Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 2996313 A1 2017/03/23

(21) 2 996 313

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2016/09/16
(87) Date publication PCT/PCT Publication Date: 2017/03/23
(85) Entrée phase nationale/National Entry: 2018/02/21
(86) N° demande PCT/PCT Application No.: US 2016/052147
(87) N° publication PCT/PCT Publication No.: 2017/049094
(30) Priorité/Priority: 2015/09/18 (US62/220,688)

(51) Cl.Int./Int.Cl. *C07K 14/00* (2006.01), *A23K 10/14* (2016.01), *C07K 14/245* (2006.01), *C07K 14/38* (2006.01), *C07K 14/39* (2006.01), *C07K 4/06* (2006.01), *C12N 15/04* (2006.01), *C12N 15/10* (2006.01), *C12N 15/52* (2006.01), *C12N 15/62* (2006.01), *C12N 15/65* (2006.01), *C12N 15/80* (2006.01), *C12N 9/16* (2006.01)

(71) Demandeur/Applicant:
AGRIVIDA, INC., US

(72) Inventeurs/Inventors:
RAAB, R. MICHAEL, US;
LAZAR, GABOR, US;
SHEN, BINZHANG, US

(74) Agent: GOUDREAU GAGE DUBUC

(54) Titre : PHYTASES MODIFIEES ET PROCEDES POUR LES UTILISER
(54) Title: ENGINEERED PHYTASES AND METHODS OF USING THE SAME

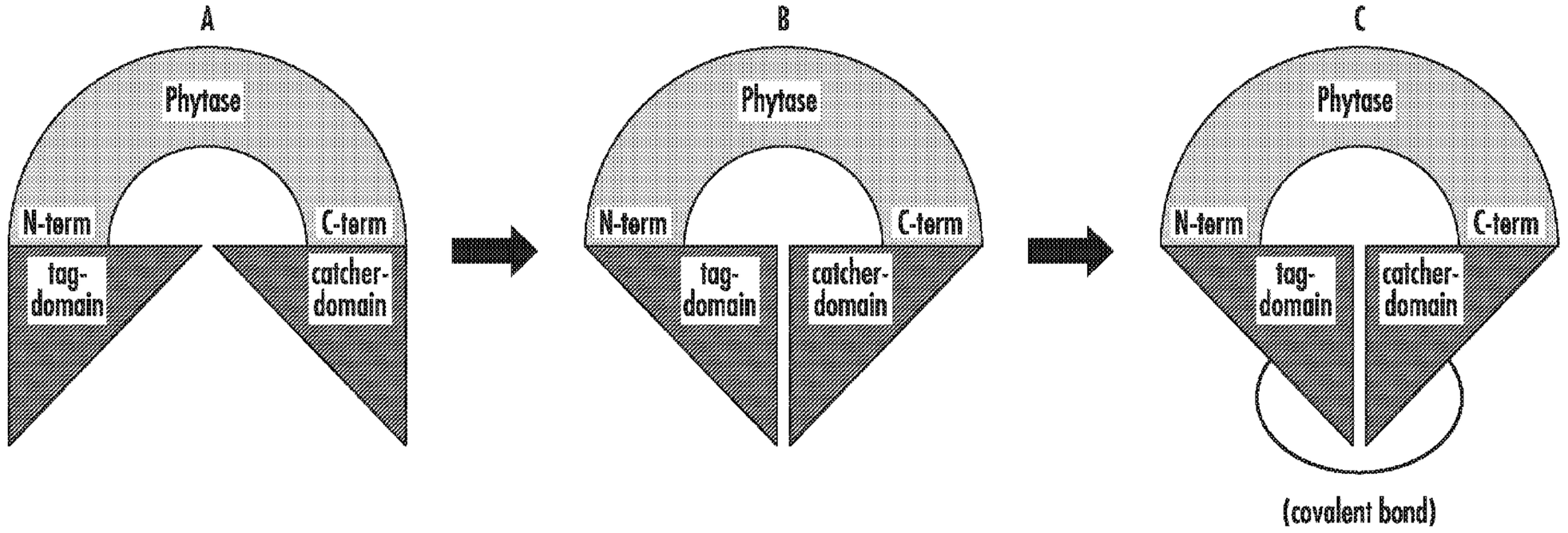

FIG. 4

(57) Abrégé/Abstract:
Methods for enhancing phytase thermal stability by fusing binding elements to target phytases are provided. Engineered phytases that include binding elements fused to target phytases to cause cyclization of the engineered phytases and enhance thermal stability of the target phytases are described. Engineered nucleic acids encoding engineered phytases and hosts engineered to express engineered nucleic acids are also provided. Methods for incorporating engineered phytases in animal feed and animal feed including the same are described.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
23 March 2017 (23.03.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/049094 A1

(51) International Patent Classification:
C07K 14/00 (2006.01) *C12N 15/04* (2006.01)
C07K 4/06 (2006.01) *C12N 15/52* (2006.01)
C07K 14/245 (2006.01) *C12N 15/80* (2006.01)
C07K 14/38 (2006.01) *C12N 15/65* (2006.01)
C07K 14/39 (2006.01) *C12N 9/16* (2006.01)
C12N 15/62 (2006.01) *A23K 10/14* (2016.01)
C12N 15/10 (2006.01)

(21) International Application Number:
PCT/US2016/052147

(22) International Filing Date:
16 September 2016 (16.09.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/220,688 18 September 2015 (18.09.2015) US

(71) Applicant: **AGRIVIDA, INC.** [US/US]; 200 Boston Avenue, Suite 2975, Medford, Massachusetts 02155 (US).

(72) Inventors: **RAAB, R. Michael**; 35 Wall Street, Arlington, Massachusetts 02476 (US). **LAZAR, Gabor**; 464 Pleasant Street, Belmont, Massachusetts 02478 (US). **SHEN, Binzhang**; 24 Tip Top Street, Boston, Massachusetts 02135 (US).

(74) Agent: **SIGAREVA, Marina A.**; Volpe and Koenig, P.C., 30 S. 17th Street, United Plaza, Philadelphia, Pennsylvania 19103 (US).

(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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: ENGINEERED PHYTASES AND METHODS OF USING THE SAME

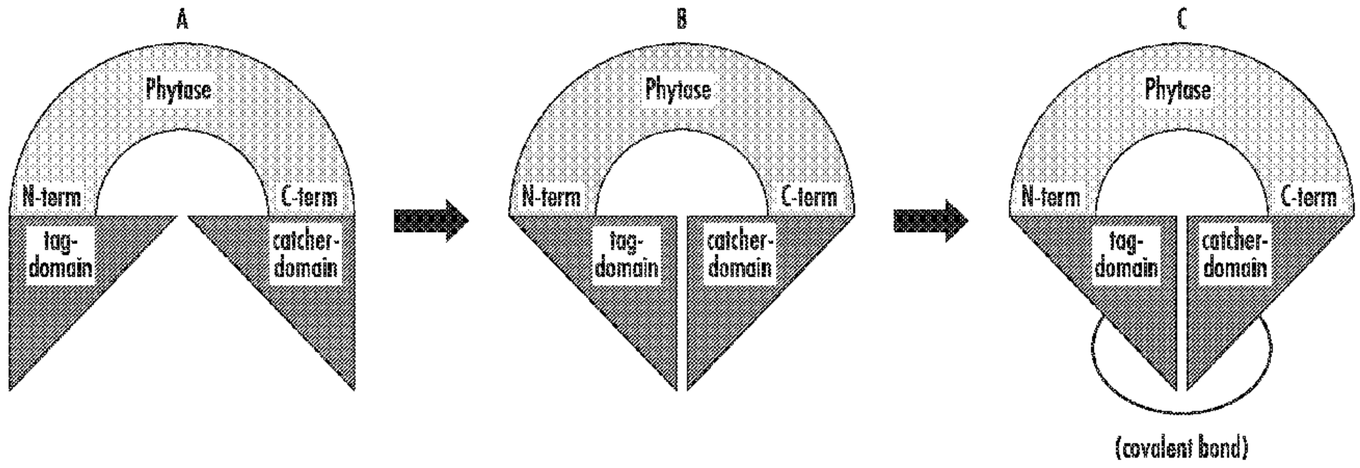

FIG. 4

(57) Abstract: Methods for enhancing phytase thermal stability by fusing binding elements to target phytases are provided. Engineered phytases that include binding elements fused to target phytases to cause cyclization of the engineered phytases and enhance thermal stability of the target phytases are described. Engineered nucleic acids encoding engineered phytases and hosts engineered to express engineered nucleic acids are also provided. Methods for incorporating engineered phytases in animal feed and animal feed including the same are described.

WO 2017/049094 A1

ENGINEERED PHYTASES AND METHODS OF USING THE SAME

CROSS REFERENCE TO RELATED APPLICATION

[0001] This application claims the benefit of U.S. provisional application No. 62/220,688 filed September 18, 2015, which is incorporated herein by reference as if fully set forth.

[0002] The sequence listing electronically filed with this application titled "Sequence Listing," created on September 18, 2016, and having a file size of 399,733 bytes is incorporated herein by reference as if fully set forth.

FIELD OF INVENTION

[0003] This disclosure relates to engineered phytase molecules that have improved thermal stability, improved or reduced gastric stability, the nucleic acids encoding the same, methods of making the same, as well as methods of using the same in industrial processing or animal feed.

[0004] This disclosure relates to transgenic plants expressing the phytases with improved thermal stability, the nucleic acids encoding the same, as well as methods of processing the transgenic plants and tissues, and producing and utilizing animal feed. The disclosure also relates to feed additives, grain and fiber processing additives that include phytases.

[0005] This disclosure relates to forms of an engineered *E. coli*-derived phytase that have been modified to improve their performance as components of feed for monogastric and ruminant animals. These modified phytases can be expressed directly in feed components such as corn grain and incorporated into animal diets, for example in mash or pelleted feeds for monogastric animals, or in silage or grain for ruminants. Diets containing these plant-expressed phytases support efficient animal growth using less phosphate than would otherwise be necessary in the absence of engineered phytase.

BACKGROUND

[0006] Phytases are a class of acid phosphatase enzymes that hydrolyze phosphates from phytate to produce free phosphate and inositol. Phytic acid (inositol hexakisphosphate), or its deprotonated form, phytate, is common in many animal feed components such as grains and legumes, and can represent a significant portion of the total phosphate content in these feeds. However, many livestock animals cannot efficiently digest phytic acid and are therefore unable to absorb the phosphate.

[0007] As a result, other forms of phosphate, such as rock phosphate or calcium phosphate, must be added to animal diets to provide this critical nutrient. Furthermore, phytic acid acts as an antinutrient in the diet, binding to proteins and chelating minerals such as iron, calcium and magnesium, which prevents their absorption. Since undigested phytic acid and excess inorganic phosphate can be excreted in the feces, they can act as a significant source of phosphate pollution in agricultural run-off. Phytase is commonly used in industrial processing and animal production. Inclusion of phytases in animal diets can alleviate the need to add inorganic phosphate, increasing the absorption of phosphate, proteins and minerals by the animal, and decreasing phosphate pollution from agricultural run-offs. When combined these effects can significantly increase the efficiency of animal growth and overall nutrition obtained from the feed they consume.

[0008] In industrial process, particularly fermentation processes, phytase is often used to hydrolyze phytate, releasing minerals and other nutrients into the fermentation, as well as enhancing starch degradation by enzymes that require cofactors sequestered by phytate (E. Khullar, J.K. Shetty, M.E. Tumbleson, V. Singh, "Use of Phytases in Ethanol Production from E-Mill Corn Processing," Cereal Chem., 88(3):223-227, 2011, which is incorporated herein by reference as if fully set forth). It is also used industrially to reduce scaling that may be associated with phytate or phosphate build-up (sometimes referred to as "beer stone"), which often occurs in fermentation or related processes. In animal production and nutrition, one

strategy for making phosphorus from phytate nutritionally available to monogastric animals is the addition of phytase to animal feeds (Jongbloed and Lenis, 1998; Onyango et al., 2005, both of which are incorporated herein by reference as if fully set forth). The use of phytase in the diets of poultry and swine has been shown to improve performance and phosphorus utilization (Baker, 2002; Nyannor et al., 2007 and 2009, both of which are incorporated herein by reference as if fully set forth). A number of phytase products are currently marketed for this use and include Natuphos™ (BASF), a phytase derived from *Aspergillus niger*, Ronozyme™ (DSM) a phytase derived from *Peniophora lycii*, and Quantum and Quantum Blue (AB Vista) phytases derived from *Escherichia coli*. The use of phytase in animal feeds allows a reduction in the amount of inorganic phosphorus added to animal feeds and has been reported to result in reductions in fecal phosphorus as high as 56% (Nahm, 2002; Sharpley *et al.*, 1994; Wodzinski and Ullah, 1996, all of which are incorporated herein by reference as if fully set forth). While phytase use in animal feed and industrial processing is beneficial, one common challenge for using microbially or plant-produced phytases in animal feed diets is their inability to maintain full activity through the conditioning, extrusion, or pelleting processes commonly used to make feed pellets. Although some enzymes have been engineered to improve their thermal stability, most lose activity during pelleting, increasing their relative costs and decreasing the efficacy of the enzyme. Therefore, enzymes with further improvements in thermal stability are needed, particularly as feed manufacturers prefer to use higher-temperature pelleting processes.

[0009] It is well known in the art that many biomolecules can be rendered inactive through exposure to high temperatures. Because proteins are ubiquitous in nature, occurring in all kingdoms of life and being present in organisms as diverse as mesophiles to extreme thermophiles, they have an enormous range of thermal stabilities. Proteins that are characterized to have low thermal stability often progress through a molecular pathway wherein their structures increase in energy, increasing molecular vibration and movement, which overcomes intramolecular bonding forces and cause the

protein to unfold. As unfolding occurs, structures within the protein are disordered, simultaneously exposing hydrophilic and hydrophobic regions and amino acids in the protein structure, and often leading to aggregation of the protein. For proteins that have low thermal stability, the unfolding process is often considerably faster than the refolding process, and in some cases may essentially be irreversible. Conversely, proteins that possess high degrees of thermal stability often have a greater degree of intramolecular bonding, which helps hold their structure together in the presence of increasing levels of thermal energy, as well as rapid rates of refolding, which can enhance a protein's ability to recover its activity when confronted by destabilizing thermal exposure. Given the broad range of thermal stabilities observed among different proteins, an opportunity exists to engineer less stable proteins to be more thermally stable. This is specifically relevant to phytases, which are often derived from mesophilic or less thermophilic organisms, and commonly struggle to maintain high levels of activity in animal feed pelleting processes, or industrial processes.

[0010] Another common challenge with producing heterologous proteins in plants, microbial cells, or other cellular production systems, is the risk that the heterologous protein poses as an allergen to humans. Any heterologously-expressed enzyme presents an allergenicity risk to those exposed to the protein through inhalation or ingestion. In order to reduce the allergenicity risk of the protein, particularly a plant-expressed protein that could be inadvertently consumed, it is desirable to engineer the phytase so that it has reduced stability when exposed to a gastric environment, an intestinal environment, or when exposed to pepsin. Reduced pepsin stability makes the protein safer as it would be readily digested in the human digestive tract if the plant material containing the engineered phytase was inadvertently ingested.

SUMMARY

[0011] In an aspect, the invention relates to an engineered phytase. The engineered phytase comprises a target phytase, a first binding element and a second binding element. The first binding element is fused to the target phytase, and the second binding element is fused the target phytase. The first binding element interacts with the second binding element to cause cyclization of the engineered phytase and enhance thermal stability of the target phytase. The first binding element is selected from the group consisting of: an intein or part thereof, a coiled-coil dimerization domain or part thereof, a tag domain, and a catcher domain. The second binding element is selected from the group consisting of: a tag domain, a catcher domain, an intein or part thereof, and a coiled-coil dimerization domain or part thereof.

[0012] In an aspect, the invention relates to an engineered nucleic acid encoding any one of the engineered phytases described herein.

[0013] In an aspect, the invention relates to an engineered nucleic acid encoding an engineered phytase. The engineered phytase comprises a target phytase, a first binding element and a second binding element. Each of the first binding element and the second binding is fused to the target phytase. The first binding element interacts with the second binding element to cause cyclization of the engineered phytase, and enhance thermal stability of the target phytase. The first binding element or the second binding element is selected from the group consisting of: a tag domain, a catcher domain, an intein or part thereof, and a coiled-coil dimerization domain or part thereof.

[0014] In an aspect, the invention relates to a vector that comprises any one of the engineered nucleic acids described herein.

[0015] In an aspect, the invention relates to a host comprising any one of the engineered phytases described herein. The host is selected from the group consisting of: a microorganism, a plant cell, a phage, a virus, a mammalian cell, and an insect cell.

[0016] In an aspect, the invention relates to a method of enhancing thermal stability of a target phytase. The method includes producing any one of the engineered phytase described herein.

[0017] In an aspect, the invention relates to a method of preparing an animal feed comprising adding any one of the engineered phytases described herein to the animal feed.

[0018] In an aspect, the invention relates to an animal feed comprising any one of the engineered phytases described herein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0019] The following detailed description of preferred embodiments of the present invention will be better understood when read in conjunction with the appended drawings. For the purpose of illustrating the invention, particular embodiments are shown in the drawings. It is understood, however, that the invention is not limited to the precise arrangements and instrumentalities shown. In the drawings:

[0020] FIG. 1 is a schematic diagram illustrating an engineered phytase with a split intein attached to the ends of the phytase coding sequence (A), binding of the split intein to cyclize the phytase using non-covalent binding (B), and the form of the cyclized phytase that results following splicing of the intein and formation of a covalent bond (C).

[0021] FIG. 2 is a schematic diagram illustrating an engineered phytase with a split intein attached to a linker that connects to the ends of the phytase coding sequence (A), binding of the split intein to cyclize the phytase using non-covalent binding (B), and the form of the cyclized phytase that results following splicing of the intein and formation of a covalent bond (C).

[0022] FIG. 3 is a schematic diagram illustrating an engineered phytase with a coiled coil domain that connects to the ends of the phytase coding sequence (A) and binding of the coiled coil domain to cyclize the phytase using non-covalent binding (B).

[0023] FIG. 4 is a schematic diagram illustrating an engineered phytase with a tag and catcher domain attached to the amino- and carboxy-termini,

respectively, of the phytase coding sequence (A) and binding of the tag and catcher domains to cyclize the phytase using non-covalent binding (B), and the form of the cyclized phytase that results following reaction of the tag-catcher domains to form a covalent bond (C).

[0024] FIG. 5 is a schematic diagram illustrating an engineered phytase with a tag and catcher domain attached to the carboxy- and amino-termini, respectively, of the phytase coding sequence (A) and binding of the tag and catcher domains to cyclize the phytase using non-covalent binding (B), and the form of the cyclized phytase that results following reaction of the tag-catcher domains to form a covalent bond (C).

[0025] FIG. 6 is a schematic diagram illustrating an engineered phytase with a tag and catcher domain attached to a linker that connects to the amino- and carboxy-termini, respectively, of the phytase coding sequence (A), and binding of the tag and catcher domains to cyclize the phytase using non-covalent binding (B), and the form of the cyclized phytase that results following reaction of the tag-catcher domains to form a covalent bond (C).

[0026] FIG. 7 is a schematic diagram illustrating an engineered phytase with a tag and catcher domain attached to a linker that connects to the carboxy- and amino-termini, respectively, of the phytase coding sequence (A), and binding of the tag and catcher domains to cyclize the phytase using non-covalent binding (B), and the form of the cyclized phytase that results following reaction of the tag-catcher domains to form a covalent bond (C).

[0027] FIG. 8 is a schematic diagram illustrating an expression vector pAG4918.

