(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
03 September 2020 (03.09.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/173817 A1

(51) International Patent Classification:
C12N 15/62 (2006.01) *C12N 9/20* (2006.01)
C07K 1/22 (2006.01) *C07K 14/465* (2006.01)

(21) International Application Number:
PCT/EP2020/054578

(22) International Filing Date:
21 February 2020 (21.02.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
19159854.9 28 February 2019 (28.02.2019) EP

(71) Applicant: **NOVOZYMES A/S** [DK/DK]; Krogshoejvej 36, 2880 Bagsvaerd (DK).

(72) Inventors: **MOURITSEN, Jeppe, Christian**; Aksel Moellers Have 15, 2.tv, 2000 Frederiksberg (DK). **ARNAU, José**; Krogshoejvej 36, 2880 Bagsvaerd (DK). **NOERREGAARD-SARUP, Rikke**; Praestelaengen 25, 1TH, 2400 Copenhagen NV (DK).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *with sequence listing part of description (Rule 5.2(a))*

WO 2020/173817 A1

(54) Title: CALCITE BINDING PROTEINS

(57) Abstract: Disclosed is proteins having calcite binding properties prepared by fusing a protein to a calcite binding sequence. Further, method of recovering such proteins from fermentation broths are disclosed and compositions comprising such proteins.

CALCITE BINDING PROTEINS

Reference to sequence listing

This application contains a Sequence Listing in computer readable form. The computer readable form is incorporated herein by reference.

FIELD OF THE INVENTION

The present invention relates to methods of preparing proteins having the ability to bind calcite. Further, the invention relates to methods of stabilizing, purifying, recovering and/or formulating proteins comprising a calcite binding sequence by binding to calcite.

BACKGROUND OF THE INVENTION

Industrial protein production is typically done by fermenting a recombinant host cell comprising a genetic construct encoding the desired protein. Such processes are well known for the production of many different proteins for example industrial enzymes and pharmaceuticals.

After the fermentation, the desired product has to be recovered from the fermentation broth, a process typically involves a number a separation and purification techniques depending on the particular protein and the purity required for the intended purpose.

In order to facilitate the recovery process, several techniques have been developed, for example providing the protein with binding sequence or a tag, a short sequence that binds to a specific binding partner, and in this way facilitate recovery. Examples of this technology includes His-tags, biotin-streptavidin and Protein A- IgG binding pairs. These solutions have found broad use in particular in research and small scale use, but have the disadvantage that the binding partner typically is not cost efficient, toxic or environmentally unfriendly; which has limited the use in large scale industrial production, in particular in productions outside the pharmaceutical area where the value of the proteins are more moderate.

Demarchi et al. https://elifesciences.org/content/5/e17092/abstract demonstrated that protein molecules in ancient Ostrich egg shells are preserved over several millions of years in the hot and humid conditions in Africa and Tanzania. It was suggested based on molecular dynamics simulations that these proteins were tightly associated to the calcite mineral of the egg shells.

SUMMARY OF THE INVENTION

The invention provides a polypeptide comprising a sequence encoding a desired polypeptide and a sequence having calcite binding properties, wherein the calcite binding sequence has the sequence of:

YHHGEEED (SEQ ID NO: 3);

LDDDDYPKG (SEQ ID NO: 4);

or having the sequence of SEQ ID NO: 3 or 4 with 1, 2, 3 or 4 amino acid substitutions.

Preferably the desired polypeptide has biological activity.

In another aspect the invention related to a method for recovering a polypeptide of the invention from an aqueous solution comprising the polypeptide of the invention comprising the steps of:

a) Providing an aqueous solution of a polypeptide according to the invention;
b) Adding calcite to the aqueous solution of the polypeptide, and mixing;
c) Optionally holding the mixture for a period to allow the polypeptide to bind to the calcite; and
d) Separating the calcite with the bound polypeptide from the mixture.

Preferably the aqueous solution in step a) is a fermentation broth.

Further, the invention relates to polynucleotides, expression constructs and vectors comprising said polynucleotide and host cells comprising these as well as methods for preparing the polypeptides of the invention using such host cells.

The invention relates to composition comprising the polypeptide of the invention, in particular comprising the polypeptide of the invention bound to calcite.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1 shows a SDS-PAGE gel; to the left is shown the molecular weight marker; lane 1 (Mark12) shows the molecular weight marker, Lane 2 shows culture supernatant from an *Aspergillus oryzae* expressing the DORA2 construct without added calcite; and lane 3 shows same culture fluid but with addition of calcite. For further details see example 2 and 3.

Figure 2 shows test tubes with culture fluid and calcite illuminated with green light, where the DsRed protein flouresses orange. To the left two test tubes with culture fluid from an *A. oryzae* strain expressing the DORA 2 construct before and after centrifucation and to the right two test tubes with culture fluid from an *A. oryzae* expressing the DORA 6 construct (Control) before and after centrifucation. For further information see example 3.

Figure 3 shows an SDS-PAGE gel disclosing the results of example 5. First and last lane is the molecular weight marker, for definition of the bands see figure 1. Lane 2 is purified T1 construct; lane 3 is purified T1 construct with addition of calcite. Lane 4 is wash solution after wash of the calcite with bound T1 construct. Lane 5 is T1 construct released from calcite with EDTA treatment.

BRIEF DESCRIPTION OF THE SEQUENCES

SEQ ID NO: 1: the amino acid sequence of Sca 1

SEQ ID NO: 2: the amino acid sequence of Sca 2

SEQ ID NO: 3: Binding sequence BS1.

SEQ ID NO: 4: Binding sequence BS2

SEQ ID NO: 5: amino acid ZEN-DsRed-BS2 (DORA 2) construct.

SEQ ID NO: 6: Primer oKRN-11.

SEQ ID NO: 7: Primer oKRN-12.

SEQ ID NO: 8: Primer oKRN-17.

SEQ ID NO: 9: Primer oKRN-14

SEQ ID NO: 10: Primer oKRN-15

SEQ ID NO: 11: Primer oKRN-18

SEQ ID NO: 12: Amino acid sequence of ZEN-DsRed (DORA6) construct.

SEQ ID NO: 13: Primer oKRN-26

SEQ ID NO: 14: Primer oKRN-25

SEQ ID NO: 15: Primer oKRN-27

SEQ ID NO: 16: Amino acid sequence of T1 construct: Lipase-BS1

SEQ ID NO: 17: Amino acid sequence of T2 construct: Lipase-BS2

SEQ ID NO: 18: Amino acid sequence of T3 construct: Lipase-BS1-BS2

SEQ ID NO: 19: Amino acid sequence of T4 construct: Lipase

DEFINITIONS

Calcite: Calcite is a carbonate mineral and the most stable polymorph of Calcium carbonate ($CaCO_3$). Calcite is a common constituent of sedimentary rocks and limestone.

Coding sequence: The term “coding sequence” means a polynucleotide, which directly specifies the amino acid sequence of a polypeptide. The boundaries of the coding sequence are generally determined by an open reading frame, which begins with a start codon such as ATG, GTG, or TTG and ends with a stop codon such as TAA, TAG, or TGA. The coding sequence may be a genomic DNA, cDNA, synthetic DNA, or a combination thereof.

Control sequences: The term “control sequences” means nucleic acid sequences necessary for expression of a polynucleotide encoding a mature polypeptide of the present invention. Each

control sequence may be native (*i.e.*, from the same gene) or foreign (*i.e.*, from a different gene) to the polynucleotide encoding the polypeptide or native or foreign to each other. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the polynucleotide encoding a polypeptide.

Calcite binding property: mean that the protein in question has the ability to bind to calcite. Calcite binding can according to the invention be determined by providing an aqueous solution of the protein in question, adding calcite to the solution and determining whether the protein has bound to calcite or not. A convenient method is to separate the solid calcite from the solution by centrifugation or filtration and determining the amount of protein in the solution or on the calcite. A polypeptide is according to the invention considered to have calcite binding properties if at least 20% of the protein is bound to calcite using such an assay, preferably at least 50%; preferably at least 60%, preferably at least 70%; preferably at least 80% or preferably at least 90%.

Expression: The term “expression” includes any step involved in the production of a polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion.

Expression vector: The term “expression vector” means a linear or circular DNA molecule that comprises a polynucleotide encoding a polypeptide and is operably linked to control sequences that provide for its expression.

Fermentation broth: The term fermentation broth is intended to mean the broth resulting from fermenting a microorganism in a substrate, preferably a liquid substrate. Thus, a fermentation broth is the primary outcome of a fermentation of a microorganism e.g. a fermentation process for producing the polypeptide of the invention. The fermentation broth may be subjected to operations to stabilize the produced fermentation product, such as a heat treatment or addition of a chemical stabilizing agent, and/or a simple solid/liquid separation process for removing the cells and insoluble remains from the substrate.

Host cell: The term "host cell" means any cell type that is susceptible to transformation, transfection, transduction, or the like with a nucleic acid construct or expression vector comprising a polynucleotide of the present invention. The term “host cell” encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication.

Nucleic acid construct: The term "nucleic acid construct" means a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or is modified to

contain segments of nucleic acids in a manner that would not otherwise exist in nature or which is synthetic, which comprises one or more control sequences.

