UPPSALA
UNIVERSITET

UPTEC X 13 018
Examensarbete 30 hp
Februari 2014

Process development for the production of a therapeutic Affibody® Molecule

Belinda Fridman

UPPSALA UNIVERSITET

Molecular Biotechnology Programme

Uppsala University School of Engineering

UPTEC X 13 018	Date of issue 2014-02

Author
Belinda Fridman

Title (English)
Process development for the production of a therapeutic Affibody® Molecule

Title (Swedish)
Processutveckling för att tillverka en Affibody®-molekyl avsedd för cancerterapi

Recently HER3, member of the epidermal growth factor receptor family (EGFR), has been found to play a crucial role in the development of resistance towards inhibitors that are given to patients with HER1- and HER2-driven cancers. As HER3 is up-regulated or over-activated in several types of human cancers, it is of outmost importance that new innovative drugs target its oncologic activity.

The Affibody® Molecule Z08698 inhibits the heregulin induced signalling of HER3 with high affinity (K_D~50 pM). As the Affibody® Molecule is small, has high solubility and outstanding folding kinetics, an effective penetration of tumour tissue is suggested together with a rationalized manufacturing process. Further coupling to an albumin binding domain (ABD) expands the plasma half-life of the molecule, hence increasing the molecule's potential of serving as a therapeutic.

A process development for production of Z08698-VDGS-ABD094 has been established, where the molecule is efficiently produced in the *E. coli* host strain BL21(DE3), through a T7 based expression system. Cultivations were performed with a fed-batch fermentation process and the conditions were further optimized in order to obtain highest expression, while avoiding undesirable modifications like gluconoylations. By employing Design of experiments in combination with multivariate data analysis, a production process resulting in ~3.5 g product/ l culture could be verified. Moreover, thermolysis was evaluated as a suitable method for cell disruption, enabling an easy and cost-effective manufacturing process of the ABD fused Affibody® Molecule.

Keywords
Affibody® Molecule, HER3, EGFR, HER2, cancer therapy, process development, DOE, multivariate data analysis, *E. coli,* ABD, plasma half-life, thermolysis, fed-batch, gluconoylation, BL21(DE3), T7 expression system

Supervisors
Finn Dunås, Affibody AB (Solna, Sweden)

Scientific reviewer
Karin Stensjö, Uppsala University

Project name	Sponsors
Language **English**	Security
ISSN 1401-2138	Classification
Supplementary bibliographical information	Pages **60**

Biology Education Centre Biomedical Center Husargatan 3 Uppsala
Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 471 4687

Process development for the production of a therapeutic Affibody® Molecule

Belinda Fridman

Populärvetenskaplig sammanfattning

På cellytan förekommer olika typer av tillväxtfaktorreceptorer som genom ett komplext signalsystem reglerar viktiga funktioner inklusive cellulär överlevnad, differentiering, migration och proliferation. Ett överuttryck eller en överaktivering av dessa typer av receptorer, kan leda till onormal celltillväxt och utveckling av cancer. Ett exempel på en receptor är HER3, som visat sig vara överuttryckt i flera av de vanligaste cancertyperna såsom bröst-, livmoderhals- och prostatacancer.

Tillsammans med Uppsala universitet och Kungliga tekniska högskolan har Affibody AB tagit fram Affibody® -molekylen Z08698, som blockerar signaleringen av HER3 med hög affinitet. Molekylen har fördelen att vara liten, vilket möjliggör effektiv penetrering av tumörvävnad, samt har en enkel karaktär, som tillåter en kostnadseffektiv storskalig produktion. Ytterligare fördelar uppnås då Affibody® -molekylen fuseras till en albumin-bindande domän, och på så vis erhåller en förlängd halveringstid i kroppen. Då fusionsproteinet ämnas ingå i kliniska studier för utvärdering som läkemedelskandidat för cancerterapi, är det av intresse att en lämplig tillverkningsprocess färdigställs.

I denna avhandling redovisas ett adekvat sätt att producera läkemedelskandidaten Z08698-VDGS-ABD094. Fusionsproteinet har uttryckts i *Escherichia coli*, där en enkel värmebehandling följt av affinitetskromatografi utförs, med avsikten att extrahera respektive specifikt rena fram produkten. Då odlingsbetingelserna för värdcellen föreslås kunna ge förändringar i uttryck av produkten, med avseende på såväl kvalitet som kvantitet, har påverkande faktorer under själva odlingsprocessen identifierats. Slutligen har uttrycket av produkten maximerats med hjälp av experimentell försöksplanering, där de verifierande odlingarna påvisade en höjning av uttrycksnivån med på ca 25 %, samtidigt som proceduren kunde kortas ner med sex timmar. Den framtagna tillverkningsprocessen för Z08698-VDGS-ABD094 presenterar således ett effektivt förfarande, där en storskalig produktion av den potentiella läkemedelskandidaten möjliggörs och förhoppningsvis banar väg för framsteg inom behandlingen av cancer.

Examensarbete 30 hp
Civilingenjörsprogrammet Molekylär bioteknik
Uppsala universitet, februari 2014

List of content

Abbreviations

6-PGLac	6-phospogluconolactone
A280	Absorbance at 280 nm
ABD	Albumin binding domain
ANOVA	Analysis of variance
CCF	Central composite fractional
CIP	Cleaning in place
DO	Dissolved oxygen
DOE	Design of experiments
DTT	Dithiothreitol
EGFR	Epidermal growth factor receptor
E-time	Expression time
HAc	Acetic acid
HCDC	High cell-density culture
HER	Human epidermal growth factor receptor
HPLC-MS	High pressure liquid chromatography mass spectrometry
HSA	Human serum albumin
IPTG	Isopropyl β-D-1-thiogalactopyranoside
I-time	Induction time
kDa	kilo Dalton
LB	Lysogeny broth
MetAP	Methionine aminopeptidase
MLP	Multiple linear regression
OD	Optical density
ORI	Origin of replication
PGL	Phospogluconolactonase
PLS	Partial least squares
pM	picomolar
PTM	Post translational modification
RCB	Research cell bank
rpm	Rates per minute
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SpG	Streptococcal protein G
TFA	Trifluoroacetic acid
WW	Wet weight

1. Introduction

1.1 Objective

1.1.1 Background to the project

Lately the human epidermal growth factor receptor 3 (HER3) has gained interest being a novel therapeutic target in human cancers. Affibody AB together with the Royal Institute of Technology (KTH, Stockholm) and Uppsala University has developed an Affibody® Molecule that with high affinity binds to HER3 and inhibits the heregulin induced signalling, thus serving as an anti-proliferative agent [1].

In comparison with the conventional antibody treatment against the receptors of the epidermal growth factor receptor (EGFR) family, the Affibody® Molecule is small, suggesting an effective penetration of tumour tissue. Further it is cysteine-free, has high solubility and impressive re-folding kinetics enabling a simplified manufacturing process [2]. As the molecule is planned to be evaluated as cancer treatment, a suitable production process for the potential therapeutic is of interest.

1.1.2 Project goals

This project aimed to develop a process for the production of the HER3- specific Affibody® Molecule Z08698-VDGS-ABD094. Recombinant expression of the molecule was carried out in *Escherichia coli,* which were cultivated to a high cell density by utilizing the conventional fed-batch fermentation process. Moreover, a suitable process analysis was established, including disintegration of cells, purification and analysis techniques in order to assess quality and quantity of the expressed protein.

By employing design of experiments (DOE) together with multivariate data analysis the cultivation protocol was further optimized with the aim of receiving expression levels of ≥2 g/l culture. Finally, the optimization was verified in a large-scale production with a working volume of 15 l.

1.2 The therapeutic target

The tyrosine kinases of the EGFR-family i.e. HER1, HER2, HER3 and HER4, are acting as regulators of cell survival, proliferation, differentiation and migration. Unfortunately, the transmembrane receptors have found to be overexpressed in many types of human cancers, and are recognized as key players driving abnormal cell growth. Consequently, they serve as important targets when producing innovative cancer therapeutics [2].

HER3 differs from the rest of the EGFRs as it possesses an inactive tyrosine kinase domain. However, the receptor has been found to dimerize with HER2, thus forming a potent signalling and oncogenic unit that is found in many HER2-driven breast cancers [1 - 3].

Moreover, oncogenic expression of HER3 has been found in ovarian, bladder, prostate and lung cancers [2].

Studies have shown that in many cases of developed resistance against HER2-specific inhibitors, HER3 plays a significant role in the progression of the disease. The continued proliferative characteristics are often explained either by upregulation of heregulin or overexpression, enhanced cell surface localisation or over-active phosphorylation of HER3 [3].

1.3 Affibody® Molecules

Antibodies for a long time have served as the main affinity protein used for life science applications. Limitations with antibodies include large size and dependence on complex patterns for glycosylation and folding, constituting of inconvenient disulphide bonds. This further suggests problems with stability and thus much more complex and expensive manufacturing processes are required. As only a small part of the antibody is utilized for antigen recognition, the same amount of affinity is suggested to be obtained with a much smaller candidate [4].

The cysteine-free and approximately 6.5 kDa small Affibody® Molecule belongs to the next generation of affinity proteins, further possessing a high stability and solubility [4]. Production of Affibody® Molecules is either performed by chemical peptide synthesis or by recombinant expression in *E. coli.* Due to its simple formation a cost effective production process is provided [5].

Originally the molecule is derived from the B-domain of the immunoglobulin binding region of the *Staphylococcal* protein A. Modifications of this domain has enabled chemically stable scaffold protein denoted the Z-domain, constituting of 58 amino acids forming a three helical bundle. Based on this structure a combinatorial randomization of 13 amino acids, positioned at helices one and two, theoretically enables targeting of any desired target protein [4]. Typically the selection of new Affibody® Molecules is carried out by displaying up to 10^{10} different variants on bacteriophages. This is a method that has proven to generate specific binders with high affinity for their targets [5]. Moreover, the Affibody® Molecules may be modified and serve as excellent scaffolds when it comes to conjugation and fusions with other molecules [1].

Several applications are suggested for the Affibody® Molecules for example for molecular imaging where it, coupled to a radionuclide, can serve as a detector of certain oncogenic cell surface receptors. This enables a powerful diagnostic tool that may further be valuable in the stratifications of patients for targeted therapy. Therapeutic Affibody® Molecules are also suggested either by direct function, acting as competitive inhibitors of cell surface receptors [1, 4] or indirect by being coupled to payload constituting of therapeutic radionuclides like ^{90}Y, ^{122}Lu, ^{131}I or ^{186}Re or small toxic protein domains. Further engineering of Affibody® Molecules, comprising of fused proteins are application dependent. Molecular

imaging requires a final product with rapid biodistribution and short plasma half-life, while therapeutic applications favour final products with long plasma half-life [5].

1.3.1 The HER3 specific Z08698

The Affibody® Molecule Z08698 targets HER3 and has the potential of being used in both therapy and diagnostics. During the construction of an affinity-matured library, this molecule showed a significantly higher affinity (K_D~50 pM) against the target, compared with the original binder. The high affinity of the molecule is desired as this feature is favourable in high-contrast molecular imaging when target expression is relatively low (~10^4 target proteins per cell), as in the case with HER3. Moreover, a high affinity also generally generates improved therapeutic efficacy [1].

In this study, the anti-HER3 Affibody® Molecule has been fused with an albumin-binding domain (ABD), in order to increase the plasma half-life (Illustration in **Fig.1**) [5]. Up to date several successful fusions with ABD have been performed, without reducing the bioactivity of the fusion partner. The native properties of serum albumin as being extravasive and accumulating in both tumour and inflammatory tissue, rather suggests favourable biodistribution of the conjugated biopharmaceutical.

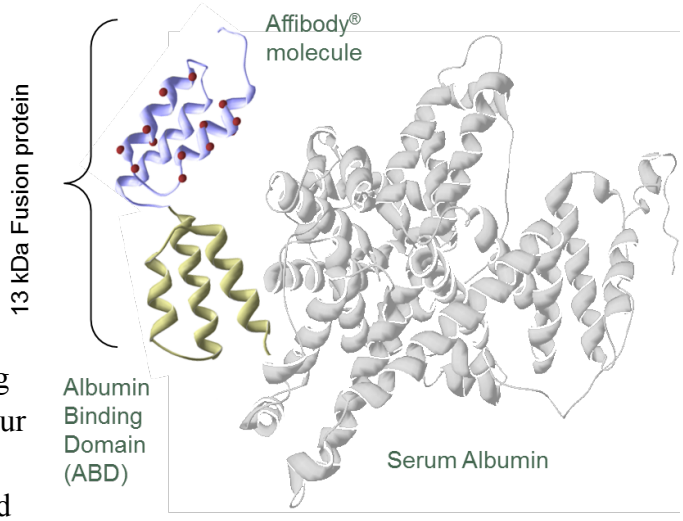


Figure 1.The Affibody® Molecule coupled to an ABD. The red dots symbolize the 13 amino acids on helices one and two that are simultaneously randomized in order to find successful binders.

The ABD has been derived from *Streptococcal* protein G (SpG), a receptor located on the bacterial cell surface, able to bind both immunoglobulin and serum albumin of various species. Originating from SpG, the smallest albumin binding unit comprising of 46 amino acids (~5 kDa) was retained and further modified for improvements [5]. After a combinatorial randomization of several amino acids of the sequence, ABD035 was found, and was shown to possess even higher affinity against human serum albumin (HSA). In order to diminish T-cell epitopes located on the molecule, deimmunization programs have been assigned, which resulted in the currently clinically evaluated ABD094 [5,6].

1.4 Process development

When producing recombinant proteins for therapy, the process development is of great importance. Steps during the process like choice of; expression system, mode of operation concerning fermentation and purification technique, must be critically considered in order to

receive a stable product with its full potency. Small changes in the manufacturing process may result in modifications, misfolding and aggregation of the product, that further can alter its clinical effect in terms of immunogenicity or activity [7].

1.4.1 Expression system

For recombinant production of a protein the host is required to possess a suitable genetic background, in combination with the plasmid that contains the target gene [8]. *E. coli* represents the most commonly used host system, especially when synthesizing therapeutic proteins without required post translational modifications [9]. The bacterial cells have the advantage of being cultivated into a high cell density, thus enabling a cost-effective product production [10]. Moreover, the genetics of *E. coli* are well-known and evolvements of plasmids, mutant strains and recombinant fusion partners have advanced its capability of expressing recombinant proteins [8].

In order for a functional expression system to be established, the host strain should stably maintain the incorporated plasmid, be deficient in harmful proteases and provide for genetic elements that are of importance for the desired expression. The most well-known *E. coli* based strain is the T7 based pET expression system (Novagen) illustrated in **Fig.2** [8]. The T7 promoter is not naturally found in *E. coli* and the corresponding polymerase holds a high specificity in combination with a transcription rate up to five times faster than the native RNA polymerase, which suggest that the T7 based pET expression system is in fact selective and efficient [8, 11].

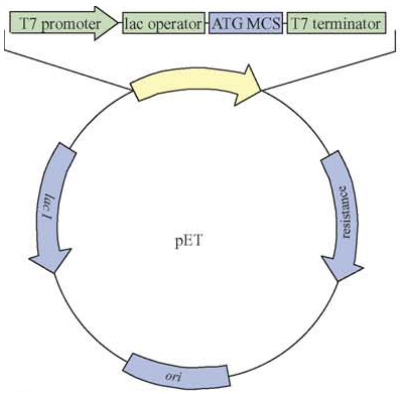
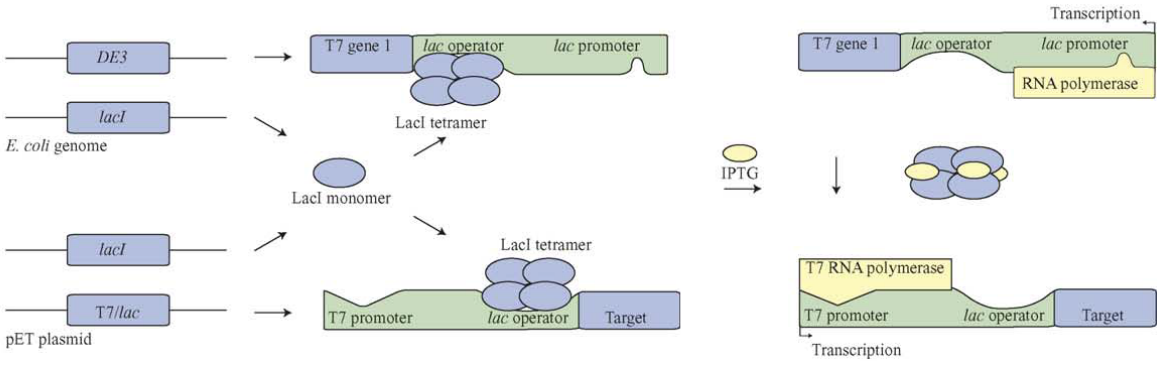


Figure 2. *T7 based pET expression system*. Expression of the target gene, located on pET-plasmid, is controlled by a T7/lac promoter. LacI (expressed from both plasmid and host genome) represses this hybrid promoter together with the lacUV5 promoter, where the latter controls the expression of the T7 RNA polymerase from the *E. coli* genome. Upon induction with IPTG, the LacI monomers are released and T7 RNA polymerase is transcribed which in turn starts the machinery expressing the target gene [8].