[0028] FIGS. 9A - 9C are schematic diagrams illustrating expression cassettes for selected engineered phytases with split inteins attached to the ends of the phytase coding sequences. FIG. 9A illustrates ZmZ27P:xGZein27ss:Gp41-1C:Phy02opt:Gp41-1N:DPNGSEKDEL:NosT. FIG. 9B illustrates ZmZ27P:Ssp DnaE-C:Phy02opt:Ssp DnaE-N:NosT. FIG. 9C illustrates ZmZ27P:xGZein27ss:Ssp DnaE-C:Phy02opt:Ssp DnaE-N: DPNGSEKDEL: NosT.

[0029] FIGS. 10A - 10H are schematic diagrams illustrating expression cassettes for selected engineered phytases with split intein attached to linkers that connect to the ends of the phytase coding sequences. FIG. 10A illustrates ZmZ27P:Ssp DnaE-C:L33-1:Phy02opt:L33-2:Ssp DnaE-N:NosT. FIG. 10B illustrates ZmZ27P:xGZein27ss: Ssp DnaE-C:L33-1:Phy02opt:L33-2:Ssp DnaE-N:DPNGSEKDEL:NosT. FIG. 10C illustrates ZmZ27P:Ssp DnaE-C:L38-1:Phy02opt:L38-2:Ssp DnaE-N:NosT. FIG. 10D illustrates ZmZ27P : xGZein27ss: Ssp DnaE-C : L38-1 : Phy02opt:L38-2:Ssp DnaE-N : DPNGSEKDEL : NosT. FIG. 10E illustrates ZmZ27P : Ssp DnaE-C:L46-1 :Phy02opt : L46-2 : Ssp DnaE-N : NosT. FIG. 10F illustrates ZmZ27P : xGZein27ss : Ssp DnaE-C : L46-1 : Phy02opt:L46-2 : Ssp DnaE-N : DPNGSEKDEL : NosT. FIG. 10G illustrates ZmZ27P : Ssp DnaE-C:L55-1 : Phy02opt : L55-2 : Ssp DnaE-N : NosT. FIG. 10H illustrates ZmZ27P : xGZein27ss : Ssp DnaE-C:L55-1 : Phy02opt:L55-2 : Ssp DnaE-N : DPNGSEKDEL : NosT.

[0030] FIG. 11 is a photograph of a gel showing expression profiles of SspDnaE-C:Phy02:SspDnaE-N constructs.

[0031] FIG. 12 is a graph illustrating the heat stability assay of Phy02.

[0032] FIGS. 13A - 13B are bar graphs illustrating heat stability of SspDnaE-C:Phy02:SspDnaE-N constructs. FIG. 13A shows enzyme activity of untreated (37 °C) and heat treated (75 °C/ 60 sec) samples. FIG. 13B shows residual phytase activity in heat pretreated samples as percentage of activity of their respective untreated control (37 °C).

[0033] FIG. 14 is a photograph of the gel showing expression profiles of SpyTag:Phy02:SpyCatcher wild type and mutated forms.

[0034] FIGS. 15A - 15B are bar graphs illustrating that SpyTag:Phy02:SpyCatcher improves heat tolerance of phytase. FIG. 15A illustrates phytase activity of heat pretreated samples. FIG. 15B illustrates retention of phytase activity of heat pretreated samples.

[0035] FIG. 16 is a graph illustrating heat pretreatment of cyclic phytases gp41-1C:linker55-1:Phy02:linker55-2:gp41-1N (closed circle) and TrxH:DPNG:gp41-1C[MTT]:linker55-1:Phy02:linker55-2:gp41-1N (closed

square) compared to the wild type enzyme Phy02 (vertical mark) and an empty vector (horizontal mark).

[0036] FIG. 17 is a bar graph illustrating phytase activity of the splicing enabled and splicing disabled (intein N125A and linker S1A) cyclic phytases gp41-1C:linker55-1 : Phy02 : linker55-2 : gp41-1N and TrxH : DPNG : gp41-1C[MTT] : linker55-1 : Phy02 : linker55-2 : gp41-1N and wild type Phy02 phytase following pretreatment at 85°C for 1 minute.

DETAILED DESCRIPTION OF EMBODIMENTS

[0037] Certain terminology is used in the following description for convenience only and is not limiting.

[0038] As used herein, "variant" refers to a molecule that retains a biological activity that is the same or substantially similar to that of the original sequence. The variant may be from the same or different species or be a synthetic sequence based on a natural or prior molecule.

[0039] An embodiment includes an engineered phytase comprising a target phytase, a first binding element and a second binding element. The first binding element may be fused to the target phytase, and the second binding element may be fused to the target phytase. The first binding element may interact with the second binding element to cause cyclization of the engineered phytase, and alter thermal stability of the target phytase.

[0040] Each of the first binding element and the second binding element may be capable of being released from the engineered phytase. The first binding element and the second binding element may be capable of being released from the engineered phytase spontaneously. The first binding element and the second binding element may be capable of being released from the engineered phytase upon exposure to a triggering condition. The triggering condition may be, but is not limited to, a triggering temperature, a triggering pH, a triggering ligand binding, a triggering light, a triggering ion, a triggering concentration of an ion, a triggering sound, a triggering compound, or a triggering concentration of a compound.

[0041] In an embodiment, the target phytase may be any phytase. As used herein, "phytase" is an enzyme capable of catalyzing the hydrolysis of phytic acid. The target phytase may be a phytase derived from a mesophilic, thermophilic, or hyperthermophilic organism. The target phytase may be a phytase derived from an eukaryotic or prokaryotic organism. The target phytase may be, but is not limited to, a phytase derived from *Escherichia coli*, *Aspergillus niger*, *Peniophora lycii*, *Neurospora crassa*, or *Schwaniomyces accidentalis*. The phytase may be modified for improved thermal stability. The thermally stable phytase may have activity when heated to a temperature of 70°C to 90°C. The thermally stable phytase may be active following exposure of a temperature of 70°C to 90°C. The target phytase may be a phytase stable to pepsin digestion, may have an increased stability in the animal digestive tract, and may be produced by a microbial host. The target phytase may be a phytase that is readily degradable by pepsin. The readily degradable phytase may completely degrade in a time period from 45 minutes to 40 minutes, from 40 minutes to 35 minutes, from 35 minutes to 30 minutes, from 30 minutes to 25 minutes, from 25 minutes to 20 minutes, from 20 minutes to 15 minutes, from 15 minutes to 10 minutes, from 10 minutes to 8 minutes, from 8 minutes to 6 minutes, from 6 minutes to 4 minutes, from 4 minutes to 2 minutes of the pepsin treatment. The time period for degradation may be in a range between any two integer value between 2 minutes and 45 minutes. The complete degradation of the phytase by pepsin may occur in 10 minutes. The target phytase may be any phytase that is sold commercially for use in animal feed.

[0042] In an embodiment, the target phytase may be the Phy02 phytase derived from *E. coli*. The Phy02 phytase may be a variant optimized for expression in plants. The variant may be a phytase having an amino acid sequence of SEQ ID NO: 53 and encoded by a codon optimized nucleic acid sequence of SEQ ID NO: 52. The variant may be a phytase having an amino acid sequence of SEQ ID NO: 219 and encoded by a codon optimized nucleic acid sequence of SEQ ID NO: 218. The target phytase may be the Nov9X phytase having an amino acid sequence of SEQ ID NO: 54. The target phytase may be the CQBscks phytase having an amino acid sequence of SEQ ID NO:

56. The target phytase may comprise, consist essentially of, or consist of an amino acid sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of SEQ ID NOS: 53, 54, and 56.

[0043] In an embodiment, a phytase of the composition may be a variant. Variants may include conservative amino acid substitutions: *i.e.*, substitutions with amino acids having similar properties. Conservative substitutions may be a polar for polar amino acid (Glycine (G, Gly), Serine (S, Ser), Threonine (T, Thr), Tyrosine (Y, Tyr), Cysteine (C, Cys), Asparagine (N, Asn) and Glutamine (Q, Gln)); a non-polar for non-polar amino acid (Alanine (A, Ala), Valine (V, Val), Thyptophan (W, Trp), Leucine (L, Leu), Proline (P, Pro), Methionine (M, Met), Phenilalanine (F, Phe)); acidic for acidic amino acid Aspartic acid (D, Asp), Glutamic acid (E, Glu)); basic for basic amino acid (Arginine (R, Arg), Histidine (H, His), Lysine (K, Lys)); charged for charged amino acids (Aspartic acid (D, Asp), Glutamic acid (E, Glu), Histidine (H, His), Lysine (K, Lys) and Arginine (R, Arg)); and a hydrophobic for hydrophobic amino acid (Alanine (A, Ala), Leucine (L, Leu), Isoleucine (I, Ile), Valine (V, Val), Proline (P, Pro), Phenylalanine (F, Phe), Tryptophan (W, Trp) and Methionine (M, Met)). Conservative nucleotide substitutions may be made in a nucleic acid sequence by substituting a codon for an amino acid with a different codon for the same amino acid. Variants may include non-conservative substitutions. A variant may have 40% phytase activity in comparison to the unchanged phytase. A variant may have at least 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% activity, or an integer between any of the two values herein, in comparison to the unchanged phytase. The phytase activity may be determined by a colorimetric enzymatic assay described in Example 6 herein.

[0044] In an embodiment, the one or more proteins having less than 100% identity to its corresponding amino acid sequence of SEQ ID NO: 53 [Phy02], SEQ ID NO: 54 [Nov9X], SEQ ID NO: 56 [CQBscks], and SEQ ID NO: 219 [Phy02opt] is a variant of the referenced protein or amino acid. In an

embodiment, an isolated protein, polypeptide, oligopeptide, or peptide having a sequence with at least 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100% identity to a protein having the sequence of any one of SEQ ID NO: 53 [Phy02], SEQ ID NO: 54 [Nov9X], SEQ ID NO: 56 [CQBscks], and SEQ ID NO: 219 [Phy02opt] along 10 to 50, 10 to 100, 10 to 150, 10 to 300, 10 to 400, 10 to 500, 10 to 600, 10 to 700, 10 to 800, 10 to 900, or 10 to all amino acids of a protein having the sequence of any of one any one of SEQ ID NO: 53 [Phy02], SEQ ID NO: 54 [Nov9X], SEQ ID NO: 56 [CQBscks] and SEQ ID NO: 219 [Phy02opt] is provided. This list of sequence lengths encompasses every full length protein in SEQ ID NO: 53 [Phy02], SEQ ID NO: 54 [Nov9X], SEQ ID NO: 56 [CQBscks], and SEQ ID NO: 219 [Phy02opt] and every smaller length within the list, even for proteins that do not include over 450 amino acids. For example, the lengths of 10 to 50, 10 to 100, 10 to 150, 10 to 300, 10 to 400, and 10 to all amino acids would apply to a sequence with 400 amino acids. A range of amino acid sequence lengths recited herein includes every length of amino sequence within the range, endpoints inclusive. The recited length of amino acids may start at any single position within a reference sequence where enough amino acids follow the single position to accommodate the recited length. The range of sequence lengths can be extended by increments of 10 to 100N amino acids, where N = an integer of ten or greater, for sequences of 1000 amino acids or larger. The fragment of the phytase may be a subsequence of the polypeptides herein that retain at least 40% activity of the phytase. The fragment may have 400, 405, or 410 amino acids. The fragments may include 20, 30, 40, 50, 100, 150, 200, 300, 400 or 410 contiguous amino acids. Embodiments also include nucleic acids encoding said amino acid sequences, and antibodies recognizing epitopes on said amino acid sequences. A less than full length amino acid sequence may be selected from any portion of one of the sequences of SEQ ID NO: 53 [Phy02], SEQ ID NO: 54 [Nov9X], SEQ ID NO: 56 [CQBscks], and SEQ ID NO: 219 [Phy02opt] corresponding to the recited length of amino acids. A less than full length amino acid sequence may be selected from a portion of any one of SEQ ID NO:

53 [Phy02], SEQ ID NO: 54 [Nov9X], SEQ ID NO: 55 [CQBscks], and SEQ ID NO: 219 [Phy02opt].

[0045] In an embodiment, the first binding element and the second binding element may be selected from the group consisting of: inteins or parts thereof, coiled-coil dimerization domains or parts thereof, and tag and catcher domains.

[0046] In an embodiment, the first binding element or the second binding element may be an intein or part thereof. The intein may be split into intein parts. The parts of the split inteins may derive from thermophilic, cis-splicing inteins. The parts of the split inteins may derive from trans-splicing inteins. The parts of the split intein may be used to bind a phytase's termini and thereby improve its thermal stability. As used herein, the term "split inteins" refers to cis- splicing inteins derived from the thermophilic organisms that can be split into trans-splicing intein pairs or parts of trans-splicing inteins. The split inteins may be identified by screening cis-splicing inteins selected from INbase based upon their sequence divergence between molecules. For INbase see Perler, F. B. (2002). InBase: the intein database. *Nucleic acids research*, *30*(1), 383-384, which is incorporated herein by reference as if fully set forth. These artificially split trans-splicing intein pairs may have canonical splicing residues at the N- and C-termini, where each new subdomain would have a net charge of at least 3.5. The artificially split trans-splicing intein pairs may include N-inteins and C-inteins. The N-inteins may be positively charged and the C-inteins may be negatively charged. The N-inteins and the C-inteins may be selected with the goal of not incorporating the internal endonuclease domain into either split intein component when an endonuclease domain was present in the cis-splicing intein precursor from which these split inteins were selected. The division points may be selected based upon sequence alignments to a miniaturized Tth intein (mTth) and the GP41-1 intein. These division points may be modified, and variants of these inteins may be used in the invention. N-inteins and C-inteins may be truncated, extended or modified for optimum performance in binding the termini of the phytase and improving thermal stability, expression, solubility,

specific activity, or gastric stability of digestion of the phytase. A methionine residue may be added to the amino terminus of the C-inteins.

[0047] In an embodiment, the first binding element may be C-intein of an intein and that the second binding element may be an N-intein of an intein. FIG. 1 illustrates that a C-intein may be connected to the N-terminus of the phytase sequence and that an N-intein may be connected to the C-terminus of the phytase sequence. The C-intein may be but is not limited to Cbu_DnaB-C, Mja_GF6P–C, Mja_Hyp1–C, Mja_IF2–C, Mja_Pol1-C, Pab_CDC211–C, Pab_IF2–C, Pab_VMA-C, Pho_IF2-C, Pho-VMA-C, Rma_DnaB–C, Sru_DnaB-C, Tag_Pol1Tsp-TYPol1–C, Ter_RIR14–C, Tko_IF2–C, Tth-HB27DnaE2–C, Gp41-1C, Gp41-1C[MTT], and Ssp DnaE-C. The N-intein may be but is not limited to Cbu_DnaB-N, Mja_GF6P–N, Mja_Hyp1–N, Mja_IF2–N, Mja_Pol1-N, Pab_CDC211–N, Pab_IF2–N, Pab_VMA-N, Pho_IF2-N, Pho-VMA-N, Rma_DnaB–N, Sru_DnaB-N, Tag_Pol1Tsp-TYPol1–N, Ter_RIR14–N, Tko_IF2–N, Tth-HB27DnaE2–N, Gp41-1N, and Ssp DnaE-N. The C-intein may comprise, consist essentially of, or consist of an amino acid sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of: SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 189, 191, and 195, and the N-intein may comprise, consist essentially of, or consist of an amino acid sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of: SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 187, and 193. The first binding element may be Cbu_DnaB-C (SEQ ID NO: 2) and the second binding element may be Cbu_DnaB-N (SEQ ID NO: 1). The first binding element may be Mja_GF6P–C (SEQ ID NO: 4) and the second binding element may be Mja_GF6P–N (SEQ ID NO: 3). The first binding element may be Mja_Hyp1–C (SEQ ID NO: 6) and the second binding element may be Mja_Hyp1–N (SEQ ID NO: 5). The first binding element may be Mja_IF2–C (SEQ ID NO: 8) and the second binding element may be Mja_IF2–N (SEQ ID NO: 7). The first binding element may be Mja_Pol1-C (SEQ ID NO: 10) and the second binding element may be Mja_Pol1–N (SEQ ID NO: 9). The first

binding element may be Pab_CDC211–C (SEQ ID NO: 12) and the second binding element may be Pab_CDC211–N (SEQ ID NO: 11). The first binding element may be Pab_IF2–C (SEQ ID NO: 14) and the second binding element may be Pab_IF2–N (SEQ ID NO: 13). The first binding element may be Pab_VMA-C (SEQ ID NO: 16) and the second binding element may be Pab_VMA-N (SEQ ID NO: 15). The first binding element may be Pho_IF2-C (SEQ ID NO: 18) and the second binding element may be Pho_IF2–N (SEQ ID NO: 17). The first binding element may be Pho_VMA-C (SEQ ID NO: 20) and the second binding element may be Pho_VMA-N (SEQ ID NO: 19). The first binding element may be Rma_DnaB–C (SEQ ID NO: 22) and the second binding element may be Rma_DnaB–N (SEQ ID NO: 21). The first binding element may be Sru_DnaB-C (SEQ ID NO: 24) and the second binding element may be Sru_DnaB-N (SEQ ID NO: 23). The first binding element may be Tag_Pol1Tsp-TYPol1–C (SEQ ID NO: 26) and the second binding element may be Tag_Pol1Tsp-TYPol1–N (SEQ ID NO: 25). The first binding element may be Ter_RIR14–C (SEQ ID NO: 28) and the second binding element may be Ter_RIR14–N (SEQ ID NO: 27). The first binding element may be Tko_IF2–C (SEQ ID NO: 30) and the second binding element may be Tko_IF2–N (SEQ ID NO: 29). The first binding element may be Tth-HB27DnaE2–C (SEQ ID NO: 32) and the second binding element may be Tth-HB27DnaE2–C (SEQ ID NO: 31). The first binding element may be Gp41-1C (SEQ ID NO: 189) and the second binding element may be Gp41-1N (SEQ ID NO: 187). The first binding element may be Gp41-1C[MTT] (SEQ ID NO: 191) and the second binding element may be Gp41-1N (SEQ ID NO: 187). The first binding element may be Ssp DnaE-C (SEQ ID NO: 195) and the second binding element may be Ssp DnaE-N (SEQ ID NO: 193).

[0048] In an embodiment, the first binding element and the second binding element may be coiled-coil dimerization domains. The coiled-coil dimerization domains may bind a target phytase's termini non-covalently. The coiled-coil domains may form stable dimers to bind the phytase's termini. The coiled-coil domains may vary in length and sequence identity, and may be optimized to improve the engineered phytase's thermal stability, specific

activity, gastric stability, gastric digestion, or heterologous expression level in a given expression host. Any coiled-coil domains may be used as the first binding element or the second binding element to bind a phytase's termini and thereby improve its thermal stability.

[0049] In an embodiment, the first binding element may be an N-coil of the coiled-coil dimerization domain and the second binding element may be a C-coil of a coiled-coil dimerization domain. FIG. 3 illustrates that an N-coil may be connected to the N-terminus of the phytase sequence and that a C-coil may be connected to the C-terminus of the phytase sequence. The N-coil may comprise, consist essentially of, or consist of an amino acid sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of SEQ ID NOS: 37 or 39, and the C-coil may comprise, consist essentially of, or consist of an amino acid sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of SEQ ID NOS: 38 or 40. The first binding element may be the cc17 N-terminal coil (SEQ ID NO: 37) and the second binding element may be the cc17 C-terminal coil (SEQ ID NO: 38). The first binding element may be the cc30 N-terminal coil (SEQ ID NO: 39) and the second binding element may be the cc30 C-terminal coil (SEQ ID NO: 40).

[0050] In an embodiment, the first binding element or the second binding element may be a tag-domain or a catcher domain. The tag- and catcher domains may bind the target phytase's termini and may create covalent bonds between the termini. The tag- and catcher domains may help in refolding of the target phytase following exposure to high temperatures, and improving phytase thermal stability. The tag- and catcher-domains may be applied to either a C-terminus or an N-terminus of the target phytase (and newly created termini if the protein sequence is rearranged to facilitate binding of the termini) and generally form a stable isopeptide bond when they react.