Operably linked: The term "operably linked" means a configuration in which a control sequence is placed at an appropriate position relative to the coding sequence of a polynucleotide such that the control sequence directs expression of the coding sequence.

Sequence identity: The relatedness between two amino acid sequences or between two nucleotide sequences is described by the parameter "sequence identity".

For purposes of the present invention, the sequence identity between two amino acid sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *J. Mol. Biol.* 48: 443-453) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, *Trends Genet.* 16: 276-277), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix. The output of Needle labeled "longest identity" (obtained using the –nobrief option) is used as the percent identity and is calculated as follows:

(Identical Residues x 100)/(Length of Alignment – Total Number of Gaps in Alignment)

For purposes of the present invention, the sequence identity between two deoxyribonucleotide sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970, *supra*) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, *supra*), preferably version 5.0.0 or later. The parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EDNAFULL (EMBOSS version of NCBI NUC4.4) substitution matrix. The output of Needle labeled "longest identity" (obtained using the –nobrief option) is used as the percent identity and is calculated as follows:

(Identical Deoxyribonucleotides x 100)/(Length of Alignment – Total Number of Gaps in Alignment)

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to proteins and peptides having the ability to bind to calcite. Further the invention relates to methods for producing and recovering such polypeptides and to compositions comprising the polypeptides of the invention.

Polypeptides of the invention

The present invention relates to polypeptides comprising a sequence encoding a desired polypeptide and a sequence having calcite binding properties, wherein the calcite binding polypeptide has the sequence of:

Binding sequence 1 (BS1): YHHGEEEED (SEQ ID NO:3);

Binding sequence 2 (BS2): LDDDDYPKG (SEQ ID NO: 4); or

having the sequence of SEQ ID NO: 3 or 4 with 1, 2, 3 or 4 amino acid substitutions.

The sequence having calcite binding properties may be located in the N-terminal of the polypeptide, in the C-terminal of the polypeptide or located internally in the polypeptide sequence, where the N-terminal or C-terminal localization is preferred.

The polypeptide of the invention may be a simple fusion i.e. a fusion that do not contain additional amino acids belonging to either the desired polypeptide or the binding motive or the fusion may comprise a linker sequence separating the desired polypeptide and the binding motive. The linker sequence may comprise one or more amino acids and has the main purpose of separating the two elements. The exact amino acid sequence of the linker sequence is not critical for the invention but it is preferred to select a sequence comprising amino acids with hydrophilic amino acids in order to secure that the linker sequence become hydrated and exposed to the surface of the polypeptide. In principle there are no limits to the size of the linker sequence, however, it is recommended that the linker sequence contains less than 50 amino acids, e.g. up to 30 amino acids, e.g. up to 20 amino acids.

In some embodiments, the linker sequence comprises a cleavage site for a site specific protease. Such a cleavage site allow the use of the binding motive e.g. during purification of the protein of the invention; and release of the desired polypeptide from the binding motive when desired. This is particularly useful if the desired polypeptide is a therapeutic polypeptide where you can benefit from the calcite binding properties during recovery and purification of the polypeptide of the invention and thereafter release the desired therapeutic polypeptide from the binding motive to obtain the therapeutic polypeptide without additional functional sequences.

The desired polypeptide may according to the invention be any polypeptide. Preferably, the desired polypeptide is a polypeptide that do not have calcite binding properties or at least only low calcite binding in comparison with the desired polypeptide fused to either BS1 or BS2. Examples of suitable desired polypeptides according to the invention include therapeutic polypeptides, such as hormones, growth factors and antibodies; and enzymes such a hydrolases, transferases and oxidases.

In a preferred embodiment, the desired polypeptide is an enzyme selected among aminopeptidase, amylase, amyloglucosidase, mannanase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease,

esterase, galactosidase, beta-galactosidase, glucoamylase, glucose oxidase, glucosidase, haloperoxidase, hemicellulase, invertase, isomerase, laccase, ligase, lipase, lyase, mannosidase, oxidase, pectinase, peroxidase, phytase, phenoloxidase, polyphenoloxidase, protease, ribonuclease, transferase, transglutaminase, lysozyme, muramidase, xanthanase or xylanase.

In some embodiments the polypeptide of the invention comprises a desired polypeptide that has the activity of the desired polypeptide when the desired polypeptide is attached with the binding motive, whereas in other embodiments the polypeptide of the invention has no or significantly less activity than the desired polypeptide alone without the binding motive. In the latter embodiments, the desired polypeptide will be active when the binding motive is removed from the desired polypeptide.

Preferably the polypeptide of the invention has at least 25% of the activity of the desired polypeptide, e.g. least 40% of the activity, e.g. at least 50% of the activity, e.g. at least 50% of the activity, e.g. at least 70% of the activity, e.g. at least 80% of the activity or at least 90% of the activity.

The polypeptide of the invention has the benefit of being able to bind to calcite with a high specificity.

This property provides significant advantages in recovery/purification and/or formulation of the polypeptide of the invention as will be described further below.

Further, the polypeptide of the invention has increased stability compared to the desired polypeptide alone, in particular when the polypeptide of the invention is bound to calcite.

The stability of such polypeptides may be expressed as the half-life under defined conditions.

Thus, in preferred embodiments the half-life of the polypeptide of the invention is preferably increased at least 10% in comparison with the desired polypeptide alone, preferably at least 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or even at least 100% in comparison with the desired polypeptide alone.

Preparing the polypeptides of the invention

The polypeptides of the invention may be prepared as a genetic fusion comprising a desired polypeptide and a binding motive where a genetic construct encoding the desired polypeptide and the binding motive in same reading frame is provided with regulatory elements capable of driving expression of the construct in the selected host cell; or it may be prepared by chemical conjugation where the desired polypeptide is attached to the binding motive using chemical attachments or conjugation after the polypeptides have been synthetized.

Methods for preparing genetic fusions and expression thereof are known in the art and the invention is not limited to any particular method, but the skilled person will appreciate that such methods

as known in the art may be used according to the invention. Further description of methods of performing genetic fusions and expression thereof is suitable host cells can be found in Ward et al. 1990 Improved production of chymosin in Aspergillus by expression as a glucoamylase-chymosin fusion, Nat Biotech 8: 435-440; Arnau et al. 2006 Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins, Prot Expr Purif 48: 1-13 , incorporated herein by reference.

The present invention also relates to polynucleotides encoding a polypeptide of the present invention, as described herein.

The techniques used to isolate or clone a polynucleotide are known in the art and include isolation from genomic DNA or cDNA, or a combination thereof. The cloning of the polynucleotides from genomic DNA can be effected, *e.g.*, by using the well-known polymerase chain reaction (PCR) or antibody screening of expression libraries to detect cloned DNA fragments with shared structural features. See, *e.g.*, Innis *et al.*, 1990, *PCR: A Guide to Methods and Application*, Academic Press, New York. Other nucleic acid amplification procedures such as ligase chain reaction (LCR), ligation activated transcription (LAT) and polynucleotide-based amplification (NASBA) may be used.

Modification of a polynucleotide encoding a polypeptide of the present invention may be necessary for synthesizing polypeptides substantially similar to the polypeptide. The term “substantially similar” to the polypeptide refers to non-naturally occurring forms of the polypeptide.

Nucleic Acid Constructs

The present invention also relates to nucleic acid constructs comprising a polynucleotide of the present invention operably linked to one or more control sequences that direct the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences.

The polynucleotide may be manipulated in a variety of ways to provide for expression of the polypeptide. Manipulation of the polynucleotide prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying polynucleotides utilizing recombinant DNA methods are well known in the art.

The control sequence may be a promoter, a polynucleotide that is recognized by a host cell for expression of a polynucleotide encoding a polypeptide of the present invention. The promoter contains transcriptional control sequences that mediate the expression of the polypeptide. The promoter may be any polynucleotide that shows transcriptional activity in the host cell including variant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

Examples of suitable promoters for directing transcription of the nucleic acid constructs of the present invention in a bacterial host cell are the promoters obtained from the *Bacillus*

amyloliquefaciens alpha-amylase gene (*amyQ*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus licheniformis* penicillinase gene (*penP*), *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus subtilis xylA* and *xylB* genes, *Bacillus thuringiensis cryIIIA* gene (Agaisse and Lereclus, 1994, *Molecular Microbiology* 13: 97-107), *E. coli lac* operon, *E. coli trc* promoter (Egon *et al.*, 1988, *Gene* 69: 301-315), *Streptomyces coelicolor* agarase gene (*dagA*), and prokaryotic beta-lactamase gene (Villa-Kamaroff *et al.*, 1978, *Proc. Natl. Acad. Sci. USA* 75: 3727-3731), as well as the *tac* promoter (DeBoer *et al.*, 1983, *Proc. Natl. Acad. Sci. USA* 80: 21-25). Further promoters are described in "Useful proteins from recombinant bacteria" in Gilbert *et al.*, 1980, *Scientific American* 242: 74-94; and in Sambrook *et al.*, 1989, *supra*. Examples of tandem promoters are disclosed in WO 99/43835.