For simplicity there are several pET plasmids commercially available that already contain important genetic elements like; origin of replication (ORI), *lacI* gene, resistance marker, hybrid promoters and multiple cloning sites. As depicted in **Fig.2**, the system requires a host with a lysogenized DE3 phage fragment. This is conveniently achieved by using the conventionally used *E. coli* strain BL21(DE3) [8]. This strain originates from the B strain BL21, which has further been modified by incorporation of the bacteriophage T7 gene1 encoding the T7 RNA polymerase, into the chromosomal DNA [11]. Moreover the original BL21 possesses valuable qualities like being robust, easy to grow, and non-pathogenic in the sense that it unlikely will survive and cause disease in a host [8].

1.4.2 Cultivation

As the productivity of cultivation, in terms of product yield, is proportional to the final cell-density, high cell-density cultures (HCDC) are of interest. For this purpose a fed-batch mode of operation is proposed during the fermentation process [10].

Fed-batch cultivation is cheap, simple and allows the culture to grow to its full potential. Here, the cells are supplied with one or more nutrients according to a controlled feeding regime, resulting in substrate limiting growth [12]. During the process it is critical that the cells are neither overfed nor underfed, as this may be harmful for the cell growth and product formation [13]. It has been shown that as nutrients of the cultivation media are represented above a certain concentration threshold, including glucose and ammonia, the cell growth is inhibited. Thus explaining why not only addition of nutrients to the start medium increases the cell density during the simple batch mode. A problem with cultivation to HCDC is the potential acetate formation, which may be caused by anaerobic or oxygen-limited conditions as well as an excessive glucose concentration. Acetate concentrations above 5 g/l at pH 7 are directly inhibiting cell growth and product formation, hypothetically by repressing synthesis of DNA, RNA, proteins and lipids [10, 13]. By using a carbon source like glucose as the limited substrate during the fermentation process, acetate formation may be avoided in the same time as the growth rate of the cells is conveniently regulated [10].

During cultivation to high cell density cultures, the dissolved oxygen (DO) typically becomes limiting. In order to battle this problem stirrer speed can be increased. Additionally pressurized conditions are of favour for the oxygen transfer, or in some cases even pure oxygen can be supplemented to the culture [10].

As recombinant proteins are highly expressed in *E. coli* they contribute to a metabolic burden of the host cell, meaning that the resources that ought to be serving the host metabolism are rather used for maintenance and expression of the foreign DNA. As a consequence the expected biomass increase is lowered during the cultivation. The metabolic burden can be considered as a stress situation for the cells that may further reprogram their gene expression machinery, down regulating genes involved in transcription, translation and amino acid synthesis [8].

1.4.3 Quality problems

When expressing recombinant proteins to higher amount, there is a risk of receiving unwanted post translational modifications (PTM) from the bacterial metabolic pathways. This must be strictly avoided when dealing with therapeutic proteins as the PTMs may affect the product in a negative way reducing its quality.

One common example of quality deterioration, observed in fermentations of *E. coli* cells, is gluconoylation. This is a modification consisting of a covalently attached intermediate from the pentose phosphate pathway i.e. 6-phospogluconolactone (6-PGLac). A suggested cause of gluconoylation is an insufficient supply of the phospogluconolactonase (PGL). As it has been observed that the given *E. coli* strain BL21(DE3) naturally express low levels of the enzyme, this modification may be expected during recombinant expression using this host. Glyconoylations can be detected with the help of HPLC-MS analysis showing a mass shift of 178 Da or 258 Da for glyconoylated product and phospoglyconoylated product, respectively [14].

Moreover, problems with the removal of the N-terminal methionine (Met) have been observed as a consequence of highly expressed proteins in bacteria. The removal of the Met normally occurs with the endogenous methionine aminopeptidases (MetAP). Retaining the amino acid may further confer immunogenic recombinant proteins, thus it is suggested that the expression takes place in suitable host/vector systems that allows complete elimination of the Met. Factors which influence the activity of the enzymes are; the N-terminal amino acid sequence, whether the expression is constitutively active or inducible and the host genotype [15].

1.5 Design of experiments (DOE)

Design of experiments is a powerful tool in the biotech industry as it, based on given experiments, confers the maximum amount of relevant information. A number of applications are suggested like; finding optimal conditions for a specific process, improving the quality of a product or determining the robustness of a certain product or process.

The methodology utilizes simultaneous variation of relevant factors, that are hypothesized to affect one or several responses, and conducts a design in which sets of carefully selected experiments are represented. As the experiments have been performed regression analysis is applied to the resulting data, which further enables a model that shows the relation between changes in the factors and the responses. After evaluation of the model interpretations can be made and it can be established which factors that are influencing the response and how they cooperate. Moreover, response contour plots, based on the model, will show the direction for best operating conditions. A flow-chart is illustrated in **Fig.3**, in order to get a quick view of the important parts of the methodology described above.

Figure 3. An overview of the basic steps assigned in DOE.

1.5.1 Model & Design

The main purpose of DOE is to find an efficient way to describe the response Y in terms of factors X_1, X_2...Xn, hence giving an opportunity of predicting new data from a new set of parameters. This is conveniently obtained by fitting the experimental data to a polynomial regression model. The most common polynomial models used in DOE are; linear, interaction and quadratic which further are formulated in **Table 1** [16].

Table 1: *Polynomial regression models of DOE where Y is response, Xn the investigated factor, βn the influence of factor Xn and ε a residual term.*

Type	**Equation**
Linear:	$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \varepsilon$
Interaction:	$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{12} X_1 X_2 + \ldots + \varepsilon$
Quadratic:	$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_{11} X_{12} + \beta_{22} X_{22} + \beta_{12} X_1 X_2 + \ldots + \varepsilon$

Depending on the choice of model a design is constructed, allowing as much information concerning a system, derived from as few experiments as possible. When using the software MODDE® (Umetrics) a design is proposed based on the experimental objective and the amount of factors, together with their levels and nature (quantitative/qualitative). The linear and interaction models enable factorial fractional designs that investigate each factor at two levels (low/high) and add an adequate number of centre point measurements. As a quadratic model is required, a composite design is of interest, which further explores each factor at three to five levels. The mentioned designs are illustrated in **Fig.4** [16].

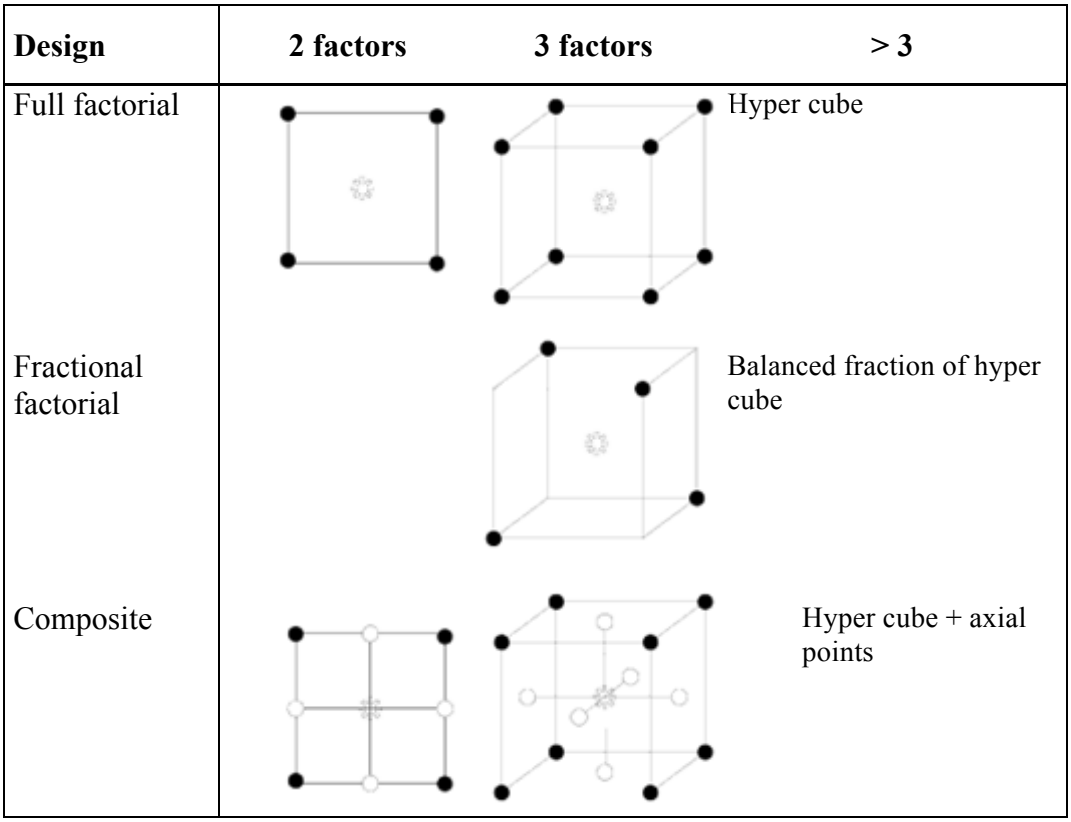


Figure 4. Full factorial, fractional factorial, and composite designs that commonly are used in DOE. The dots represent the suggested measurements for each factor (dimension) and the center point is denoted by a snowflake.

1.5.2 Experimental objective

Depending on the problem formulation, there are three major objectives in DOE; screening; optimization and robustness testing.

1.5.2.1 Screening

Screening is often used as a first step in an investigation of finding the optimal operating conditions. This gives an overview of which factors that are the most dominating ones and further what ranges of them that will allow the optimal response/responses to be encountered. For this approach a linear or interaction model is used.

1.5.2.2 Optimization

The next step in the investigation constitutes of exploring how the set of factors affect the response, either in a negative or positive fashion. In order to approximate the factors and response true relation, a regression model of the quadratic type is of interest as this is flexible and most probably enables the location of the optima [16].

1.5.2.3 Robustness testing

As the optimal operating conditions have been established, it is often of interest to investigate how stable the system is. This is done by applying a small variation of the influential factors, around their set point (for example found during optimization). If the response is insensitive

to small changes the system is declared as robust. Further it is explored with what allowed variation the response stays satisfactory. As narrow ranges of the factors are being explored with this approach, departures from linearity are unlikely, thus suggesting a linear model [16, 17].

1.5.3 Regression analysis

As the experiments have been carried out the data is fitted to the model. When using the software MODDE® this is done with either multiple linear regression (MLR) or partial least squares (PLS). The latter is preferred when dealing with more complex systems.

After fitting the data to the model the model should be evaluated before usage. The most important diagnostic tool for evaluating the model consists of the parameters R^2 and Q^2. Here the former is a measurement of how well the regression model fits the data, while the latter indicates how well the model can predict data. Preferably both R^2 and Q^2 should be close to 1 and their difference rather not more than 0.2-0.3. Further, the calculated model validity should be higher than 0.25 and the reproducibility higher than 0.5 for a good model.

The parameter Q^2 is suggested to be a more useful indicator since the main goal for the model is to predict new data. For a good versus an excellent model, generally $Q^2 > 0.5$ and $Q^2 > 0.9$ respectively, are to be expected. Unfortunately the limits are different depending on the experimental objective. For screening a $Q^2 > 0.1$ is considered enough and for robustness testing a Q^2 near zero is ideal, as this means that there is an extremely weak relationship between factors and response, thus indicating that the system is robust [16].

Another important diagnostic tool is the analysis of variance (ANOVA) that deals with estimations of variability in the response data. Here two F-tests are performed in order to explore differences in the estimates. The evaluation is made by means of probability scores. The first test aims to declare the significance of the regression model and here $p < 0.05$ shows that the test is satisfactory. Secondly, the lack of fit is explored where $p > 0.05$ is satisfactory and reflects a sufficiently low model error and further good fit to data [16].

2. Materials and methods

2.1 Expression system

The HER3-targeting molecule Z08698-VDGS-ABD094 was intracellularly expressed in *E. coli,* upon induction with IPTG.

2.1.1 Strain & vector

A vector denoted pAY03190 was constructed at Affibody AB, by means of a cloning strategy using restriction enzymes and ligases. The vector illustrated in **Fig.5** contains; resistance marker, ORI and a *lacI* gene in addition to the target gene. As the *E. coli* host strain BL21(DE3) was selected, the vector and host together form a T7 based expression system, as depicted in **Fig.2.** Since the recombinant protein is a potential therapeutic, the components were strictly prohibited to contain animal derived contents.

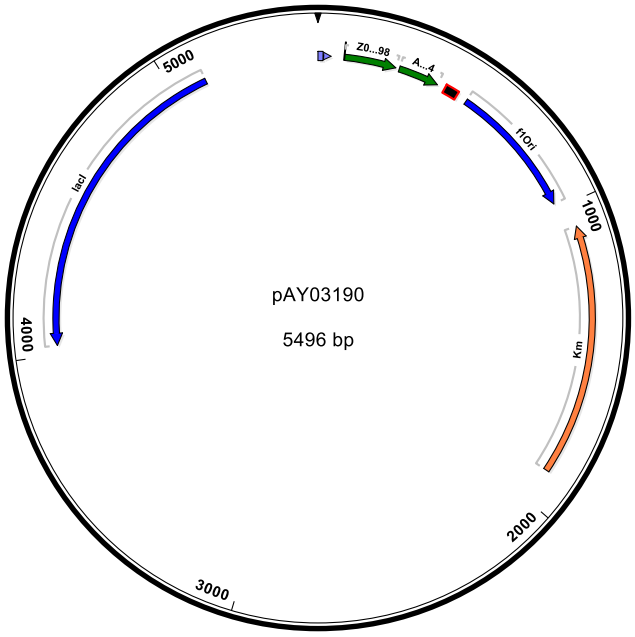


Figure 5. The vector pAY03190: target gene encoding Z_{08698}-VDGS-ABD_{094} (green), ORI (blue), Kanamycin resistance K_m (orange) and *lacI* gene (blue)

2.1.2 Transformation

The electrocompetent BL21(DE3) cells were thawed on ice. Next 1 µl of prepared plasmid solution containing the construct pAY03190 (concentration: 66 ng/µl) was aseptically added and the bacterial suspension was further incubated on ice for a couple of minutes. As 50 µl of the mixture was transferred to a cooled 1 mm cuvette, the cells were pulsed in the electroporator MicroPulser™ (Bio Rad). Quickly about 1 ml of sterile Select APS™ LB-medium (detailed information in **Appendix A**) was added and the whole content of the cuvette was transferred back to the sterile microcentrifuge tube, with the remaining bacterial suspension. The tube was incubated in 37 ºC at 175 rpm for about 45-60 min in the incubation

shaker Multitron® (Infors AG). Finally, 50 µl of the transformed cells were spread on an animal-component free culture plate with Select APS™ LB-medium and ~0.1 % kanamycin, followed by incubation overnight in 37 ºC.

To be able to pick only one colony from the plate, there is an anticipation of the colonies not being grown too compact together. However, if that was the case, the streaking technique was applied, where simply one colony is transferred to a new plate with a sterile inoculation loop.

2.2 Research Cell Bank (RCB)

The construction of an RCB was made for practical reasons and in order to minimize the batch-to-batch variations.

A sterile 300 ml Tunair® shake flask was filled with 100 ml of Select APS™ LB-medium and 100 µl 50 mg/ml kanamycin solution, after which one of the isolated colonies from the culture plate was inoculated to the medium. The flask was incubated in the incubation shaker Multitron® (Infors AG) at 37 ºC and 175 rpm. After about two hours, 150 ml of the cultivation was transferred to a new sterile Tunair® shake flask. This allowed aseptic growth of the cultivation in the new flask, while the cell-density with regular intervals could be measured in the original flask. As the cultivation had reached a satisfying cell density in terms of OD_{600} = 0.8 – 1.0, it was incubated on ice. To reach a final concentration of 15 % of glycerol, 35 ml of the cultivation was mixed with 15 ml of ice-cold sterile 50 % glycerol. The mixture was distributed 1 ml to each of 34 sterile 1.5 ml microcentrifuge tubes, which finally were stored at −80 ºC.

2.3 Cultivation

The transformed *E. coli* cells were cultivated to high cell-density in order to receive as much of the product, i.e. Affibody® Molecule, as possible [10]. Detailed information about the solutions and medium components used are found in **Appendix A**. Further a summary of the performed cultivations are listed in **Appendix C.**

2.3.1 Inoculum

One tube from the RCB was thawed on ice while 100 ml of defined shake flask medium was prepared in a 300 ml Tunair® flask. Preparation of the medium was done in a Laminar Air-Flow (LAF) bench in order to sustain sterility. The flask was filled with sterile Milli-Q-water whereupon solutions of 10 × (YNB + Glucose) and 10 × (Phosphate + Citrate) were added and mixed properly. Finally, the medium was inoculated with 10 µl of RCB and thereafter incubated at 37 ºC at 175 rpm, using the incubation shaker Multitron® (Infors AG).

