 JM109	F′ *traD36 proA*$^+$ *B*$^+$ *lacIq* Δ*(lacZ)M15*/Δ*(lac-proAB) glnV44e14*$^-$ *gyrA96 recA1 relA1 endA1 thi hsdR17*	Yanisch-Perron et al. (1985)
E.coli JM109/pLEI45.1	Expression strain for His$_6$-Amo45	This study
E.coli JM109/pLEI45.1/pGro7	Expression strain for His$_6$-Amo45 and GroEL-GroES	This study
E.coli JM109/pLEI90.1	Expression strain for mobile-loop-Amo45	This study
E.coli JM109/pLEI90.1/pGro7	Expression strain for mobile-loop-Amo45 and GroEL-GroES	This study

multiple sequence alignments were done using ClustalX software.

The method for the retrieval of full-length genes has been described (Labes et al. 2008; Turner et al. 2005). In short, the flanking regions, upstream and downstream the fragments obtained with the degenerate primers, were amplified from the corresponding genomic DNA in a series of nested PCR reactions. In the first reaction, one gene-specific, 5′-biotin-labeled primer and one arbitrary primer (Arb1 or Arb2), targeting the unknown flanking sequence were used and a nested gene-specific primer downstream of the previous one and a primer (Arb3) targeting the 5′ consensus sequence of the previously used arbitrary primer was used in the second PCR. Table II lists the primers used for the retrieval of the *amo45* gene. The resulting PCR products were cloned and sequenced as described above and the sequences were assembled with the existing *amo45* gene fragment.

Construction of expression vectors

Recombinant DNA techniques, for example, plasmid preparation, agarose gel electrophoresis were performed by conventional methods (Sambrook et al. 1989).

The plasmids for *amo45* expression were derived from the L-rhamnose inducible expression vector pJOE5751.1. The vector contains a single *Nde*I site between the *rha*P_{BAD} promoter and *His*$_6$*-eGFP* sequence, a single *Bam*HI site between *His*$_6$*-tag* and *eGFP* and a single *Hin*dIII site at the C-terminal end of *eGFP*. To construct the *mobile-loop-eGFP*

Table II. Primers used in this study.

Primer	Nucleotide sequences	Application
GH57-ECY-f	GAGGATCGCAAACAACTGCaartgytayyt	Screening
GH57-GWFF-r	GGTGCTCGCCGAAGaaraaccanwc	Screening
amo45-1bio	GGATCCAGAGGGAGTATGGTATGGTTG	Retrieval
amo45-3	TGGTGGCCATAGCTCCCCTTCCT	Retrieval
amo45-2bio	CCTGGAGCTGTCCTCACGGTGTGGAA	Retrieval
amo45-4	TGGAAAGATGGAGGGATCACTGTG	Retrieval
amo45-6-bio	TACACTCCCAAAGACATAGCCGTTACTG	Retrieval
amo45-8	CCTCATGTACCTTGCTGATGTGGAAA	Retrieval
amo45-10-bio	TCCTTGACGCAGAATGCTGATCCT	Retrieval
amo45-12	AACCGTGGAGGAGATTCTCA	Retrieval
Arb1	GGCCACGCGTCGACTAGTACNNNNNNNNNNGATAT	Retrieval
Arb2	GGCCACGCGTCGACTAGTACNNNNNNNNNNACGCC	Retrieval
Arb3	GGCCACGCGTCGACTAGTAC	Retrieval
S8481	TATGGAAGTTGAAACCAAATCTGCTGGTGGTATCGTTCTGACCGGTTCTGCTGCTGCGG	synthesis of mobile-loop
S8482	GATCCCGCAGCAGCAGAACCGGTCAGAACGATACCACCAGCAGATTTGGTTTCAACTTCCA	synthesis of mobile-loop
S7422	AAAAAATGATCAATAATATTTCACGGTCACTTTTA	PCR for *amo45*
S7423	AAAAAAAAGCTTTCATAGTGTGGC	PCR for *amo45*