Examples of suitable promoters for directing transcription of the nucleic acid constructs of the present invention in a filamentous fungal host cell are promoters obtained from the genes for *Aspergillus nidulans* acetamidase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger* or *Aspergillus awamori* glucoamylase (*glaA*), *Aspergillus oryzae* TAKA amylase, , *Aspergillus oryzae* or *Aspergillus niger* translation elongation factor 1 (tef1), *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Fusarium oxysporum* trypsin-like protease (WO 96/00787), *Fusarium venenatum* amyloglucosidase (WO 00/56900), *Fusarium venenatum* Daria (WO 00/56900), *Fusarium venenatum* Quinn (WO 00/56900), *Rhizomucor miehei* lipase, *Rhizomucor miehei* aspartic proteinase, *Trichoderma reesei* beta-glucosidase, *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, *Trichoderma reesei* endoglucanase I, *Trichoderma reesei* endoglucanase II, *Trichoderma reesei* endoglucanase III, *Trichoderma reesei* endoglucanase V, *Trichoderma reesei* xylanase I, *Trichoderma reesei* xylanase II, *Trichoderma reesei* xylanase III, *Trichoderma reesei* beta-xylosidase, and *Trichoderma reesei* translation elongation factor, as well as the NA2-tpi promoter (a modified promoter from an *Aspergillus* neutral alpha-amylase gene in which the untranslated leader has been replaced by an untranslated leader from an *Aspergillus* triose phosphate isomerase gene; non-limiting examples include modified promoters from an *Aspergillus niger* neutral alpha-amylase gene in which the untranslated leader has been replaced by an untranslated leader from an *Aspergillus nidulans* or *Aspergillus oryzae* triose phosphate isomerase gene); and variant, truncated, and hybrid promoters thereof. Other promoters are described in U.S. Patent No. 6,011,147.

In a yeast host, useful promoters are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* galactokinase (GAL1), *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH1, ADH2/GAP), *Saccharomyces cerevisiae* triose phosphate isomerase (TPI), *Saccharomyces cerevisiae* metallothionein (CUP1), and *Saccharomyces cerevisiae* 3-phosphoglycerate kinase. Other useful promoters for yeast host cells are described by Romanos *et al.*, 1992, *Yeast* 8: 423-488.

The control sequence may also be a transcription terminator, which is recognized by a host cell to terminate transcription. The terminator is operably linked to the 3'-terminus of the polynucleotide encoding the polypeptide. Any terminator that is functional in the host cell may be used in the present invention.

Preferred terminators for bacterial host cells are obtained from the genes for *Bacillus clausii* alkaline protease (*aprH*), *Bacillus licheniformis* alpha-amylase (*amyL*), and *Escherichia coli* ribosomal RNA (*rrnB*).

Preferred terminators for filamentous fungal host cells are obtained from the genes for *Aspergillus nidulans* acetamidase, *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* glucoamylase, *Aspergillus niger* alpha-glucosidase, *Aspergillus oryzae* TAKA amylase, *Fusarium oxysporum* trypsin-like protease, *Trichoderma reesei* beta-glucosidase, *Trichoderma reesei* cellobiohydrolase I, *Trichoderma reesei* cellobiohydrolase II, *Trichoderma reesei* endoglucanase I, *Trichoderma reesei* endoglucanase II, *Trichoderma reesei* endoglucanase III, *Trichoderma reesei* endoglucanase V, *Trichoderma reesei* xylanase I, *Trichoderma reesei* xylanase II, *Trichoderma reesei* xylanase III, *Trichoderma reesei* beta-xylosidase, and *Trichoderma reesei* translation elongation factor.

Preferred terminators for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase, *Saccharomyces cerevisiae* cytochrome C (CYC1), and *Saccharomyces cerevisiae* glyceraldehyde-3-phosphate dehydrogenase. Other useful terminators for yeast host cells are described by Romanos *et al.*, 1992, *supra*.

The control sequence may also be an mRNA stabilizer region downstream of a promoter and upstream of the coding sequence of a gene which increases expression of the gene.

Examples of suitable mRNA stabilizer regions are obtained from a *Bacillus thuringiensis cryIIIA* gene (WO 94/25612) and a *Bacillus subtilis* SP82 gene (Hue *et al.*, 1995, *Journal of Bacteriology* 177: 3465-3471).

The control sequence may also be a leader, a nontranslated region of an mRNA that is important for translation by the host cell. The leader is operably linked to the 5'-terminus of the polynucleotide encoding the polypeptide. Any leader that is functional in the host cell may be used.

Preferred leaders for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase, amyB (neutral amylase), triose phosphate isomerase (tpiA) and *Aspergillus nidulans tpiA*.

Suitable leaders for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* 3-phosphoglycerate kinase, *Saccharomyces cerevisiae* alpha-factor, and *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP).

The control sequence may also be a polyadenylation sequence, a sequence operably linked to the 3'-terminus of the polynucleotide and, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence that is functional in the host cell may be used.

Preferred polyadenylation sequences for filamentous fungal host cells are obtained from the genes for *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* glucoamylase, *Aspergillus niger* alpha-glucosidase *Aspergillus oryzae* TAKA amylase, and *Fusarium oxysporum* trypsin-like protease.

Useful polyadenylation sequences for yeast host cells are described by Guo and Sherman, 1995, *Mol. Cellular Biol.* 15: 5983-5990.

The control sequence may also be a signal peptide coding region that encodes a signal peptide linked to the N-terminus of a polypeptide and directs the polypeptide into the cell's secretory pathway. The 5'-end of the coding sequence of the polynucleotide may inherently contain a signal peptide coding sequence naturally linked in translation reading frame with the segment of the coding sequence that encodes the polypeptide. Alternatively, the 5'-end of the coding sequence may contain a signal peptide coding sequence that is foreign to the coding sequence. A foreign signal peptide coding sequence may be required where the coding sequence does not naturally contain a signal peptide coding sequence. Alternatively, a foreign signal peptide coding sequence may simply replace the natural signal peptide coding sequence in order to enhance secretion of the polypeptide. However, any signal peptide coding sequence that directs the expressed polypeptide into the secretory pathway of a host cell may be used.

Effective signal peptide coding sequences for bacterial host cells are the signal peptide coding sequences obtained from the genes for *Bacillus* NCIB 11837 maltogenic amylase, *Bacillus licheniformis* subtilisin, *Bacillus licheniformis* beta-lactamase, *Bacillus stearothermophilus* alpha-amylase, *Bacillus stearothermophilus* neutral proteases (*nprT*, *nprS*, *nprM*), and *Bacillus subtilis prsA*. Further signal peptides are described by Simonen and Palva, 1993, *Microbiological Reviews* 57: 109-137.

Effective signal peptide coding sequences for filamentous fungal host cells are the signal peptide coding sequences obtained from the genes for *Aspergillus niger* neutral amylase, *Aspergillus niger* glucoamylase, *Aspergillus oryzae* TAKA amylase, *Humicola insolens* cellulase, *Humicola insolens* endoglucanase V, *Humicola lanuginosa* lipase, and *Rhizomucor miehei* aspartic proteinase.

Useful signal peptides for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* alpha-factor and *Saccharomyces cerevisiae* invertase. Other useful signal peptide coding sequences are described by Romanos *et al.*, 1992, *supra*.

The control sequence may also be a propeptide coding sequence that encodes a propeptide positioned at the N-terminus of a polypeptide. The resultant polypeptide is known as a proenzyme or propolypeptide (or a zymogen in some cases). A propolypeptide is generally inactive and can be converted to an active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding sequence may be obtained from the genes for *Bacillus subtilis* alkaline protease (*aprE*), *Bacillus subtilis* neutral protease (*nprT*), *Myceliophthora thermophila* laccase (WO 95/33836), *Rhizomucor miehei* aspartic proteinase, and *Saccharomyces cerevisiae* alpha-factor.

Where both signal peptide and propeptide sequences are present, the propeptide sequence is positioned next to the N-terminus of a polypeptide and the signal peptide sequence is positioned next to the N-terminus of the propeptide sequence.

It may also be desirable to add regulatory sequences that regulate expression of the polypeptide relative to the growth of the host cell. Examples of regulatory sequences are those that cause expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory sequences in prokaryotic systems include the *lac*, *tac*, and *trp* operator systems. In yeast, the ADH2 system or GAL1 system may be used. In filamentous fungi, the *Aspergillus niger* glucoamylase promoter, *Aspergillus oryzae* TAKA alpha-amylase promoter, and *Aspergillus oryzae* glucoamylase promoter, *Trichoderma reesei* cellobiohydrolase I promoter, and *Trichoderma reesei* cellobiohydrolase II promoter may be used. Other examples of regulatory sequences are those that allow for gene amplification. In eukaryotic systems, these regulatory sequences include the dihydrofolate reductase gene that is amplified in the presence of methotrexate, and the metallothionein genes that are amplified with heavy metals. In these cases, the polynucleotide encoding the polypeptide would be operably linked to the regulatory sequence.

Expression Vectors

The present invention also relates to recombinant expression vectors comprising a polynucleotide of the present invention, a promoter, and transcriptional and translational stop signals. The various nucleotide and control sequences may be joined together to produce a recombinant expression vector that may include one or more convenient restriction sites to allow for insertion or substitution of the polynucleotide encoding the polypeptide at such sites. Alternatively, the polynucleotide may be expressed by inserting the polynucleotide or a nucleic acid construct comprising the polynucleotide into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

The recombinant expression vector may be any vector (*e.g.*, a plasmid or virus) that can be conveniently subjected to recombinant DNA procedures and can bring about expression of the

polynucleotide. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vector may be a linear or closed circular plasmid.