After 18 h a cell density in terms of OD_{600} = 4.2- 4.5 was to be expected. As the incubator possesses a timer, the inoculum could be completed whenever it was time to inoculate the fermenters. When using the 20 l bioreactor the amount of prepared shake flask medium was doubled and incubated within two Tunair® flasks.

2.3.2 Medium

Since the purpose is to develop a process for the production of a therapeutic molecule, it is of interest that the batch to batch variation is minimal and that the components of the medium are well defined. Hence, a defined animal-component free medium is used, thus avoiding contaminants as viruses and prions derived from animals.

The medium was prepared by dissolving ammonium sulphate and the solution of 20 × (phosphate + citrate) in Milli-Q-water. This was followed by sterilization in autoclave or with the online sterilization program coupled to the 20 l fermenter. Further, the evaporated water was compensated for by the addition of Milli-Q-water whereupon solutions of; glucose, trace elements, magnesium sulphate and kanamycin were added during agitation and aseptic conditions.

2.3.3 Fed-Batch

In order to receive a reproducible HCDC, a fed-batch process was used for cultivation. Thus the carbon source is growth limiting, enabling a specific growth rate of the bacterial cells [12, 13].

Serving as a suitable carbon source, a 60 % glucose solution was used as substrate and supplemented by a controlled feeding regime. The applied glucose feed profile depicted in **Table 2**, is based on earlier successful studies performed at Affibody AB. However, the feed was slightly modified by increasing the feed 10 % half an hour before induction. In this way the growth rate is expected to suddenly rise at the point of induction, resulting in higher product yield. As the intrinsic glucose consumption was measured, the glucose feed deviation was calculated for each cultivation. This feature was also added as an uncontrolled factor in the investigation comprising the optimization of the cultivation.

Table 2. *Glucose feed used for cultivation*

Time (h)	Feed (g/(l×h))
0	0
2.99	0
3.00	1.5
4.00	1.5
5.00	2.0
6.00	3.0
7.00	4.4
8.00	6.7
9.00	10.0
10.00	15.0
50.00	15.0

2.3.4 Fermenters

The cultivations were performed in two different fermenters depending on the desired amount of cultivation. When larger cultivation volumes were needed, as in the case of producing the reference and the final cultivations comprising the found optimal conditions, the 20 l fermenter (Belach Bioteknik) was used. As the multivariate screenings and optimization studies were performed, the GRETA multi-fermenter system spanning 6 × 1 litre fermenters (Belach Bioteknik) were more suitable.

The fermenters were respectively coupled to a computer where stirrer speed, agitation, glucose feed, pH and air-flow could be monitored and controlled through the InControl Phantom software (Belach Bioteknik). In order to sustain the desired pH in the cultivation, an alkali-pump was used for regulation, adding 25 % NH_4OH. Initially, the temperature was set to 37 ºC and thereafter programmed to reach the decided cultivation temperature one hour before induction. Breox FMT30 anti-foam agent was added to the fermenters in order to diminish foaming.

2.3.4.1 20 l scale

Medium was prepared directly in the fermenter, as described above and in **Appendix A**. Instead of adding the solution of glucose together with the rest of the medium components, it was added with the online filling method, in the same time enabling a calibration of the glucose feed pump. The DO-electrode, pH-sensor and pressure holder were connected to the fermenter along with the alkali- and substrate pump. A profile for the agitation was set, starting with 300 rpm and finally changing to the maximum speed of 1200 rpm, two hours before maximum glucose feed. Along with the set aeration of 15 l/min and pressure of 0.3 bars, the dissolved oxygen (DO) concentration is expected to be sustained at a minimal level of 30 % saturation.

2.3.4.2 GRETA multi-fermenter system

The medium was prepared according to former depiction. Further the fermenters were automatically filled with 600 ml of each, through the InControl Phantom software. The profile for agitation was programmed into the system, where the speed was changed from 300 rpm to the maximal 1500 rpm, two hours before maximal glucose feed. Aeration of 1 l/min was started and the desired pH was set. In order to sustain a minimal DO-level of 30 %, oxygen was supplemented through a PID- regulator.

2.3.5 Induction

Expression of the Affibody® molecule was induced by the addition of Isopropyl-β-D-thiogalactopyranoside (IPTG). The cultures were provided with 0.5 ml IPTG/l. As the GRETA multi-fermenter system holds an automatic induction feature, this could be used when induction was necessary at inconvenient hours.

2.3.6 Harvest

After 7.5 or 10 hours post induction, the cultivation was automatically set to cool down with agitation at 300 rpm. Regulation as pH, aeration, glucose feed and possible oxy-flow and pressure holder was turned off. The fermenter was emptied and the cultivation was weighed and next centrifuged at 15,900 × *g* and 4 °C for 25 min. The supernatant was further discarded while the pellet was weighed and refrigerated at −20 °C.

2.3.7 Analysis

During the fermentation process samples were collected in falcon tubes according to the demand for cultivation times. After rigorous vortexing of the mash, a cell density analysis was made. Additionally, an OD1-sample was prepared for each sample, which was further analysed with SDS-PAGE. From the remaining mash, samples were prepared for the adjacent process analysis as described in **Section 2.4**.

2.3.7.1 Cell density

To confirm cell-density the optic density of the cultures were measured off-line at 600 nm (OD_{600}), with the cell density meter CO8000 (WPA, Cambridge UK). The samples were appropriately diluted with 0.9 % (w/v) sodium chloride solution, in order to allow the device to work in its operational range of $0.1 \leq OD_{600} \geq 1.0$.

The purpose of the OD1-samples was to standardize the amount of bacterial cell pellet for adjacent analysis. Samples were prepared by acquiring the amount of culture (x) according to the **Eq. 1** below, where OD_{600} corresponds to the measured cell-density. This was followed by centrifugation at 16,060 × *g* for 10 min, whereupon the supernatant could be carefully removed. The standardization means that by resuspending the pellet with 1 ml of water, an $OD_{600} = 1$ will be obtained.

$$x = \frac{1000}{OD_{600}} \ (\mu l) \qquad (1)$$

2.3.7.2 SDS-PAGE

In order to evaluate the product expression an SDS-PAGE analysis was performed on a NUPAGE™ 4-12% bis-tris Gel (Life Technologies). The OD1-samples were suspended with 150-200 µl CelLytic™ B cell lysis reagent and placed on a shaker for 15-20 min, followed by centrifugation at 16,060 × *g*. The supernatant attained was expected to contain the soluble product, while the pellet which was resuspended with Milli-Q-water could contain insoluble product. Both samples were respectively mixed with 4 × LDS sample preparation buffer and DTT (Life Technologies) according to manufacturer's recommendations. Finally, the samples were loaded onto the gel together with Novex® sharp pre-stained protein standard (Life Technologies) and run at 200 V for about 35 min.

The gel was dyed with a Coomassie-staining solution for approximately an hour, whereupon de-staining was performed with a 10 % ethanol – and 10 % acetic acid (HAc) solution for about 3-5 hours.

2.4 Process analysis

This section aims to describe the process analysis used for the project. Included are a suitable lysis method, purification technique and strategy for analysing the quantity and quality of the product.

2.4.1 Lysis

In order to release the intracellularly expressed proteins, thermolysis will be evaluated as a potential lysis method. Previous experiments have confirmed heat-treatment being an excellent tool for disrupting the cell wall, moreover leaving the thermostable protein partially purified. Thermolysis is further suitable for industrial-scale processes as the treatment can be performed while the cells are remained in the fermenter mash [18].

Initially, the thermostability of the peptide was investigated by suspending the cultivated cells with 6.3 ml of 20 mM Tris-Hydrochloride (Tris-HCl) buffer (pH 7.5), followed by heat-treatment within a water bath having temperatures of 85, 90 or 95 °C for 5 or 10 min, respectively. As comparison, a bacterial pellet was suspended with 150 µl CelLytic™ B cell lysis buffer (Sigma-Aldrich) per OD1-sample and further placed on a shaker for 15-20 minutes. All samples, both chemically and thermally lysed, were centrifuged at 16,060 × *g* in 4 ºC for 5 min. The supernatant obtained was further analysed with SDS-PAGE according to description above.

2.4.2 Purification

Affibody AB has designed a matrix, specific for the ABD conjugated to the Affibody® Molecule, which serves as an excellent stationary phase for affinity chromatography. Here 0.4 ml of the anti-ABD gel matrix was packed onto NAP-5 columns (GE Healthcare). In order to avoid saturation of the matrix, it is critical that no more than 1 mg product/0.1 ml of matrix is run on each column per occasion [19]. With precautions in mind the maximum loaded amount of product is 2 mg. Due to the recommendations above, and as a previous cultivation (121010B) resulted in 45 mg product/g pellet, the samples obtained from the cultivations were prepared to contain around 44 mg of bacterial pellet.

As it was hypothesized that the wet weight would be hard to physically measure for each sample, the wet weight was decided to be estimated based on its OD_{600}-value. The found correlation between the wet weight (WW) and cell-density is represented in **Eq. 2** and moreover described in **Appendix E**.With the estimated wet weight it could further be established how much of each sample was to be loaded onto the column.

$$WW = 1.0833 \times OD_{600} \ \ (g/kg) \qquad (2)$$

As the cell lysate with the soluble Affibody® Molecule has been loaded onto the column, the ABD is covalently bond to the matrix, thus leaving the product retained in the column. In order to remove unbound particles, the columns were washed with 1 × Tris/saline/Tween-

buffer (TST) followed by addition of 5 mM NH_4Ac, which further lowered the pH before elution. Since a sufficiently low pH causes a relocation of the molecules charges, resulting in the release from the stationary phase, elution was conveniently achieved with 0.1 M HAc, pH = 2.88. The entire procedure of purification is described in **Appendix F** including steps like regeneration, washing, elution and "cleaning in place" (CIP).

2.4.3 Protein quantification

After purification, a spectrophotometer NanoDrop ND-1000 was employed to determine the absorption at 280 nm, thus registering the proportion of protein in the sample. Each sample was measured three times in order to receive reliable data. The average absorption coefficient (Abs280) was calculated and further converted to the corresponding product concentration (C) according to **Eq. 3.**

$$C = 0.626 \times Abs280 \quad (mg/ml) \tag{3}$$

2.4.4 Quality analysis with HPLC-MS

The quantified product was further analysed with the HPLC-MS 1100 series (Agilent Technologies) using the column Zorbax 300SB-C8 (Agilent Technologies), in order to uncover the potential modifications addressed in **Section 1.5.3**. A mobile phase consisting of a gradient of buffer A: 0.1 % trifluoroacetic acid (TFA) in Milli-Q-water and buffer B: 0.1% TFA in acetonitrile was employed, where buffer A was increased from 10 to 70%, in contrast with a corresponding decrease of buffer B, in 25 minutes.

After analysis the deconvolution tool of the HPLC-MS 1100 series software was used, including 4 peaks per set with an abundance cut-off = 5 %, in order to obtain a spectrogram with the explored masses. The product has an expected mass of 11,939.3 Da. Moreover the potential modifications of phoshogluconoylation, gluconoylation and retained N-terminal methionine are expected to confer with a mass increase of +258 Da, +178 Da and +131 Da respectively.

2.5 Optimization of cultivation protocol

2.5.1 Factors and response

The aim of this part of the project is to investigate how the responses i.e. the expression can be maximized, in the same time attaining the best possible quality of the molecule. Based on previous studies at Affibody AB, the expected factors to influence these desired responses were chosen to be*; temperature, pH, induction- and expression time*. Here *induction time* refers to the cultivation time passed when inducing the expression of the peptide with IPTG, while *expression time* refers to how long time the peptide is being expressed. In order to find the interesting factor ranges, in which the optimization can take place, several attempts were carried out and are described below.

2.5.2 Reference cultivation

The reference cultivation 130212B was performed with a working volume of 15 l in the 20 l fermenter. Conditions were set to reflect the optimal conditions (illustrated in **Table 3**), found for the similar construct pAY02023 during a former investigation at Affibody AB. Further this cultivation will serve as a reference when comparing with the upcoming ones.

Table 3. *Conditions during reference cultivation*

Factor	Setting
Temperature (ºC)	33
pH	6.95
Induction time (h)	20
Expression time (h)	7.5

2.5.3 Finding the experimental area for optimization

To be able to find the intervals of the given factors, in were the given responses can be optimized, cultivations were made in the GRETA multi-fermenter system (130212H1-130212H6). Here, two of the reactors were run with the same given parameters as for the reference cultivation, found in **Table 3** above. The parameter settings of the remaining reactors were also based on the previous investigation mentioned above where two of the reactors were run with only low levels and two with only high levels of the given factors. In order to receive as much data as possible from the cultivations, samples were taken from each after 5, 7.5 and 10 h, respectively. The parameter settings for the six bioreactors are represented in **Table 4**.

Table 4. *Only low, centre-point and high levels of the given factors used for the cultivations in the GRETA multi-fermenter system, consisting of the bioreactors; H1, H2, H3, H4, H5 and H6.*

Factor	Low: H1/H2	Centre: H3/H4	High: H5/H6
Temperature (ºC)	30	33	36
pH	6.70	6.95	7.20
Induction time (h)	18	20	22

2. 6 DOE & Multivariate data analysis

2.6.1 Experimental design of multivariate screening to improve thermolysis

In order to improve the thermolysis, factors anticipated to affect the yield of the peptide were investigated through a multivariate screening. The factors surveyed were; the amount of *dilution* with Tris-HCl buffer per gram bacterial pellet, *temperature* and *time* of treatment in water bath. Appropriate levels of the factors are represented in **Table 5**, which further were implemented in the software MODDE® 9.1. The ranges of the chosen factors were based on a previous screening (**Appendix G**).

A full factorial design with three centre points was executed, resulting in 17 experiments. To investigate how short the treatment can be held, the experimental protocol was further extended with treatment times of 1, 2, 4 and 5 min in the centre point (25 ml/g, 91 ºC). The experiments were divided into three runs, depending on temperature, and were performed as represented in **Appendix H**.

Table 5. *Ranges of the quantitative factors in the multivariate screening*

Factor	Low	Centre	High
Dilution (ml/g)	20	25	30
Temperature (ºC)	89	91	93
Time (min)	2	4	6

After heat-treatment the samples were directly cooled down on ice. As comparison, a chemically lysed sample was prepared by suspending the bacterial pellet with CelLytic™ B cell lysis buffer (Sigma-Aldrich), using the proportions 1 ml/0.1 g cells. Next the bacterial suspension was placed on a shaker for 15-20 min. All the samples were centrifuged at 16,060 × *g* at 4 ºC for 5 min and the supernatant was obtained and purified, before protein quantification. In order to receive significant results the samples were carefully prepared, enabling the same amount of pellet being analysed in each purification process.

2.6.2 Experimental design of robustness testing of thermolysis

The robustness of the thermolysis was investigated by creating an experimental design in the software MODDE® 9.1. Since the aim is to investigate whether small fluctuations in the influenced factors are affecting the response significantly, the ranges of the factors should be narrow and a linear approach such as fractional factorial design resolution III is desired [16].

The ranges of factors are displayed in **Table 6** and are based on the process analysis established in **Section 3.1.2** and own speculations around possible variations. By implementing the design and the chosen factors in the software, a worksheet comprising of seven experiments was obtained (illustrated in **Appendix H**). The samples were processed as earlier described, followed by quantification of product.

Table 6. *Ranges of quantitative factors in robustness testing of thermolysis*			
Factor	**Low**	**Centre**	**High**
Dilution (ml/g)	29	30	31
Temperature (ºC)	89	90	91
Time (min)	2.5	3.0	3.5

2.6.3 Experimental design optimizing cultivation

In order to find the optimal cultivation conditions generating the highest expression levels and the best quality of the product, an optimization was performed in the software MODDE® 9.1. Based on previous experiments the low and high levels of the given factors were selected and implemented in the software according to **Table 7**. The glucose feed deviation was included and defined as an uncontrolled factor.

Table 7. *Factor ranges in the multivariate optimization*			
Factor	**Low**	**Centre**	**High**
Temperature (ºC)	33	35	37
pH	6.70	6.95	7.20
Induction time (h)	10	12	14
Expression time (h)	2.5	5.0	7.5

When optimizing a quadratic model is of interest, ensuring the discovery of the optima. Therefore, the central composite fractional design (CCF) was chosen [16] and a suggested experimental protocol, a so-called worksheet was obtained from the software. Since it is easy to take samples from an already started cultivation, the worksheet was manually extended, ending up with all the levels of expression time, namely 2.5, 5 and 7.5 h for all of the proposed cultivations.