plasmid pLEI88.1, the oligonucleotides encoding the amino acid sequence (EVETKSAGGIVLTG-SAAA) of the *E.coli* GroES mobile-loop and its complement were synthesized by Eurofins MWG Operon with codon optimization for *E.coli* (Donnelly et al. 2001). The oligonucleotides are described in Table II. Restriction sites for *Nde*I and a *Bam*HI up- and downstream the mobile-loop sequence were used for integration of the complementary oligonucleotides into pJOE5751.1. Hereby the His_6*-tag* sequence was deleted and replaced with the mobile-loop sequence. The host strain for cloning and transformation was *E.coli* JM109. The sequence of the constructed plasmid pLEI88.1, containing *mobile-loop-eGFP*, was confirmed by DNA sequencing.

To construct the plasmids for His_6*-amo45* and *mobile-loop-amo45* expression, the vectors pJOE5751.1 (His_6*-eGFP*) and pLEI88.1 (*mobile-loop-eGFP*) were cut with the restriction enzymes *Bam*HI and *Hin*dIII to eliminate the *eGFP* gene. The *amo45* gene was amplified by PCR with Phusion Hot Start II High-Fidelity DNA Polymerase (Thermo Scientific) using two primers. A *Bcl*I and a *Hin*dIII digestion site in the primer sequences enabled the insertion of the *amo45* PCR product (2030 bp) into pJOE5751.1 and pLEI88.1. The new plasmids for expression of His_6*-amo45* and *mobile-loop-amo45* were designated pLEI45.1 and pLEI90.1, respectively.

Gene expression

Overnight cultures of *E.coli* JM109/pLEI45.1, JM109/pLEI45.1/pGro7, JM109/pLEI90.1, and JM109 pLEI90.1/pGro7 were diluted 100-fold in 10 ml LB (Lysogeny Broth) medium containing 100 μg/ml^{-1} ampicillin and grown at 37°C to an optical density (OD_{600}) of 0.4. Induction of *amo45* expression was achieved by adding L-rhamnose. Expression of *groEL groES* genes under control of the arabinose promoter $araP_{BAD}$ on plasmid pGro7 was achieved by addition of L-arabinose. Induction was done in two different ways, simultaneously by adding both sugars at the same time and in a successive way by adding first arabinose and later rhamnose. In the simultaneous protocol, the overnight culture was diluted 100-fold in LB containing ampicillin and 25 μg ml^{-1} chloramphenicol and at an OD_{600} of 0.4, 0.2% rhamnose and 0.05% L-arabinose were added at once and the cells were further incubated at 30°C for 4.5 h. In the successive induction protocol 0.1% arabinose was added at OD_{600} of 0.3 and the cells incubated for 3 h at 30°C. The cells were harvested, washed by centrifugation, and diluted to OD_{600} of 0.4. Then 0.2% rhamnose was added and the cells were incubated for another 4.5 h. For better comparison, JM109/pLEI45.1 and pLEI90.1 without pGro7 were treated with arabinose and rhamnose according to the successive induction protocol.

Finally, the cells were harvested by centrifugation and cell pellets corresponding to 10 OD_{600} units were suspended in 1 ml 0.1 mM potassium buffer (pH 6.5) and lysed by ultrasonication. The lysate was centrifuged for 15 min at 16,000 × g in a bench top centrifuge. The supernatant represented the soluble protein crude extract, while the precipitate, resuspended in the same buffer, represented the insoluble protein fraction.