The vector may be an autonomously replicating vector, *i.e.*, a vector that exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, *e.g.*, a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one that, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids that together contain the total DNA to be introduced into the genome of the host cell, or a transposon, may be used.

The vector preferably contains one or more selectable markers that permit easy selection of transformed, transfected, transduced, or the like cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like.

Examples of bacterial selectable markers are *Bacillus licheniformis* or *Bacillus subtilis dal* genes, or markers that confer antibiotic resistance such as ampicillin, chloramphenicol, kanamycin, neomycin, spectinomycin, or tetracycline resistance. Suitable markers for yeast host cells include, but are not limited to, ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3. Selectable markers for use in a filamentous fungal host cell include, but are not limited to, *adeA* (phosphoribosylaminoimidazole-succinocarboxamide synthase), *adeB* (phosphoribosyl-aminoimidazole synthase), *amdS* (acetamidase), *argB* (ornithine carbamoyltransferase), *bar* (phosphinothricin acetyltransferase), *hph* (hygromycin phosphotransferase), *niaD* (nitrate reductase), *pyrG* (orotidine-5'-phosphate decarboxylase), *sC* (sulfate adenyltransferase), *lysF*, *metF*, *metG*, *met H* and *trpC* (anthranilate synthase), as well as equivalents thereof. Preferred for use in an *Aspergillus* cell are *Aspergillus nidulans* or *Aspergillus oryzae amdS* and *pyrG* genes and a *Streptomyces hygroscopicus bar* gene. Preferred for use in a *Trichoderma* cell are *adeA*, *adeB*, *amdS*, *hph*, and *pyrG* genes.

The selectable marker may be a dual selectable marker system as described in WO 2010/039889. In one aspect, the dual selectable marker is an *hph-tk* dual selectable marker system.

The vector preferably contains an element(s) that permits integration of the vector into the host cell's genome or autonomous replication of the vector in the cell independent of the genome.

For integration into the host cell genome, the vector may rely on the polynucleotide's sequence encoding the polypeptide or any other element of the vector for integration into the genome by homologous or non-homologous recombination. Alternatively, the vector may contain additional polynucleotides for directing integration by homologous recombination into the genome of the

host cell at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, the integrational elements should contain a sufficient number of nucleic acids, such as 100 to 10,000 base pairs, 400 to 10,000 base pairs, and 800 to 10,000 base pairs, which have a high degree of sequence identity to the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding polynucleotides. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

Also techniques developed for increasing site specific integration in filamentous fungi such as the use of the FLP based integration in *A. niger* and/or *A. oryzae* are useful according to the present invention.

For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. The origin of replication may be any plasmid replicator mediating autonomous replication that functions in a cell. The term “origin of replication” or “plasmid replicator” means a polynucleotide that enables a plasmid or vector to replicate *in vivo*.

Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, and pACYC184 permitting replication in *E. coli*, and pUB110, pE194, pTA1060, and pAMß1 permitting replication in *Bacillus*.

Examples of origins of replication for use in a yeast host cell are the 2 micron origin of replication, ARS1, ARS4, the combination of ARS1 and CEN3, and the combination of ARS4 and CEN6.

Examples of origins of replication useful in a filamentous fungal cell are AMA1 and ANS1 (Gems *et al.*, 1991, *Gene* 98: 61-67; Cullen *et al.*, 1987, *Nucleic Acids Res.* 15: 9163-9175; WO 00/24883). Isolation of the AMA1 gene and construction of plasmids or vectors comprising the gene can be accomplished according to the methods disclosed in WO 00/24883.

More than one copy of a polynucleotide of the present invention may be inserted into a host cell to increase production of a polypeptide. An increase in the copy number of the polynucleotide can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the polynucleotide where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the polynucleotide, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, *e.g.*, Sambrook *et al.*, 1989, *supra*).

Host cells

The host cell may be a prokaryot or an eukaryot.

The prokaryotic host cell may be any Gram-positive or Gram-negative bacterium. Gram-positive bacteria include, but are not limited to, *Bacillus*, *Clostridium*, *Enterococcus*, *Geobacillus*, *Lactobacillus*, *Lactococcus*, *Oceanobacillus*, *Staphylococcus*, *Streptococcus*, and *Streptomyces*. Gram-negative bacteria include, but are not limited to, *Campylobacter*, *E. coli*, *Flavobacterium*, *Fusobacterium*, *Helicobacter*, *Ilyobacter*, *Neisseria*, *Pseudomonas*, *Salmonella*, and *Ureaplasma*.

The bacterial host cell may be any *Bacillus* cell including, but not limited to, *Bacillus alkalophilus*, *Bacillus altitudinis*, *Bacillus amyloliquefaciens*, *B. amyloliquefaciens* subsp. *plantarum*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus methylotrophicus*, *Bacillus pumilus*, *Bacillus safensis, Bacillus stearothermophilus*, *Bacillus subtilis*, and *Bacillus thuringiensis* cells.

The bacterial host cell may also be any *Streptococcus* cell including, but not limited to, *Streptococcus equisimilis, Streptococcus pyogenes, Streptococcus uberis*, and *Streptococcus equi subsp. Zooepidemicus* cells.

The bacterial host cell may also be any *Streptomyces* cell including, but not limited to, *Streptomyces achromogenes, Streptomyces avermitilis, Streptomyces coelicolor, Streptomyces griseus,* and *Streptomyces lividans* cells.

The introduction of DNA into a *Bacillus* cell may be effected by protoplast transformation (see, *e.g.*, Chang and Cohen, 1979, *Mol. Gen. Genet.* 168: 111-115), competent cell transformation (see, *e.g.*, Young and Spizizen, 1961, *J. Bacteriol.* 81: 823-829, or Dubnau and Davidoff-Abelson, 1971, *J. Mol. Biol.* 56: 209-221), electroporation (see, *e.g.*, Shigekawa and Dower, 1988, *Biotechniques* 6: 742-751), or conjugation (see, *e.g.*, Koehler and Thorne, 1987, *J. Bacteriol.* 169: 5271-5278). The introduction of DNA into an *E. coli* cell may be effected by protoplast transformation (see, *e.g.*, Hanahan, 1983, *J. Mol. Biol.* 166: 557-580) or electroporation (see, *e.g.*, Dower *et al.*, 1988, *Nucleic Acids Res.* 16: 6127-6145). The introduction of DNA into a *Streptomyces* cell may be effected by protoplast transformation, electroporation (see, *e.g.*, Gong *et al.*, 2004, *Folia Microbiol.* (*Praha*) 49: 399-405), conjugation (see, *e.g.*, Mazodier *et al.*, 1989, *J. Bacteriol.* 171: 3583-3585), or transduction (see, *e.g.*, Burke *et al.*, 2001, *Proc. Natl. Acad. Sci. USA* 98: 6289-6294). The introduction of DNA into a *Pseudomonas* cell may be effected by electroporation (see, *e.g.*, Choi *et al.*, 2006, *J. Microbiol. Methods* 64: 391-397) or conjugation (see, *e.g.*, Pinedo and Smets, 2005, *Appl. Environ. Microbiol.* 71: 51-57). The introduction of DNA into a *Streptococcus* cell may be effected by natural competence (see, *e.g.*, Perry and Kuramitsu, 1981, *Infect. Immun.* 32: 1295-1297), protoplast transformation (see, *e.g.*, Catt and Jollick, 1991, *Microbios* 68: 189-207), electroporation (see, *e.g.*, Buckley *et al.*, 1999, *Appl. Environ. Microbiol.* 65: 3800-3804), or

conjugation (see, *e.g.*, Clewell, 1981, *Microbiol. Rev.* 45: 409-436). However, any method known in the art for introducing DNA into a host cell can be used.

The host cell may also be a eukaryote, such as a mammalian, insect, plant, or fungal cell.

The host cell may be a fungal cell. "Fungi" as used herein includes the phyla Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota as well as the Oomycota and all mitosporic fungi (as defined by Hawksworth *et al.*, *In*, *Ainsworth and Bisby's Dictionary of The Fungi*, 8th edition, 1995, CAB International, University Press, Cambridge, UK).

The fungal host cell may be a yeast cell. "Yeast" as used herein includes ascosporogenous yeast (*Endomycetales*), basidiosporogenous yeast, and yeast belonging to the *Fungi Imperfecti* (*Blastomycetes*). Since the classification of yeast may change in the future, for the purposes of this invention, yeast shall be defined as described in *Biology and Activities of Yeast* (Skinner, Passmore, and Davenport, editors, *Soc. App. Bacteriol. Symposium Series* No. 9, 1980).

The yeast host cell may be a *Candida, Hansenula, Kluyveromyces, Pichia, Saccharomyces, Schizosaccharomyces,* or *Yarrowia* cell, such as a *Kluyveromyces lactis, Saccharomyces carlsbergensis, Saccharomyces cerevisiae, Saccharomyces diastaticus, Saccharomyces douglasii, Saccharomyces kluyveri, Saccharomyces norbensis, Saccharomyces oviformis, or Yarrowia lipolytica* cell.