To get a trustworthy investigation, four centre point measurements were included in the design. These cultivations were further distributed to take place in different bioreactors in order to consider the experimental variation. Finally, the worksheet comprised of 54 experiments, which could be divided into 18 different cultivations and further three runs of cultivations performed in the GRETA multi-fermenter system. The final experimental protocol is found in **Appendix H.**

3. Results

3.1 Establishment of process analysis

3.1.1 Test of thermostability

The thermostability of Z08698-VDGS-ABD094 was investigated by dissolving the bacterial pellet in 20 mM Tris-HCl buffer, followed by heat-treatment in water bath. It was shown that thermolysis can be used for the given Affibody® molecule and that this lysis technique is nearly as efficient as CelLytic. This is confirmed with the SDS-PAGE analysis from which the gel is illustrated in **Fig.6**. Thermolysis further allows native proteins of the *E. coli* to be denatured and precipitated which may simplify adjacent downstream processes.

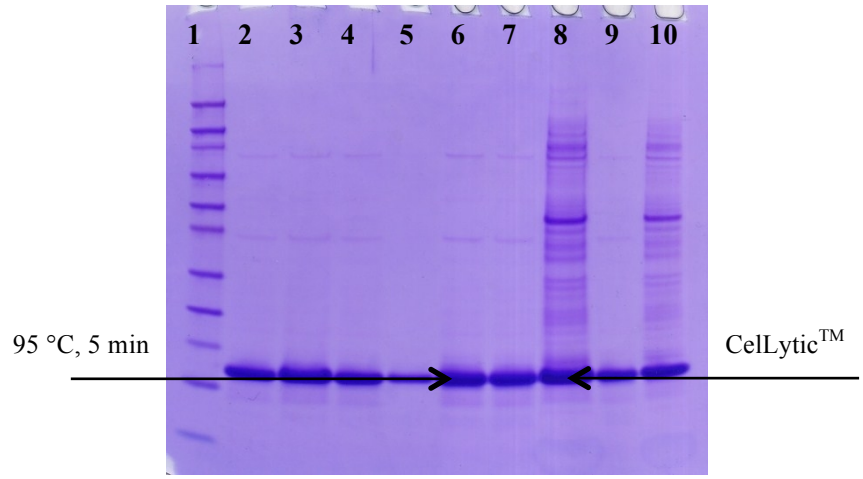


Figure 6. Expression analysis of thermolysed samples from cultivation 121010B. Well 1: Ladder, well 2: 85 °C 5 min, well 3: 85 °C 10 min, well 4: 90 °C 5 min, well 5: 90 °C 10 min, well 6: 95 °C 5 min, well 7: 95 °C 10 min, well 8: CelLytic™, well 9: 95 °C 10 min 11.3 µl, well 10: CelLytic™ 11.3 µl

3.1.2 Multivariate screening of thermolysis

Low and high levels of the factors; *dilution*, *temperature* and *time* were combined in the software MODDE® 9.1 in order to get an efficient experimental design. As the experiments were run, the A_{280} coefficient (later transformed to product concentration) from spectrophotometry analysis was reported to the software. The data were fitted with multiple linear regression (MLR) and the model was further treated for improvements.

As it was found that the temperature was not a contributing factor, but rather disturbing the model, it was neglected. The remaining factors having an impact on the yield of the product are represented in the coefficient plot illustrated in **Fig.7.**

The model attained supports for significance as $R^2 = 0.64$ and $Q^2 = 0.40$, recall from **Section 1.5.3**. Additionally, the model validity = 0.62 and reproducibility = 0.82 are satisfying. The data is represented in the summary plot in **Fig.8** below.

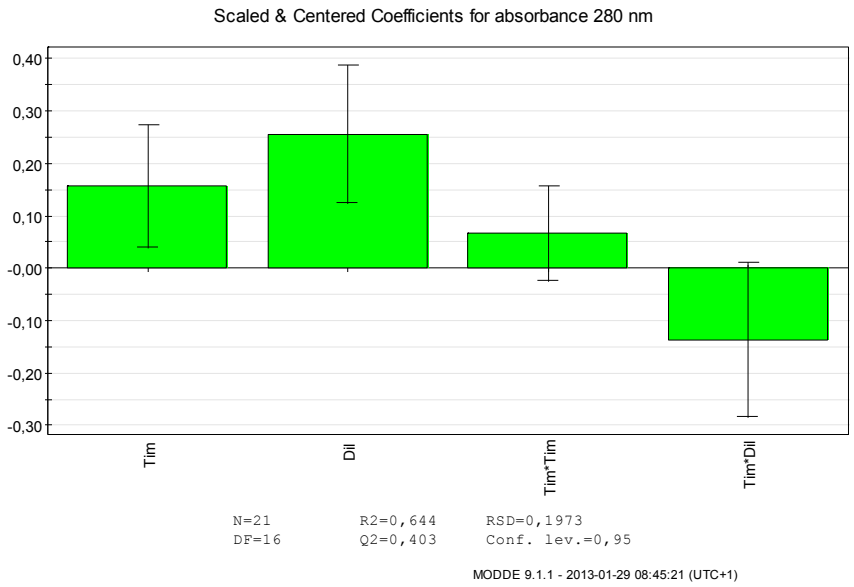


Figure 7. The coefficient plot with the model terms; Time, Dilution, Time*Time and Time*Dilution from left to right. MODDE®

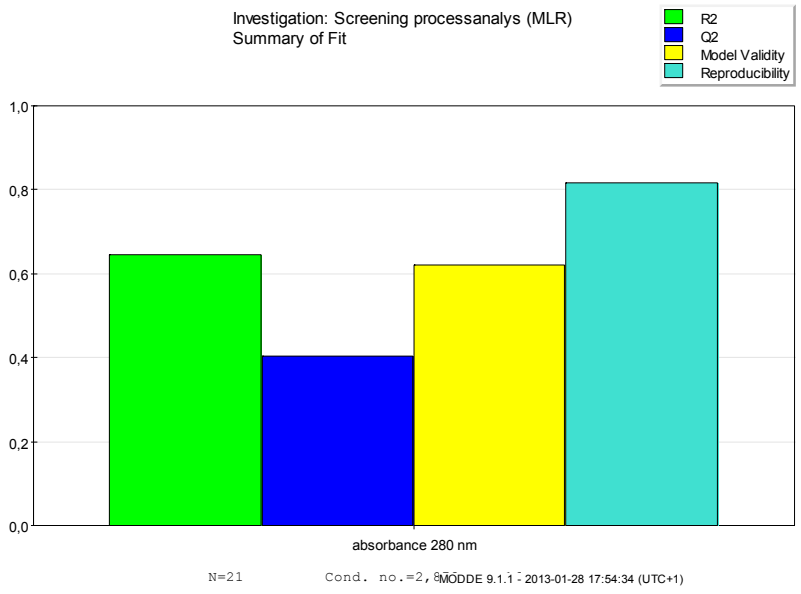


Figure 8. The summary plot representing a significant model. From left to right the staples illustrate the R^2, Q^2, Model validity and Reproducibility. MODDE®

Based on the model, a contour plot, shown in **Fig.9**, was obtained, representing the factor settings giving the highest expected absorbance at 280 nm and thus highest product concentration. In favour of the response, high levels of dilution and time are desired. However, as there was no significant difference in response when reducing the treatment time to *3 min* instead of 6 min, the lower alternative was preferred. As the temperature was not a contributing factor in the ranges of 89-93 ºC, this factor is held at a level of 90 ºC.

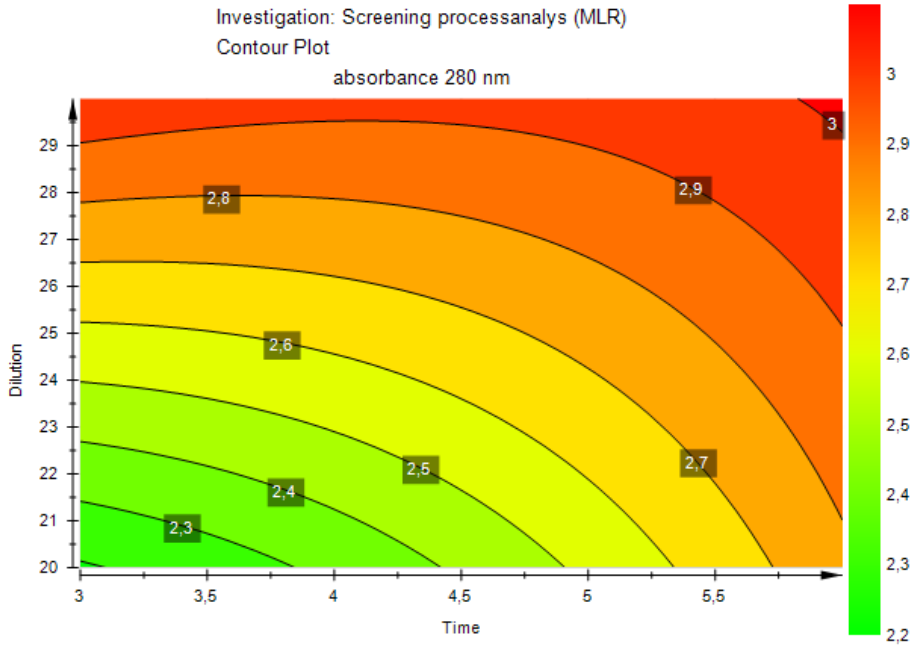


Figure 9. Contour plot illustrating the contributing factors and their ranges of interest, according to the desired response. The color scheme illustrates the A_{280}-value and x-axis and y-axis represent the time in min and dilution in ml/g respectively. MODDE®

Finally the obtained thermolysis protocol represented in **Table 8** was subjected to the predictive function of the software. With the given settings an expected A280 of 2.97 was suggested, thus resulting in an expected product concentration of 1.86 mg/ml, according to **Eq. 3**. This may be compared with the product concentration of 1.85 mg/ml, obtained when lysing the cells with CelLytic™ B cell lysis buffer.

Table 8. *Final thermolysis protocol*

Factor	**Setting**
Dilution (ml/g)	30
Temperature (°C)	90
Time (min)	3

3.1.3 Robustness testing of thermolysis

As the samples were processed according to the prescribed worksheet, the results of the absorbance at 280 nm were reported to the investigation in MODDE® 9.1. The model was fitted to the data with the expectation of receiving a weak relationship between the factors and response, thus indicating a non-significant model and a robust system. MLR was used for fitting and a weakly significant model was attained with $Q^2 = 0.25$, illustrated in **Fig.10**. Additionally, the p-value = 0.35 of the regression implies of an insignificant model.

The minimum desired value of absorbance was set to 2.4, corresponding to about 1.5 mg product/ml of sample. This criterion was fulfilled for all the experiments, shown in the replicate plot, **Fig.10.**

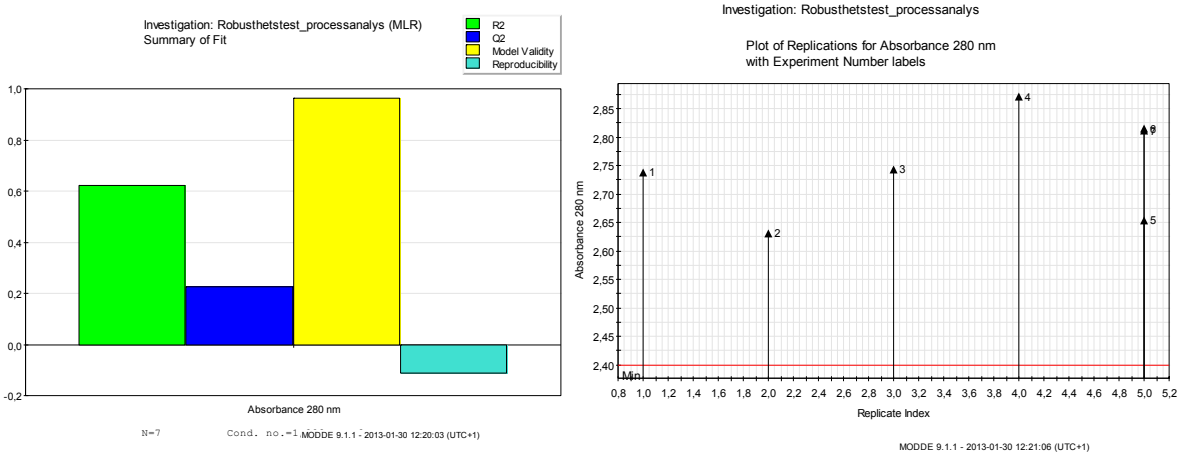


Figure 10. The following graphs are obtained from the robustness testing of the thermolysis; summary plot (left) confirms the desired insignificant model (Q2<0.5), with staples representing R^2, Q^2, Model validity and reproducibility. The replicate plot (right) further shows that the set criteria of obtaining at least 1.5 mg product/ml sample (red line), is fulfilled for all experiments (black arrows). MODDE®

To test the robustness of the current protocol of thermolysis, a Monte Carlo simulation was carried out simulating random disturbances to the factors. As the criterion of obtaining at least 1.5 mg/ml product was not accomplished within the originally set specifications represented in **Table 6** (**Section 2.6.2**), the regions were further adjusted.

Robustness was acquired for the set conditions; allowing the temperature to vary ± 0.5 ºC, time ± 6 s and dilution ± 0.1 ml. Defaults per million (DPMO) of 590 was obtained, meaning that only 0.059 % of the disturbances conferred a response located outside the threshold and moreover indicates that the current lysis technique is robust. As a result, the protocol of thermolysis was extended with the requirements for robustness, represented in **Table 9**.

Table 9. *Final thermolysis protocol supporting the requirements for robustness.*

Factor	**Setting**	**Allowed variation**
Dilution (ml/g)	30	± 0.1
Temperature (°C)	90	± 0.5
Time (min)	3	± 0.1

3.2 Optimization

3.2.1 Pre-studies enabling the experimental area of interest for optimization

The factors; temperature, pH, induction time and expression time were chosen to be investigated during cultivation, with the purpose of optimizing expression together with the quality of the Affibody® Molecule. As the impact of only low, mid-point and high levels of the factors were explored according to **Table 4** (**Section 2.5.3**), the cultivations with merely high levels of the factors were of favour for the expression levels. Thus the settings; temperature 36 ºC, pH 7.2, induction time 22 h and expression time 10 h result in the highest expression levels of **3.73 g/l** product. This is shown in **Fig.11**, where the results of the reference cultivation is also illustrated. From the reference cultivation, **2.89 g/l** product was obtained after 7.5 h of expression.

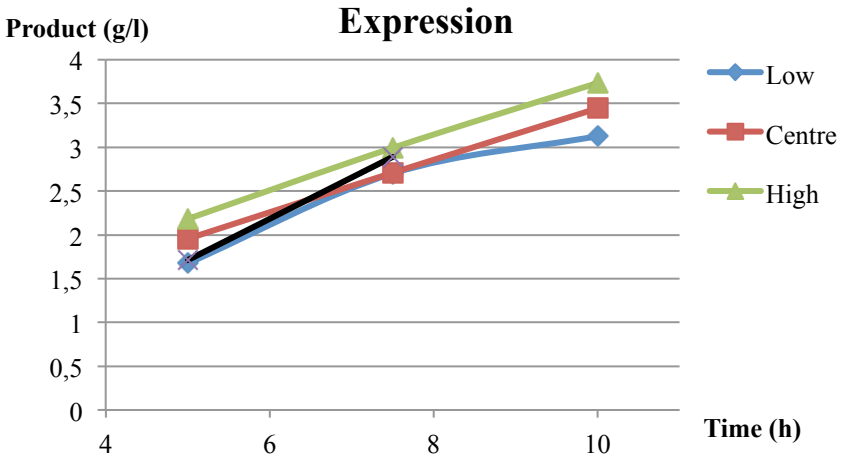


Figure 11. Expression levels of cultivations 120212H1-120212H6 when either low levels (Temp = 30°C, pH = 6.7, I-time = 18), centre points (Temp=33°C, pH = 6.95, I-time = 20) or high levels (Temp = 36°C, pH = 7.2, I-time = 22) were set. Additionally the expression levels obtained from the reference cultivation 130212B are represented as a comparison.

After evaluation of the quality of the product in terms of modifications such as gluconoylation and retained N-terminal methionine, further experiments were planned aiming for improvements. As high induction time and expression time were found to give higher amounts of gluconoylated product, lower levels of these two factors were presumed to improve the quality.

In order to get a quick assessment of the theory, new cultivations were performed in the GRETA multi-fermenter system (130308H1-130308H6). The centre-point levels of the factors temperature and pH (that is 33 ºC, 6.95) were retained, in the same time varying the induction time from 12-22 h and taking samples after 2.5, 5, 7.5, 10, 12.5 and 15 h.

As a result it was found that, when reducing the induction time, the expression levels were increased (**Fig.12**) and in the same time the product modifications were decreased (**Fig.13**).