[0051] In an embodiment, the first binding element may be a tag domain or a catcher domain. The second binding element may be a tag domain or a catcher domain. The domain selected as the first binding element

may differ from the domain selected as the second binding element. FIG. 4 illustrates that a tag-domain may be connected to the N-terminus of the phytase sequence and that a catcher domain may be connected to the C-terminus of the phytase sequence. FIG. 5 illustrates that a catcher domain may be connected to the N-terminus of the phytase sequence and that a tag domain may be connected to the C-terminus of the phytase sequence. The tag domain may comprise, consist essentially of, or consist of an amino acid sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of SEQ ID NOS: 33 or 34. The catcher domain may comprise, consist essentially of, or consist of an amino acid sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of SEQ ID NOS: 35 or 36. The first binding element may be Phy_catcher1-C (SEQ ID NO: 36) and the second binding element may be Phy_tag1-N (SEQ ID NO: 33). The first binding element may be Phy_tag1-C (SEQ ID NO: 34) and the second binding element may be Phy_catcher1-N (SEQ ID NO: 35).

[0052] To further facilitate binding of the phytase termini using a first binding element or the second binding element, an embodiment provides the engineered phytase that comprises one or more linkers. The one or more linkers may be a first linker and a second linker. The engineered phytase may comprise a first linker. The engineered phytase may comprise a second linker. The engineered phytase may comprise a first linker and a second linker. The first linker may be contiguous with and between the first binding element and the target phytase. The second linker may be contiguous with and between the target phytase and the second binding element. The first linker or the second linker may be a peptide sequence placed contiguously between the target phytase and the first binding element or the second binding element. When using a split intein, either, or both, of the amino-intein (N-intein) and carboxy-intein (C-intein) portions of the split intein may be connected to the first linker or the second linker and to the termini of the target phytase. In naming the linkers, the convention of proceeding an N-linker with a prefix of "N-" was adopted, which denotes that an N-linker would attach to the C-

terminus of a desired binding element and the N-terminus of the phytase. Likewise, the convention of appending the suffix "-C" to the end of the names of the C-linkers was used, which denotes that a C-linker attaches to the C-terminus of the phytase and the N-terminus of a desired binding element.

[0053] In an embodiment, the first linker may be an N-linker and the second linker may be a C-linker. For example, FIG. 2 illustrates that a C-intein may be connected to an N- linker that connects to the N-terminus of the phytase sequence and that an N-intein may be connected to a C-linker that connects to the C-terminus of the phytase sequence. FIGS. 6 and 7 illustrate examples where a tag-domain and catcher-domain may be connected to the phytase using either a linker to the amino- or carboxy-terminus of the phytase. FIG. 6 illustrates that a tag domain may be connected to an N-linker that connects to the N-terminus of the phytase sequence and that a catcher domain may be connected to a C-linker that connects to the C-terminus of the phytase sequence. FIG. 7 illustrates that a catcher domain may be connected to an N-linker that connects to the N-terminus of the phytase sequence and that a tag domain may be connected to a C-linker that connects to the C-terminus of the phytase sequence. The first linker or the second linker may be useful in positioning the first binding element or the second binding element to enhance their binding and thereby enhance overall thermal stability of the resulting engineered phytase. The length (defined as at least one amino acid long), flexibility or rigidity, isoelectric point, structure, hydrophobicity, and sequence of the first linker or the second linker may vary depending upon the target phytase and the binding elements used to engineer the target phytase. The first linker or the second linker, or both, may be used for improving the thermal stability, expression level, pepsin digestibility, pepsin stability, or specific activity of the engineered phytase relative to the engineered phytase using identical binding elements but lacking the first linker or the second linker.

[0054] Variants of the first linker or the second linker may also be used. The first linker or the second linker may be initially used in the engineered phytase, and subsequently amino acids may be substituted to improve the

thermal stability, expression level, specific activity, pepsin stability, or pepsin digestibility of the engineered phytase. The first linker or the second linker may be highly flexible and largely unstructured peptide sequences. The first linker or the second linker may be rigid. The first linker or the second linker may form ordered structures. The ordered structures may be but are not limited to helices or coils, beta-sheets, or other domains. The first linker or the second linker may include a domain that slows down the rate of unfolding of the enzyme or improves the rate of refolding following exposure of the enzyme to higher temperatures. The first linker or the second linker may include a domain or structure that increases the thermal stability of the engineered phytase. The first linker or the second linker may contain another enzyme, or peptide sequence possessing enzymatic activity.

[0055] The first linker or the second linker may be easily modified and optimized for performance with any particular target phytase and molecular structure through mutagenesis techniques including site directed mutagenesis, deletion, insertion, or other methods. The variations of the first linker or the second linker may be constructed by moving an amino acid in the sequence from the N-terminus of an N-linker to the C-terminus of a C-linker, or from the C-terminus of a C-Linker to the N-terminus of an N-linker. The first linker or the second linker may be used to attach an intein molecular structure to the phytase. If intein splicing is desired, the N-terminus of the N-linker must be either a serine, threonine, or cysteine amino acid residue in most cases in order to facilitate intein splicing. Furthermore, it is known that some inteins have preferred insertion site motifs, and when using these linkers with a given intein, it may be beneficial to incorporate either the native insertion site motif, or a preferred insertion site motif, into the linker. See Apgar et al., 2012, A predictive model of intein insertion site for use in the engineering of molecular switches, *PloS one*, 7(5), e37355, which is incorporated herein by reference as if fully set forth.

[0056] In an embodiment, the first linker may comprise, consist essentially of, or consist of a sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a sequence selected from the

group consisting of: SEQ ID NOS: 41, 43, 45, 47, 48, 50, and 51 and the second linker may comprise, consist essentially of, or consist of a sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a sequence selected from the group consisting of: SEQ ID NOS: 42, 44, 46, 49, 50, and 51. The first linker may be L33-1 linker (N-linker) (SEQ ID NO: 41) and the second linker may be L33-2 linker (C-linker) (SEQ ID NO: 42). The first linker may be L38-1 linker (N-linker) (SEQ ID NO: 43) and the second linker may be L38-2 linker (C-linker) (SEQ ID NO: 44). The first linker may be L46-1 linker (N-linker) (SEQ ID NO: 45) and the second linker may be L46-2 linker (C-linker) (SEQ ID NO: 46). The first linker may be L55-1.1 linker (N-linker) (SEQ ID NO: 47) and the second linker may be L55-2 linker (C-linker) (SEQ ID NO: 49). The first linker may be L55-1 linker (N-linker) (SEQ ID NO: 48) and the second linker may be L55-2 linker (C-linker) (SEQ ID NO: 49). The first linker may be Phy_taglink (N-linker) (SEQ ID NO: 50) and the second linker may be Phy_catcherlink (C-linker) (SEQ ID NO: 51). The first linker may be Phy_catcherlink (N-linker) (SEQ ID NO: 51) and the second linker may be Phy_taglink (C-linker) (SEQ ID NO: 50). The thermal stability of the engineered phytase may be enhanced. The phytase activity may be stable at a temperature in a range from 70°C to 90°C. The temperature may be 70°C, 75°C, 80°C, 85°C, 90°C, 70°C to 75°C, 70°C to 80°C, 70°C to 85°C, 70°C to 90°C, or less than 90°C. The engineered phytase modified for thermal stability may be produced by standard molecular biological techniques and then screened. The engineered phytase may be subjected to mutation and then screened for thermal stability. Screening systems that can be utilized may include lambda phage, yeast, or other expression systems that allow production of the protein and/or testing of its physical and/or functional characteristics. From a population of engineered proteins, candidates may be isolated and may be further analyzed. Further analysis may include DNA sequencing, functional assays, structural assays, enzyme activity assays, and monitoring changes in thermal stability, or structure in response to elevated temperature conditions.

[0057] In an embodiment, the engineered phytase may comprise, consist essentially of or consist of an amino acid sequence having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of: SEQ ID NO: 58 [Cbu_DnaB-C:Phy02 :Cbu_DnaB-N (#12 Phy02C)], SEQ ID NO: 60 [Mja_GF6P-C:Phy02: Mja_GF6P-N (#44 Phy02C)], SEQ ID NO: 62 [Mja_Hyp1S–N:Phy02: Mja_Hyp1S-C (#46 Phy02C)], SEQ ID NO: 64 [Mja_IF2–N:Phy02 :Mja_IF2-C (#47 Phy02C)], SEQ ID NO: 66 [Mja_Pol1–C :Phy02: Mja_Pol1–N (#50 Phy02C)], SEQ ID NO: 68 [Pab_CDC211-C :Phy02: Pab_CDC211–N (#79 Phy02C)], SEQ ID NO: 70 [Pab_IF2-C:Phy02:Pab_IF2-N (#81 Phy02C)], SEQ ID NO: 72 [Pab_VMA–C:Phy02:Pab_VMA-N (#92 Phy02C)], SEQ ID NO: 74 [Pho_IF2–C:Phy02:Pho_IF2-N (#103 Phy02C)], SEQ ID NO: 76 [Pho_VMA–C:Phy02:Pho_VMA-N (#110 Phy02C)], SEQ ID NO: 78 [Rma_DnaB–C:Phy02:Rma_DnaB-N (#116 Phy02C)], SEQ ID NO: 80 [Sru_DnaB–C:Phy02:Sru_DnaB-N (#123 Phy02C)], SEQ ID NO: 82 [Tag_Pol1_TspTYPol1–C:Phy02 :Tag_Pol1_TspTYPol1–N (#128 Phy02C)], SEQ ID NO: 84 [Ter_RIR14–C:Phy02:Ter_RIR4-N (#135 Phy02C)], SEQ ID NO: 86 [Tko_IF2-C:Phy02:Tko_IF-N (#143 Phy02C)], SEQ ID NO: 88 [Tth-HB27_DnaE2–C:Phy02:Tth-HB27_DnaE2–N (#150 Phy02C)], SEQ ID NO: 90 [Ssp_DnaE–C:Phy02:Ssp_DnaE-N (#225 Phy02C)], SEQ ID NO: 92 [Gp411–C:Phy02 :Gp411-N (#230 Phy02C)], SEQ ID NO: 93 [Gp411-C:P hy02r14:Gp411-N], SEQ ID NO: 95 [Phy02C-27:SspDnaE (SSp_DnaE-C : L33-1: Phy02: L33-2 : Ssp_DnaE-N)], SEQ ID NO: 97 [Phy02C-32:SspDnaE (SSp_DnaE-C:L38-1: Phy02 : L38-2 : Ssp_DnaE-N)], SEQ ID NO: 99 [Phy02C-40: SspDnaE (SSp_DnaE-C : L46-1: Phy02 : L46-2 : Ssp_DnaE-N)], SEQ ID NO: 101 [Phy02C-49:SspDnaE (SSp_DnaE-C : L55-1 : Phy02: L55-2 : Ssp DnaE-N)], SEQ ID NO: 103 [Phy02-33:cc17 (cc17-N: L33-1-Phy02-L33-2 :cc17-C)], SEQ ID NO: 105 [Phy02-38: cc17 (cc17-N: L38-1-Phy02-L38-2 :cc17-C)], SEQ ID NO: 107 [Phy02-46: cc17 (cc17-N: L46-1 -Phy02- L46-2 :cc17-C)], SEQ ID NO: 109 [Phy02-55: cc17 (cc17-N: L55-1 -Phy02- L55-2 :cc17-C)], SEQ ID NO: 111 [Phy02-33:cc30 (cc30-N: L33-1 -Phy02- L33-2 :cc30-C)], SEQ ID NO: 113 [Phy02-38: cc30 (cc30-N: L38-1 -Phy02- L38-2 :cc30-C)], SEQ ID NO: 115

[Phy02-46: cc30 (cc30-N: L46-1 -Phy02- L46-2 :cc30-C)], SEQ ID NO: 117 [Phy02-55: cc30 (cc30-N: L55-1 -Phy02- L55-2 :cc30-C)], SEQ ID NO: 119 [Tag-Domain :Taglink1 :Phy02 :Catcherlink1: Catcher], SEQ ID NO: 201 [gp41-1C:L55-1:Phy02:L55-2:gp41-1N(#1 gp41-Phy02)], SEQ ID NO: 203 [gp41-1C[MTT]:L55-1:Phy02:L55-2:gp41-1N(#2 gp41-Phy02)], SEQ ID NO: 205 [TrxH:DPNG:gp41-1C[MTT]:L55-1:Phy02:L55-2:gp41-1N(#1 TrxH-Phy02)], and SEQ ID NO: 207 [TrxH:DPNG:gp41-1C[MTT]:L46-1:Phy02:L46-2:gp41-1N (#2 TrxH-Phy02)].

[0058] Determining percent identity of two amino acid sequences or two nucleic acid sequences may include aligning and comparing the amino acid residues or nucleotides at corresponding positions in the two sequences. If all positions in two sequences are occupied by identical amino acid residues or nucleotides then the sequences are said to be 100% identical. Percent identity can be measured by the Smith Waterman algorithm (Smith TF, Waterman MS 1981 "Identification of Common Molecular Subsequences," *Journal of Molecular Biology* 147: 195 -197, which is incorporated by reference in its entirety as if fully set forth).

[0059] In an embodiment, an engineered nucleic acid encoding any one of the engineered phytases described herein is provided. The sequence encoding the target phytase may have at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of: SEQ ID NO: 52 [Phy02], SEQ ID NO: 55 [CQBscks], SEQ ID NO: 185 [Nov9X], and SEQ ID NO: 218[Phy02opt].

[0060] In an embodiment, the engineered nucleic acid may include a sequence that encodes the first binding element, or the second binding element. The engineered nucleic acid may comprise a sequence encoding a C-intein of an intein. The engineered nucleic acid may comprise, consist essentially of, or consist of a sequence having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100% identity to SEQ ID NO: 143 [Cbu_DnaB-C], SEQ ID NO: 145 [Mja_GF6P–C], SEQ ID NO: 147 [Mja_Hyp1–C], SEQ ID NO: 149 [Mja_IF2–C], SEQ ID NO: 151 [Mja_Pol1-C], SEQ ID NO: 153 [Pab_CDC211–C], SEQ ID NO: 155 [Pab_IF2–C], SEQ ID NO: 157 [Pab_VMA-

C], SEQ ID NO: 159 [Pho_IF2-C], SEQ ID NO: 161 [Pho-VMA-C], SEQ ID NO: 163 [Rma_DnaB–C], SEQ ID NO: 165 [Sru_DnaB-C], SEQ ID NO: 167 [Tag_Pol1Tsp-TYPol1–C], SEQ ID NO: 169 [Ter_RIR14–C] SEQ ID NO: 171 [Tko_IF2–C], SEQ ID NO: 173 [Tth-HB27DnaE2–C], SEQ ID NO: 188 [Gp41-1C], SEQ ID NO: 190 [Gp41-1C[MTT]], and SEQ ID NO: 194 [Ssp DnaE-C]. The engineered nucleic acid may comprise a sequence encoding an N-intein of an intein. The engineered nucleic acid may comprise, consist essentially of, or consist of a sequence having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100% identity to a sequence selected from the group consisting of: SEQ ID NO: 142 [Cbu_DnaB-N], SEQ ID NO: 144 [Mja_GF6P–N], SEQ ID NO: 146 [Mja_Hyp1–N], SEQ ID NO: 148 [Mja_IF2–N], SEQ ID NO: 150 [Mja_Pol1-N], SEQ ID NO: 152 [Pab_CDC211–N], SEQ ID NO: 154 [Pab_IF2–N], SEQ ID NO: 156 [Pab_VMA-N], SEQ ID NO: 158 [Pho_IF2-N], SEQ ID NO: 160 [Pho-VMA-N], SEQ ID NO: 162 [Rma_DnaB–N], SEQ ID NO: 164 [Sru_DnaB-N], SEQ ID NO: 166 [Tag_Pol1Tsp-TYPol1–N], SEQ ID NO: 168 [Ter_RIR14–N], SEQ ID NO: 170 [Tko_IF2–N], SEQ ID NO: 172 [Tth-HB27DnaE2–N], SEQ ID NO: 186 [Gp41-1N], and SEQ ID NO: 192 [Ssp DnaE-N].

[0061] The engineered nucleic acid may comprise a sequence encoding an N-coil of the coiled-coil dimerization domain. The engineered nucleic acid may comprise, consist essentially of, or consist of a sequence having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100% identity to a sequence of SEQ ID NO: 178 [cc17 N-terminal coil] or SEQ ID NO: 180 [cc30 N-terminal coil]. The engineered nucleic acid may comprise a sequence encoding a C-coil of the coiled-coil dimerization domain. The engineered nucleic acid may comprise, consist essentially of, or consist of a sequence having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100% identity to SEQ ID NO: 179 [cc17 N-terminal coil] or SEQ ID NO: 181 [cc30 N-terminal coil].

[0062] The engineered nucleic acid may comprise a sequence encoding a tag domain. The engineered nucleic acid may comprise, consist essentially of, or consist of a sequence having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95,

96, 97, 98, 99, or 100% identity to a sequence of SEQ ID NO: 174 [Phy_tag1-N] or SEQ ID NO: 176 [Phy_tag1-C]. The engineered nucleic acid may comprise a sequence encoding a catcher domain. The engineered nucleic acid may comprise, consist essentially of, or consist of a sequence having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100% identity to a sequence of SEQ ID NO: 176 [Phy_catcher1-N] or SEQ ID NO: 177 [Phy_catcher1-C].

[0063] In an embodiment, the engineered nucleic acid may include a sequence that encodes an N-linker or a C-linker. The engineered nucleic acid may comprise a sequence having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of: SEQ ID NO: 120 [L33-1 linker; N-linker], SEQ ID NO: 122 [L38-1 linker; N-linker], SEQ ID NO: 124 [L46-1 linker; N-linker], SEQ ID NO: 126 [L55-1 linker; N-linker] and SEQ ID NO: 188 [L55-1.1 linker; N-linker]. The engineered nucleic acid may comprise a sequence having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of: SEQ ID NO: 121 [L33-2 linker; C-linker], SEQ ID NO: 123 [L38-2 linker; C-linker], SEQ ID NO: 125 [L46-2 linker; C-linker], and SEQ ID NO: 127: [L55-2 linker; C-linker]. The engineered nucleic acid may include sequences of other linkers. The engineered nucleic acid may include sequences of a taglinker or a catcherlinker, or both. The engineered nucleic acid may comprise a sequence having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of: SEQ ID NO: 183 [Phy_taglink1] or SEQ ID NO: 184 [Phy_catcherlink1].