The fungal host cell may be a filamentous fungal cell. "Filamentous fungi" include all filamentous forms of the subdivision Eumycota and Oomycota (as defined by Hawksworth *et al.*, 1995, *supra*). The filamentous fungi are generally characterized by a mycelial wall composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as *Saccharomyces cerevisiae* is by budding of a unicellular thallus and carbon catabolism may be fermentative.

The filamentous fungal host cell may be an *Acremonium, Aspergillus, Aureobasidium, Bjerkandera, Ceriporiopsis, Chrysosporium, Coprinus, Coriolus, Cryptococcus, Filibasidium, Fusarium, Humicola, Magnaporthe, Mucor, Myceliophthora, Neocallimastix, Neurospora, Paecilomyces, Penicillium, Phanerochaete, Phlebia, Piromyces, Pleurotus, Schizophyllum, Talaromyces, Thermoascus, Thielavia, Tolypocladium, Trametes,* or *Trichoderma* cell.

For example, the filamentous fungal host cell may be an A*spergillus awamori, Aspergillus foetidus, Aspergillus fumigatus, Aspergillus japonicus, Aspergillus nidulans, Aspergillus niger, Aspergillus oryzae, Bjerkandera adusta, Ceriporiopsis aneirina, Ceriporiopsis caregiea, Ceriporiopsis gilvescens, Ceriporiopsis pannocinta, Ceriporiopsis rivulosa, Ceriporiopsis subrufa, Ceriporiopsis subvermispora, Chrysosporium inops, Chrysosporium keratinophilum, Chrysosporium lucknowense, Chrysosporium merdarium, Chrysosporium pannicola, Chrysosporium queenslandicum, Chrysosporium tropicum, Chrysosporium zonatum, Coprinus cinereus, Coriolus hirsutus,*

Fusarium bactridioides, Fusarium cerealis, Fusarium crookwellense, Fusarium culmorum, Fusarium graminearum, Fusarium graminum, Fusarium heterosporum, Fusarium negundi, Fusarium oxysporum, Fusarium reticulatum, Fusarium roseum, Fusarium sambucinum, Fusarium sarcochroum, Fusarium sporotrichioides, Fusarium sulphureum, Fusarium torulosum, Fusarium trichothecioides, Fusarium venenatum, Humicola insolens, Humicola lanuginosa, Mucor miehei, Myceliophthora thermophila, Neurospora crassa, Penicillium purpurogenum, Phanerochaete chrysosporium, Phlebia radiata, Pleurotus eryngii, Thielavia terrestris, Trametes villosa, Trametes versicolor, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei, or *Trichoderma viride* cell.

Fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known *per se*. Suitable procedures for transformation of *Aspergillus* and *Trichoderma* host cells are described in EP 238023, Yelton *et al.*, 1984, *Proc. Natl. Acad. Sci. USA* 81: 1470-1474, and Christensen *et al.*, 1988, *Bio/Technology* 6: 1419-1422. Suitable methods for transforming *Fusarium* species are described by Malardier *et al.*, 1989, *Gene* 78: 147-156, and WO 96/00787. Yeast may be transformed using the procedures described by Becker and Guarente, *In* Abelson, J.N. and Simon, M.I., editors, *Guide to Yeast Genetics and Molecular Biology*, *Methods in Enzymology*, Volume 194, pp 182-187, Academic Press, Inc., New York; Ito *et al.*, 1983, *J. Bacteriol.* 153: 163; and Hinnen *et al.*, 1978, *Proc. Natl. Acad. Sci. USA* 75: 1920.

Method for chemical attachments or conjugation are also known in the art and the present invention is not limited to any particular such method. Further description of methods for performing chemical attachments or conjugations as known in the art can be used according to the invention.

Recovering the polypeptides of the invention

After preparation of the polypeptide of the invention it is typically recovered from the mixture wherein it has been produced, e.g. if the polypeptide was prepared in a fermentation process the polypeptide is usually recovered from the fermentation broth using a number of separation steps such as filtration, centrifugation, precipitation, solubilization; until the desired purity has been achieved.

The polypeptide of the invention may when bound to calcite be released from the calcite using one of the methods 1) lowering the pH of a solution comprising the polypeptide of the invention bound to calcite to a pH value that favors the release of CO_2 from dissolved $CaCO_3$; 2) addition of strong chelators, such as EDTA; that have the ability to bind Ca^{2+} and thereby dissolve the

calcite; and 3) release the polypeptide of the invention by titrating with the binding sequence as synthetic peptide.

Also if the polypeptide of the invention has been prepared in a chemical attachment or conjugation the polypeptide is recovered from the reaction mixture using usual separation procedures.

The ability of the polypeptide of the invention to bind calcite provides a particularly useful property for the recovery process.

Thus, the invention provides a method for recovery of the polypeptide of the invention from a fermentation broth, comprising optionally removing the solids, such as cells, cell debris and remaining solids of the substrate; from the fermentation broth, mixing the fermentation broth with sufficient calcite to bind all the polypeptide of the invention, followed by a solid-liquid separation process where the calcite with the polypeptide of the invention attached is separated from the rest of the fermentation broth.

The invention also provides a method for recovery of the polypeptide of the invention from a chemical reaction mixture, comprising mixing the reaction mixture with sufficient calcite to bind all the polypeptide of the invention, followed by a solid-liquid separation process where the calcite with the polypeptide of the invention attached is separated from the rest of the reaction mixture.

The methods of the invention provides a fast and highly efficient separation of the polypeptide of the invention from the fermentation broth or the reaction mixture where it has been produced.

In this way, the method of the invention is similar to the affinity-based separation methods as known in the art, in particular in the analytical or small scale production, using affinity pairs such as his-tag, Biotin-streptavidin, antibody-antigen-based purification, but the present invention has the benefit of using calcite as the material that the protein binds to, which is very cheap and nontoxic compared with other binding materials, used in the prior art for releasing tagged polypeptides of the prior art, meaning that the method of the invention is readily useable in large scale industrial recovery processes, e.g. for fermentation processes in a volume above 5,000 liters, such as above 10,000 liter, such as above 25,000 liters, such as above 50,000 liters or even above 100,000 liters and the purification process is amenable to non-column based separation (i.e., using batch or expanded bed adsorption).

Compositions comprising the polypeptide of the invention

The present invention also relates to compositions comprising a polypeptide of the present invention.

In a preferred embodiment, the composition of the invention comprises the polypeptide of the invention attached to calcite.

The compositions may comprise a polypeptide of the present invention as the major enzymatic component, *e.g.*, a mono-component composition. Alternatively, the compositions may comprise multiple enzymatic activities, such as one or more (*e.g.*, several) enzymes selected from the group consisting of hydrolase, isomerase, ligase, lyase, oxidoreductase, or transferase, *e.g.*, an alpha-galactosidase, alpha-glucosidase, aminopeptidase, amylase, beta-galactosidase, beta-glucosidase, beta-xylosidase, carbohydrase, carboxypeptidase, catalase, cellobiohydrolase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, endoglucanase, esterase, glucoamylase, invertase, laccase, lipase, mannosidase, mutanase, oxidase, pectinolytic enzyme, peroxidase, phytase, polyphenoloxidase, proteolytic enzyme, ribonuclease, transglutaminase, or xylanase.

The compositions may be prepared in accordance with methods known in the art and may be in the form of a liquid or a dry composition.

EXAMPLES

Materials and Methods

General methods of PCR, cloning, ligation nucleotides etc. are well-known to a person skilled in the art and may for example be found in in "Molecular cloning: A laboratory manual", Sambrook et al. (1989), Cold Spring Harbor lab., Cold Spring Harbor, NY; Ausubel, F. M. et al. (eds.); "Current protocols in Molecular Biology", John Wiley and Sons, (1995); Harwood, C. R., and Cutting, S. M. (eds.); "DNA Cloning: A Practical Approach, Volumes I and II", D.N. Glover ed. (1985); "Oligonucleotide Synthesis", M.J. Gait ed. (1984); "Nucleic Acid Hybridization", B.D. Hames & S.J. Higgins eds (1985); "A Practical Guide To Molecular Cloning", B. Perbal, (1984).

PCR amplification

All PCR amplifications was performed in a volume of 100 µL containing 2.5 units Taq polymerase, 100 ng of pSO2, 250 nM of each dNTP, and 10 pmol of each of the two primers de-scribed above in a reaction buffer of 50 mM KCl, 10 mM Tris-HCl pH 8.0, 1.5 mM MgCl2.
Amplification was carried out in a Perkin-Elmer Cetus DNA Termal 480, and consisted of one cycle of 3 minutes at 94°C, followed by 25 cycles of 1 minute at 94°C, 30 seconds at 55°C, and 1 minute at 72°C.

Expressing and recovering polypeptides from Aspergillus

Aspergillus oryzae COLS1300 (*amyA, amyB, amyC, alpA, nprA, kusA, niaD, niiA, amdS*+) was created from *A. oryzae* PFJ0220 (EP2147107B1) by deleting the promoter and 5' part of both the nitrite reductase (*niiA*) gene and nitrate reductase (*niaD*) gene, thereby inactivating their

expression (Nielsen M. L. *et al.* 2006, *Efficient PCR-based gene targeting with a recyclable marker for Aspergillus nidulans*, Fungal Genetics and Biology vol. 43: 54–64).