SDS-PAGE and enzyme assay

The soluble and insoluble protein extracts were qualitatively analyzed by 10% SDS-PAGE. The protein concentration was determined with the Bio-Rad protein assay by measuring the absorbance at 595 nm (A_{595}) with a spectrophotometer (Pharmacia Biotech) using bovine serum albumin as standard. The enzyme activity of Amo45 was determined by measuring the absorbance at A_{405} resulting from the cleavage of the α-glucosidic linkage of 2-chloro-4-nitrophenyl-α-D-maltotrioside (CNP-G3) by Amo45 and the release of the chromophore (2-chloro-4-nitrophenol) (Winn-Deen et al. 1988; Lorentz 1999). The reaction mixture consisted of 450 μl 0.1 mM potassium buffer (pH6.5), 25 μl enzyme sample containing Amo45 and 25 μl 10 mM CNP-G3 as substrate. The sample and buffer were incubated at first for 5 min at 85°C and then the substrate was added to start the reaction. The reaction was stopped by adding 1 ml 0.4 M Na-borate (pH9.8). The enzyme activity (U mg^{-1} protein) was defined as the release of 1 μM chromophore per min at 85°C using extinction coefficients of $\varepsilon_{405,pH10}$ 17×10^3 M^{-1} cm^{-1} for 2-chloro-4-nitrophenol.

Results

Retrieval of the amo45 gene

A set of degenerate CODEHOP primers targeted against sequences encoding conserved regions in enzymes of the GH family 57 was used to amplify gene fragments from environmental DNA prepared from hot spring samples (65°C, pH6) enriched with starch. BlastX analysis of one of the sequenced fragments indicated similarities with a protein from the hyperthermophilic bacterium *Thermocrinis albus,* annotated as GH of family 57. The complete gene, designated *amo45*, was retrieved in one step for the 5' part and three steps for the 3'part. The predicted protein, Amo45 showed a high degree of homology (98% identity, eight amino acid residues exchanges)

to the uncharacterized GH family protein of *Thermoocrinus albus* DSM14484 (Genbank accession number YP_003474157.1) which was annotated during genome sequencing of *Thermocrinus albus* DSM 14484. NCBI conserved domain search in both protein sequences revealed for the amino acid residues 1–288 an N-terminal putative catalytic domain of the GH family 57 (cd10797), followed by a domain of unknown function (DUF3536) (spanning amino acid residues 296–483) and an extended C-terminal region (amino acid residues 483–673) with no hits for conserved domains. Within the N-terminal putative GH57 catalytic domain of both proteins three amino acid exchanges were found whereas the catalytic amino acids His5, His7, and Glu139 were conserved in both proteins.

Cloning and expression of the amo45 gene

In a first attempt at heterologous expression of the native *amo45* in *E.coli,* the gene was inserted into the L-rhamnose inducible expression vector pJOE3075 (Stumpp et al. 2000). The protein production in *E.coli* JM109 was low and most of the protein was found in the insoluble fraction. Even expression at lower temperatures down to 22°C and/or using *E.coli* TOP10 as host did not prevent the formation of inclusion bodies (data not shown). To improve the expression, the gene was inserted in another L-rhamnose inducible expression vector pJOE5751.1 and thereby N-terminally fused to a *His$_6$-tag* sequence of the vector to allow His$_6$-Amo45 production. Rhamnose induction of JM109/pLEI45.1, harboring *His$_6$-amo45,* resulted in amounts of protein which were easily detected on SDS-PAGE. Again, most protein was in the insoluble fraction (Figure 1). The amount of soluble enzyme was sufficient to test its amylolytic activity. No activity was detected for the tested polymeric substrates starch, pullulan and amylopectin (data not shown) and only cleavage of the α-glucosidic linkage of 2-chloro-4-nitrophenyl-α-D-maltotrioside (CNP-G3) was found. Preliminary characterization of the His$_6$-Amo45 enzyme revealed a temperature optimum at 85°C. Crude extract of JM109/pLei45.1 exhibited a specific enzyme activity of His$_6$-Amo45 of about 12 mU mg^{-1} (Figure 3). To improve the solubility of His$_6$-Amo45 the effect of coexpression of chaperonin GroES–GroEL on the folding of His$_6$-Amo45 was studied. The genes *groEL* and *groES* were provided on the L-arabinose inducible vector pGro7. When the cells of JM109/pLEI45.1/pGro7 were simultaneously induced with rhamnose and arabinose, the specific enzyme activity of His$_6$-Amo45 in crude cell extract increased slightly to 15 mU mg^{-1}. A prominent protein band at a size of approximately 57 kDa on SDS-PAGE indicated a dominant production of GroEL (Figure 1). To exclude a competitive expression of *His$_6$-amo45* and chaperones, the pGro7 was first induced for 3 h