[0064] In an embodiment, the engineered nucleic acid may comprise a sequence having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of: SEQ ID NO: 57 [Cbu_DnaB-C:Phy02:Cbu_DnaB-N (#12 Phy02C)], SEQ ID NO: 59 [Mja_GF6P-C:Phy02:Mja_GF6P-N (#44 Phy02C)], SEQ ID NO: 61 [Mja_Hyp1S–N:Phy02 : Mja_Hyp1S-C (#46 Phy02C)], SEQ ID NO: 63 [Mja_IF2–N:Phy02: Mja_IF2-C (#47 Phy02C)], SEQ ID NO: 65 [Mja_Pol1–C:Phy02:Mja_Pol1–N (#50 Phy02C)], SEQ ID NO: 67 [Pab_CDC211-

C:Phy02:Pab_CDC211–N (#79 Phy02C)], SEQ ID NO: 69 [Pab_IF2-C:Phy02 :Pab_IF2-N(#81 Phy02C)], SEQ ID NO: 71 Pab_VMA–C:Phy02:Pab_VMA-N (#92 Phy02C)], SEQ ID NO: 73 [Pho_IF2–C:Phy02:Pho_IF2-N (#103 Phy02C)], SEQ ID NO: 75 [Pho_VMA–C:Phy02 :Pho_VMA-N (#110 Phy02C)], SEQ ID NO: 77 [Rma_DnaB–C:Phy02: Rma_DnaB-N (#116 Phy02C)], SEQ ID NO: 79 [Sru_DnaB–C:Phy02:Sru_DnaB-N (#123 Phy02C)], SEQ ID NO: 81 [Tag_Pol1_TspTYPol1–C:Phy02:Tag_Pol1_TspTYPol1–N (#128 Phy02C)], SEQ ID NO: 83 [Ter_RIR14–C:Phy02:Ter_RIR4-N (#135 Phy02C)], SEQ ID NO: 85 [Tko_IF2-C:Phy02:Tko_IF-N (#143 Phy02C)], SEQ ID NO: 87 [Tth-HB27_DnaE2–C:Phy02:Tth-HB27_DnaE2–N (#150 Phy02C)], SEQ ID NO: 89 [Ssp_DnaE–C:Phy02:Ssp_DnaE-N (#225 Phy02C)], SEQ ID NO: 91 [Gp411–C:Phy02:Gp411-N(#230 Phy02C)], SEQ ID NO: 94 [Phy02C-27:SspDnaE (SSp_DnaE-C:L33-1: Phy02: L33-2:Ssp_DnaE-N)], SEQ ID NO: 96 [Phy02C-32:SspDnaE (SSp_DnaE-C: L38 -1: Phy02 : L38-2 : Ssp_DnaE-N)], SEQ ID NO: 98 [Phy02C-40: SspDnaE (SSp_DnaE-C:L46-1: Phy02 : L46-2:Ssp_DnaE-N)], SEQ ID NO: 100 Phy02C-49:SspDnaE (SSp_DnaE-C: L55-1 : Phy02: L55-2: Ssp DnaE-N)], SEQ ID NO: 102 [Phy02-33:cc17 (cc17-N: L33-1-Phy02-L33-2 :cc17-C)], SEQ ID NO: 104 [Phy02-38: cc17 (cc17-N: L38-1-Phy02-L38-2 :cc17-C)], SEQ ID NO: 106 Phy02-46: cc17 (cc17-N: L46-1 -Phy02- L46-2 :cc17-C)], SEQ ID NO: 108 [Phy02-55: cc17 (cc17-N: L55-1 -Phy02- L55-2 :cc17-C)], SEQ ID NO: 110 [Phy02-33:cc30 (cc30-N: L33-1 -Phy02- L33-2 :cc30-C)], SEQ ID NO: 112 [Phy02-38: cc30 (cc30-N: L38-1 -Phy02- L38-2 :cc30-C)], SEQ ID NO: 114 [Phy02-46: cc30 (cc30-N: L46-1 -Phy02- L46-2 :cc30-C)], SEQ ID NO: 116 Phy02-55: cc30 (cc30-N: L55-1 -Phy02- L55-2 :cc30-C)], SEQ ID NO: 118 [Tag-Domain:Taglink1:Phy02:Catcherlink1:Catcher], SEQ ID NO: 128 [ZmZ27P : Gp411C:Phy02opt : Gp411N:NosT (#1Phy02opt)], SEQ ID NO: 129 [Z27P : xGZein27ss : Gp411-C : Phy02opt : Gp411-N : DPNGSEKDEL : NosT (#2Phy02opt)], SEQ ID NO: 130 [ZmZ27P : Ssp_DnaE–C : Phy02opt : Ssp_DnaE–N :N osT (#3Phy02op)t], SEQ ID NO: 131 [mZ27P : xGZein27ss : Ssp_DnaE–C : Phy02opt:Ssp_DnaE–N : DPNGSEKDEL : NosT (#4Phy02op)t], SEQ ID NO: 132 [ZmZ27P : Ssp_DnaE : L33-1 :P hy02opt : L33-2 : NosT (SSp_DnaE-C : L33-1 : Phy02opt : L33-2 : Ssp_DnaE-N) #5Phy02opt, SEQ ID

NO: 133 [ZmZ27P : xGZein27ss : Ssp_DnaE : L33-1 : Phy02opt : L33-2 : DPNGSEKDEL : NosT(#6Phy02opt]), SEQ ID NO: 200 [gp41-1C : L55-1 : Phy02 : L55-2 : gp41-1N (#1 gp41-Phy02)], SEQ ID NO: 202 [gp41-1C[MTT] : L55-1 : Phy02:L55-2 : gp41-1N (#2 gp41-Phy02)], SEQ ID NO: 204 [TrxH:DPNG :gp41-1C[MTT] : L55-1 : Phy02 :L55-2 : gp41-1N (#1 TrxH-Phy02)], and SEQ ID NO: 206 [TrxH : DPNG : gp41-1C [MTT] : L46-1 : Phy02 : L46-2 : gp41-1N (#2 TrxH-Phy02)].

[0065] The engineered nucleic acid may be included in the expression cassette. The expression cassette may include at least one regulatory element. The regulatory element may be operably connected to the engineered nucleic acid. In this context, operably linked means that the regulatory element imparts its function on the nucleic acid. The regulatory element may be selected from the group consisting of: a promoter, a signal peptide, a C-terminal extension and a terminator. For example, a regulatory element may be a promoter, and the operably linked promoter would control expression of the engineered nucleic acid.

[0066] The expression of an engineered nucleic acid encoding an engineered phytase from the expression cassette may be under the control of a promoter which provides for transcription of the nucleic acid in a plant. The promoter may be a constitutive promoter or, tissue specific, or an inducible promoter. A constitutive promoter may provide transcription of the nucleic acid throughout most cells and tissues of the plant and during many stages of development but not necessarily all stages. An inducible promoter may initiate transcription of the nucleic acid sequence only when exposed to a particular chemical or environmental stimulus. A tissue specific promoter may be capable of initiating transcription in a particular plant tissue. Plant tissue may be, but is not limited to, a stem, leaves, trichomes, anthers, cob, seed, endosperm, or embryo. The constitutive promoter may be, but is not limited to the Cauliflower Mosaic Virus (CAMV) 35S promoter, the Cestrum Yellow Leaf Curling Virus promoter (CMP), the actin promoter, or the Rubisco small subunit promoter. The tissue specific promoter may be the maize globulin promoter (ZmGlb1), the rice glutelin promoter (prGTL), the maize gamma zein

promoter (ZmZ27), or the maize oleosin promoter (ZmOle). The signal peptide may be but is not limited to a maize gamma zein 27 signal peptide or a rice glutelin B4 signal peptide. The C-terminal extension may be buts is not limited to HvVSD (from the *Hordeum vulgare* vacuolar sorting determinant (Cervelli et al., 2004)) or SEKDEL (Endoplasmic reticulum retention signal; (Arakawa, Chong, & Langridge, 1998; Haq, Mason, Clements, & Arntzen, 1995; Korban, 2002; Munro & Pelham, 1987)). The terminator may be but is not limited to a NOS (from the *Agrobacterium tumefaciens* nopaline synthase gene) terminator or a maize globulin 1 terminator.

[0067] The promoter may be a maize zein 27 promoter. The maize zein 27 promoter (ZmZ27P) may be encoded by a sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of: SEQ ID NO: 137. The signal peptide may be a maize zein 27 signal peptide. The maize zein 27 signal peptide (xGZein27ss) may be encoded by a sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of: SEQ ID NO: 138. The C-terminal extension may be SEKDEL (SEQ ID NO: 140). The SEKDEL may be encoded by a sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a sequence of SEQ ID NO: 139. The terminator may be a NOS terminator. The NOS terminator (NosT) may be encoded by a sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of: SEQ ID NOS: 141.

[0068] In an embodiment, a vector comprising any one the engineered nucleic acids or expression cassettes described herein is provided.

[0069] Any one of the engineered phytase described herein may be expressed in a host. The host may be but is not limited to a microorganism, a plant cell, a phage, a virus, a mammalian cell, or an insect cell. In an embodiment, any one of the engineered phytases may be produced in a plant or plant tissue. The engineered phytases may be produced upon introduction into the plant genome of any one more of the engineered nucleic acids described herein. The engineered nucleic acid may encode the engineered phytase enzyme or fragment thereof. The engineered nucleic acid may be an

expression cassette that directs the plant to express one or more engineered phytases. The methods of introduction of engineered nucleic acids into the plants are known in the art. The method may be transformation of the plant with a vector that includes engineered nucleic acids encoding the one or more of the engineered phytases. The one or more engineered phytases may be isolated from the plant or plant tissue. The one or more engineered phytases expressed in a transgenic plant herein may have activity when exposed to a temperature in the range of 70°C to 90°C, endpoints inclusive. The temperature may be 70°C, 75°C, 80°C, 85°C, 90°C, 70°C to 75°C, 70°C to 80°C, 70°C to 85°C, 70°C to 90°C, or less than 90°C. The one or more engineered phytases may be produced in any transgenic plant.

[0070] In an embodiment, a host comprising any one of the engineered nucleic acids described herein is provided. The host may be but is not limited to a microorganism, a plant cell, a phage, a virus, a mammalian cell, or an insect cell.

[0071] The host may be a transgenic plant or part thereof including an engineered nucleic acid encoding any one or more of the engineered phytases described herein is provided. As used herein, the transgenic plant may refer to a whole transgenic plant or a part thereof. The part may be but is not limited to one or more of leaves, stems, flowers, buds, petals, ovaries, fruits, or seeds. The part may be callus from a transgenic plant. A transgenic plant may be regenerated from parts of a transgenic plant. A transgenic plant may be a product of sexual crossing of a first transgenic plant and a second transgenic plant or a non-transgenic plant where the product plant retains an engineered nucleic acid introduced to the first transgenic plant. An embodiment provides a progeny of any one of the transgenic plants described herein.

[0072] In an embodiment a method of enhancing thermal stability of the target phytase is provided. One mechanism to improve the thermal stability of a target phytase may be to bind its N- and C-termini together in a way that restricts movement of the termini. Restricting movement of the termini may increase the energy necessary for unfolding of the target phytase, as well as

facilitate refolding of the target phytase. Binding of the ends of the target phytase may occur through both intramolecular covalent and non-convalent bonds. It is understood that binding the N- and C-termini of the target phytase may occur specifically in a reaction between the first amino acid of the target phytase and the last amino acid of the target phytase, or between any amino acid in between, such that the reaction between the amino acids improves thermal stability of the target phytase. Likewise, more than two amino acids may be involved in the binding of the termini, especially when the binding either completely or partially uses non-covalent bonds. A variety of intramolecular bonds may be useful for binding the termini of the target phytase including cysteine bonds, peptide bonds, isopeptide bonds, amide bonds, hydrogen bonds, and others. Thus, the method may include producing an engineered phytase by fusing a first binding element, and a second binding element to a target phytase. Within the engineered phytase, the first binding alement may interact with the second binding element. The first binding element may interact with the second binding element to cause cyclization of the engineered phytase. The cyclization of the engineered phytase may alter thermal stability of the target phytase. The first binding element or the second binding element may be any one of the inteins or parts thereof, coiled-coil dimerization domains or parts thereof, tags and catcher domains described herein.

[0073] The step of engineering may include making an expression construct that includes a nucleic acid encoding the engineered phytase.

[0074] The step of making the expression construct may include analyzing the molecular structures that are useful for binding a target phytase's termini and, or, catalyzing a reaction to create a covalent bond between a target phytase's termini. A variety of intramolecular bonds may be useful for binding the termini of the protein including cysteine bonds, peptide bonds, isopeptide bonds, amide bonds, hydrogen bonds, and others. The step of engineereing may include selecting molecular structures that can be used to facilitate either, or both, the formation of covalent or non-covalent bonds within the phytase molecule to improve its thermal stability. These structures

may include inteins, tag and catcher domains, coiled coil domains, and other affinity domains. See Perler et al., 1994, Protein splicing elements: inteins and exteins--a definition of terms and recommended nomenclature. Nucleic acids research, 22(7), 1125; Gogarten et al., 2002, Inteins: structure, function, and evolution. Annual Reviews in Microbiology, 56(1), 263-287; Perler, 2002, InBase: the intein database. Nucleic acids research, 30(1), 383-384; Schoene et al., 2014, SpyTag/SpyCatcher cyclization confers resilience to boiling on a mesophilic enzyme. Angewandte Chemie International Edition, 53(24), 6101-6104.; Zakeri et al., 2012, Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. Proceedings of the National Academy of Sciences, 109(12), E690-E69; US Application 14/774,954, "Use of Dimerization Domains for Temperature Regulation of Enzyme Activity," all of which are incorporated herein by reference as if fully set forth. The molecular structures may be assessed for their ability to bind the phytases termini and form covalent or non-covalent bonds along the phytases termini or at point near the termini. The molecular structures may be used as a first binding element or the second binding element in the method described herein. The molecular structures may be a split intein attached to the termini of a target phytase that may bind its amino-intein and carboxy-intein components together, effectively binding the termini of the phytase, but may not react to form either an isopeptide or peptide bond. Likewise, in some cases, the intein may react to form an isopeptide or peptide bond, in the latter case, releasing the intein segments that were bond to the phytase and leaving a fully cyclized phytase. In each of these cases, the engineered phytase may be tested for improvements in thermal stability relative to the form of the phytase prior to engineering.

[0075] The step of making the expression construct may include making variations of the sequences encoding engineered phytases. The variants of the engineered phytases may be created, screened, and developed further. There are many techniques known in the art for modifying DNA sequences and the corresponding protein sequences they encode. Mutagenesis techniques that may be useful in this regard include site directed mutagenesis, saturating

mutagenesis (where each amino acid is individually substituted at each position in the protein sequence, and improved variants are selected and combined), random mutagenesis, domain swapping or exchange, and others. Additionally, small deletions, or insertions, may be beneficial when optimizing the sequences for thermal stability, specific activity, host expression, gastric stability or gastric digestibility.

[0076] The method may further include linking a nucleic acid that encodes the first binding element, or the second binding element to the nucleic acid encoding the terminus of the target phytase in such a position that effects interaction of the binding elements and causes cyclization of the target phytase. The binding elements may be portions of a split inteins. The first binding element may be a C-intein of an intein. The second binding element may be an N-intein of an intein. FIG. 1 shows that when the C-intein is fused to the N-terminus of the phytase, and the N-intein is fused to the C-terminus of the phytase (A structure), the C-intein associates with the N-intein (B structure). FIG. 1 also shows that following the association, the inteins splice and the termini of the phytase get connected by the covalent bond (C structure), and that the phytase is cyclized. The cyclized phytase shown in structure C may have an enhanced thermal stability compared to the phytases shown in structures A or B. The structure B shown in FIG. 1 is an intermediate structure having association of the C- and N-inteins. The association without splicing may stabilize the engineered phytase. However, this stabilization may not be permanent and may be lost at the dissociation temperature. On the other hand, when association of the C- and N-inteins progresses to splicing, a stable covalent bond may link the termini of the engineered phytase and a permanent structure (C) may be produced that has high thermal stability.

[0077] FIG. 2 shows that the C-intein and the N-intein may be connected to the phytase termini via linkers. As shown in FIG. 2, the N-linker is placed between the C-intein and the N-terminus of the phytase and the C-linker is placed between the N-intein and the C-terminus of the phytase (A structure). When the C-intein associates with the N-intein (B structure), the

inteins splice, and the N-linker gets connected to the C-linker by the covalent bond causing cyclization of the phytase (C structure). The cyclized phytase shown in structure C may have an enhanced thermal stability compared to the phytase shown in structures A or B of FIG. 2.

[0078] The binding elements may be coiled-coil dimerization domains. The first binding element may be an N-coil. The second binding element may be a C-coil. Referring to FIG. 3, the N-coil and C-coil dimerization domains may be fused to the N-terminus and C-terminus of the target protein (A structure). When domains associate, the phytase together with the associated domains form cyclic structure (B structure) which has an enhanced thermal stability compared to the A structure shown in FIG. 3. Coiled coil dimerization domains may be tailored to dissociate at a specific temperature or remain stably associated at high temperature. The stability of coiled coils is proportional with the number of heptad repeats and the correct pairing of the hydrophobic and ionic residues (Lau et al, 1984; Woolfson DN, 2005; Parry at al. 2008, all of which are incorporated herein by reference as if fully set forth). The larger coil interface may increase the strength of dimerization of the coiled coil and may be used to stabilize target proteins above their melting point without covalent linkage.

[0079] The binding elements may be tag- and catcher domains. The first binding element may be a tag domain. The second binding element may be a catcher domain. FIG. 4 shows the tag- domain may be fused to the N-terminus of the target phytase, and the catcher-domain may be fused to the C-terminus of the target phytase (A structure). When domains associate (B structure), they get linked by a covalent bond and form a cyclic structure together with the target phytase (C structure) which has an enhanced thermal stability compared to the phytase shown in A structure of FIG. 4. FIG. 5 shows that tag- and catcher domains are interchangeable and that the catcher- domain may be fused to the N-terminus of the target phytase, and the tag-domain may be fused to the C-terminus of the target phytase. FIGS. 6 and 7 show that tag- and catcher domains may be connected to the phytase termini via linkers. The cyclic structures (C structures) shown in FIGS. 4 - 7 may have enhanced

thermal stability compared to non-cyclized target phytases shown in these figures.

[0080] The step of engineering may further include contacting a host with an expression construct. The expression construct may include any one of the engineered nucleic acids described herein. The expression construct may be inserted in a transformation vector. The transformation vector may be used to transform the host. The transformation may be but is not limited to an *Agrobacterium* - mediated transformation, electroporation with a plasmid DNA, a DNA uptake, a biolistic transformation, a virus-mediated transformation, or a protoplast transformation. The transformation may be any other transformation procedure suitable for a particular host. The method may include selecting the host cell that includes the engineered nucleic acid and expresses the chimeric protein. The method may include regenerating the host cell into a multicellular organism. The method may include multiplying the host cell to obtain a plurality of the host cells that include the engineered nucleic acid and expresses the engineered phytase. The thermal stability of the target phytase may be enhanced.

[0081] In an embodiment, an animal feed that includes any one of the engineered phytases described herein is provided. The term “animal feed” refers to any food, feed, feed composition, preparation, additive, supplement, or mixture suitable and intended for intake by animals for their nourishment and growth. The engineered phytases include in the animal feed may be active in the gastrointestinal or rumen environment of animals. The engineered phytases included the animal feed may be a phytase that is stable to pepsin digestion. The animal may be a monogastric animal. The animal may be a ruminant animal. The monogastric animal may be but is not limited to a chicken, a turkey, a duck, a swine, a fish, a cat, or a dog. The ruminant animal may be but is not limited to cattle, a cow, a sheep, a horse, or a goat. The engineered phytases may be active after preparation of the animal feed. The temperatures which feeds are exposed to during ensiling may be within range of 20°C to 70°C. The temperatures which feeds are exposed to during pelleting may be within range of 70°C to 130°C. The engineered phytases may have

improved thermal stability and may retain activity after being exposed to high temperatures during feed pelleting.