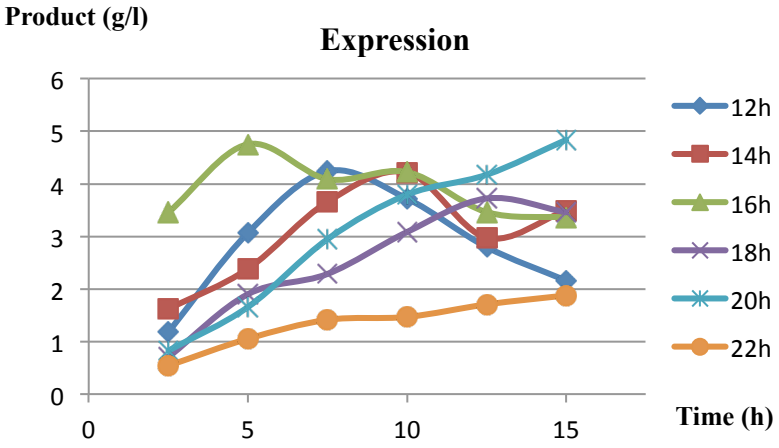


Figure 12. Expression levels for cultivation 130308H1-130308H6 when temp = 33 °C, pH = 6.95 while induction time was varied from 12-22 h.

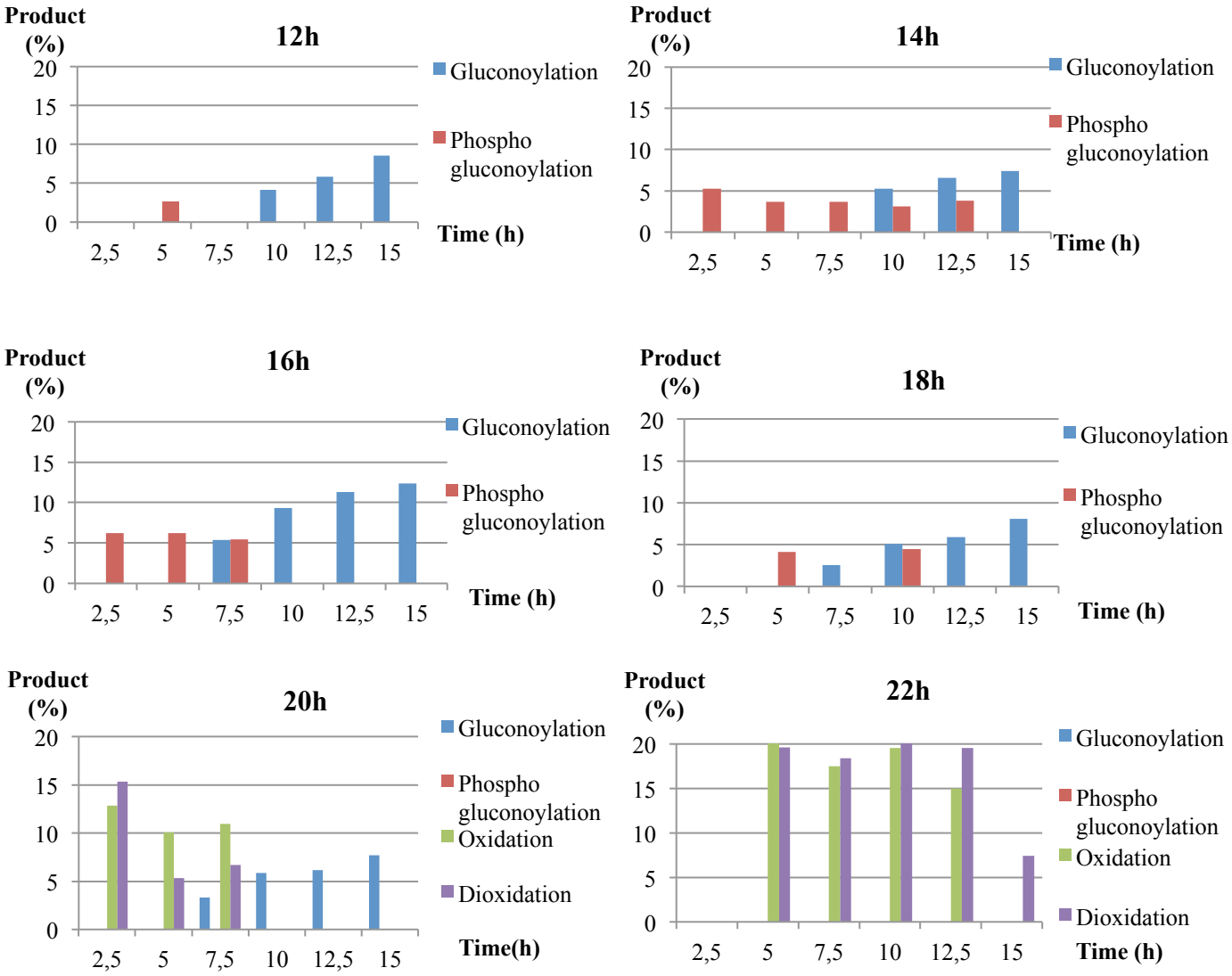


Figure 13. An illustration of the modifications when varying the induction time from 12-22 h. Here also unwanted deteriorations like oxidations [19] were detected when employing I-time = 20 or 22 h.

Since the higher values of the factors generated the highest expression (**Fig.11**) and the lower induction time and expression time resulted in fewer modifications together with higher expression levels (**Fig.12** and **Fig.13,** respectively), the optimization was decided to take place in the experimental area according to **Table 7** (**Section 2.6.3**).

3.2.2 Multivariate optimization

3.2.2.1 Model evaluation

As the three runs of cultivations in the GRETA multi-fermenter system was completed, the results of the quantity and quality of the product were reported to the multivariate data analysis software. The final worksheet with the response of each experiment is shown in **Appendix H**. Here, the response gluconoylation refers to the amount (%) of both phospho-gluconoylation and gluconoylation, meaning that the modifications are not represented individually. (No oxidations were presumed to be discovered with the given values of induction time, a theory that was proven in the HPLC-MS analysis.)

The data was fitted with PLS whereupon two models were calculated; one describing the quantity and the other describing the quality of the product. The models were further treated for improvements and the final results are illustrated in the summary plot (**Fig.14)** and the coefficient plots (**Fig.15)**.

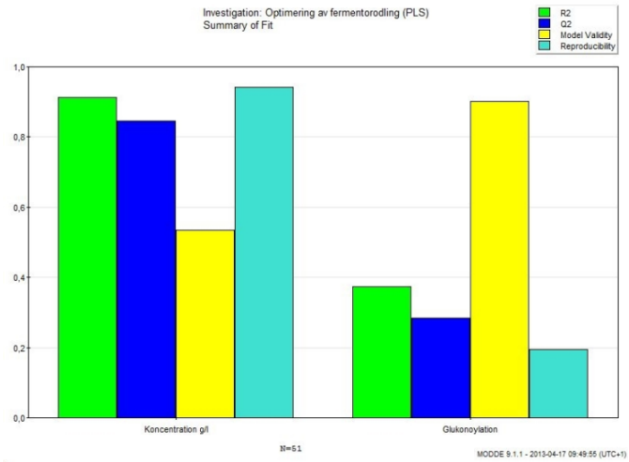


Figure 14. The summary plots for the product concentration (left) and modifications in terms of gluconoylation (right) with staples for R^2, Q^2, Model validity and reproducibility from left to right. MODDE®

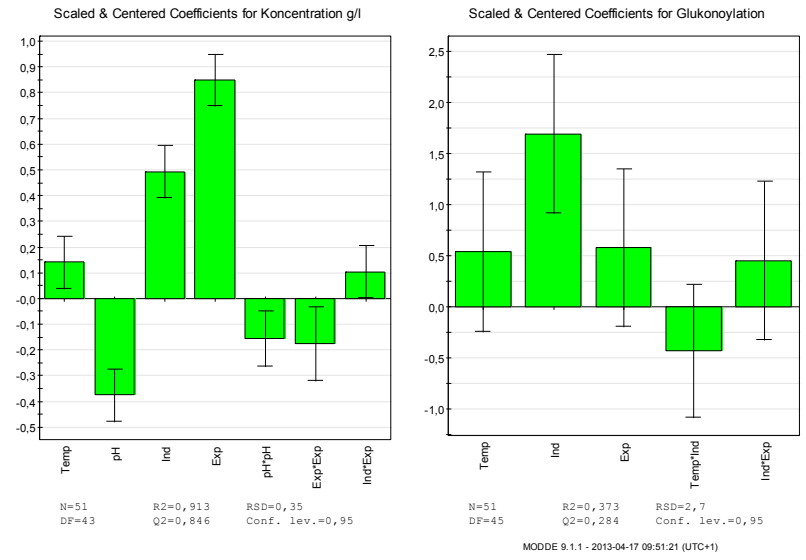


Figure 15. Coefficient plots for the responses. Product concentration with the influencing terms; Temp, pH, I-time, E-time, pH^2, E-$time^2$ and I-time*E-time from left to right. For gluconoylation the terms were Temp, I-time, E-time, Temp*I-time and I-time*E-time. MODDE®

The model for gluconoylation was found to be weakly significant with $R^2 = 0.373$ and $Q^2 = 0.284$, despite several attempts of adjustments. Because of this the optimal conditions, in favour of the quality of the product, cannot be assessed with absolute certainty. However, the coefficient plot in **Fig.15** implies that the induction time seems to be an influential factor.

Based on the model, the contour plot in **Fig.16** was created. Here, it is implied that low temperature, high pH and low induction time expect to minimize gluconoylations. Though may this suggestion only serve as a guidance, as the model did not reach the level of significance.

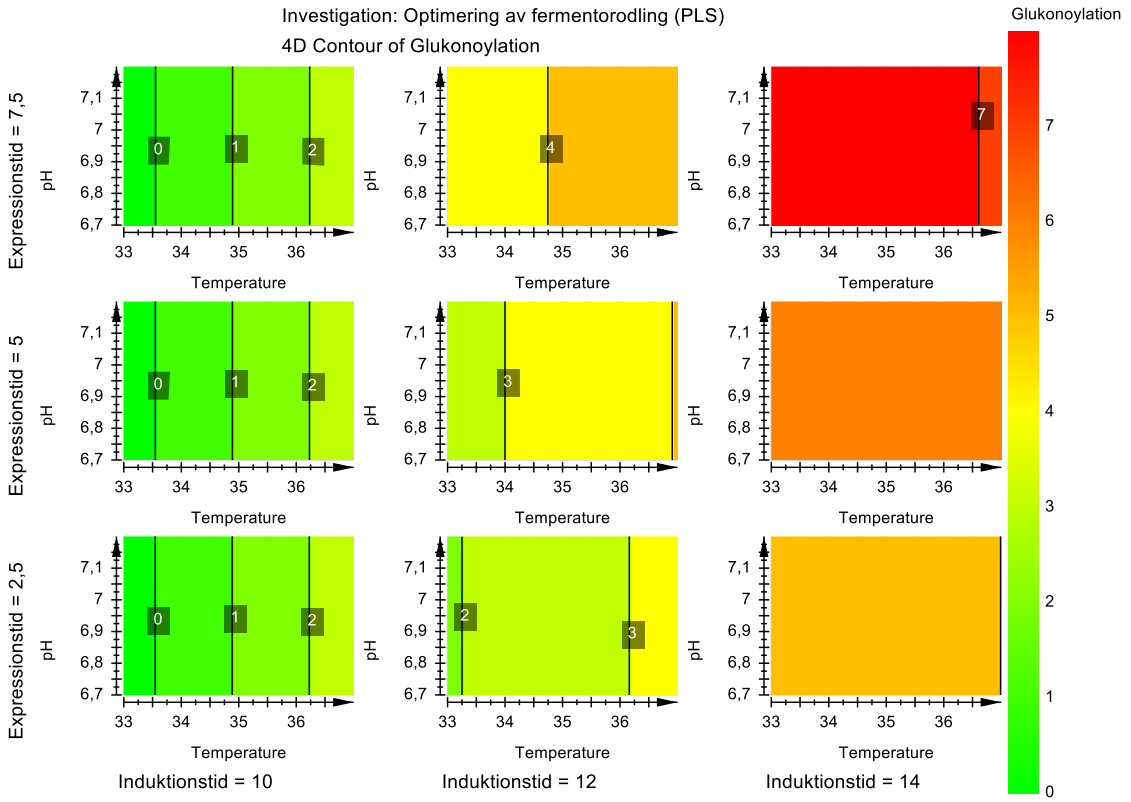


Figure 16. The contour plot based on the model for gluconoylations. The major x-axis represents the I-time 2.5, 5 and 7.5h from left to right, while the major y-axis represents the E-time 2.5, 5 and 7.5h from bottom to top. The color scheme represents the amount of glyconoylated product in %. MODDE®

At first the model describing the product concentration also turned out to be weakly significant. As a consequence the uncontrolled factor glucose feed was excluded together with three experiments, originating from one unsuccessful cultivation (130410H2 found in **Appendix D**). After careful selection of the model factors, represented in the coefficient plot **Fig.15**, the model reached level of significance with $Q^2 = 0.85$ and $R^2 = 0.91$. Since the fitting method MLR did not contribute to a desirable condition number (3-5), the method was further exchanged to PLS giving the satisfying condition number of 4.38. The ANOVA statistics suggests for a significant regression model with p-value = 0.00 (<0.05) together with a lack of fit p-value = 0.16 (>0.05) that further suggests a good fit of data model.

Based on the results a contour plot, illustrated in **Fig.17**, was calculated. The highest product concentration is obtained in the upper right corner of the plot. Thus low levels of pH and high levels of temperature, induction time and expression time are suggested.

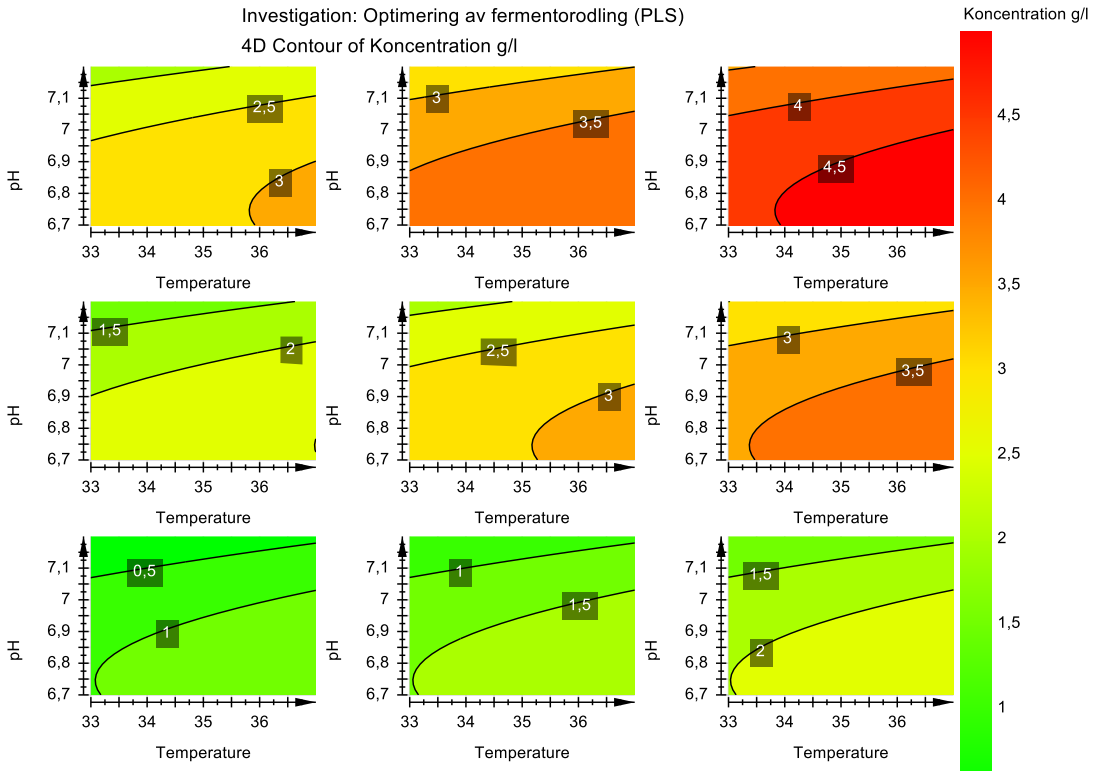


Figure 17. The contour plot based on the model for product concentration (g/l). The major x-axis represents the I-time 2h, 5h and 7.5h from left to right, while the major y-axis represents the E-time 2.5, 5 and 7.5h from bottom till top. Further the concentration is indicated by g/l with the color scheme. MODDE®

3.2.2.2 Optimization

Since the model describing the quality of the product could not guarantee absence of modifications, this response was decided to be neglected. It was also found that when taking account to the quality, the expression levels would be dramatically decreased i.e. approximately 2.5 g/l compared with the expected 4.5 g/l (according to **Fig.17).**

As a result the final optimization was carried out only in regard to the expression levels of the product. The optimization function of the software was used in order to find the eight best conditions, where the settings; temperature = 36.04 ºC, pH = 6.70, induction time = 14.00 h and expression time 7.50 h gave the highest expression of **4.70 g/l**. Moreover a second optimization was performed by picking new coordinates from this optimal point. Here the conditions (36, 6.7, 14, 7.5) and (37, 6.7, 14, 7.5) were included, generating the highest expression of **4.80 g/ l** with the latter.