Aspergillus oryzae COLS1300 transformation was done as described by Christensen et al.; Biotechnology 1988 6 1419-1422. In short, *A.oryzae* mycelia were grown in a rich nutrient broth. The mycelia were separated from the broth by filtration. The enzyme preparation Glucanex® (Novozymes, Bagsværd DK) was added to the mycelia in osmotically stabilizing buffer such as 1.2 M $MgSO_4$ buffered to pH 5.0 with sodium phosphate. The suspension was incubated for 60 minutes at 37degrees C with agitation. The protoplast was filtered through mira-cloth to remove mycelial debris. The protoplast was harvested and washed twice with STC (1.2 M sorbitol, 10 mM $CaCl_2$, 10 mM Tris-HCl pH 7.5). The protoplasts were finally re-suspended in 200-1000 µL STC.

Transformation was done essentially as described in WO 2017/191210 examples 9-11, with DNA inserts encoding the polypeptide of interest. Selection is based on the ability to grow on Sucrose Medium with nitrate. After 5-7 days of growth at 37 degrees C, stable transformants appeared as vigorously growing and sporulating colonies. Transformants were purified once through conidiospores.

Sucrose Medium

1M Sucrose, 0.18 µM Na_2B4O_7, 2.3 µM $CuSO_4$, 4.7 µM $FeSO_4$, 4.7 µM $MnSO_4$, 3.6 µM Na_2MoO_4, 45 µM $ZnSO_4$, 7 mM KCl, 4.3 mM $MgSO_4$, 11.2 mM KH_2PO_4

Sucrose Medium with nitrate

Sucrose medium containing 10 mM $NaNO_3$

Nunc tube fermentation

Shake flask containing 10 ml YPD medium (2 g/l yeast extract, 2 g/l peptone, and 2% glucose) were inoculated with spores from a transformant/heterokaryon or diploid strain and incubated at 30 degrees C, 200 rpm for 4 days. The fermentation broth was used directly or recovered as described in the examples.

SDS-page

SDS gel used was Criterion™ XT precast gels, 10% Bis-Tris, from BIO-RAD and was run and stained with Coomassie blue as recommend by the manufactory.

Example 1, Construction of ZEN-DsRed constructs:

ZEN-DsRed::sca.1 (DORA2) (SEQ ID NO: 5)

MRTRSTISTPNGITWYYEQEGTGPDVVLVPDGLGECQMFDSSVSQIAAQGFRVTT-
FDMPGMSRSAKAPPETYTEVTAQKLASYVISVLDALDIKHATVWGCSSGASTVVALLLGYPDRI
RNAMCHELPTKLLDHLSNTAVLEDEEISKILANVMLNDVSGGSEAWQAMGDEVHARLHKNY-
PVWARGYPRTIPPSAPVKDLEALRGKPLDWTVGAATPTESFFDNIVTATKAGVNIGLLPGMHFP
YVSHPDVFAKYVVETTQKHLGNSAM-
DNTEDVIKEFMQFKVRMEGSVNGHYFEIEGEGEGKPYEGTQTAKLQVTKGGPLPFAWDILSPQ
FQYGSKAYVKHPADIPDYMKLSFPEGFTWERSMNFEDGGVVEVQQDSSLQDGT-
FIYKVKFKGVNFPADGPVMQKKTAGWEPSTEKLYPQDGVLKGEISHALKLKDGGHYTCDFKTV
YKAKKPVQLPGYYYVDSKLDITNHNEDYTVVEQYEHAEARHSGSQGYSALDDDDYPKG*

The purpose of producing a construct containing ZEN-DsRed and SEQ ID NO: 4 is to examine the effect of the tag on affinity purification. ZEN-DsRed is used as the screening marker for its ability to be secreted from the A. Oryzae cell, as well as its fluorescent properties.

Flank1 was created by PCR using primers: oKRN-11 (5'-ccagaccagcagaggagataatactct, SEQ ID NO: 6) and oKRN-12 (5'- ggtgcggccgcccccagttg, SEQ ID NO: 7) on template plasmid AT1770. Resulting in a fragment of **3625 bp.**

Flank2 was created by PCR using primers: oKRN-17 (5'- GGCTACTCCGCTCTCGATGAC-GATGACTACCCCAAGGGCtgaacctggcggtagacaatcaat, SEQ ID NO: 8) and oKRN-14 (5'-caaggatacctacagttattcgaaacct, SEQ ID NO: 9) on template plasmid pAT1771. Resulting in fragment of **3596 bp.**

Insert was created by PCR using primers: oKRN-15 (5'- attatatacacaactgggggcggccgcac-cATGCGGACTAGGTCG, SEQ ID NO: 10) and oKRN-18 (5'-GTAGTCATCGTCATCGAGAGCG-GAGTAGCCcactggctcccgct, SEQ ID NO: 11) on template plasmid pAT1111. Resulting in fragment of **1542 bp.**

The fragments described above were used in Aspergillus transformation as described in the method section above.

Control, ZEN-DsRed (DORA6) (SEQ ID NO: 12)

MRTRSTISTPNGITWYYEQEGTGPDVVLVPDGLGECQMFDSSVSQIAAQGFRVTT-
FDMPGMSRSAKAPPETYTEVTAQKLASYVISVLDALDIKHATVWGCSSGASTVVALLLGYPDRI
RNAMCHELPTKLLDHLSNTAVLEDEEISKILANVMLNDVSGGSEAWQAMGDEVHARLHKNY-
PVWARGYPRTIPPSAPVKDLEALRGKPLDWTVGAATPTESFFDNIVTATKAGVNIGLLPGMHFP
YVSHPDVFAKYVVETTQKHLGNSAM-
DNTEDVIKEFMQFKVRMEGSVNGHYFEIEGEGEGKPYEGTQTAKLQVTKGGPLPFAWDILSPQ

FQYGSKAYVKHPADIPDYMKLSFPEGFTWERSMNFEDGGVVEVQQDSSLQDGT-
FIYKVKFKGVNFPADGPVMQKKTAGWEPSTEKLYPQDGVLKGEISHALKLKDGGHYTCDFKTV
YKAKKPVQLPGYYYVDSKLDITNHNEDYTVVEQYEHAEARHSGSQ

The purpose of producing a construct containing ZEN-DsRed without a calcite binding sequence is to have a control when comparing protein properties of tagged vs. non- tagged proteins. ZEN-DsRed is used as the screening marker for its ability to be secreted from the *A. oryzae* cell, as well as its fluorescent properties.

Flank1 was created by PCR using primers: oKRN-11 (5'-ccagaccagcagaggagataatactct, SEQ ID NO: 6) and oKRN-26 (5'- CGACCTAGTCCGCATggtgcggccgccccc, SEQ ID NO: 13) on template plasmid AT1770. Resulting in a fragment of **3640 bp.**

Flank2 was created by PCR using primers: oKRN-25 (5'- cacagcgggagccagacctggcggtaga-caatcaatc, SEQ ID NO: 14) and oKRN-14 (5'- caaggatacctacagttattcgaaacct, SEQ ID NO: 9) on template plasmid pAT1771. Resulting in fragment of **3569 bp**

Insert was created by PCR using primers: oKRN-15 (5'- attatatacacaactgggggcggccgcac-cATGCGGACTAGGTCG, SEQ ID NO: 10) and oKRN-27 (5'- tgtctaccgccaggtctggctcccgct, SEQ ID NO: 15) on template plasmid pAT1111. Resulting in fragment of **1509 bp.**

The fragments described above were used in Aspergillus transformation as described in the method section above.

Example 2 cultivation and expression of DsRed in transformants.

Spores from the transformants DORA2 and DORA6 were transferred to culture tubes containing YP-medium + 10% sucrose and incubated on a shake table (200 rpm) at 30°C for 4 days.

After the incubation the supernatants of all transformants were red indicating secretion of DsRed to the supernatant. SDS-PAGE analysis of the supernatants disclosed bands of the expected size, 56 kDa (see figure 1) and peptide analysis by MS confirmed the sequence of the expressed polypeptides (Data not shown).

Example 3 binding to calcite

100 µl of the supernatants of the supernatants prepared in example 2 was transferred to two test tubes and 50 mg powdered calcite was added. After mixing the calcite was pelletize by centrifugation and a sample was run on a SDS-PAGE analysis together with the supernatant which had not been contacted to calcite.

The SDS-PAGE gel clearly show that for DORA2, the construct having the binding sequence SEQ ID NO: 3, the band at 56 kDa representing the construct was removed from the supernatant after the contact with calcite, whereas the band was not affected in DORA6, the control without the binding sequence (see figure 1).

Another 100µl of the supernatants prepared in example 2 was transferred to two test tubes. One of the test tuber for each of the supernatants were mixed with 50 mg powdered calcite and the calcite was pelletized by centrifugation. All tubes were illuminated with green light, where the DsRed protein fluoresces red. The experiment showed that for the DORA 2 construct the red colour was bound and pelletized with calcite, whereas no binding or pelletizing was seen in the control DORA6 (See figure 2).