[0082] In an embodiment, the animal feed may further include a feed supplement. The feed supplement may be any plant material. The plant material may be a non-transgenic plant or an engineered plant. The plant material may include an engineered plant or a mutant plant. The plant material may be a grain that contains starch. The plant material may be a grain that contains fiber. The plant material may be a chemically treated forage. The feed supplement may be a mineral. The mineral may be a trace mineral. The mineral may be a macro mineral. The mineral may be rock phosphate or a phosphate salt. The mineral may be calcium phosphate. The feed supplement may be at least one vitamin. The at least one vitamin may be a fat-soluble vitamin. The feed supplement may be an amino acid. The feed supplement may include one or more exogenous enzymes. The one or more exogenous enzymes may include a hydrolytic enzyme. The hydrolytic enzyme may be an enzyme classified under EC3.4 as hydrolase. The hydrolytic enzymes may be but are not limited to xylanases, mannanases, carbohydrases, proteases, peptidases, glucanases, cellulases, lipases, phospholipases, pectinases, galactosidases, laccases, amylases, hemicellulases, or cellobiohydrolases. The hydrolytic enzymes may be expressed in the engineered plants or parts thereof included in the feed supplement. The feed supplement may include purified hydrolytic enzymes. The feed supplements may be but are not limited to growth improving additives, coloring agents, flavorings, stabilizers, limestone, stearine, starch, saccharides, fatty acids, or a gum. The coloring agents may be carotenoids. The carotenoids may be but are not limited to cantaxanthin, beta-carotene, astaxanthin, or lutein. The fatty acids may be polyunsaturated fatty acids. The polyunsaturated fatty acids may include but are not limited to arachidonic acid, docosohexaenoic acid (DHA), eicosapentaenoic acid (EPA) or gamma-linoleic acid. The plant material may be a non-transgenic plant or part thereof. The plant material may include at least one component selected from the group consisting of: barley, wheat, rye, oat, corn, rice, triticale, beet, sugar beet, spinach, cabbage,

quinoa, corn meal, corn pellets, corn oil, distillers grains, forage, wheat meal, wheat pellets, wheat grain, barley grain, barley pellets, soybean meal, soybean oilcake, lupin meal, rapeseed meal, sorghum grain, sorghum pellets, rapeseed, sunflower seed, and cotton seed.

[0083] The feed supplement may include at least one component selected from the group consisting of: soluble solids, fat and vermiculite, limestone, plain salt, DL-methionine, L-lysine, L-threonine, COBAN®, vitamin premix, dicalcium phosphate, selenium premix, choline chloride, sodium chloride, and mineral premix. The feed supplement may include fish meal, fish oil, bone meal, feather meal and animal fat. The feed supplement may include yeast or yeast extract.

[0084] In an embodiment, a method of preparing an animal feed is provided. The method may include producing any one of the engineered phytases described herein by any one of the methods described herein.

[0085] An embodiment provides a method of producing an animal feed. The method may include mixing any one of the transgenic plants or parts thereof described herein, or the progeny thereof with plant material. The transgenic plant may be a progeny of the transgenic plant. The engineered nucleic acid(s) may be included in a genetic construct(s) or an expression cassette(s). The method may comprise making any transgenic plant herein. The transgenic plant or its progeny may be the plant, in which phytase levels may be increased by the method herein. The method may further include pelletizing the mixture. The method may further include adding a feed supplement to the mixture. The feed supplement may include at least one exogenous enzyme. The at least one exogenous enzyme may be a hydrolase selected from the group consisting of: xylanase, mannanase, protease, glucanase, and cellulase. Preparing the animal feed may include combining one or more transgenic plant herein with any other feed supplement.

[0086] An expression cassette having an engineered nucleic acid encoding an engineered phytase in a plant in may be expressed at any point in the methods. The engineered nucleic acid may be expressed prior to the step of step of mixing the plant. The engineered nucleic acid may be expressed

during the step of pelletizing the plant. The expression may be induced. Upon the expression of the nucleic acid(s), the transgenic plant may have an increased level of an engineered phytase compared to the level of a phytase in a non-genetically engineered plant of the same genetic background but lacking the one or more expression cassettes.

[0087] The engineered phytase may be isolated, purified and added to the animal feed as a pure phytase. The engineered phytase may be isolated from the intact host organism and added to the animal feed as a phytase composition. The engineered phytase may be added to the animal feed in admixture with other feed supplements. The transgenic plant including the engineered phytase or the purified engineered phytase may be combined with other feed supplements to form premixes.

[0088] An animal feed may be produced as mash feed. The animal feed may be produced as pellets. The milled feed stuffs may be mixed with the premix that includes any one of the transgenic plants that include an engineered phytase. The engineered phytase may be a phytase stable to pepsin digestion. The milled stuffs may include the plant material and the feed supplements described herein. The feed supplements may include one or more exogenous enzymes described herein. Enzymes may be added as liquid or solid formulations. For mash feed, a solid or liquid enzyme formulation may be added before or during the mixing step. For pelleted feed, the enzyme preparation may be added before or after the pelleting step. The phytase may be included in premix. The premix may also include vitamins and trace minerals. Macro minerals may be added separately to animal feedstock.

[0089] In an embodiment, a method of enhancing thermal stability of a target phytase is provided. The method may include producing a transgenic plant that includes an engineered nucleic acid encoding the phytase. The engineered nucleic acid may include any one the sequences described herein. The phytase may be thermally stable upon exposure to temperatures in the range of 70°C to 90°C, endpoints inclusive. The phytase may be thermally stable upon exposure to temperatures in the range of 70°C to 90°C, endpoints inclusive. The phytase may be thermally stable upon exposure to

temperatures in the range from 70°C, 75°C, 80°C, 85°C, 90°C, 70°C to 75°C, 70°C to 80°C, 70°C to 85°C, 70°C to 90°C, or less than 90°C. The thermally stable phytase may be a phytase that is stable to pepsin digestion.

[0090] The following list includes particular embodiments. The list, however, is not limiting and does not exclude the embodiments otherwise described herein or alternate embodiments.

EMBODIMENTS

1. An engineered phytase comprising a target phytase, a first binding element and a second binding element, wherein each of the first binding element and the second binding is fused to the target phytase, the first binding element interacts with the second binding element to cause cyclization of the engineered phytase, and enhance thermal stability of the target phytase, wherein each of the first binding element and the second binding element is selected from the group consisting of: a tag domain, a catcher domain, an intein or part thereof, and a coiled-coil dimerization domain or part thereof.

2. The engineered phytase of embodiment 1, wherein upon the interaction, each of the first binding element and the second binding element is capable of being released from the engineered phytase spontaneously.

3. The engineered phytase of any one or both of embodiments 1 or 2, wherein upon the interaction, each of the first binding element and the second binding element is capable of being released from the engineered phytase upon exposure to a triggering condition.

4. The engineered phytase of embodiment 3, wherein the triggering condition is selected from the group consisting of triggering temperature, a triggering pH, a triggering ligand binding, a triggering light, a triggering ion, a triggering concentration of an ion, a triggering sound, a triggering compound, or a triggering concentration of a compound.

5. The engineered phytase of any one or more of the preceding embodiments, wherein the first binding element or the second binding element is fused to the N-terminus or the C-terminus of the target phytase.

6. The engineered phytase of any one or more of the preceding embodiments, wherein the N-terminus of the second binding element is linked to and contiguous with the C-terminus of the target phytase.

7. The engineered phytase of any one or more of the preceding embodiments, wherein the C-terminus of the first binding element is linked to and contiguous with the N-terminus of the target phytase, and the N-terminus of the second binding element is linked to and contiguous with the C-terminus of the target phytase.

8. The engineered phytase of any one or more of the preceding embodiments, wherein the target phytase is selected from the group consisting of phytases derived from *Escherichia coli, Aspergillus niger*, *Peniophora lycii, Neurospora crassa,* and *Schwaniomyces accidentalis*.

9. The engineered phytase of any one or more of the preceding embodiments, wherein the target phytase comprises an amino acid sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of SEQ ID NOS: 53 - 54, 56, and 219.

10. The engineered phytase of any one or more of the preceding embodiments, wherein the first binding element is a C-intein of an intein and the second binding element is an N-intein of an intein.

11. The engineered phytase of any one or more of the preceding embodiments, wherein the C-intein comprises an amino acid sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of: SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 189, 191, and 195, and the N-intein comprises an amino acid sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of: SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 187, and 193.

12. The engineered phytase of any one or more of embodiments 1 - 9, wherein the first binding element is a C-coil of the coiled-coil dimerization

domain and the second binding element is an N-coil of a coiled-coil dimerization domain.

13. The engineered phytase of embodiment 12, wherein the C-coil comprises an amino acid sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of SEQ ID NOS: 38 or 40, and the N-coil comprises an amino acid sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of SEQ ID NOS: 37 or 39.

14. The engineered phytase of any one or more of embodiments 1 - 9, wherein each of the first binding element and the second biding element comprises a tag domain or a catcher domain, wherein the domain selected as the first binding element differs from the domain selected as the second binding element.

15. The engineered phytase of embodiment 14, wherein the tag domain comprises an amino acid sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of SEQ ID NOS: 33 or 34.

16. The engineered phytase of embodiment 14 , wherein and the catcher domain comprises an amino acid sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of SEQ ID NOS: 35 or 36.

17. The engineered phytase of any one or more of preceding embodiments further comprising a first linker and a second linker, wherein the first linker is contiguous with and between the first binding element and the target phytase and the second linker is contiguous with and between the target phytase and the second binding element.

18. The engineered phytase of embodiment 17, wherein the first linker comprises a sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a sequence selected from the group consisting of: SEQ ID NOS: 41, 43, 45, 47, 48, 50, and 51, and the second linker comprises a sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or

100% identity to a sequence selected from the group consisting of: SEQ ID NOS: 42, 44, 46, 49, 50, and 51.

19. The engineered phytase of any one or more of the preceding embodiments comprising an amino acid sequence having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of: SEQ ID NOS: 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 201, 203, 205, and 207.

20. The engineered phytase of any one or more of the preceding embodiments, wherein the phytase activity is stable at a temperature in a range from 70°C to 90°C.

21. The engineered phytase of any one or more of the preceding embodiments, wherein the engineered phytase is expressed in a host selected form the group consisting of a microorganism, a plant cell, a phage, a virus, a mammalian cell, and an insect cell.

22. An engineered nucleic acid encoding the engineered phytase of any one or more of the preceding embodiments.

23. An engineered nucleic acid encoding an engineered phytase comprising a target phytase, a first binding element and a second binding element, wherein each of the first binding element and the second binding is fused to the target phytase, the first binding element interacts with the second binding element to cause cyclization of the engineered phytase, and enhance thermal stability of the target phytase, and each of the first binding element and the second binding element is selected from the group consisting of: a tag domain, a catcher domain, an intein or part thereof, and a coiled-coil dimerization domain or part thereof.

24. The engineered nucleic acid of embodiment 23, wherein upon the interaction, each of the first binding element and the second binding element is capable of being released from the engineered phytase spontaneously.

25. The engineered nucleic acid of any one or both of embodiments 23 or 24, wherein upon the interaction, each of the first binding element and the second

binding element is capable of being released from the engineered phytase upon exposure to a triggering condition.

26. The engineered nucleic acid of embodiment 25, wherein the triggering condition is selected from the group consisting of triggering temperature, a triggering pH, a triggering ligand binding, a triggering light, a triggering ion, a triggering concentration of an ion, a triggering sound, a triggering compound, or a triggering concentration of a compound.

27. The engineered nucleic acid of any one or more of embodiments 23 - 26, wherein the first binding element or the second binding element is fused to the N-terminus or the C-terminus of the target phytase.

28. The engineered nucleic acid of any one or more of embodiments 23 - 27, wherein the N-terminus of the second binding element is linked to and contiguous with the C-terminus of the target phytase.

29. The engineered nucleic acid of any one or more of embodiments 23 - 28, wherein the C-terminus of the first binding element is linked to and contiguous with the N-terminus of the target phytase, and the N-terminus of the second binding element is linked to and contiguous with the C-terminus of the target phytase.

30. The engineered nucleic acid of any one or more of embodiments 23 - 29 comprising a sequence encoding the target phytase selected from the group consisting of phytases derived from *Escherichia coli, Aspergillus niger, Peniophora lycii, Neurospora crassa,* and *Schwaniomyces accidentalis*.

31. The engineered nucleic acid of any one or more embodiments 23 - 30 comprising a sequence encoding the target phytase and having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of SEQ ID NOS: 52, 55, 185, and 218.

32. The engineered nucleic acid of any one or more of embodiments 23 - 31 comprising the sequence encoding the first binding element, wherein the first binding element is a C-intein of an intein.

33. The engineered nucleic acid of any one or more of embodiments 23 - 32 comprising the sequence encoding the second binding element, wherein the second binding element is an N-intein of an intein.

34. The engineered nucleic acid of embodiment 32 comprising the sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of: SEQ ID NOS: 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 188, 190, and 194.

35. The engineered nucleic acid of embodiment 33 comprising the sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of: SEQ ID NOS: 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 186, and 192.

36. The engineered nucleic acid of any one or more of embodiments 23 - 31 comprising the sequence encoding the first binding element, wherein the first binding element is a C-coil of the coiled-coil dimerization domain.

37. The engineered nucleic acid of any one or more of embodiments 23 - 31 and 36 comprising the sequence encoding the second binding element, wherein the second binding element is an N-coil of a coiled-coil dimerization domain.

38. The engineered nucleic acid of embodiment 36 comprising the sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of SEQ ID NOS: 179 or 181.

39. The engineered nucleic acid of embodiment 37 comprising the sequence with at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of SEQ ID NOS: 178 or 180.

40. The engineered nucleic acid of any one or more of embodiments 23 - 31 comprising the sequence encoding the first binding element, wherein the first binding element is a tag domain or a catcher domain.

41. The engineered nucleic acid of any one or more of embodiments 23 - 31 and 40 comprising the sequence encoding the second binding element, wherein the second binding element is a tag domain or a catcher domain, and wherein

the sequence selected as the second binding element differs from the sequence selected as the first binding element.

42. The engineered nucleic acid of any one or both of embodiments 40 and 41 comprising a sequence encoding the tag domain and having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of SEQ ID NOS: 174 or 175.

43. The engineered nucleic acid of any one or more of embodiments 40 - 42 comprising a sequence encoding the catcher domain and having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence of SEQ ID NOS: 176 or 177.

44. The engineered nucleic acid of any one or more of embodiments 23 - 43 further comprising a sequence encoding a first linker and a sequence encoding a second linker, wherein the first linker is contiguous with and between the first binding element and the target phytase and the second linker is contiguous with and between the target phytase and the second binding element.

45. The engineered nucleic acid of embodiment 44 comprising a sequence encoding the first linker and having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a sequence selected from the group consisting of: SEQ ID NOS: 120, 122, 124, 126, 182, 183, and 184, and a sequence encoding the second linker and having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a sequence selected from the group consisting of: SEQ ID NOS: 121, 123, 125, 127, 183 and 184.

46. The engineered nucleic acid of any one or more of embodiments 23 - 45 comprising a sequence having at least 70, 72, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% identity to a reference sequence selected from the group consisting of: SEQ ID NOS: 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 128 - 133, 200, 202, 204 and 206.

47. The engineered nucleic acid of any one or more of embodiments 23 - 46 comprising a sequence encoding the engineered phytase having stable phytase activity at a temperature in a range from 70°C to 90°C.

48. The engineered nucleic acid of any one or more of embodiments 23 - 47 expressed in a host is selected form the group consisting of a microorganism, a plant cell, a phage, a virus, a mammalian cell, and an insect cell.

49. The engineered nucleic acid of embodiment 48, wherein the host is the plant cell.

50. A vector comprising the engineered nucleic acid encoding the engineered phytase of any one or more of embodiments 1 - 21.

51. A vector comprising the engineered nucleic acid of any one or more of embodiments 23 - 48.

52. A host comprising the engineered phytase of any one or more of embodiments 1 - 21 or the engineered nucleic acid of any one or more of embodiments 23 - 48, wherein the host is selected from the group consisting of: a microorganism, a plant cell, a phage, a virus, a mammalian cell, and an insect cell.

53. A method of enhancing thermal stability of a target phytase comprising producing the engineered phytase of any one or more of embodiments 1 - 21.

54. An animal feed comprising an engineered phytase of any one or more of embodiments 1 - 21.

55. The animal feeds of embodiment 54 further comprising a feed supplement.

56. The animal feed of embodiment 55, wherein the feed supplement is plant material.

57. The animal feed of embodiment 56, wherein the plant material is a non-transgenic plant or an engineered plant.

58. The animal feed of any one or more of embodiments 54 - 57, wherein the feed supplement includes one or more exogenous enzymes.

59. The animal feed of embodiment 58, wherein the one or more exogenous enzymes includes a hydrolytic enzyme selected from the group consisting of: xylanase, endoglucanase, cellulase, protease, glucanase, amylase and mannanase.

60. The animal feed of any one or more of embodiments 54 - 59, wherein the plant material includes at least one component selected from the group

consisting of: corn meal, corn pellets, wheat meal, wheat pellets, wheat grain, barley grain, barley pellets, soybean meal, soybean oilcake, sorghum grain and sorghum pellets.

61. The animal feed of any one or more of embodiments 55 - 60, wherein the feed supplement includes at least one component selected from the group consisting of: soluble solids, fat and vermiculite, limestone, plain salt, DL-methionine, L-lysine, L-threonine, COBAN®, vitamin premix, dicalcium phosphate, selenium premix, choline chloride, sodium chloride, and mineral premix.

62. A method of preparing an animal feed comprising adding the engineered phytase of any one or more of embodiments 1 - 21 to the animal feed.

63. The method of embodiment 62 further comprising pelletizing the mixture.

64. The method of any one or both of embodiments 62 or 63 further comprising adding a feed supplement to the mixture.

65. The method of embodiment 64, wherein the feed supplement includes at least one exogenous enzyme.

66. The method of embodiment 65, wherein the at least one exogenous enzyme is a hydrolase selected from the group consisting of: xylanase, mannanase, protease, glucanase, and cellulase.

67. A method of promoting the release of inorganic phosphate from a phytic acid or phytate in an animal comprising feeding an animal with an animal feed comprising the engineered phytase of any one or more of embodiments 1 - 22.

68. The method of embodiment 31 further comprising preparing the animal feed according to a method of any one or more of embodiments 62 - 66.

69. The method of any one or both of embodiments 67 or 68, wherein the animal is a monogastric animal or a ruminant animal.

70. A cyclized phytase comprising the engineered phytase of any one or more of embodiments 1, 5 - 9, and 12 - 21, wherein the first binding element is bound to the second binding element.

71. A cyclized phytase comprising the engineered phytase of any one or more of embodiments 1 - 10, and 13 - 21, wherein upon interaction, the first binding element and the second binding element are released from the engineered phytase, and the N-terminus of the target phytase and the C-terminus of the target phytase are linked.

[0091] Further embodiments herein may be formed by supplementing an embodiment with one or more elements from any one or more other embodiments herein, and/or substituting one or more elements from one embodiment with one or more elements from one or more other embodiments herein.

[0092] EXAMPLES

[0093] The following non-limiting examples are provided to illustrate particular embodiments. The embodiments throughout may be supplemented with one or more details from one or more examples below, and/or one or more elements from an embodiment may be substituted with one or more details from one or more examples below.

[0094] ***Example 1. Descriptions of genetic elements for improving phytase thermal stability***

[0095] *Molecular structures or domains for improving phytase thermal stability.* Among the molecular structures that are useful for binding a protein's termini and, or, catalyzing a reaction to create a covalent bond between a protein's termini, are inteins, and tag- and catcher-domains.