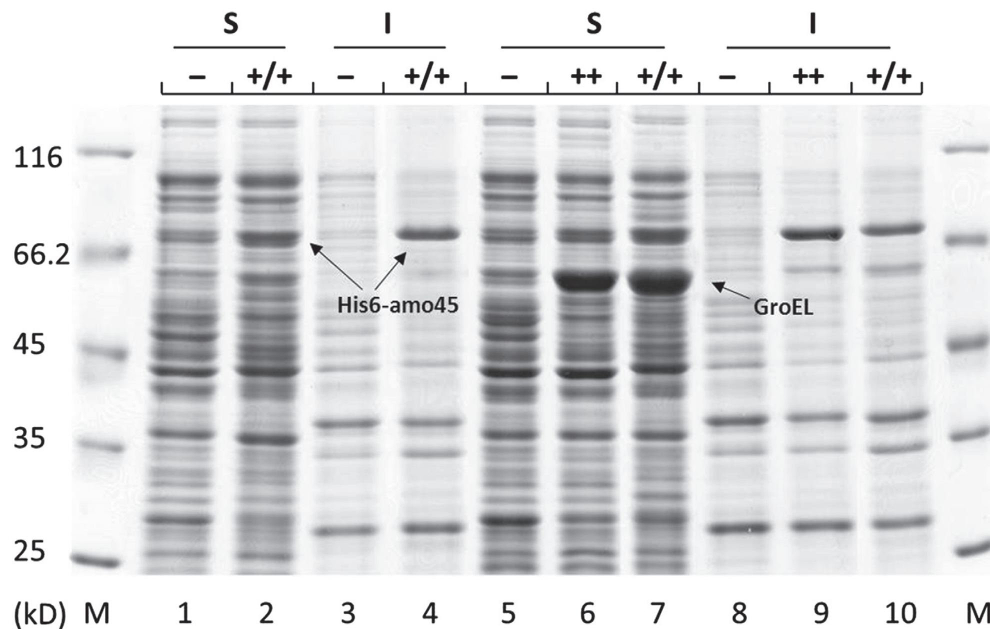

Figure 1. Expression of *His$_6$-amo45* with and without co-expression of *groEL–groES*, analyzed by SDS-PAGE. Lanes 1–4: Crude extract of JM109 pLEI45.1 (*His$_6$-amo45*). Lanes 5–10: Crude extract of JM109 pLEI45.1/pGro7. The −, ++ and +/+ indicate uninduced, with arabinose and rhamnose simultaneously induced (4.5 h), and successively induced (3 h + 4.5 h) cell cultures, respectively. The M, S, and I labeled the molecular weight marker, soluble, and insoluble fractions, respectively.

before the His_6-Amo45 production was started (see Material and Methods). This resulted in a much less-dominant protein band at 57 kDa. On the other hand, the chaperone co-expression increased the His_6-Amo45 enzyme activity 4-fold to 46 mU mg^{-1} in the crude cell extract compared to the absence of GroEL–GroES overproduction. Obviously, a higher amount of GroEL–GroES in the cells improved the folding of His_6-Amo45.

Construction and expression of mobile-loop-amo45

The *amo45* gene was introduced into the constructed expression vector pLEI88.1 to replace the *eGFP* and to generate pLEI90.1 (*mobile-loop-amo45*). Induction of JM109 pLEI90.1 with rhamnose resulted in a specific enzyme activity of about 60 mU mg^{-1}. This is in accordance with SDS-PAGE analysis (Figure 2) which showed a lower amount of insoluble protein. Unfortunately a higher amount of soluble protein could not be demonstrated due to overlap with an *E.coli* protein. By simultaneous expressing of *groEL–groES* and *mobile-loop-amo45* in JM109/pLEI90.1/pGro7, the specific enzyme activity decreased to 35 mU mg^{-1} (Figure 3). It was assumed that this effect might be due to the dominant chaperone production (Figure 2). Therefore, *groEL–groES* and *mobile-loop-amo45* were induced one after the other and the specific activity returned to 62 mU mg^{-1}. Thus, under these conditions additional expression of *groEL–groES* had no beneficial effect on the folding of mobile-loop-Amo45. Additionally, all of the non-induced samples showed a specific enzyme activity of about 1–2 mU mg^{-1} and clearly indicated that the production of the Amo45-variants only occurred after induction with rhamnose.