Example 4. Lipase constructs

Four lipase constructs were prepared and expressed in *Aspergillius oryzae* using the methods described above.

T1: Lipase-BS1 contained the sequence:

EVSQDLFNQFNLFAQYSAAAYCGKNNDAPAGTQITCTGNACPEVEKADATFLYSFED
SGVGDVTGFLALDNTNKLIVLSFRGSRSIENWIQNLNFDLKEINDICSGCRGHDGFTSS
WRSVADTLRQKVEDAVREHPDYRVVFTGHSLGGALATVAGADLRGNGYDIDVFSYG
APRVGNRAFAEFLTVQTGGTLYRITHTNDIVPRLPPRQFGYSHSSPEYWIKSGTLVPVR
RRDIVKIEGIDATGGNNQPNIPDAPAHLWYFGLIGTCLGSGSGSGSGSG**YHHGEEEEDVW**

(SEQ ID NO: 16)

where the first part is the lipase sequence, the underlined sequence is a linker sequence separating the lipase sequence from the calcite binding sequence BS1 (SEQ ID NO: 3) in bold.

T2: Lipase-BS2 contained the sequence:

EVSQDLFNQFNLFAQYSAAAYCGKNNDAPAGTQITCTGNACPEVEKADATFLYSFED
SGVGDVTGFLALDNTNKLIVLSFRGSRSIENWIQNLNFDLKEINDICSGCRGHDGFTSS
WRSVADTLRQKVEDAVREHPDYRVVFTGHSLGGALATVAGADLRGNGYDIDVFSYG

APRVGNRAFAEFLTVQTGGTLYRITHTNDIVPRLPPRQFGYSHSSPEYWIKSGTLVPVR
RRDIVKIEGIDATGGNNQPNIPDAPAHLWYFGLIGTCLGSGSGSGSGSG**YSALDDDDYPKG**

(SEQ ID NO: 17)

where the first part is the lipase sequence, the underlined sequence is a linker sequence separating the lipase sequence from the calcite binding sequence BS2 (SEQ ID NO: 4) in bold.

T3: Lipase-BS1-BS2 contained the sequence:

EVSQDLFNQFNLFAQYSAAAYCGKNNDAPAGTQITCTGNACPEVEKADATFLYSFED
SGVGDVTGFLALDNTNKLIVLSFRGSRSIENWIQNLNFDLKEINDICSGCRGHDGFTSS
WRSVADTLRQKVEDAVREHPDYRVVFTGHSLGGALATVAGADLRGNGYDIDVFSYG
APRVGNRAFAEFLTVQTGGTLYRITHTNDIVPRLPPRQFGYSHSSPEYWIKSGTLVPVR
RRDIVKIEGIDATGGNNQPNIPDAPAHLWYFGLIGTCLGSGSGSGSGSG**YHH-**
GEEEEDVWGSG**YSALDDDDYPKG**

(SEQ ID NO: 18)

where the first part is the lipase sequence, the underlined sequence is linker sequences separating the lipase sequence from the calcite binding sequence BS1 (SEQ ID NO: 3) and BS2 (SEQ ID NO: 4) in bold.

T4: Lipase contained the sequence:

EVSQDLFNQFNLFAQYSAAAYCGKNNDAPAGTQITCTGNACPEVEKADATFLYSFED
SGVGDVTGFLALDNTNKLIVLSFRGSRSIENWIQNLNFDLKEINDICSGCRGHDGFTSS
WRSVADTLRQKVEDAVREHPDYRVVFTGHSLGGALATVAGADLRGNGYDIDVFSYG
APRVGNRAFAEFLTVQTGGTLYRITHTNDIVPRLPPRQFGYSHSSPEYWIKSGTLVPVR
RRDIVKIEGIDATGGNNQPNIPDAPAHLWYFGLIGTCL

(SEQ ID NO: 19)

and consisted only of the lipase sequence without any linker and calcite binding sequences.

Supernatants comprising T1-T4 was prepared and the expected mass of the proteins confirmed by LC-MS, confirming the identity and integrity of the constructs.

The supernatants comprising T1-T4 were mixed powdered calcite and it was shown that T1-T3 had the calcite binding properties. Further it was demonstrated that the proteins could be released from calcite by treating with EDTA.

Example 5. Lipase construct T1 maintain activity after binding and release from calcite

Purified T1 as prepared in example 4 was contacted with calcite and mixed. The mixture was centrifuged and the supernatant removed. The calcite was washed once with a buffer solution at pH 6.0.

After the wash solution was removed and T1 was released from the calcite by treating the calcite with the bound protein with excess of 0.5 M EDTA pH 8.0, 0.2 M NaCl.

All fractions was analysed on SDS-PAGE (see figure 3). The gel show Molecular weight marker (first and lase lane), purified T1 solution, absence of purified T1 in the solution after addition of calcite, wash solution and T1 released after EDTA treatment.

Example 6 Lipase activity

The fractions of Example 5 was analysed for lipase activity using a standard lipase assay. It was shown that the purified T1 had lipase activity, and further it was confirmed that the EDTA fraction also had lipase activity, which confirms that T1 was capable of binding to calcite, and subsequent be released from calcite by EDTA treatment and still maintained lipase activity.

CLAIMS

1. A polypeptide comprising a sequence encoding a desired polypeptide and a sequence having calcite binding properties, wherein the calcite binding sequence has the sequence of:

YHHGEEED (SEQ ID NO: 3);

LDDDDYPKG (SEQ ID NO: 4);

or having the sequence of SEQ ID NO: 3 or 4 with 1, 2, 3 or 4 amino acid substitutions.

2. The polypeptide of claim 1, wherein the desired polypeptide does not have calcite binding properties.

3. The polypeptide of claim 1 or 2, wherein the sequence having calcite binding properties is located in the N-terminal or in the C-terminal of the polypeptide.

4. The polypeptide of any of claims 1 to 3, wherein the desired polypeptide is an enzyme selected among selected among aminopeptidase, amylase, amyloglucosidase, mannanase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, esterase, galactosidase, beta-galactosidase, glucoamylase, glucose oxidase, glucosidase, haloperoxidase, hemicellulase, invertase, isomerase, laccase, ligase, lipase, lyase, mannosidase, oxidase, pectinase, peroxidase, phytase, phenoloxidase, polyphenoloxidase, protease, ribonuclease, transferase, transglutaminase, lysozyme, muramidase, xanthanase or xylanase.

5. The polypeptide of claim 4, having at least at least 25% enzyme activity in comparison with the enzyme activity of the desired polypeptide without the sequence having calcite binding properties, e.g. least 40% enzyme activity, e.g. at least 50% enzyme activity, e.g. at least 60% enzyme activity, e.g. at least 70% enzyme activity, e.g. at least 80% enzyme activity or at least 90% enzyme activity in comparison with the enzyme activity of the desired polypeptide without the sequence having calcite binding properties.

6. A polynucleotide encoding a polypeptide according to any of the claims 1-5.

7. An expression construct or expression vector comprising the polynucleotide of claim 6.

8. A host cell comprising the polynucleotide of claim 6 or the expression construct or expression vector of claim 7.

9. A method for preparing the polypeptide according to any of the claims 1-5, comprising the steps of:

a) Providing an aqueous solution of a polypeptide according to any of the claims 1-5;
b) Adding powdered calcite to the aqueous solution of the polypeptide, and mixing;
c) Optionally holding the mixture for a period to allow the polypeptide to bind to the calcite; and

d) Separating the calcite with the bound polypeptide from the mixture.

10. The method of claim 9 further comprising the step of:

e) Recovering the polypeptide from the calcite.

11. The method of claim 9 or 10, wherein the aqueous solution in step a) is a fermentation broth comprising the polypeptide.

Figure 1

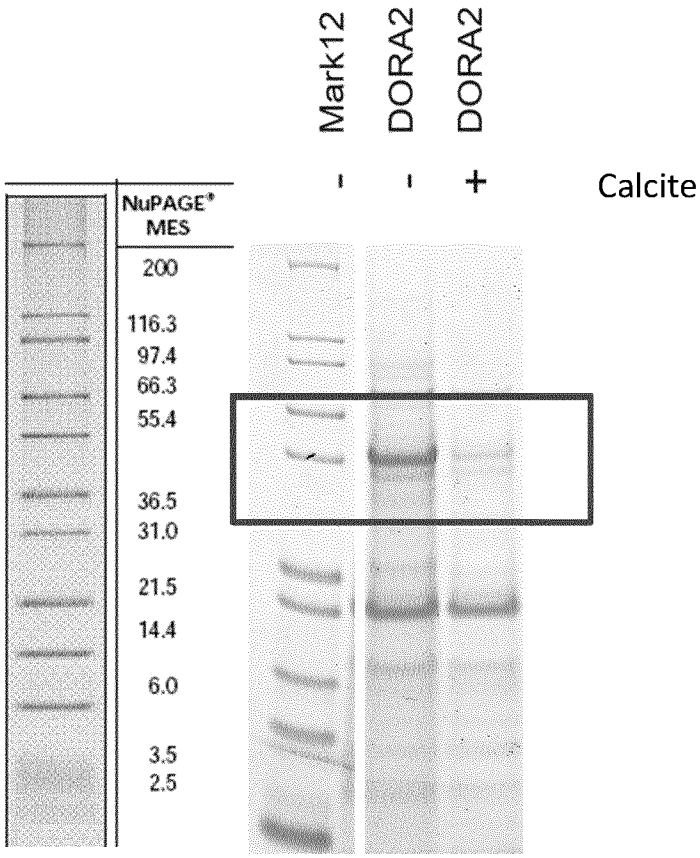

Mark12
DORA2
DORA2
-
-
+
Calcite
NuPAGE® MES
200
116.3
97.4
66.3
55.4
36.5
31.0
21.5
14.4
6.0
3.5
2.5