[0096] *Inteins.* While any split intein may be used in this invention to bind a phytase's termini and thereby improve its thermal stability, a set of inteins derived from thermophilic, cis-splicing inteins was used. This set was assembled by screening a set of 157 cis-splicing inteins selected from INbase based upon their sequence divergence between molecules. For INbase see Perler, F. B. (2002). InBase: the intein database. *Nucleic acids research*, *30*(1), 383-384, which is incorporated herein by reference as if fully set forth. Cis-splicing inteins from thermophilic organisms were selected and divided into trans-splicing intein pairs. These artificially split inteins were required to have canonical splicing residues at the N- and C-termini, where each new

subdomain would have a net charge of at least 3.5. This resulted in 18 split inteins, of which all N-inteins are positively charged and C-inteins are negatively charged. N- and C-terminal domains were selected with the goal of not incorporating the internal endonuclease domain into either split intein component (that is, either the N-intein or the C-intein) when an endonuclease domain was present in the cis-splicing intein precursor from which these split inteins were selected. Division points were then selected based upon sequence alignments to a miniaturized Tth intein (mTth) and the GP41-1 intein. A methionine residue was added to the amino terminus of the C-inteins in the set below. The sequences of trans-splicing inteins are shown in Table 1 as follows:

Table 1. Sequences of the Trans-splicing Inteins

SEQ ID NO Amino Acid	SEQ ID NO Nucleic Acid	SEQUENCE DESCRIPTION
1	142	Cbu_DnaB–N (#12-N)
2	143	Cbu_DnaB–C (#12-C)
3	144	Mja_GF6P–N (#44-N)
4	145	Mja_GF6P–C (#44-C)
5	146	Mja_Hyp1–N (#46-N)
6	147	Mja_Hyp1–C (#46-C)
7	148	Mja_IF2–N (#47-N)
8	149	Mja_IF2–C (#47-C)
9	150	Mja_Pol1–N (#50-N)
10	151	Mja_Pol1-C (#50-C)
11	152	Pab_CDC211–N (#79-N)
12	153	Pab_CDC211–C (#79-C)
13	154	Pab_IF2–N (#81-N)
14	155	Pab_IF2–C (#81-C)
15	156	Pab_VMA-N (#92-N)
16	157	Pab_VMA-C (#92-C)
17	158	Pho_IF2–N (#103-N)

SEQ ID NO Amino Acid	SEQ ID NO Nucleic Acid	SEQUENCE DESCRIPTION
18	159	Pho_IF2-C (#103-C)
19	160	Pho_VMA-N (#110-N)
20	161	Pho_VMA-C (#110-C)
21	162	Rma_DnaB–N (#116-N)
22	163	Rma_DnaB–C (#116-C)
23	164	Sru_DnaB-N (#123-N)
24	165	Sru_DnaB-C (#123-C)
25	166	Tag_Pol1Tsp-TYPol1–N (#128-N)
26	167	Tag_Pol1Tsp-TYPol1–C (#128-C)
27	168	Ter_RIR14–N (#135-N)
28	169	Ter_RIR14–C (#135-C)
29	170	Tko_IF2–N (#143-N)
30	171	Tko_IF2–C (#143-C)
31	172	Tth-HB27DnaE2–N (#150-N)
32	173	Tth-HB27DnaE2–C (#150-C)
187	186	gp41-1N
189	188	gp41-1C
191	190	gp41-C[MTT]
193	192	Ssp DnaE-N
195	194	Ssp DnaE-C

[0097] *Tag- and catcher domains*. Tag- and catcher-domain can create covalent bonds between the protein's termini and are used to help in refolding of the protein following exposure to high temperatures. The sequences of the tag- and catcher domains are shown in Table 2 as follows.

Table 2. Sequences of the Tag-Catcher Domains

SEQ ID NO Amino Acid	SEQ ID NO Nucleic Acid	SEQUENCE DESCRIPTION
33	174	Phy_tag1-N
34	175	Phy_tag1-C

35	176	Phy_catcher1-N
36	177	Phy_catcher1-C

[0098] *Coiled- coil dimerization domains.* A set of coiled-coil domains may be used as described in Table 3 and illustrated in FIG. 3. The sequences of coiled coil domains are shown in Table 3.

Table 3. Sequences of the Coiled-Coil Domains

SEQ ID NO Amino Acid	SEQ ID NO Nucleic Acid	SEQUENCE DESCRIPTION
37	178	cc17 N-terminal coil
38	179	cc17 C-terminal coil
39	180	cc30 N-terminal coil
40	181	cc30 C-terminal coil

[0099] The coiled-coil cc17 was designed for heat stability, forms dimers at elevated temperatures, which are stable up to at least 60 °C. Conversely, the coiled-coil cc30 forms dimers at temperatures <30°C and begins to dissociate at temperatures around 50°C.

[00100] *Linkers.* Linkers vary in both sequence composition and length. The sequences of the linkers are shown in Table 4.

Table 4. Sequences of the Linkers

SEQ ID NO Amino Acid	SEQ ID NO Nucleic Acid	SEQUENCE DESCRIPTION
41	120	L33-1 linker (N-linker)
42	121	L33-2 linker (C-linker)
43	122	L38-1 linker (N-linker)
44	123	L38-2 linker (C-linker)
45	124	L46-1 linker (N-linker)
46	125	L46-2 linker (C-linker)
47	182	L55-1.1 linker (N-linker)
48	126	L55-1 linker (N-linker)
49	127	L55-2 linker (C-linker)
50	183	Phy_taglink
51	184	Phy_catcherlink
199	198	DPNG linker

[00101] An engineered phytase constructed with a selection of molecular structures and with any desired linker, if necessary, that possesses increased thermal stability may be stable to pepsin digestion, as might be used in a microbial product to increase its stability in the animal, or it may be readily degraded (in less than 30 minutes, or less than 10 minutes) by pepsin to decrease its potential allergenicity.

[00102] *Target Phytases.* Although any phytase can be used as the target phytase of the invention, one target phytase for expression in plants is the

Phy02 phytase variant derived from *E. coli*. The *E. coli* codon optimized sequence (Phy02opt) of the enzyme, without a signal sequence, leader, or first methionine is given below.

[00103] The sequences of the target phytases are shown in Table 5.

Table 5. Sequences of the Target Phytases

SEQ ID NO Amino Acid	SEQ ID NO Nucleic Acid	SEQUENCE DESCRIPTION
53	52	Phy02
219	218	Phy02opt
54	185	Nov9X
56	55	CQBscks

[00104] ***Example 2. Creating an engineered phytase using inteins directly attached to the phytase***

[00105] Genes encoding engineered, or cyclized, phytase molecules are constructed using standard recombinant DNA and molecular biology techniques (Ausubel, Current Methods in Molecular Biology) that are known in the art. Alternatively, fully synthetic genes can be ordered and obtained directly from the design of a specified enzyme sequence. Such synthetic DNA sequences can be obtained from a vendor, codon optimized for expression in any particular organism (microbial, plant, mammalian, *et cetera*), and comprising any desirable restriction sites that may facilitate cloning and expression.

[00106] The DNA sequence of the phytase (Phy02, SEQ ID NO. 52, was used as the target phytase in this example, but could be substituted by other phytases) without the signal sequence, was fused to DNA sequences encoding the trans-splicing intein portions to create a linear molecule encoding the C-intein at the amino terminus of the molecule, whose carboxy terminus was fused directly to the amino terminus of the Phy02 phytase, and with the N-intein's amino terminus fused directly to the carboxy terminus of the Phy02 phytase (C-intein:Phy02:N-intein) as described in FIG. 1. FIG. 1 illustrates an engineered phytase with a split intein attached to the ends of the phytase coding sequence (A), binding of the split intein to cyclize the phytase using

non-covalent binding (B), and the form of the cyclized phytase that results following splicing of the intein and formation of a covalent bond (C). Constructs were cloned between the EcoRI and XhoI sites of the pETDuet I expression vector and transformed into the Shuffle T7 *E. coli* host (NEB). One skilled in the art would be knowledgeable of the requirements for intein splicing and would understand that an appropriate amino acid is necessary at the junction between the C-intein and amino-terminus of the target phytase to facilitate intein splicing. See Apgar et al., 2012, A predictive model of intein insertion site for use in the engineering of molecular switches. *PloS one*, *7*(5), e37355; Xu, M. Q., & Perler, F. B., 1996, The mechanism of protein splicing and its modulation by mutation. *The EMBO journal*, *15*(19), 5146, both of which are incorporated herein by reference as if fully set forth. Whether this single amino acid is considered a linker or as part of the target phytase, is not a critical point of differentiation in this example. In this example, the addition of the single serine amino acid at the N-terminus of the Phy02 phytase, could be considered a linker between the C-intein and target phytase with a length of one amino acid. This single amino acid serine linker can be substituted by a threonine or a cysteine. Nucleotide sequences of the constructs are listed below. Nucleotide sequences of trans-splicing C-intein and N-intein are capitalized, a splicing essential serine (agc) has been added to the N-terminus of the phytase sequences are in bold character, sequences of the Phy02 phytase are in lower case underlined characters.

> The nucleotide sequence encoding Cbu_DnaB-C:Phy02:Cbu_DnaB-N (#12 Phy02C) [Amino Acid (AA)_SEQ ID NO: 58] is as follows:

ATGTCGGACCTGTTCTGGGATAGGATCGTGTCGATTGAGGAGAAGGGGTCTGAGGAGGTCTACGAT
CTCACAGTTCCAAAGTACGCTTCTTGGCTCGCGGATGGGGTTGTTTCACATAAT**agc**gcccaatc
ggaaccggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgaccaa
atttacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctgggtga
actgaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggtggc
agatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtcga
cgaacgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattaccgt
tcatacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctgtca
gctggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattttac
cggtcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatctggc
gctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagt
ctccgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttttct

gctgcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtggaa
cacgctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacgttc
gcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagcaagc
gtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaatct
gggcggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggtga
actggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggtctt
tcagaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagtgaa
gctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttaccca
gattgttaatgaagcacgcatcccggcttgtagtctgTGCGTGACAGGGGACACTCTCATCTGCCT
CGCTGACGGGCGCCGCGTTCCTATTCAGGATCTCGTGGGGCATTCGCCGGAGGTTATTGCGGTCGA
CGATAAGGGCCGCCTCGTTTGCGCTAAGTCAGAGGTCATCTGGAAGGTCGGCGAGCGGTCCGTTTT
CGAGATCAAGCTGGCTTCCGGGAGGAGCATTAAGGCTACCGCTGAGCACAGGCTCCTGGCGTTCAA
GGGCTGGAGGCATGTTAAGGACTTCAAAGTGGGGGATAGGCTCGCTATTGCTCACTAA (SEQ ID
NO: 57)

> The nucleotide sequence encoding Mja_GF6P-C:Phy02:Mja_GF6P-N (#44 Phy02C) [AA_SEQ ID NO: 60] is as follows:

ATGGCGGATATCGTTTGGACGAAGTTCAAGATTGAGGAGGTGGAGAGCGATGTTGAGTATGTGTAC
GATCTGGAGGTGGAGGACTACCACAACTTCATTGGCAATCTCATCATCAACCACAAC**agc**gccca
atcggaaccggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgac
caaatttacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctggg
tgaactgaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggt
ggcagatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgt
cgacgaacgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattac
cgttcatacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctg
tcagctggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattt
taccggtcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatct
ggcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaa
agtctccgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttt
tctgctgcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtg
gaacacgctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacg
ttcgcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagca
agcgtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaa
tctgggcggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcgg
tgaactggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggt
ctttcagaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagt
gaagctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttac
ccagattgttaatgaagcacgcatcccggcttgtagtctgTGCCTGCACCCTGACACATACGTTAT
TCTCCCTGACGGGCGCATGAAGAAGATTTCGGAGATTGATGAGGATGAGGTTCTCTCAGTCAACTT
CGAGGACCTGAAGCTCTACAATAAGAAGATCAAGAAGTTCAAGCACAAGGCTCCGAAGATCCTCTA
CAAGATTAAGACCGCGTTCTCCGAGCTCATCACCACGGGCGAGCATAAGCTGTTCGTGGTCGAGAA
CGGGAAGATCGTCGAGAAGTGCGTTAAGGACCTCAATGGCAGCGAGCTGATCGGGGTTGTGAGGTA
A (SEQ ID NO: 59)

> The nucleotide sequence encoding Mja_Hyp1S–N:Phy02:Mja_Hyp1S-C (#46 Phy02C) [AA_SEQ ID NO: 62] is as follows:

ATGGGGAATTACCTGTACGCTCCCATCATTAGGATCGGCCGGGAGTACTACGACGGGTTCGTCTAC
AATCTGGAGGTGGAGGATGACTCTTCATACGTTACAGTCTCAGGCACTCTGCACAAC**agc**gccca
atcggaaccggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgac
caaatttacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctggg
tgaactgaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggt

ggcagatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgt
cgacgaacgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattac
cgttcatacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctg
tcagctggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattt
taccggtcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatct
ggcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaa
agtctccgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttt
tctgctgcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtg
gaacacgctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacg
ttcgcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagca
agcgtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaa
tctgggcggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcgg
tgaactggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggt
ctttcagaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagt
gaagctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttac
ccagattgttaatgaagcacgcatcccggcttgtagtctgTGCGTTCCGCCTGACACTCTGCTCAT
CCTGGAGAATGGGTTCAAGCGCATCGTGGACATTAAGGTCGGGGACAAGGTCCTGACGCACGAGAA
CCGGTTCAAGAAGGTTGAGAAGGTGTACAAGCGCAGGTACATCGGCGACATCATTAAGATTAAGGT
GCGCTACTTCCCAGAGGAGATCATTCTCACCCCAGAGCACCCTGTCTACGCTATCAAGACGGAGAA
GAGGTGCGATGGCTCTCATGGGATCTGCAAGTTCAACTGCCTCACACAGTACACTAATCCTTCATG
CAAGAAGCGGTACCGCAAGTACAAGAGGGAGTGGATCATTGCCAAGGACCTGAAGGTCGGCGATGT
GATCGTCTACCCGATTCCCAACTAA (SEQ ID NO: 61)

> The nucleotide sequence encoding Mja_IF2–N:Phy02:Mja_IF2-C (#47 Phy02C) [AA_ SEQ ID NO: 64] is as follows:

ATGAACATTGCGTTCGTCGAGGTTGAGGATGTCGAGATCATTGACTACGATGGCTACGTTTACGAT
CTCACAACAGAGACTCATAACTTCATTGCTAATGGCATCGTGGTTCATAAT**agc**gcccaatcgga
accggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgaccaaatt
tacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctgggtgaact
gaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggtggcaga
tggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtcgacga
acgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattaccgttca
tacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctgtcagct
ggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattttaccgg
tcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatctggcgct
gaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagtctc
cgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttttctgct
gcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtggaacac
gctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacgttcgcg
cgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagcaagcgta
tggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaatctggg
cggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggtgaact
ggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggtctttca
gaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagtgaagct
gaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttacccagat
tgttaatgaagcacgcatcccggcttgtagtctgTGCCTGATGCCGCATGAGAAGGTGCTGACGGA
GTACGGGGAGATTAAGATTGAGGACCTGTTCAAGATCGGGAAGGAGATCGTGGAGAAGGACGAGCT
CAAGGAGATCAGGAAGCTGAATATTAAGGTGCACACTCTCAACGAGAATGGCGAGATCAAGATCAT
TAACGCCCCATACGTGTGGAAGCTCAAGCATAAGGGGAAGATGATCAAGGTCAAGCTGAAGAACTG
GCACTCGATCACCACGACACCGGAGCATCCCTTCCTGACCAACAATGGCTGGATCAAGGCGGAGAA
TATTAAGAAGGGGATGTATGTGGCTATCCCTCGCTAA (SEQ ID NO: 63)

> The nucleotide sequence encoding Mja_Pol1–C:Phy02:Mja_Pol1–N (#50 Phy02C)[AA_SEQ ID NO: 66] is as follows:

ATGTACGGGTTCTACGACCTCGACGATGTGTGCGTCTCACTGGAGTCCTACAAGGGCGAGGTGTAC
GATCTCACTCTGGAGGGCAGGCCTTACTACTTCGCCAATGGCATCCTCACTCATAAT**agc**gccca
atcggaaccggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgac
caaatttacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctggg
tgaactgaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggt
ggcagatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgt
cgacgaacgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattac
cgttcatacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctg
tcagctggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattt
taccggtcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatct
ggcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaa
agtctccgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttt
tctgctgcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtg
gaacacgctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacg
ttcgcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagca
agcgtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaa
tctgggcggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcgg
tgaactggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggt
ctttcagaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagt
gaagctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttac
ccagattgttaatgaagcacgcatcccggcttgtagtctg**TGCCATCCAAAGGGGACAAAGGTCGT**
GGTCAAGGGCAAGGGCATCGTGAATATTGAGGACGTTAAGGAGGGGAATTACGTTCTCGGCATCGA
CGGCTGGCAGAAGGTTAAGAAGGTCTGGAAGTACGAGTACGAGGGCGAGCTCATTAACGTTAATGG
GCTGAAGTGCACACCGAACCACAAGATCCCCCTCCGCTACAAGATTAAGCATAAGAAGATCAACAA
GAACGATTACCTGGTGAGGGACATCTACGCGAAGTCGCTCCTGACCAAGTTCAAGGGCGAGGGGAA
GCTCATCCTGTGCAAGTAA(SEQ ID NO: 65)

> The nucleotide sequence encoding Pab_CDC211-C:Phy02:Pab_CDC211–N (#79 Phy02C) [AA_SEQ ID NO: 68] is as follows:

ATGTCCGTGAGCTGGGACGAGGTCGCGGAGATCCTGGAGTACGAGCCAAAGGATCCTTGGGTCTAC
GATCTGCAGGTTCCAGGCTACCACAACTTCCTCGCTAATGGCATCTTCGTTCATAAT**agc**gccca
atcggaaccggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgac
caaatttacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctggg
tgaactgaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggt
ggcagatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgt
cgacgaacgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattac
cgttcatacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctg
tcagctggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattt
taccggtcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatct
ggcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaa
agtctccgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttt
tctgctgcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtg
gaacacgctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacg
ttcgcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagca
agcgtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaa
tctgggcggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcgg
tgaactggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggt
ctttcagaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagt
gaagctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttac
ccagattgttaatgaagcacgcatcccggcttgtagtctg**TGCGTGGATTACGAGACTGAGGTCGT**

GCTGGGGAATGGGGAGCGGAAGAAGATCGGGGAGATCGTGGAGCGGGCTATTGAGGAGGCTGAGAA
GAACGGCAAGCTCGGGCGGGTTGACGATGGCTTCTACGCTCCGATCGACATTGAGGTCTACTCGCT
CGATCTGGAGACCCTCAAGGTTCGGAAGGCGCGGGCAAATATCGCGTGGAAGCGCACAGCTCCAAA
GAAGATGATGCTGGTGAAGACTAGGGGCGGGAAGCGCATTAGGGTCACCCCGACGCACCCCTTCTT
CGTTCTGGAGGAGGGCAAGGTGGCTATGAGGAAGGCCCGGGACCTGGAGGAGGGCAACAAGATCGC
CACGATTGAGTAA (SEQ ID NO: 67)

> The nucleotide sequence encoding Pab_IF2-C:Phy02:Pab_IF2-N (#81 Phy02C)

[AA_SEQ ID NO: 70] is as follows:

ATGACGCTGGTGTTCATCCCCGTTGAGAATGTGGAGGAGGAGGAGTACGACGGCTACGTTTACGAT
CTCACTACGGAGACTCATAACTTCATTGCTAATGGCATCCTCGTTCATAAT**agc**gcccaatcgga
accggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgaccaaatt
tacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctgggtgaact
gaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggtggcaga
tggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtcgacga
acgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattaccgttca
tacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctgtcagct
ggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattttaccgg
tcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatctggcgct
gaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagtctc
cgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttttctgct
gcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtggaacac
gctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacgttcgcg
cgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagcaagcgta
tggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaatctggg
cggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggtgaact
ggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggtctttca
gaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagtgaagct
gaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttacccagat
tgttaatgaagcacgcatcccggcttgtagtctgTGCCTCCTCCCTGATGAGAAGGTCGTGGTTCC
CTCGGTCGGGTTCGTGACACTCAAGGAGCTGTTCGAGACGGCTTCCAAGGTCGTGGAGCGCGACGA
TGAGAAGGAGATCAGGGAGCTCGACGAGCGGATTACCAGCGTTAACGGCGATGGGAAGACGGGCCT
GGTCAAGGCCTCCTACGTGTGGAAGGTTAGGCACAAGGGCAAGGTCATCCGGGTCAAGCTCAAGAA
TTGGCACGGCGTTACAGTGACTCCGGAGCATCCCTTCCTCACCACGAAGGGGTGGAAGAGGGCTGA
CCAGCTGAGGCCAGGCGATTACGTCGCGGTTCCTAGGTAA (SEQ ID NO: 69)