3.3 Verification in a large- scale production

The optimal conditions found for cultivation, were tried out in a large-scale production. Further thermolysis was scaled up allowing approximately 1 kg of bacterial cell pellet being lysed in a simple procedure.

3.3.1 Cultivation

Two cultivations, 130417B and 130424B, were performed in the 20 l fermenter, with a working volume of 15 l as described earlier. The optimal conditions found for expression of the Z08698-VDGS-ABD094 were applied, namely temperature = 37 °C, pH = 6.7 and induction time = 14 h. Samples were taken from the cultivations after 2.5, 5, 7.5 and 10 h of expression time and further processed according to earlier description. According to the optimization the optimal expression time would be 7.5 h. Despite this expectation approximately 6 l of culture was harvested after both 7.5 and 10 h, respectively.

As the inoculum for the first cultivation did not reach the desired cell density of OD_{600} = 4.2-4.5, the profiles for glucose feed and stirrer speed along with the time for induction, was set to take place 1 hour later. In this way the cell density was expected to be sufficient at the time for induction. Since the expression levels of the product turned out to be quite similar for these cultivations 7.5 h after induction: 3.55 and 3.43 g/ l respectively, this approach was regarded as successful. Hence, the average expression of Z08698-VDGS-ABD094 obtained from both cultivations was **3.49 g/l**. These results are illustrated, together with the expression levels obtained from the reference cultivation 130212B, in **Fig.18.**

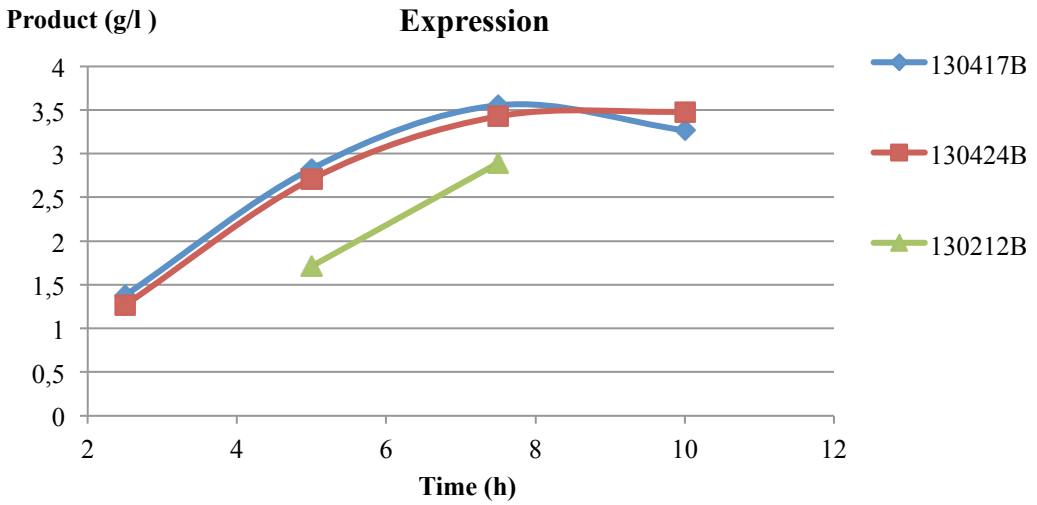


Figure 18. Expression levels of Z0898-VDGS-ABD094 attained from the large-scale cultivations 130417B and 130424B together with results from the reference 130212B.

3.3.2 Large-scale thermolysis in the fermenter

A pellet of approximately 1 kg, obtained from cultivations 130417B and 130424B, was subjected to a large-scale thermolysis in the 20 l fermenter. The fermenter was filled with 19 l of 20 mM Tris-HCl buffer (pH 7.5) and the temperature was set to 93 °C, as the temperature is expected to decrease drastically upon suspension of the frozen pellet. When the set temperature was reached, the bacterial cell pellet was suspended in the buffer during agitation. The lid was further assembled together with the pressure holder and condenser.

Shortly after the desired temperature of 90 °C was attained, the treatment time of 3 min was initiated. Next, the fermenter was set to cool down, whereupon the cell lysate was filled within e-flasks that were positioned in ice-water bath, enabling a quick cool down. As the cell lysate had reached a temperature of 25 °C, it was further weighed followed by centrifugation at 15,900 × *g* for 25 min. Finally, the supernatant was weighted and refrigerated at −20 °C.

A sample constituting of 1.3 ml of the supernatant was purified through the NAP-5 columns, as earlier described. Moreover, the final product concentration was estimated to be **2.78 g/l** culture.

4. Discussion

4.1 Process analysis

Initially, a suitable process analysis was developed, in order to detect changes by means of quality and quantity of the product during the optimization event. As earlier mentioned, small changes in the manufacturing process, may confer a great impact on the product. The Z08698-VDGS-ABD094 was intracellularly expressed in *E. coli*, and thus the cultivated cells had to be disrupted for releasing the product. For this purpose thermolysis was proposed. The strategy would be convenient for the company by means of cost and practical reasons.

Even if the Affibody® Molecules by nature are very robust and stable, the approach of heat-treatment has not always been successful for earlier constructs. As the procedure allows the native proteins to be precipitated, there is a risk of the expressed product being trapped inside one of these complexes. Hence, heat-treatment was initially evaluated for releasing the Z08698-VDGS-ABD094. The bacterial pellet, attained from cultivation 121010B, was therefore suspended in 20 mM Tris-HCl buffer, according to general practice at the company (6.3 ml/g pellet), and further heat-treated in a water-bath for 5 min and 10 min, respectively. The SDS-PAGE analysis revealed that the product remained soluble after applying temperatures of up to 95 °C. It was also found that the product yield was highly comparable with the results obtained when using CelLytic™.

The thermolysis procedure was further improved by employing DOE and multivariate data analysis, ending up being as efficient as its competitor CelLytic™. The results in **Section 3.1.2** reveal that the product concentration was 1.85 g/ml in the purified sample when using CelLytic™. Comparably, the model for thermolysis predicted a product concentration of 1.86 g/ml, when applying the found thermolysis protocol on the same amount of pellet (cultivation 121010B).

As reproducibility is important when producing therapeutics the robustness of the current thermolysis protocol was tested. The result is again represented in **Table 10** below. In order for the protocol to be robust the criteria for minimal expected concentration had to be somewhat lowered, more precisely to 1.5 g/ml, which either way is found to be an acceptable product yield. The recommended limits for factor variation were also found to be manageable, at least when preparing the cultivation samples in a smaller scale. During the small-scale process an additional calibrated thermometer was employed in water-bath, whereas the amount of added buffer and treatment time was strictly controlled with pipette and stopwatch, respectively.

Table 10. *Final thermolysis protocol supporting the requirements for robustness.*

Factor	Setting	Allowed variation
Dilution (ml/g)	30	± 0.1
Temperature (°C)	90	± 0.5
Time (min)	3	± 0.1

During development of the process analysis it was found that some intermediate steps are of utter importance like for example; the cell lysate has to be cooled down directly on ice after heat-treatment and should further be kept cold during centrifugation and storage before purification. In this way the proteases are expected to be non-functional and thus digestion of the product is avoided.

The uncertainties of the developed process analysis are proposed to lie in the approximation of the wet weight based on the cell-density. First of all, the cell density meter CO8000 (WPA, Cambridge UK) was found to be rather inconsistent during the measurements. The measurements should have been performed in replicates for better results. On the other hand this would have been quite inconvenient with the large amount of samples prepared within this project. Secondly, a heterogenic sample could contribute to an inaccurate measurement. This problem was counteracted during sampling by rigorous vortexing of the mash before each OD_{600} measurement. At last, the employed correlation between cell density and wet weight (described in **Appendix E**) may not reflect their natural relationship. Generally there are uncertainties when deciding the wet weight for cultivations and therefore often dry weights are calculated instead, as they seem more certain. Moreover, the employed correlation was improved, as more and more cultivations were performed along the project. This may further contribute to that the amount of pellet, in each sample, have been incorrectly calculated. As a consequence, the dilution of the samples (before thermolysis) may have been inaccurately done, conferring inaccurate data and further insignificant models for the multivariate analysis. Hopefully, the fluctuations in the data have been correctly adjusted by means of the regression analyses and statistics used in MODDE®. Finally, as biological systems in fact are varying by nature, all types of fluctuations are pretty hard to exclude.

4.2 Cultivation

A clear focus in this project has been to create an optimized cultivation procedure that exhibit a high reproducibility. As proposed in Jenzsch M *et al*. 2009 reproducibility is very important when producing therapeutics, and thus their manufacturing processes should be strictly controlled. Consequently, for example a fed-batch process has been applied, allowing the bacteria to grow along a predefined curve. Moreover, a defined medium was used during cultivations and each inoculum has been prepared from the same RCB, in order to minimize the batch to batch variation. By means of the final large cultivations 130417B and 130424B, the reproducibility of the final cultivation protocol was found to be rather high. Between the two, the expression level deviation of the product was ~3 %.

During the reference cultivation 130212B, a medium component analysis was performed. Here it was uncovered that some components could be somewhat adjusted in order to reduce risk of precipitation, in the same time saving money and protecting the environment. The suggestion is to reduce the amount of copper with 50 %, magnesium 20 % and manganese with 30 %. It is also proposed that the iron should be increased with 10 %. A more detailed description of the analysis is described in **Appendix B**.

4.2.1 Optimized cultivation conditions

As the aim of the project was to find cultivation conditions where the product concentration could be maximized, meanwhile the modifications were minimized, the optimization may have been somewhat falsely directed from the beginning. During the optimization of the two responses, the model for gluconoylation was found to be non-significant. Consequently, the optimal conditions, in regard to minimized modifications, could not be confirmed. Moreover, the optimization continued only in respect to the product concentration. In order to affirm the optima of the product, it is suggested that the optimization model is further expanded with higher levels of induction and expression time (16-18 h and 10-12.5 h, respectively). Affibody AB now has decided to continue recombinant expression using another *E. coli* host strain, where sufficient amounts of PGL are expressed and as a consequence the problems with gluconoylations are eliminated.

The optimization performed in this project contributed to a 25 % increase of the expression levels of Z08698-VDGS-ABD094, when comparing with the reference cultivation. Moreover, as the aim was to reach levels of 2 g/l culture, an expression of 3.5 g/l is regarded as a satisfying achievement. Further, the cultivation protocol was made more efficient as the induction time was reduced to 14 h compared with 20 h, enabling a 6 h shorter cultivation time. Based on the developed cultivation procedure a convenient scheme could be constructed as depicted in **Fig.19.** This allows cultivation, with a working volume of 15 l, to be performed within ~2.5 working days. The example is based on the availability of an incubation shaker with timer features, like for example Multitron® (Infors, AG).

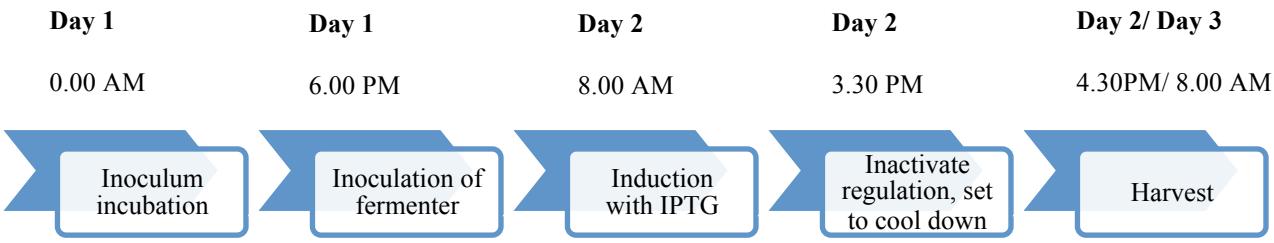


Figure 19. Scheme of the developed cultivation process including the time-points for each planned step.

According to **Appendix I,** the expression of a similar construct (Z08698-$(G4S)_4$-ABD094-$(G4S)_4$-Z08698) was found to follow the same pattern as the construct used in this project. Interestingly the expression levels are though much lower which is suggested to be explained by thermolysis not being a suitable method for releasing this larger homodimer. Theoretically the larger size enables a higher risk of the product being captured in between complexes of native precipitated proteins. When it is time for establishment of a suitable process analysis applicable to this construct, CelLytic™ is instead suggested for small scale production, while homogenisation may be suitable for large scale production.

4.3 Large-scale production

Generally, cell disruption techniques like sonication, homogenisation, or the chemical lysis agent CelLytic™ are employed at Affibody AB. Neither of these methods is particularly suitable for large-scale production. There were though possibilities of connecting a homogenisator to the 20 l fermenter, but this approach was suggested to be rather inconvenient. Moreover, the conventionally used CelLytic™ is quite expensive costing around 7000 SEK/l, and only allows maximum 25 g of cells being suspended at once [20]. As 10 ml of the agent is preferably used per gram of cells, according to earlier specifications, the total cost would be 70 SEK/g pellet. Hence, this approach is most certainly not cost-effective.

A more preferable alternative would be to use heat-treatment for disrupting the cells. In *Ren X et al. 2007* a procedure is described in which a thermostable protein is released by heating at 80 °C, in fermentation broth in situ. Perhaps would this be a possible alternative for the production of this Affibody® Molecule as well, but in this project a slightly different approach was applied. Instead a bacterial cell pellet (~1 kg) was suspended in Tris-HCl buffer, in the 20 l fermenter. The suspension was further heated to 90 °C for about 3 min with the online sterilization program of the fermenter. Next, the important step of finding an appropriate way to efficiently cool down the cell lysate was solved. Here it is recommended to quickly empty the fermenter and catch the cell lysate in large flasks that simultaneously are stirred and stored on ice-water baths during filling.

It is hypothesized that when using the 20 l fermenter for thermolysis, it is important that the reactor is filled almost entirely, in order for the temperature to be evenly distributed in the tank. If not, parts of the bacterial suspension may be subjected to too high temperatures that may further precipitate the product.

Moreover the improved thermolysis protocol recommends a dilution of 30 ml buffer per g pellet which may not always be possible for practical reasons. There has to be a consideration from case to case, whether this high volume of buffer is necessary. During the large-scale thermolysis performed in this project, approximately 19 ml/g of pellet was used. This further enabled a product yield of 2.86 g/l culture. Theoretically a culture with a working volume of 15 l ends up with approximately 18 l as a consequence of the fed-batch approach, and thus is expected to provide ~50 g of product. As the average wet weight of the cultures that the employed pellet originated from was ~86.62 g/kg, a total cell pellet of 1.56 kg is expected. These cells may further be easily thermolysed as described earlier, either all at once thus reducing the dilution, or in two rounds. The latter suggestion is recommended when the expected product yield is ≥2.8 g/l as the dilution factor was found to be a significant contributor of the product yield during thermolysis.

5. Conclusion

The project aimed to develop an optimal cultivation procedure, enabling recombinant expression of the HER3 specific Z08698-VDGS-ABD094 in *E. coli*. Initially, a suitable process analysis was identified including thermolysis for cell disruption, affinity chromatography with an anti-ABD matrix for purification, ending up with NanoDrop and HPLC-MS analysis for determining the concentration and quality of the product. Further, the thermolysis protocol was improved by applying DOE and multivariate data analysis in the software MODDE® 9.1, resulting in the thermolysis procedure being as efficient as CelLytic™, by means of product yield. As a consequence, thermolysis is suggested to be carried out by suspending each gram of bacterial pellet with 30 ml of 20 mM Tris-HCl buffer (pH = 7.5), followed by heat-treatment in a water-bath of 90 °C for 3 min.

An optimization of the cultivation process was performed in MODDE® 9.1, so that the highest expression of the ABD coupled Affibody® Molecule could be obtained, while modifications like gluconoylations could be minimized. Unfortunately, the amount of gluconoylations could only be reduced and not eliminated completely due to attributes of the host strain, later resulting in the cultivating conditions only being adjusted in favour of the expression levels of the molecule. By employing a temperature of 37 °C, pH at 6.7, induction- and expression time corresponding to 14 and 7.5 hours, respectively, during fermentation process on a 15 l scale, a product concentration of ~3.5 g/l was obtained. Thus, the optimization of the cultivating conditions has increased the expression levels with 25% and further reduced the cultivation time with 6 hours, when comparing with starting conditions.

Finally thermolysis was evaluated in a large-scale production allowing 1 kg of bacterial pellet being suspended in 19 l of Tris-HCl buffer, followed by heat-treatment in the fermenter using the on-line sterilisation program. Originating from the bacterial pellet obtained with the optimized conditions, a product concentration of ~2.8 g/l was received. In order to maximize product yield, it is suggested that the cell lysate is being cooled down quickly after treatment. Moreover, it is proposed that thermolysis is carried out directly after cultivation, fermentation mash in situ, allowing an additionally rationalized manufacturing process of Z08698-VDGS-ABD094.