Discussion

Metagenomic techniques opened the way to the vast and unexploited gene pool of uncultured microorganisms. Nevertheless, the process from identifying a gene in environmental habitats to the production of highly active enzymes is quite often long and tedious. Hurdles to overcome are incomplete genes, DNA sequencing errors, erroneous PCR, wrongly determined translational starts, low expression levels, lack of expected enzyme activities deduced from sequence comparisons, lack of special cofactors and last but not least, the correct folding of the protein during overexpression in the new host. Some of the problems can be solved by providing a strong promoter and optimized translation initiation region and maybe by resynthesizing the gene optimized for the host codon usage. In the case of *amo45,* protein production was considerably increased in *E.coli* by providing the rhamnose-inducible promoter and an optimized translation initiation region, which was achieved by fusion of *amo45* with the *His_6-tag* sequence of the vector. To improve the correct folding of the protein, genes encoding various chaperones were co-expressed. No positive effect was seen with DnaK, DnaJ, and GrpE (data not shown) in contrast to GroEL–

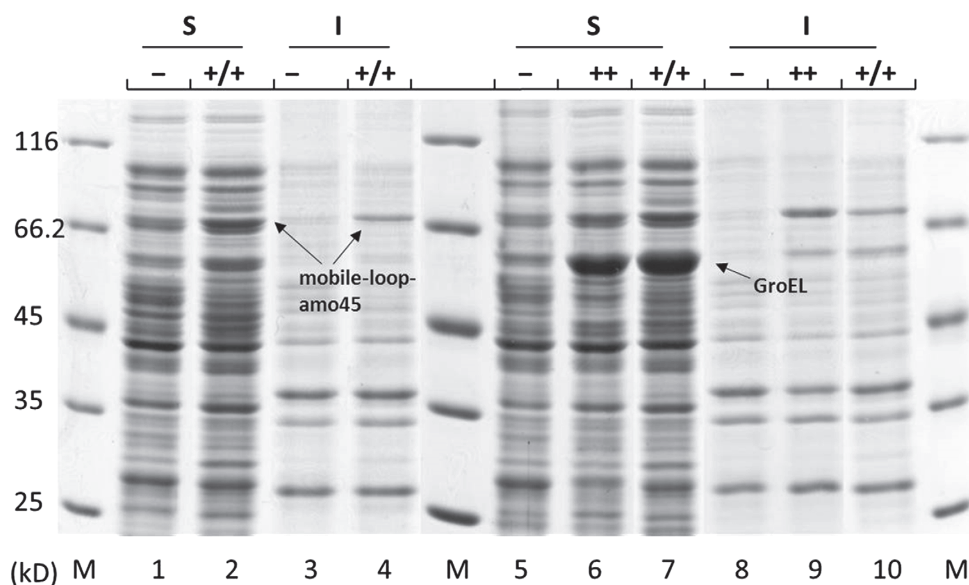

Figure 2. Expression of *mobile-loop-amo45* with and without co-expression of *groEL-groES*, analyzed by SDS-PAGE. Lanes 1–4: Crude extract of JM109 pLEI90.1 (*mobile-loop-amo45*). Lanes 5–10: Crude extract of JM109 pLEI90.1 pGro7. The −, ++ and +/+ indicate uninduced, with arabinose and rhamnose simultaneously induced (4.5 h), and successively induced (3 h + 4.5 h) cell cultures, respectively. The M, S, and I labeled the molecular weight marker, soluble and insoluble fractions, respectively.