> The nucleotide sequence encoding Pab_VMA–C:Phy02:Pab_VMA-N (#92 Phy02C) [AA_SEQ ID NO:72] is as follows:

ATGACCCATGTTCTGTTCGACGAGATCGTGGAGATTCGGTACATCTCCGAGGGCCAGGAGGTGTAC
GACGTTACTACGGAGACTCATAATTTCATTGGGGGCAACATGCCTACTCTGCTCCACAAC**agc**gc
ccaatcggaaccggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctcc
gaccaaatttacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagct
gggtgaactgaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcct
ggtggcagatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctga
tgtcgacgaacgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccat
taccgttcatacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgt
ctgtcagctggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctga
ttttaccggtcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaa
tctggcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaact

gaaagtctccgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaat
ttttctgctgcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatca
gtggaacacgctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggc
acgttcgcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaa
gcaagcgtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggc
aaatctgggcggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccggg
cggtgaactggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcct
ggtctttcagaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcga
agtgaagctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttt
tacccagattgttaatgaagcacgcatcccggcttgtagtctgTGCGTGGACGGGGACACTCTCGT
GCTGACAAAGGAGTTCGGGCTCATCAAGATCAAGGACCTCTACAAGATTCTGGACGGCAAGGGGAA
GAAGACAGTGAACGGCAATGAGGAGTGGACAGAGCTGGAGAGGCCAATCACTCTGTACGGCTACAA
GGACGGGAAGATCGTCGAGATTAAGGCTACCCACGTTTACAAGGGCTTCTCCGCCGGGATGATCGA
GATTCGGACCCGCACGGGCCGCAAGATTAAGGTCACGCCCATCCATAAGCTCTTCACAGGCAGGGT
TACTAAGAATGGGCTGGAGATCCGGGAGGTCATGGCCAAGGACCTCAAGAAGGGCGATCGGATCAT
TGTGGCGAAGTAA (SEQ ID NO: 71)

> The nucleotide sequence encoding Pho_IF2–C:Phy02:Pho_IF2-N (#103 Phy02C)[AA_SEQ ID NO: 74] is as follows:

ATGAACTTCGTTTTCCTGCCGGTGGAGAAGATCGAGGAGTTCGAGTACGATGGCTACGTCTACGAC
GTTACTACAGAGACTCATAATTTCATTGCTAATGGCATCCTCGTTCATAAT**agc**gcccaatcgga
accggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgaccaaatt
tacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctgggtgaact
gaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggtggcaga
tggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtcgacga
acgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattaccgttca
tacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctgtcagct
ggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattttaccgg
tcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatctggcgct
gaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagtctc
cgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaattttttctgct
gcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtggaacac
gctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacgttcgcg
cgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagcaagcgta
tggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaatctggg
cggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggtgaact
ggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggtctttca
gaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagtgaagct
gaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttacccagat
tgttaatgaagcacgcatcccggcttgtagtctgTGCCTGCTGCCGGAGGAGCGGGTTATTCTGCC
TGACTACGGGCCTATTACTCTGGAGGAGCTCTTCAATATGACAAAGGAGACAGTGTTCAAGGACGA
GGAGAAGGAGGTCCGGAAGCTCGGCATCCGCATGCCAGTGGCTGGCGTCGATGGGCGGGTGCGCCT
GCTGGAGGGCCCCTACGTTTGGAAGGTGCGCTACAAGGGGAAGATGCTCAGGGTCAAGCTGAAGGA
CTGGCACAGCGTGGCTGTCACACCAGAGCATCCCTTCCTCACCACGCGGGGCTGGGTGCGCGCTGA
CCAGCTGAAGCCCGGGGATTACGTTGCCGTGCCAAAGTAA (SEQ ID NO: 73)

> The nucleotide sequence encoding Pho_VMA–C:Phy02:Pho_VMA-N (#110 Phy02C) [AA_SEQ ID NO: 76] is as follows:

ATGCAGCATATCATTTTCGACGAGGTCATCGATGTCAGGTACATTCCGGAGCCCCAGGAGGTGTAC
GATGTTACTACAGAGACTCATAATTTCGTGGGGGGCAACATGCCAACTCTGCTCCACAAT**agc**gc
ccaatcggaaccggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctcc
gaccaaatttacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagct

gggtgaactgaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcct
ggtggcagatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctga
tgtcgacgaacgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccat
taccgttcatacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgt
ctgtcagctggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctga
ttttaccggtcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaa
tctggcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaact
gaaagtctccgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaat
ttttctgctgcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatca
gtggaacacgctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggc
acgttcgcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaa
gcaagcgtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggc
aaatctgggcggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccggg
cggtgaactggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcct
ggtctttcagaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcga
agtgaagctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttt
tacccagattgttaatgaagcacgcatcccggcttgtagtctgTGCGTGGACGGGGACACACTGGT
GCTGACAAAGGAGTTCGGGCTCATCAAGATCAAGGAGCTCTACGAGAAGCTGGACGGCAAGGGGCG
CAAGATTGTGGAGGGCAACGAGGAGTGGACCGAGCTGGAGAAGCCAATCACGGTCTACGGCTACAA
GGACGGGAAGATCGTTGAGATTAAGGCCACCCACGTTTACAAGGGCGTGTCCAGCGGGATGGTCGA
GATCAGGACCCGGACGGGCCGGAAGATCAAGGTGACGCCGATTCACCGCCTGTTCACAGGCAGGGT
CACTAAGGACGGGCTGATCCTCAAGGAGGTCATGGCTATGCATGTTAAGCCCGGCGATAGGATCGC
CGTGGTCAAGTAA (SEQ ID NO: 75)

> The nucleotide sequence encoding Rma_DnaB–C:Phy02:Rma_DnaB-N (#116 Phy02C) [AA_SEQ ID NO:78] is as follows:

ATGTCAGACGTCTACTGGGATCCGATCGTTTCCATTGAGCCCGACGGCGTTGAGGAGGTGTTCGAT
CTCACTGTTCCAGGGCCACATAACTTCGTTGCTAATGACATCATTGCTCATAAT**agc**gcccaatc
ggaaccggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgaccaa
atttacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctgggtga
actgaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggtggc
agatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtcga
cgaacgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattaccgt
tcatacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctgtca
gctggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattttac
cggtcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatctggc
gctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagt
ctccgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttttct
gctgcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtggaa
cacgctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacgttc
gcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagcaagc
gtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaatct
gggcggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggtga
actggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggtctt
tcagaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagtgaa
gctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttaccca
gattgttaatgaagcacgcatcccggcttgtagtctgTGCCTCGCGGGGGACACTCTCATTACACT
GGCTGACGGGCGGCGGGTTCCTATTCGGGAGCTGGTCTCGCAGCAGAATTTCTCGGTCTGGGCGCT
GAACCCGCAGACGTACAGGCTGGAGAGGGCTCGGGTCTCCCGGGCCTTCTGCACAGGCATCAAGCC
CGTTTACAGGCTGACCACGAGGCTCGGGAGGAGCATTAGGGCTACTGCTAATCACCGCTTCCTGAC
CCCACAGGGCTGGAAGAGGGTGGACGAGCTCCAGCCTGGGGATTACCTGGCTCTCCCAAGGTAA
(SEQ ID NO: 77)

> The nucleotide sequence encoding Sru_DnaB–C:Phy02:Sru_DnaB-N (#123 Phy02C) [AA_SEQ ID NO: 80] is as follows:

ATGTGGCGGATGACCGGCATCGATGTCGAGCCCGACGGCGTTGGGGATTACTTCGGCTTCACTCTG
GATGGCAATGGGCGCTTCCTCCTCGGGGATGGCACTGTTACTCATAAT**agc**gcccaatcggaacc
ggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgaccaaatttac
gcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctgggtgaactgac
cccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggtggcagatgg
tctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtcgacgaacg
tacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattaccgttcatac
gcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctgtcagctgga
tgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattttaccggtca
ctaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatctggcgctgaa
acgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagtctccgc
cgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttttctgctgca
gcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtggaacacgct
gctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacgttcgcgcgc
cacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagcaagcgtatgg
cgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaatctgggcgg
tgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggtgaactggt
tttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggtctttcagac
cctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagtgaagctgac
cctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttacccagattgt
taatgaagcacgcatcccggcttgtagtctgTGCCTCGGGAAGGGGACACCGGTTATGATGTACGA
TGGGCGGACAAAGCCAGTGGAGAAGGTGGAGGTCGGGGACAGGCTCATGGGGGACGATGGCAGCCC
AAGGACGGTGCAGTCGCTGGCCAGGGGGAGGGAGCAGATGTACTGGGTCCGCCAGAAGAGGGGCAT
GGACTACAGGGTTAACGAGAGCCACATCCTCTCGCTGAAGAAGTCTAGGAGGGAGGGCGCCCGCGA
CAGGGGGTCAATCGCGGATATTTCCGTCCGCGACTAA (SEQ ID NO: 79)

> The nucleotide sequence encodingTag_Pol1_TspTYPol1 –C:Phy02: Tag_Pol1_TspTYPol1–N (#128 Phy02C) [AA_SEQ ID NO: 82] is as follows:

ATGAATTCTTTCTACAATCTGTCAACCTTCGAGGTGTCATCCGAGTACTACAAGGGCGAGGTCTAC
GATCTCACTCTGGAGGGCAATCCTTACTACTTCGCCAATGGCATCCTCACACATAAT**agc**gccca
atcggaaccggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgac
caaatttacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctggg
tgaactgaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggt
ggcagatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgt
cgacgaacgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattac
cgttcatacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctg
tcagctggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattt
taccggtcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatct
ggcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaa
agtctccgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttt
tctgctgcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtg
gaacacgctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacg
ttcgcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagca
agcgtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaa
tctgggcggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcgg
tgaactggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggt
ctttcagaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagt
gaagctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttac
ccagattgttaatgaagcacgcatcccggcttgtagtctgTGCCATCCTGCGGACACTAAGGTCAT
CGTGAAGGGCAAGGGCATCGTTAATATCTCGGACGTGAAGGAGGGGGACTACATTCTCGGCATCGA

CGGCTGGCAGCGGGTCAAGAAGGTTTGGAAGTACCACTACGAGGGCAAGCTCATCAACATTAATGG
GCTGAAGTGCACGCCGAACCATAAGGTTCCCGTGGTCACAGAGAATGACAGGCAGACTCGCATCAG
GGATTCCCTCGCCAAGAGCTTCCTGTCGGGCAAGGTCAAGGGGAAGATCATTACCACGAAGTAA
(SEQ ID NO: 81)

> The nucleotide sequence encoding Ter_RIR14–C:Phy02:Ter_RIR4-N (#135 Phy02C) [AA_SEQ ID NO: 84] is as follows:

ATGTCGAAGTGCGTCCTCAACTACTCGCCCTACAAGATCGAGTCTGTTAATATTGGCGCTGTGTGC
GACTACAGCTACGATTTCGCCATCGAGGGCATCAATGATAATGACTCTTGGTACTGGCAGGGGGCT
CTCAAGTCTCACAAC**agc**gcccaatcggaaccggaactgaaactggaaagtgtggttattgtgtc
tcgtcatggcgttcgcgctccgaccaaatttacgcagctgatgcaagatgtcaccccggacgcctt
ctatacgtggccggtgaagctgggtgaactgaccccgcgtggcggtgaactgatcgcctatctggg
tcactactggcgtcagcgcctggtggcagatggtctgctgccgaaaaagggctgcccgcagagcgg
tcaagttgcaattatcgctgatgtcgacgaacgtacccgcaaaacgggtgaagcatttgcggccgg
tctggcaccggattgcgccattaccgttcatacgcaggcagataccagctctccggacccgctgtt
caacccgctgaaaaccggcgtctgtcagctggatgtcgcgcaagtgacggacgccattctggaacg
tgcaggcggttccatcgctgattttaccggtcactaccagacggcattccgtgaactggaacgcgt
tctgaactttccgcagtcaaatctggcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgac
ccaagccctgccgagtgaactgaaagtctccgccgacaatgtgtcactgaccggcgcatggtcact
ggcttcgatgctgacggaaatttttctgctgcagcaagcacagggtatgccggaaccgggttgggg
tcgtatcaccgattcgcatcagtggaacacgctgctgagcctgcacaatgcgcagttcgacctgct
gcaacgtaccccggaagtggcacgttcgcgcgccacgccgctgctggatctgattaaaaccgctct
gacgccgcatccgccgcagaagcaagcgtatggcgtgaccctgccgacgagcgttctgtttatcgc
gggtcacgacaccaacctggcaaatctgggcggtgctctggaactgcagtggaccctgccgggtca
accggataacacgccgccgggcggtgaactggttttcgaacgttggcgtcgcctgagcgacaattc
tcagtggatccaagttagcctggtctttcagaccctgcagcaaatgcgcgataaaaccccgctgtt
cctgaacacgccgccgggcgaagtgaagctgaccctggcgggttgcgaagaacgtaacgcccaggg
catgtgttctctggcaggttttacccagattgttaatgaagcacgcatcccggcttgtagtctgTG
CCTGGACAAGACGGCTCTGCGGATTTTCAATCAGGGGCTGCTCTACGCGGATGAGGTCGTGACACC
GGGCTCGGGGGAGACAGTCGGCCTCGGGCTGACGGTCAGGAACGGCATCGGGGCGTCCACAGCCAT
TGCGAATCAGCCGATGGAGCTGGTTGAGATCAAGCTCGCTAACGGCCGGAAGCTGCGCATGACCCC
TAATCACCGGATGTCCGTGAAGGGCAAGTGGATTCATGCCTGCAACCTCAAGCCGGGGATGCTCCT
GGACTACAGCATCGGCGAGTACCAGAAGCGCGAGGACACCCTCCTGATTCCTCTCTAA (SEQ ID NO: 83)

> The nucleotide sequence encoding Tko_IF2-C:Phy02:Tko_IF-N (#143 Phy02C) [AA_SEQID NO: 86] is as follows:

ATGAATCTCGTCTTCATCCCGGTTGAGGACATTGAGGAGTTCGAGTACGAGGGCTACGTTTACGAC
GTTACTACAGAGACTCATAATTTCGTTGCTAATGGCATCCTCGTTCATAAT**agc**gcccaatcgga
accggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgaccaaatt
tacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctgggtgaact
gaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggtggcaga
tggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtcgacga
acgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattaccgttca
tacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctgtcagct
ggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattttaccgg
tcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatctggcgct
gaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagtctc
cgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttttctgct
gcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtggaacac
gctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacgttcgcg
cgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagcaagcgta

tggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaatctggg
cggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggtgaact
ggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggtctttca
gaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagtgaagct
gaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttacccagat
tgttaatgaagcacgcatcccggcttgtagtctgTGCCTGCTGCCGGATGAGAAGGTTATTCTCCC
TGAGCATGGGCCTATTACACTCAAGGGGCTCTTCGATCTCGCTAAGGAGACAGTCGTGGCTGACAA
CGAGAAGGAGATCCGCAAGCTGGGCGCCAAGCTCACCATTGTGGGCGAGGATGGGAGGCTCAGGGT
CCTGGAGAGCCCATACGTTTGGAAGGTGCGGCACCGCGGCAAGATGCTGAGGGTCAAGCTCAAGAA
CTGGCACTCAGTGTCCGTCACGCCAGAGCATCCCTTCCTGACCACGCGGGGCTGGGTGCGCGCTGA
CCAGCTCAAGCCGGGGGATTACGTTGCGGTGCCCAGGTAA (SEQ ID NO: 85)

> The nucleotide sequence encoding Tth-HB27_DnaE2–C:Phy02:Tth-HB27_DnaE2–N (#150 Phy02C) [AA_SEQ ID NO: 88] is as follows:

ATGGCTGAGGTTTACTGGGATCGCGTCGAGGCGGTTGAGCCGCTCGGCGAGGAGGAGGTCTTCGAT
CTCACTGTGGAGGGCACTCATACTTTCGTTGCGGAGGATGTTATCGTTCATAATagcgcccaatc
ggaaccggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgaccaa
atttacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctgggtga
actgaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggtggc
agatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtcga
cgaacgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattaccgt
tcatacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctgtca
gctggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattttac
cggtcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatctggc
gctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagt
ctccgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttttct
gctgcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtggaa
cacgctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacgttc
gcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagcaagc
gtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaatct
gggcggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggtga
actggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggtctt
tcagaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagtgaa
gctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttaccca
gattgttaatgaagcacgcatcccggcttgtagtctgTGCCTGCCTGCGCGGGCTAGGGTCGTGGA
TTGGTGCACAGGGCGGGTCGTTCGGGTCGGGGAGATCGTTAGGGGGGAGGCTAAGGGCGTCTGGGT
GGTCTCCCTGGACGAGGCTAGGCTGAGGCTCGTTCCAAGGCCTGTTGTGGCTGCTTTCCCAAGCGG
CAAGGCTCAGGTGTACGCTCTGAGGACCGCTACGGGCAGGGTGCTGGAGGCGACAGCTAACCACCC
AGTCTACACTCCAGAGGGCTGGAGGCCACTGGGGACCCTCGCTCCTGGCGACTACGTCGCTCTGCC
AAGGTAA (SEQ ID NO: 87)

> The nucleotide sequence encoding Ssp_DnaE–C:Phy02:Ssp_DnaE-N (#225 Phy02C) [AA_SEQ ID NO: 90] is as follow:

ATGGTTAAGGTGATTGGAAGACGTTCTCTTGGTGTTCAAAGGATCTTCGATATCGGATTGCCACAA
GACCACAACTTTCTTCTCGCTAATGGTGCCATCGCTGCCAATagcgcccaatcggaaccggaact
gaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgaccaaatttacgcagct
gatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctgggtgaactgaccccgcg
tggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggtggcagatggtctgct
gccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtcgacgaacgtacccg
caaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattaccgttcatacgcaggc
agataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctgtcagctggatgtcgc
gcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattttaccggtcactacca

gacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatctggcgctgaaacgcga
aaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagtctccgccgacaa
tgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttttctgctgcagcaagc
acagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtggaacacgctgctgag
cctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacgttcgcgcgccacgcc
gctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagcaagcgtatggcgtgac
cctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaatctgggcggtgctct
ggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggtgaactggttttcga
acgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggtctttcagaccctgca
gcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagtgaagctgaccctggc
gggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttacccagattgttaatga
agcacgcatcccggcttgtagtctgTGCCTTTCTTTCGGAACTGAGATCCTTACCGTTGAGTACGG
ACCACTTCCTATTGGTAAGATCGTTTCTGAGGAAATTAACTGCTCAGTGTACTCTGTTGATCCAGA
AGGAAGAGTTTACACTCAGGCTATCGCACAATGGCACGATAGGGGTGAACAAGAGGTTCTGGAGTA
CGAGCTTGAAGATGGATCCGTTATTCGTGCTACCTCTGACCATAGATTCTTGACTACAGATTATCA
GCTTCTCGCTATCGAGGAAATCTTTGCTAGGCAACTTGATCTCCTTACTTTGGAGAACATCAAGCA
GACAGAAGAGGCTCTTGACAACCACAGACTTCCATTCCCTTTGCTCGATGCTGGAACCATCAAGTA
A (SEQ ID NO: 89)

>The nucleotide sequence encoding Gp411–C:Phy02:Gp411-N (#230 Phy02C) [AA_SEQ ID NO: 92] is as follows:

ATGATGCTGAAGAAAATTCTGAAGATCGAAGAACTGGATGAACGTGAACTGATTGACATCGAAGTT
AGCGGCAACCATCTGTTTTACGCGAATGACATTCTGACCCACAACagcgcccaatcggaaccgga
actgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgaccaaatttacgca
gctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctgggtgaactgacccc
gcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggtggcagatggtct
gctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtcgacgaacgtac
ccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattaccgttcatacgca
ggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctgtcagctggatgt
cgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattttaccggtcacta
ccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatctggcgctgaaacg
cgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagtctccgccga
caatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttttctgctgcagca
agcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtggaacacgctgct
gagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacgttcgcgcgccac
gccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagcaagcgtatggcgt
gaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaatctgggcggtgc
tctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggtgaactggtttt
cgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggtctttcagaccct
gcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagtgaagctgaccct
ggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttacccagattgttaa
tgaagcacgcatcccggcttgtagtctgTGTCTGGACCTGAAAACGCAAGTGCAAACCCCGCAAGG
CATGAAGGAAATCTCAAACATCCAAGTCGGTGACCTGGTGCTGTCGAATACCGGCTATAACGAAGT
GCTGAATGTTTTTCCGAAGAGCAAAAAGAAATCTTACAAGATCACGCTGGAAGATGGCAAGGAAAT
TATTTGCAGCGAAGAACATCTGTTCCCGACCCAGACGGGCGAAATGAATATCTCCGGCGGTCTGAA
AGAAGGCATGTGTCTGTACGTCAAGGAATAA (SEQ ID NO: 91)

[00107] One skilled in the art will appreciate that many variations on these sequences can be created, screened, and developed further. There are many techniques known in the art for modifying DNA sequences and the

corresponding protein sequences they encode. Mutagenesis techniques that would be useful in this regard include site directed mutagenesis, saturating mutagenesis (where each amino acid is individually substituted at each position in the protein sequence, and improved variants are selected and combined), random mutagenesis, domain swapping or exchange, and others. Additionally, small deletions, or insertions, may be beneficial when optimizing the sequences for thermal stability, specific activity, host expression, gastric stability or gastric digestibility.