6. Future studies

- Make a large scale purification of the cell lysate obtained after large-scale heat-treatment, thus confirming a final large scale production of Z08698-VDGS-ABD094.

- Perform recombinant expression of the anti-HER3 Affibody® Molecule in another host, were the gluconoylation problems can be avoided for example an *E. coli* host with a restored production of the PGL enzyme.

- Expand the model obtained from optimization with experiments spanning higher levels of induction and expression time, in order to affirm the optima of the product concentration.

- Compare the large-scale heat-treatment with for example homogenisation.

- Evaluate whether CelLytic™ contribute with a higher product yield of the homodimer (Appendix I) compared with a thermolysis procedure.

- Try to perform thermolysis directly after cultivation, fermentation mash in situ.

7. Acknowledgments

I would like to offer my special thanks to my supervisor Finn Dunås for giving me the opportunity to do my master thesis at Affibody AB. Throughout the project he has supported me, given feed-back and provided valuable knowledge. Most importantly Finn Dunås has allowed me to work independently, permitting me to make my own mistakes and discovering the ups and downs being a scientist, which will be an invaluable experience for my future career. My special thanks are extended to the staff at Affibody AB for giving me a nice stay at the company and for assisting me with their expertise. Especially I am grateful for all the help that Per Jonasson, Pär Eklund, Olof Widmark and Leena Rinnevuo have given me in the lab. Additionally, I thank Karin Stensjö at Uppsala University for the critical reviewing of this master thesis.

Finally I wish to declare my gratitude to my fiancé Roger Röjsäter for always supporting me and understanding the dedication to my work. I also thank my family for believing in me and allowing me to make my own choice of career, without any pressure.

References

1. Malm M, Kronqvist Nina, Lindberg H, Gudmundsdotter L, Bass T, Frejd F Y, Höidén-Guthenberg I, Varasteh Z, Orlova A, Tolmachev V, Ståhl S, Löfblom J. 2013. Inhibiting HER3-mediated tumor cell growth with Affibody® Molecules engineered to low picomolar affinity by position-directed error prone PCR-like diversification. PLOS ONE 8: 1-13

2. Kronqvist N, Malm M, Göstring L, Gunneriusson E, Nilsson M, Höidén-Guthenberg I, Gedda L, Frejd F Y, Ståhl S, Löfblom J. 2011. Combining phage and staphylococcal surface display for generation of ErbB3-specific Affibody® Molecules. Protein Engineering, Design & Selection 24: 385-396

3. Göstring L, Malm M, Höidén-Guthenberg I, Frejd F Y, Ståhl S, Löfblom J, Gedda L. 2012. Cellular effects of HER3 specific Affibody® Molecules. PLOS ONE 7:1-9

4. Löfblom J, Feldwisch J, Tolmachev V, Carlsson J, Ståhl S, Frejd F Y. 2010. Affibody® Molecules: Engineered proteins for therapeutic, diagnostic and biotechnological applications. FEBS letters 584: 2670-2680

5. Feldwisch J, Tolmachev V. 2012. Engineering of Affibody® Molecules for therapy and diagnostics. Methods in Molecular Biology 899:103-126

6. Frejd F Y. 2012. Half-life extension by binding to albumin through an albumin binding domain. In: Kontermann R (eds). Therapeutic proteins: strategies to modulate their plasma half-lives, p.269-283. Wiley-VCH Verlag GmbH & Co.

7. Hincal F. 2009. An introduction to safety-issues in biosimilars/follow-on biopharmaceuticals. Journal of Medical Chemical, Biological and Radiological Defense 7: 1-18

8. Sorensen H P, Mortensen K K. 2005. Advanced strategies for recombinant protein expression in Escherichia coli. Journal of Biotechnology 115: 113-128

9. Jenzsch M, Gnoth S, Kleinschmidt M, Simutis R, Lubbert A. 2006. Improving the batch-to-batch reproducibility in microbial cultures durin recombinant protein production by guiding the process along a predefined total biomass profile. Bioprocess and Biosystem Engineering 29: 315-321

10. Lee S Y. 1996. High cell-density culture of Escherichia coli. TIBTECH 14:98-105

11. Studier W, Moffatt B. 1986. Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-level Expression of Cloned Genes. Journal of Molecular Biology 189:113-130

12. Lenzak J L, Schmidell W, Falcao de Aragou G M. 2013. High-cell-density culture strategies for polyhydroxyalkanoate production: a review. Industrial Microbiology and Biotechnology 40:275-286

13. Lee J, Lee S Y, Park S, Middelberg P.J. A. 1999. Control of fed-batch fermentations. Biotechnology Advances 17: 29-48

14. Aon JC, Caimi R J, Taylor A H, Lu Q, Oluboyede F, Dally J, Kessler M D, Kerrigan J J, Lewis T s, Wysocki L A, Patel P s. 2008. Suppressing posttranslational gluconoylation of heterologous proteins by metabolic engineering of Escherichia coli. Applied and environmental microbiology 74:950-958

15. Vassilera-Atanassova A, Mironova R, Nacheva G, Ivanov Ivan. 1999. N-terminal methionine in recombinant proteins expressed in two different *Escherichia coli* strains. Journal of Biotechnology 69: 63-67

16. Eriksson L, Johansson E, Kettaneh-Wold N, Wikström C, Wold S. 2008. Design of experiments principles and applications. 3 vol. Umetrics Academy

17. Umetrics. 2009. Tutorial robustness (HPLC)

18. Ren X,, Yu D, Han S, Feng Y. 2007. Thermolysis of recombinant Escherichia coli for recovering a thermostable ezyme. Biochemical Engineering Journal 33: 94-98

19. Jonasson Per, Affibody AB

20. Sigma-Aldrich Co. 2013. B7435 Sigma CelLytic™ B Cell Lysis Reagent: http://www.sigmaaldrich.com/catalog/product/sigma/b7435?lang=en&region=SE, 2013-05-26

Appendix A: Medium

Table 11. *Composition of the animal-component free Select APS™ LB-medium*	
Chemical	**Concentration (g/l)**
Soy Hydrolysate	2.5
Yeast Extract	12.5
Sodium Chloride	5.0

Table 12. *Composition of 100 ml of defined shake flask medium*	
Chemical	**Amount**
Milli-Q-water	80 ml
10 × (Phosphate + Citrate)	10 ml
K_2HPO_4	70 g/l
$Na_3C_6H_5O_7 * H_2O$	10 g/l
10 × (YNB + Glucose)	10 ml
YNB (Yeast Nitrogen Base)	67 g/l
$C_6H_{12}O_6 * H_2O$	55 g/l

Table 13. *Composition of 5 l of defined cultivation medium*	
Chemical	**Amount**
Milli-Q-water	4625 ml
$(NH_4)_2SO_4$	18.75 g
20 × (Phosphate + Citrate)	375 ml
K_2HPO_4	40 g/l
KH_2PO_4	60 g/l
$Na_3C_6H_5O7 * 2 H_2O$	25 g/l
Glucose (60% solution)	6.832 ml
Trace element solution	7.5 ml
$FeCl_3*6 H_2O$	35 g/l
$ZnSO_4 * 7 H_2O$	10.56 g/l
$CuSO_4 * 5 H_2O$	2.64 g/l
$MnSO_4 * H_2O$	13.2 g/l
$CaCl_2 * 2 H_2O$	13.84 g/l
1.217 M Magnesium Sulphate solution	25 ml
$MgSO_4 * 7 H_2O$	300 g/l
Kanamycin (50 mg/ml)	5 ml

Appendix B. Medium component analysis

Table 14. *Tabulation of medium component analysis.*
$Sample_{start}$ and $Sample_{end}$ are medium samples taken pre and post cultivation from130212B. Analysis of the samples was performed at ALS Scandinavia AB. Initial concentration refers to the assumed concentration based on the medium recipe and the deviation is thus the difference between the initial concentration and the measured concentration of $Sample_{start}$. Consumption refers to the proportion consumed ($Sample_{start}$ - $Sample_{end}$ in relation to the concentration of $Sample_{start}$

Component	Initial concentration (mg/kg)	$Sample_{start}$ (mg/kg)	Deviation (%)	$Sample_{end}$ (mg/kg)	Consumption (%)
Ca	5.66	<30		<30	
Co	0.00	<0.004		<0.004	
Cu	1.01	1.06	5.16	0.7	33.96
Fe	10.85	10.9	0.48	0.787	92.78
Mg	147.93	113	-23.61	33.5	70.35
Mn	6.44	4.3	-33.21	1.73	59.77
P	1557.40	1430	-8.18	292	79.58
Zn	4.34	3.29	-24.16	0.694	78.91

Table 15. *Tabulation of the suggested adjustment of the medium and the resulting expected consumption based on **Table 13.***

Component	Decrease (%)	New concentration (mg/kg)	Expected consumption (mg/kg)	Expected consumption (%)
Cu	-50.00	0.53	0.36	67.92
Fe	10.00	11.99	10.11	84.35
Mg	-20.00	90.40	79.50	87.94
Mn	-30.00	3.01	2.57	85.38

Appendix C. Summary of cultivations

Table 16. *All cultivations are listed below with their corresponding expression levels of Z08698-VDGS-ABD094, obtained after 7.5 h of cultivations post induction.*

Cultivation	V (l)	Temp. (°C)	pH	I-time (h)	Cult. Time (h)	OD_{600}	Expr. (g/l)	Description
121010B	20	37	-	(OD_{600}=2.3)	-	8.6	-	*Performed by supervisor*
130212B	18	33	6.95	20	27.5	136	2.89	*Reference*
130212H1	0.6	30	6.70	18	25.5	130	2.71	*Finding factor ranges Before optimization*
130212H2	0.6	30	6.70	18	25.5	120	2.69	
130212H3	0.6	33	6.95	20	27.5	122	2.50	
130212H4	0.6	33	6.95	20	27.5	140	2.92	
130212H5	0.6	36	7.20	22	29.5	130	3.48	
130212H6	0.6	36	7.20	22	29.5	110	2.51	
130308H1	0.6	35	6.95	12	19.5	88	4.24	*Exploring the expression times´ effect on modifications*
130308H2	0.6	35	6.95	14	21.5	75	3.66	
130308H3	0.6	35	6.95	16	23.5	122	4.09	
130308H4	0.6	35	6.95	18	25.5	100	2.29	
130308H5	0.6	35	6.95	20	27.5	140	2.95	
130308H6	0.6	35	6.95	22	29.5	113	1.41	
130318H1	0.6	33	6.70	10	17.5	65	2.96	*Optimization round 1*
130318H2	0.6	37	6.70	10	17.5	64	3.01	
130318H3	0.6	33	7.20	10	17.5	30	1.30	
130318H4	0.6	37	7.20	10	17.5	30	1.83	
130318H5	0.6	35	6.95	10	17.5	49	2.85	
130318H6	0.6	35	6.95	12	19.5	73	3.87	
130404H1	0.6	35	6.95	12	19.5	72	3.87	*Optimization round 3*
130404H2	0.6	35	6.95	12	19.5	69	3.34	
130404H3	0.6	35	7.20	12	19.5	51	2.05	
130404H4	0.6	33	6.95	12	19.5	88	3.48	
130404H5	0.6	35	6.70	12	19.5	80	3.71	
130404H6	0.6	37	6.95	12	19.5	86	4.45	
130410H1	0.6	33	6.70	14	21.5	99	4.24	*Optimization round 2*
130410H2	0.6	37	6.70	14	21.5	23	0.85	
130410H3	0.6	33	7.20	14	21.5	100	3.89	
130410H4	0.6	37	7.20	14	21.5	78	4.07	
130410H5	0.6	35	6.95	12	19.5	90	3.73	
130410H6	0.6	35	6.95	14	21.5	78	3.80	
130417B	15	37	6.70	14	21.5	77	3.55	*Large-scale production 1*
130424B	15	37	6.70	14	21.5	75	3.43	*Large-scale production 2*

Appendix D. Optimization cultivations

Table 17. *Below the conditions for cultivation during optimization are listed together with the found expression levels (Expr.) and amount of modified product (Mod.). Gluc. Dev refers to the glucose deviation found for the fermenter after cultivation, which further were considered as an uncontrolled factor during optimization*

Fermenter & date	Temp (°C)	pH	I-time (h)	E-time (h)	Gluc. Dev (%)	OD_{600}	Expr. (g/l)	Mod. (%)
130318								
H1	33	6.70	10	2.5	-1.4	40	0.90	0
	33	6.70	10	5.0		54	2.14	0
	33	6.70	10	7.5		65	2.96	0
H2	37	6.70	10	2.5	-9.2	33	0.99	3.43
	37	6.70	10	5.0		44	2.59	0
	37	6.70	10	7.5		64	3.01	0
H3	33	7.20	10	2.5	3.9	22	0.39	0
	33	7.20	10	5.0		28	1.03	0
	33	7.20	10	7.5		30	1.31	0
H4	37	7.20	10	2.5	-3.3	22	0.89	5.99
	37	7.20	10	5.0		33	1.78	12.05
	37	7.20	10	7.5		30	1.83	
H5	35	6.95	10	2.5	-3.0	27	1.03	0
	35	6.95	10	5.0		35	1.83	0
	35	6.95	10	7.5		49	2.85	0
H6	35	6.95	12	2.5	-4.9	50	1.49	0
	35	6.95	12	5.0		62	2.37	0
	35	6.95	12	7.5		72	3.86	0
130410								
H1	33	6.70	14	2.5	12.1	72	2.38	7.86
	33	6.70	14	5.0		86	3.51	6.4
	33	6.70	14	7.5		99	4.24	4.76
H2	37	6.70	14	2.5	6.9	11	0.82	3.74
	37	6.70	14	5.0		16	0.90	11.32
	37	6.70	14	7.5		23	0.84	15.54
H3	33	7.20	14	2.5	16.8	77	1.07	4.34
	33	7.20	14	5.0		88	2.44	5.61
	33	7.20	14	7.5		100	3.89	5.03
H4	37	7.20	14	2.5	10.3	65	1.38	4.11
	37	7.20	14	5.0		80	2.73	4.74
	37	7.20	14	7.5		78	4.07	11.09
H5	35	6.95	12	2.5	13.0	49	0.87	4.65
	35	6.95	12	5.0		70	3.07	0
	35	6.95	12	7.5		90	3.73	3.73
H6	35	6.95	14	2.5	4.5	62	2.41	5.08
	35	6.95	14	5.0		71	3.39	4.47
	35	6.95	14	7.5		78	3.80	9.71

Fermenter & date	Temp (°C)	pH	I-time (h)	E-time (h)	Gluc. Dev (%)	OD_{600}	Expr. (g/l)	Mod. (%)
H1	35	6.95	12	2.5	-0.9	55	1.06	4.19
	35	6.95	12	5.0		70	2.80	4.96
	35	6.95	12	7.5		72	3.87	6.47
H2	35	6.95	12	2.5	-6.7	60	1.12	3.79
	35	6.95	12	5.0		65	2.74	0
	35	6.95	12	7.5		69	3.34	4.57
H3	35	7.20	12	2.5	4.7	27	1.34	0
	35	7.20	12	5.0		30	1.64	0
	35	7.20	12	7.5		51	2.05	6.84
H4	33	6.95	12	2.5	-2.1	63	1.30	3.95
	33	6.95	12	5.0		74	2.88	4.09
	33	6.95	12	7.5		88	3.48	3.62
H5	35	6.70	12	2.5	1.0	53	1.51	3.45
	35	6.70	12	5.0		79	2.98	0
	35	6.70	12	7.5		80	3.71	4.18
H6	37	6.95	12	2.5	-10.4	57	1.59	0
	37	6.95	12	5.0		79	3.49	3.51
	37	6.95	12	7.5		86	4.45	4.61

Appendix E. Relation between OD_{600} and wet weight

Based on the wet weight and OD_{600} measurements of a few selected cultivations listed in **Table 18**, a linear relation between the two could be estimated according to **Fig.20.** Thus, the wet weight of each sample during optimization could be predicted according to the linear function obtained in the same figure; y = 1.0833x, where *y* is the estimated wet weight and *x* the measured OD_{600}.