Figure 2

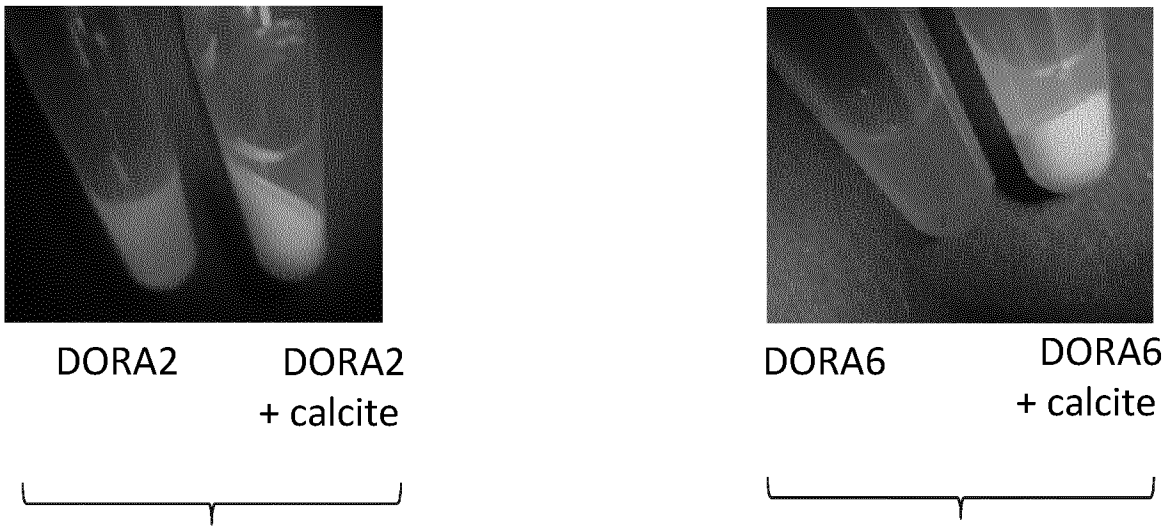
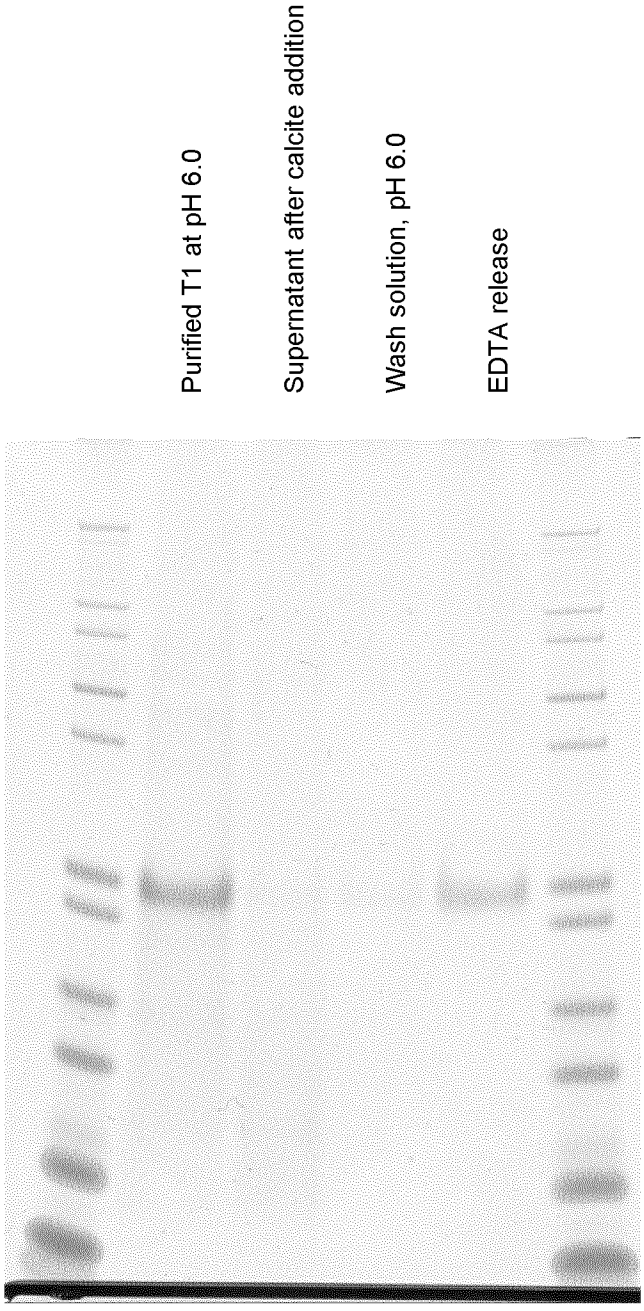


Figure 3


Purified T1 at pH 6.0
Supernatant after calcite addition
Wash solution, pH 6.0
EDTA release

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2020/054578

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/62 C07K1/22 C12N9/20 C07K14/465
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, BIOSIS, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MANN K ET AL: "Ostrich (Struthio camelus) eggshell matrix contains two different C-type lectin-like proteins. Isolation, amino acid sequence, and posttranslational modifications", BIOCHIMICA ET BIOPHYSICA ACTA (BBA) - PROTEINS & PROTEOMICS, ELSEVIER, NETHERLANDS, vol. 1696, no. 1, 14 January 2004 (2004-01-14), pages 41-50, XP004483644, ISSN: 1570-9639, DOI: 10.1016/J.BBAPAP.2003.09.006 page 44; figure 5 page 46; figure 7 page 49, right-hand column, paragraph 3 ----- -/--	1-11

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 April 2020	12/05/2020

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Le Cornec, Nadine

3

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2013/091801 A1 (SCHAEFER KALK GMBH & CO KG [DE]; EPPLE MATTHIAS [DE] ET AL.) 27 June 2013 (2013-06-27) page 3 - page 13; claims; tables 1-3	1-11
X	BEATRICE DEMARCHI ET AL: "Protein sequences bound to mineral surfaces persist into deep time", ELIFE, vol. 5, 27 September 2016 (2016-09-27), page e17092, XP055687766, DOI: 10.7554/eLife.17092 cited in the application page 6, paragraph 3 page 7; table 2	1-11
A	US 2009/029902 A1 (CUNNINGHAM SCOTT D [US] ET AL) 29 January 2009 (2009-01-29) claims	1-11
A	SHIRLY BORUKHIN ET AL: "Screening the Incorporation of Amino Acids into an Inorganic Crystalline Host: the Case of Calcite", ADVANCED FUNCTIONAL MATERIALS, vol. 22, no. 20, 23 October 2012 (2012-10-23), pages 4216-4224, XP055330954, DE ISSN: 1616-301X, DOI: 10.1002/adfm.201201079 page 4218, right-hand column, paragraph 2.3 page 4219; figure 3 page 4221 - page 4223, right-hand column, paragraph 1	1-11
A	OREN SHUR ET AL: "A designed, phase changing RTX-based peptide for efficient bioseparations", BIOTECHNIQUES, vol. 54, no. 4, 1 April 2013 (2013-04-01), pages 197-206, XP055688027, US ISSN: 0736-6205, DOI: 10.2144/000114010 the whole document	1-11

-/--

3

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WANG ET AL: "Recombinant perlucin nucleates the growth of calcium carbonate crystals: Molecular cloning and characterization of perlucin from disk abalone, Haliotis discus discus", COMPARATIVE BIOCHEMISTRY AND PHYSIOLOGY PART B: BIOCHEMISTRY AND MOLECULAR BIOLOGY, ELSEVIER, AMSTERDAM, NL, vol. 149, no. 2, 18 October 2007 (2007-10-18), pages 354-361, XP022402730, ISSN: 1096-4959 the whole document -----	1-11
A	EVA WEBER ET AL: "GFP Facilitates Native Purification of Recombinant Perlucin Derivatives and Delays the Precipitation of Calcium Carbonate", PLOS ONE, vol. 7, no. 10, 1 January 2012 (2012-01-01), pages e46653-e46653, XP055056303, ISSN: 1932-6203, DOI: 10.1371/journal.pone.0046653 the whole document -----	1-11
X	DATABASE UniProt [Online] 28 February 2018 (2018-02-28), "SubName: Full=transcription factor BIM2-like isoform X4 {ECO:0000313|RefSeq:XP_016547008.1};", XP55687992, retrieved from EBI accession no. UNIPROT:A0A2G2YCV4 Database accession no. A0A2G2YCV4 sequence -----	1-3,6

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2020/054578

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2013091801 A1	27-06-2013	NONE	
US 2009029902 A1	29-01-2009	NONE	