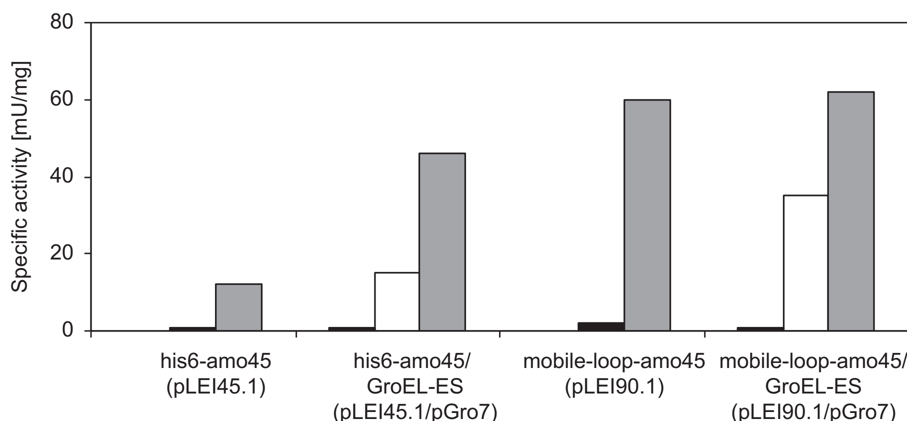

Figure 3. Average specific enzyme activity of soluble crude extract samples containing His_6-Amo45 or mobile-loop-Amo45 measured by using synthetic substrate CNP-G3. The black, white, and gray columns present the specific activity of uninduced, with arabinose and rhamnose simultaneously induced (4.5 h), and successively induced (3 h + 4.5 h) cell cultures, respectively.

GroES, which clearly assisted in the heterologous soluble expression of *His_6-amo45*. However, it was important to induce *groEL–groES* and *His_6-amo45* expression in a two-step protocol and not simultaneously. Obviously, the extremely high expression rate of the chaperones is in competition to the expression of *His_6-amo45* and reduces optimal protein production.

Instead of co-expression of the *groEL–groES* genes we fused the mobile-loop sequence originating from *groES* to the N-terminal end of *amo45*. Despite the different N-terminal regions of *mobile-loop-amo45* and *His_6-amo45* the overall Amo45 protein production was about the same. On the other hand, this new N-terminal extension increased the amount of soluble Amo45 by 5-fold relative to that of His_6-tagged Amo45, when expressed under the same conditions. Most likely, the mobile loop, originating from GroES, increased the affinity of nascent peptide sequence of Amo45 for GroEL and hence the assistance of folding of Amo45 by GroEL improved. However, as pointed out by Donnelly et al. (2001) it seems also possible that this peptide extension facilitates initial steps in the folding process. Previously it had been shown that the mobile loop of GroES improved the expression of a gene encoding the highly toxic protein Bax when *groEL–groES* was expressed simultaneously (Donnelly et al. 2001). Without co-expression of *groEL–groES*, the mobile-loop-Bax production increased just slightly. This is in contrast to our results, where simultaneous co-expression with chaperones had a negative effect on the *mobile-loop-amo45* expression. However expression of *His_6-amo45* under successive co-expression of *groEL–groES* increased the amount of soluble protein. Obviously, under the conditions used here for the *mobile-loop-amo45* expression, there is already sufficient GroEL–GroES in the cell to fully support folding of Amo45 and any high production of additional proteins is detrimental.

Declaration of interest: The authors report no declarations of interest. The authors alone are responsible for the content and writing of the paper.

The work was supported by the EU FP7 grant 265992 AMYLOMICS.

References

Altenbuchner J, Mattes R. 2005. *Escherichia coli*. In: Gellissen, G. ed. Production of recombinant proteins. Weinheim: Wiley-VCH, pp 7–43.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res 37:D233–238.

Donnelly MI, Stevens PW, Stols L, Su SX, Tollaksen S, Giometti C, Joachimiak A. 2001. Expression of a highly toxic protein, Bax, in *Escherichia coli* by attachment of a leader peptide derived from the GroES cochaperone. Protein Expr Purif 22:422–429.

Grantham R, Gautier C, Gouy M, Mercier R, Pave A. 1980. Codon catalog usage and the genome hypothesis. Nucleic Acids Res 8:r49–r62.