Table 18. *Measured OD_{600} and wet weights of selected cultivations*

Cultivation	OD_{600}	Wet weight (g/kg)
Origo	0	0
130318H1	74	80
130318H2	66	65.2
130318H3	30	38.6
130318H4	34	44.7
130318H5	43	60.8
130318H6	72	87.7
130404H1	100	87.8
130404H2	89	84.8
130404H3	40	42.9
130404H4	88	109.8
130404H5	103	101.9
130404H6	82	93.3
130410H1	108	124.6
130410H2	26	30.7
130410H3	135	149.2
130410H4	113	112.7
130410H5	100	119.1
130410H6	97	104
130417B	81	87.36

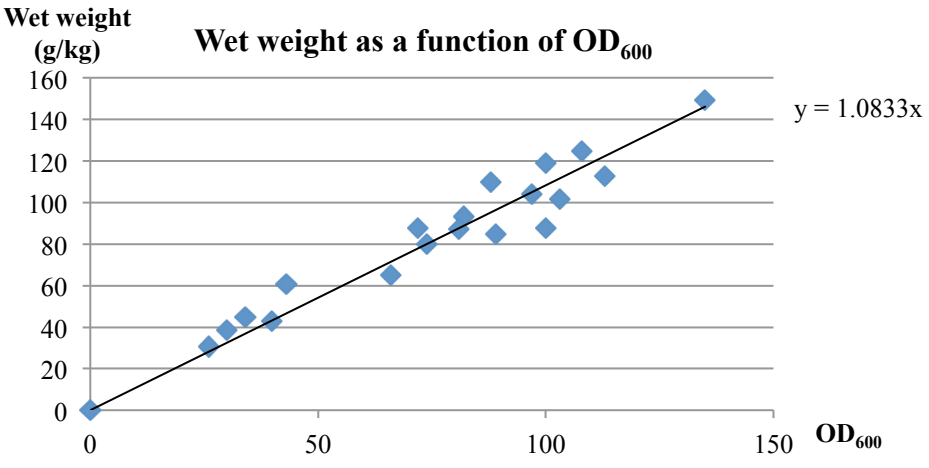


Figure 20. Linear correlation between OD_{600} and wet weight based on selected cultivations in Table 18.

Appendix F. Purification protocol

The different parts during purifications are listed below. They consist of *regeneration of columns* which have to be performed each time before usage, *wash & elution* and finally the procedure of *cleaning in place* that enables the columns to be reused. Column volume (CV) refers to the volume of the stationary phase used, in this case 0.4 ml of anti-ABD matrix.

Regeneration of columns

- 4 CV 1 × TST
- 4 CV 0.1 M HAc, pH = 2.88
- 4 CV 1 × TST

Wash & Elution

- Loading of sample
- Wash with 5 CV 1 × TST
- Wash with 5 CV 5 mM NH_4AC, pH =5 .5
- Elution with 0.1 M HAc;
    - Add 200 µl and discard the flow through
    - Add 4 CV and collect the eluent in a microcentrifuge tube.

Cleaning in place (CIP)

- 4 CV 1 × TST
- 3 CV 0.5 M NaOH
- 5 CV 1 × TST
- Store cold in 4 CV 20 % EtOH in 1 × TST buffer

Appendix G. Pre-screening of thermolysis

Below a screening investigation of thermolysis is presented. As it was found that an improper DOE had been utilized for the purpose, a new one was constructed with the intention of fitting the received data to an interaction model. The already gained responses were implemented in the new more suitable design comprising of a full factorial design.

The quantitative factors that were investigated were dilution of the bacterial pellet with 20 mM Tris-HCl buffer, temperature in water bath during heat-treatment and the duration of the treatment. Moreover, a qualitative factor was inspected, if cooling the samples on ice directly after treatment was necessary or if room temperature was enough. The response explored constituted of the measured absorbance at 280 nm, which further could be transformed in terms of product concentration (**Eq. 3**). Low, centre point and high levels of the given factors are represented in **Table 19.**

Table 19. *Factor ranges during screening of thermolysis*

Factor	Low	Centre	High
Dilution (ml/g)	5.00	6.25	7.50
Temperature (°C)	93	95	97
Time (min)	6.0	7.5	9.0
Cooling (RT/Ice)	RT	Ice	Ice

As the data was fitted to the model with MLR, the result was a weakly significant model with $Q^2 = 0.604$ and $R^2 = 0.249$ (represented in **Fig.21**). The model validity represented in the same figure is though below 0.5, which further raises concerns to the model. The ANOVA statistics suggests that the F-test for regression was fulfilled with p-value = 0.021 and thus the obtained model can be regarded as statistically significant. According to the coefficient plot in **Fig.21**, the factors dilution and cooling are the most influential ones. Moreover cooling down in RT seems to have a negative impact on the response while cooling down on ice is of favour of the response.

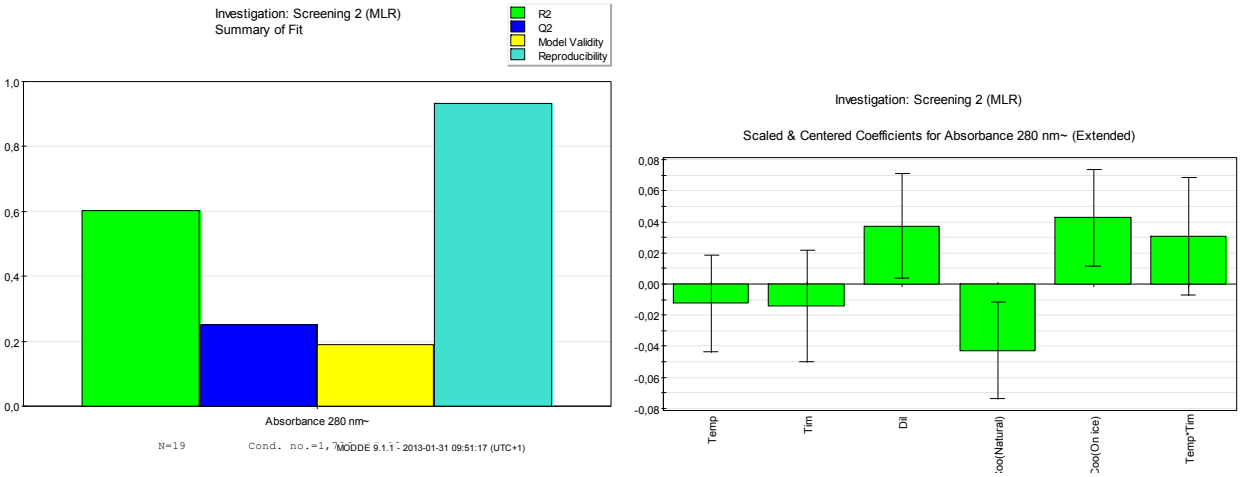


Figure 21 The summary plot (left) with staples representing R^2, Q2, Model validity and reproducibility. The coefficient plot illustrates the model terms from left to right comprising of; Temp, Time, Dilution, Cooling (RT), cooling (ice) and Temp*Time

Based on the model a contour plot was constructed, depicted in **Fig.22**. Interpretation of the model concluded that low levels of time and temperature together with high levels of dilution are suggested. Further the qualitative factor cooling seems to have a great impact on the product concentration, where cooling on ice is the most preferable alternative.

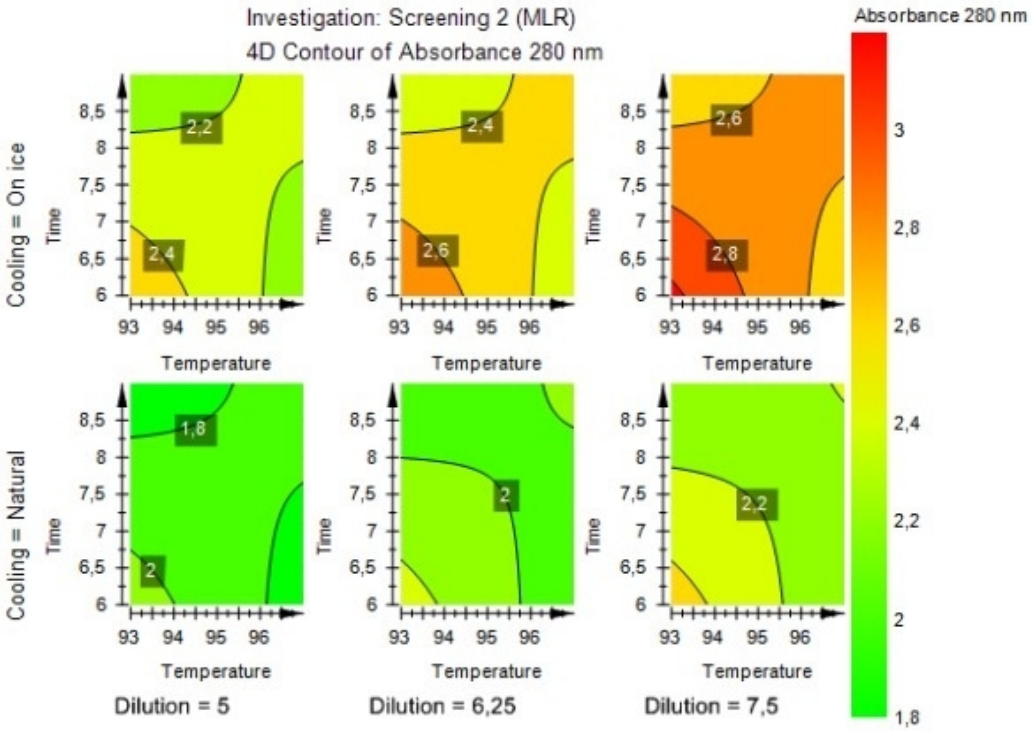


Figure 22. Contour plot where the upper row of plots are based on cooling down on ice while the bottom row represents cooling down in RT. The Major x-axis constitutes of the factor dilution, where columns left to right is 5, 6.25 and 7 ml/g respectively. The response A280 is implicated by the color scheme where green refers to lower and red higher A280.

Appendix H. Worksheet from MODDE®

Below, the worksheets from each investigation are listed. The three experiments in the worksheet for the optimization investigation that are marked in grey are the ones that got excluded from the model as their originating cultivation was unsuccessful.

Table 20. *Worksheet from the screening investigation of thermolysis*

		Factors		***Response***
Exp No	Temp °C	Dilution (ml/g)	Time (min)	A280
1	89	20	3.0	2.083
2	93	20	3.0	2.283
3	89	30	3.0	3.190
4	93	30	3.0	2.710
5	89	20	6.0	2.730
6	93	20	6.0	2.927
7	89	30	6.0	2.930
8	93	30	6.0	3.175
9	89	25	4.5	2.801
10	93	25	4.5	2.789
11	91	20	4.5	2.675
12	91	30	4.5	3.257
13	91	25	3.0	2.679
14	91	25	6.0	2.740
15	91	25	4.5	2.733
16	91	25	4.5	2.471
17	91	25	4.5	2.651
18	91	25	1.0	2.704
19	91	25	2.0	2.715
20	91	25	4.0	2.419
21	91	25	5.0	2.759

Table 21. *Worksheet obtained for the robustness testing of thermolysis*

		Factors		***Response***
Exp No	Temp °C	Time (min)	Dilution (ml/g)	A280
1	89	2.5	31	2.739
2	91	2.5	29	2.632
3	89	3.5	29	2.744
4	91	3.5	31	2.871
5	90	3.0	30	2.653
6	90	3.0	30	2.815
7	90	3.0	30	2.811

Table 22. *Worksheet for the optimization investigation of cultivation conditions*

	Factors				***Response***	
Exp No	Temperature	pH	I-time (h)	E- time (h)	Concentration g/l	Gluconoylation
1	33	6.70	10	2.5	0.901536	0
2	37	6.70	10	2.5	0.987741	0
3	33	7.20	10	2.5	0.393178	0
4	37	7.20	10	2.5	0.890528	5.99
5	33	6.70	14	2.5	2.37907	7.86
6	37	6.70	14	2.5	0.820644	3.74
7	33	7.20	14	2.5	1.06681	4.34
8	37	7.20	14	2.5	1.37889	4.11
9	33	6.70	10	7.5	2.9611	0
10	37	6.70	10	7.5	3.00854	3.43
11	33	7.20	10	7.5	1.30538	0
12	37	7.20	10	7.5	1.83025	0
13	33	6.70	14	7.5	4.24176	4.76
14	37	6.70	14	7.5	0.844707	8.63
15	33	7.20	14	7.5	3.88546	5.03
16	37	7.20	14	7.5	4.06932	11.09
17	33	6.95	12	5.0	2.87904	4.09
18	37	6.95	12	5.0	3.49017	3.51
19	35	6.70	12	5.0	2.98039	0
20	35	7.20	12	5.0	1.64307	0
21	35	6.95	10	5.0	1.83256	0
22	35	6.95	14	5.0	3.391	4.47
23	35	6.95	12	2.5.	1.49359	0
24	35	6.95	12	7.5	3.86587	0
25	35	6.95	12	5.0	2.36553	0
26	35	6.95	12	5.0	3.06683	0
27	35	6.95	12	5.0	2.79521	4.96
28	33	6.70	10	5.0	2.13646	0
29	37	6.70	10	5.0	2.58584	0
30	33	7.20	10	5.0	1.02547	0
31	37	7.20	10	5.0	1.77664	12.08
32	33	6.70	14	5.0	3.51276	6.4
33	37	6.70	14	5.0	0.897267	6.62
34	33	7.20	14	5.0	2.43917	5.61
35	37	7.20	14	5.0	2.73093	4.74
36	33	6.95	12	2.5	1.30415	3.95
37	33	6.95	12	7.5	3.48013	3.62
38	37	6.95	12	2.5	1.58918	0
39	37	6.95	12	7.5	4.45288	4.61
40	35	6.70	12	2.5	1.508	3.45
41	35	6.70	12	7.5	3.70939	4.18

Exp No	Temperature	pH	I-time (h)	E- time (h)	Concentration g/l	Gluconoylation
42	35	7.20	12	2.5	1.33638	0
43	35	7.20	12	7.5	2.04741	6.84
44	35	6.95	10	2.5	1.02646	0
45	35	6.95	10	7.5	2.84995	0
46	35	6.95	14	2.5	2.4074	5.08
47	35	6.95	14	7.5	3.79999	9.71
48	35	6.95	12	2.5	0.861895	4.65
49	35	6.95	12	7.5	3.73076	3.73
50	35	6.95	12	2.5	1.05957	4.19
51	35	6.95	12	7.5	3.86645	9.28
52	35	6.95	12	2.5	1.12102	3.79
53	35	6.95	12	5.0	2.73833	0
54	35	6.95	12	7.5	3.33971	4.57

Appendix I. Optimization tested on similar construct

In order to determine how general the optimized cultivation conditions are, cultivations were performed based on the set conditions, while expressing a similar construct. This similar construct comprised of a homodimer of the anti-HER3 Affibody® Molecule, denoted Z08698-$(G4S)_4$-ABD094-$(G4S)_4$-Z08698.

Transformation of the construct to electro competent *E. coli* cells, and further RCB production, was made according to earlier depiction for the monomer. The cultivations were performed in the GRETA multifermenter system, as describes earlier, and the reactors were set to reflect the conditions that generated the worst, medium-well and best expression levels of Z08698-VDGS-ABD094 after 7.5 h of cultivation. The conditions are listed in **Table 23**.

Table 23. *Cultivation conditions in each reactor*

Fermenter	Temperature (°C)	pH	Induction time (h)	Resulting expression for monomer (g/l)
H1	33	7.20	10	1.30538
H2	37	7.20	10	1.83025
H3	35	6.70	12	3.70939
H4	35	6.95	12	3.73076
H5	37	6.95	12	4.45288
H6	37	6.70	14	*Not found during optimization, but suggested as the best according to model*

Samples were collected after 2.5, 5, 7.5 and 10 h, respectively, and processed as earlier described. The Expression levels found are illustrated in **Fig.23**. It was found that the homodimer showed the same pattern of expression concerning the worst, medium-well and best expression levels of the monomer. However, the expression levels were dramatically reduced, implying that heat-treatment may not be a suitable lysis method for releasing the homodimer from the cells after cultivation.

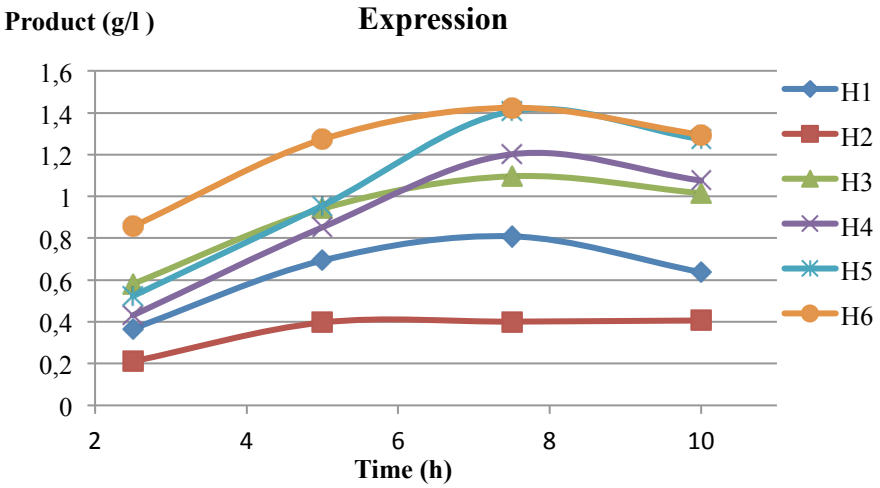


Figure 23. Expression levels of Z08698-$(G4S)_4$-ABD094-$(G4S)_4$-Z08698 according to the cultivation conditions depicted in **Table 23**.