Hartl FU. 1996. Molecular chaperones in cellular protein folding. Nature 381:571–579.

Hartl FU, Hayer-Hartl M. 2002. Molecular chaperones in the cytosol: from nascent chain to folded protein. Science 295:1852–1858.

Horwich AL, Fenton WA, Chapman E, Farr GW. 2007. Two families of chaperonin: physiology and mechanism. Annu Rev Cell Dev Biol 23:115–145.

Ikemura T. 1981. Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system. J Mol Biol 151:389–409.

Labes A, Karlsson EN, Fridjonsson OH, Turner P, Hreggvidson GO, Kristjansson JK, Holst O, Schonheit P. 2008.

Novel members of glycoside hydrolase family 13 derived from environmental DNA. Appl Environ Microbiol 74: 1914–1921.

Landry SJ, Zeilstra-Ryalls J, Fayet O, Georgopoulos C, Gierasch LM. 1993. Characterization of a functionally important mobile domain of GroES. Nature 364:255–258.

Lorentz K, Gutschow B, Renner F. 1999. Evaluation of a direct alpha-amylase assay using 2-chloro-4-nitrophenyl-alpha-D-maltotrioside. Clin Chem Lab Med 37:1053–1062.

Nishihara K, Kanemori M, Kitagawa M, Yanagi H, Yura T. 1998. Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*. Appl Environ Microbiol 64: 1694–1699.

Palomo M, Pijning T, Booiman T, Dobruchowska JM, van der Vlist J, Kralj S, Planas A, Loos K, Kamerling JP, Dijkstra BW, van der Maarel MJ, Dijkhuizen L, Leemhuis H. 2010. Thermus thermophilus glycoside hydrolase family 57 branching enzyme: crystal structure, mechanism of action, and products formed. J Biol Chem 286:3520–3530.

Richardson A, Landry SJ, Georgopoulos C. 1998. The ins and outs of a molecular chaperone machine. Trends Biochem Sci 23:138–143.

Richardson A, van der Vies SM, Keppel F, Taher A, Landry SJ, Georgopoulos C. 1999. Compensatory changes in GroEL/Gp31 affinity as a mechanism for allele-specific genetic interaction. J Biol Chem 274:52–58.

Rose TM, Schultz ER, Henikoff JG, Pietrokovski S, McCallum CM, Henikoff S. 1998. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. Nucleic Acids Res. 26:1628–1635.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

Sigler PB, Xu Z, Rye HS, Burston SG, Fenton WA, Horwich AL. 1998. Structure and function in GroEL-mediated protein folding. Annu Rev Biochem 67:581–608.

Stumpp T, Wilms B, Altenbuchner J. 2000. Ein neues, L-Rhamnose-induzierbares Expressionssystem für *Escherichia coli*. Biospectrum 6:33–36.

Thomas JG, Ayling A, Baneyx F. 1997. Molecular chaperones, folding catalysts, and the recovery of active recombinant proteins from *E. coli*. To fold or to refold. Appl Biochem Biotechnol 66:197–238.

Turner P, Labes A, Fridjonsson OH, Hreggvidson GO, Schonheit P, Kristjansson JK, Holst O, Karlsson EN. 2005. Two novel cyclodextrin-degrading enzymes isolated from thermophilic bacteria have similar domain structures but differ in oligomeric state and activity profile. J Biosci Bioeng 100:380–390.

Waugh DS. 2005. Making the most of affinity tags. Trends Biotechnol 23:316–320.

Wegerer A, Sun T, Altenbuchner J. 2008. Optimization of an *E. coli* L-rhamnose-inducible expression vector: test of various genetic module combinations. BMC Biotechnol 8:2.

Winn-Deen ES, David H, Sigler G, Chavez R. 1988. Development of a direct assay for alpha-amylase. Clin Chem 34:2005–2008.

Xu Z, Horwich AL, Sigler PB. 1997. The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. Nature 388:741–750.

Yanisch-Perron C, Vieira J, Messing J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119.