[00108] In this particular example, when it is desired to fuse the inteins directly to the termini of the target phytase without adding another serine amino acid, because the target phytase sequence, Phy02 (SEQ ID NO: 53), begins with AQSEPELKLE... (SEQ ID NO: 134), it is readily apparent that in each of the sequences provided in this example, the added serine amino acid (...S...) between the carboxy terminus of the C-intein (...HN), and the amino terminus of the phytase (AQSEPELKLE... (SEQ ID NO: 134)), would not be necessary if the first two amino acids alanine and glutamine (AQ) of the phytase sequence was deleted (resulting in SEPELKLE... (SEQ ID NO: 135), and the first serine at the resulting amino terminus of the phytase sequence (SEPELKLE... (SEQ ID NO: 135)) was used as the serine to facilitate intein splicing. If it is desired to reassemble the entire target phytase sequence (including the deleted alanine and glutamine) during binding of the termini, the alanine and glutamine removed from the amino terminus of the phytase sequence, can be added to the carboxy terminus of the phytase sequence, right at the junction with the N-intein. In this way, the entire native sequence of the phytase will be reassembled following the intein splicing reaction, with no apparent rearrangement of the target phytase sequence. Likewise, even if the inteins bind to cyclize the protein, but do not splice, the added alanine and glutamine will be in a position spatially similar to where it would have been had it been left at the amino terminus of the phytase following binding of the termini.

[00109] This technique, of removing amino-terminal amino acid residues from the phytase and adding them in sequence to the carboxy terminus, can be

extended and applied to any desired intein insertion point in the target phytase. This provides a general algorithm and technique for facilitating intein-based binding and, or, cyclization of the target phytase. For example, if the termini of the target phytase are spatially too distant to enable effective binding of the termini using inteins, tag-catcher domains, coiled coil domains, or other molecular structures, then a new set of termini can be selected by moving amino acids from the amino terminus and adding them in sequence to the carboxy terminus of the target phytase, and adding the molecular structures to the newly selected termini.

[00110] To illustrate the rearrangement technique described above, the final protein sequence of Gp411–C:Phy02:Gp411-N (#230 Phy02C) could be rearranged as follows (Phy02 (in bold) amino acid string ***AQSEPELKLESVVIV*** (SEQ ID NO: 136) is moved from its N-terminal to its C-terminal). The amino acid sequence of Gp411-C:Phy02r14:Gp411-N is as follows:

MMLKKILKIEELDERELIDIEVSGNHLFYANDILTHN**SRHGVRAPTKFTQ** 50
LMQDVTPDAFYTWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKKGC 100
PQSGQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNP 150
LKTGVCQLDVAQVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNL 200
ALKREKQDESASLTQALPSELKVSADNVSLTGAWSLASMLTEIFLLQQAQ 250
GMPEPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKT 300
ALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELQWTLPGQPDNT 350
PPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLFLNTPPGEVKL 400
TLAGCEERNAQGMCSLAGFTQIVNEARIPACSLA*QSEPELKLESVVIV*CL 450
DLKTQVQTPQGMKEISNIQVGDLVLSNTGYNEVLNVFPKSKKKSYKITLE 500
DGKEIICSEEHLFPTQTGEMNISGGLKEGMCLYVKE* (SEQ ID NO: 93)

[00111] ***Example 3. Creating an engineered phytase using inteins linked to the phytase***

[00112] Similar to Example 2, engineered, or cyclized, phytases can be constructed using linker sequences as illustrated in FIG. 2. FIG. 2 illustrates an engineered phytase with a split intein attached to a linker that connects to the ends of the phytase coding sequence (A), binding of the split intein to cyclize the phytase using non-covalent binding (B), and the form of the cyclized phytase that results following splicing of the intein and formation of a covalent bond (C). Such molecules can be made as described in Example 2, using known recombinant DNA and molecular biology methods, or by directly

ordering the DNA sequences that encode these engineered phytases. Sample linker sequences are listed in Example 1 and were used to construct the following engineered phytases, where the intein sequences are capitalized, the linker sequences are italicized underlined lower case font, and the phytase sequence is lower case and not italicized.

>The amino acid and nucleotide sequence encoding Phy02C-27:SspDnaE (SSp_DnaE-C: L33-1: Phy02: L33-2:Ssp_DnaE-N) are as follows:

ATGGTTAAGGTGATTGGAAGACGTTCTCTTGGTGTTCAAAGGATCTTCGATATCG
GATTGCCACAAGACCACAACTTTCTTCTCGCTAATGGTGCCATCGCTGCCAAT*agc*
*gggggtggcagtggaggcggttcgaccccgcagtccgcatttgcc*gcccaatcggaaccggaactgaaactggaaagtgt
ggttattgtgtctcgtcatggcgttcgcgctccgaccaaatttacgcagctgatgcaagatgtcaccccggacgccttctatac
gtggccggtgaagctgggtgaactgaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctg
gtggcagatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtcgacgaacgtacc
cgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattaccgttcatacgcaggcagataccagctctc
cggacccgctgttcaacccgctgaaaaccggcgtctgtcagctggatgtcgcgcaagtgacggacgccattctggaacgtgc
aggcggttccatcgctgattttaccggtcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaa
atctggcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagtctccgcc
gacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttttctgctgcagcaagcacagggtatgc
cggaaccgggttggggtcgtatcaccgattcgcatcagtggaacacgctgctgagcctgcacaatgcgcagttcgacctgct
gcaacgtaccccggaagtggcacgttcgcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccg
cagaagcaagcgtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaatctgggc
ggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggtgaactggttttcgaacgttgg
cgtcgcctgagcgacaattctcagtggatccaagttagcctggtctttcagaccctgcagcaaatgcgcgataaaaccccgc
tgttcctgaacacgccgccgggcgaagtgaagctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctc
tggcaggttttacccagattgttaatgaagcacgcatcccggcttgtagtct*gggtggcgggagcggtggagggagtgggg*
*gcggt*TGCCTTTCTTTCGGAACTGAGATCCTTACCGTTGAGTACGGACCACTTCCTA
TTGGTAAGATCGTTTCTGAGGAAATTAACTGCTCAGTGTACTCTGTTGATCCAGA
AGGAAGAGTTTACACTCAGGCTATCGCACAATGGCACGATAGGGGTGAACAAGA
GGTTCTGGAGTACGAGCTTGAAGATGGATCCGTTATTCGTGCTACCTCTGACCAT
AGATTCTTGACTACAGATTATCAGCTTCTCGCTATCGAGGAAATCTTTGCTAGGC
AACTTGATCTCCTTACTTTGGAGAACATCAAGCAGACAGAAGAGGCTCTTGACAA
CCACAGACTTCCATTCCCTTTGCTCGATGCTGGAACCATCAAGTAA (SEQ ID NO: 94)

MVKVIGRRSLGVQRIFDIGLPQDHNFLLANGAIAANSGGGSGGGSTPQSA 50
FAAQSEPELKLESVVIVSRHGVRAPTKFTQLMQDVTPDAFYTWPVKLGEL 100
TPRGGELIAYLGHYWRQRLVADGLLPKKGCPQSGQVAIIADVDERTRKTG 150
EAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDVAQVTDAILER 200
AGGSIADFTGHYQTAFRELERVLNFPQSNLALKREKQDESASLTQALPSE 250
LKVSADNVSLTGAWSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLL 300
SLHNAQFDLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSV 350
LFIAGHDTNLANLGGALELQWTLPGQPDNTPPGGELVFERWRRLSDNSQW 400
IQVSLVFQTLQQMRDKTPLFLNTPPGEVKLTLAGCEERNAQGMCSLAGFT 450
QIVNEARIPACSLGGGSGGGSGGGCLSFGTEILTVEYGPLPIGKIVSEEI 500
NCSVYSVDPEGRVYTQAIAQWHDRGEQEVLEYELEDGSVIRATSDHRFLT 550
TDYQLLAIEEIFARQLDLLTLENIKQTEEALDNHRLPFPLLDAGTIK* SEQ ID NO: 95)

>The amino acid and nucleotide sequence encoding Phy02C-32:SspDnaE (SSp_DnaE-C:L38-1: Phy02 : L38-2:Ssp_DnaE-N) are as follows:

>ATGGTTAAGGTGATTGGAAGACGTTCTCTTGGTGTTCAAAGGATCTTCGATATCGGATTGCCACA
AGACCACAACTTTCTTCTCGCTAATGGTGCCATCGCTGCCAAT*agcggtggctcgtcagggagtac*
*gacaaccacgcgtatcaccccgcaatctgcgttcgct*gcccaatcggaaccggaactgaaactgga
aagtgtggttattgtgtctcgtcatggcgttcgcgctccgaccaaatttacgcagctgatgcaaga
tgtcaccccggacgccttctatacgtggccggtgaagctgggtgaactgaccccgcgtggcggtga
actgatcgcctatctgggtcactactggcgtcagcgcctggtggcagatggtctgctgccgaaaaa
gggctgcccgcagagcggtcaagttgcaattatcgctgatgtcgacgaacgtacccgcaaaacggg
tgaagcatttgcggccggtctggcaccggattgcgccattaccgttcatacgcaggcagataccag
ctctccggacccgctgttcaacccgctgaaaaccggcgtctgtcagctggatgtcgcgcaagtgac
ggacgccattctggaacgtgcaggcggttccatcgctgattttaccggtcactaccagacggcatt
ccgtgaactggaacgcgttctgaactttccgcagtcaaatctggcgctgaaacgcgaaaagcagga
tgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagtctccgccgacaatgtgtcact
gaccggcgcatggtcactggcttcgatgctgacggaaatttttctgctgcagcaagcacagggtat
gccggaaccgggttggggtcgtatcaccgattcgcatcagtggaacacgctgctgagcctgcacaa
tgcgcagttcgacctgctgcaacgtaccccggaagtggcacgttcgcgcgccacgccgctgctgga
tctgattaaaaccgctctgacgccgcatccgccgcagaagcaagcgtatggcgtgaccctgccgac
gagcgttctgtttatcgcgggtcacgacaccaacctggcaaatctgggcggtgctctggaactgca
gtggaccctgccgggtcaaccggataacacgccgccgggcggtgaactggttttcgaacgttggcg
tcgcctgagcgacaattctcagtggatccaagttagcctggtctttcagaccctgcagcaaatgcg
cgataaaaccccgctgttcctgaacacgccgccgggcgaagtgaagctgaccctggcgggttgcga
agaacgtaacgcccagggcatgtgttctctggcaggttttacccagattgttaatgaagcacgcat
cccggcttgtagtctg*caaaacacgtttagccaggggagtagctcgggatcc*TGCCTTTCTTTCGG
AACTGAGATCCTTACCGTTGAGTACGGACCACTTCCTATTGGTAAGATCGTTTCTGAGGAAATTAA
CTGCTCAGTGTACTCTGTTGATCCAGAAGGAAGAGTTTACACTCAGGCTATCGCACAATGGCACGA
TAGGGGTGAACAAGAGGTTCTGGAGTACGAGCTTGAAGATGGATCCGTTATTCGTGCTACCTCTGA
CCATAGATTCTTGACTACAGATTATCAGCTTCTCGCTATCGAGGAAATCTTTGCTAGGCAACTTGA
TCTCCTTACTTTGGAGAACATCAAGCAGACAGAAGAGGCTCTTGACAACCACAGACTTCCATTCCC
TTTGCTCGATGCTGGAACCATCAAGTAA (SEQ ID NO: 96)

MVKVIGRRSLGVQRIFDIGLPQDHNFLLANGAIAANSGGSSGSTTTTRIT 50
PQSAFAAQSEPELKLESVVIVSRHGVRAPTKFTQLMQDVTPDAFYTWPVK 100
LGELTPRGGELIAYLGHYWRQRLVADGLLPKKGCPQSGQVAIIADVDERT 150
RKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDVAQVTDA 200
ILERAGGSIADFTGHYQTAFRELERVLNFPQSNLALKREKQDESASLTQA 250
LPSELKVSADNVSLTGAWSLASMLTEIFLLQQAQGMPEPGWGRITDSHQW 300
NTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTL 350
PTSVLFIAGHDTNLANLGGALELQWTLPGQPDNTPPGGELVFERWRRLSD 400
NSQWIQVSLVFQTLQQMRDKTPLFLNTPPGEVKLTLAGCEERNAQGMCSL 450
AGFTQIVNEARIPACSL*QNTFSQGSSSGS*CLSFGTEILTVEYGPLPIGKI 500
VSEEINCSVYSVDPEGRVYTQAIAQWHDRGEQEVLEYELEDGSVIRATSD 550
HRFLTTDYQLLAIEEIFARQLDLLTLENIKQTEEALDNHRLPFPLLDAGT 600
IK* (SEQ ID NO: 97)

>The amino acid and nucleotide sequence encoding Phy02C-40: SspDnaE (SSp_DnaE-C:L46-1: Phy02 : L46-2:Ssp_DnaE-N) are as follows:

>ATGGTTAAGGTGATTGGAAGACGTTCTCTTGGTGTTCAAAGGATCTTCGATATCGGATTGCCACA
AGACCACAACTTTCTTCTCGCTAATGGTGCCATCGCTGCCAAT*agcgcctttgca*gcccaatcgga
accggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgaccaaatt
tacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctgggtgaact
gaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggtggcaga

tggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtcgacga
acgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattaccgttca
tacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctgtcagct
ggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattttaccgg
tcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatctggcgct
gaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagtctc
cgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttttctgct
gcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtggaacac
gctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacgttcgcg
cgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagcaagcgta
tggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaatctggg
cggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggtgaact
ggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggtctttca
gaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagtgaagct
gaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttacccagat
tgttaatgaagcacgcatcccggcttgtagtctg*ggtgcagctccagcggccgcaccggctaaaca*
ggaagcggcagctccggctcctgcagcgaaggcggaagcaccggccgcagctcctgcggcaaaagc
*gaccccgcag*TGCCTTTCTTTCGGAACTGAGATCCTTACCGTTGAGTACGGACCACTTCCTATTGG
TAAGATCGTTTCTGAGGAAATTAACTGCTCAGTGTACTCTGTTGATCCAGAAGGAAGAGTTTACAC
TCAGGCTATCGCACAATGGCACGATAGGGGTGAACAAGAGGTTCTGGAGTACGAGCTTGAAGATGG
ATCCGTTATTCGTGCTACCTCTGACCATAGATTCTTGACTACAGATTATCAGCTTCTCGCTATCGA
GGAAATCTTTGCTAGGCAACTTGATCTCCTTACTTTGGAGAACATCAAGCAGACAGAAGAGGCTCT
TGACAACCACAGACTTCCATTCCCTTTGCTCGATGCTGGAACCATCAAGTAA (SEQ ID NO: 98)

MVKVIGRRSLGVQRIFDIGLPQDHNFLLANGAIAANSAFAAQSEPELKLE 50
SVVIVSRHGVRAPTKFTQLMQDVTPDAFYTWPVKLGELTPRGGELIAYLG 100
HYWRQRLVADGLLPKKGCPQSGQVAIIADVDERTRKTGEAFAAGLAPDCA 150
ITVHTQADTSSPDPLFNPLKTGVCQLDVAQVTDAILERAGGSIADFTGHY 200
QTAFRELERVLNFPQSNLALKREKQDESASLTQALPSELKVSADNVSLTG 250
AWSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQR 300
TPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSVLFIAGHDTNLAN 350
LGGALELQWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQ 400
MRDKTPLFLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARIPACS 450
LGAAPAAAPAKQEAAAPAPAAKAEAPAAAPAAKATPQCLSFGTEILTVEY 500
GPLPIGKIVSEEINCSVYSVDPEGRVYTQAIAQWHDRGEQEVLEYELEDG 550
SVIRATSDHRFLTTDYQLLAIEEIFARQLDLLTLENIKQTEEALDNHRLP 600
FPLLDAGTIK* (SEQ ID NO: 99)

>The amino acid and nucleotide sequence encoding Phy02C-49:SspDnaE (SSp_DnaE-C:L55-1:Phy02: L55-2: Ssp DnaE-N) are as follows:

ATGGTTAAGGTGATTGGAAGACGTTCTCTTGGTGTTCAAAGGATCTTCGATATCGGATTGCCACAA
GACCACAACTTTCTTCTCGCTAATGGTGCCATCGCTGCCAAT*agcgcagccgaagccgctgcgaag*
gaggcagctgcgaaagaagcggctgcaaaagaagcggcagctaaggctttgaatacccccgcaatcg
*gctttcgctg*ccaatcggaaccggaactgaaactggaaagtgtggttattgtgtctcgtcatggc
gttcgcgctccgaccaaatttacgcagctgatgcaagatgtcaccccggacgccttctatacgtgg
ccggtgaagctgggtgaactgaccccgcgtggcggtgaactgatcgcctatctgggtcactactgg
cgtcagcgcctggtggcagatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgca
attatcgctgatgtcgacgaacgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccg
gattgcgccattaccgttcatacgcaggcagataccagctctccggacccgctgttcaacccgctg
aaaaccggcgtctgtcagctggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggt
tccatcgctgattttaccggtcactaccagacggcattccgtgaactggaacgcgttctgaacttt
ccgcagtcaaatctggcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctg
ccgagtgaactgaaagtctccgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatg

ctgacggaaatttttctgctgcagcaagcacagggtatgccggaaccgggttggggtcgtatcacc
gattcgcatcagtggaacacgctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtacc
ccggaagtggcacgttcgcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcat
ccgccgcagaagcaagcgtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgac
accaacctggcaaatctgggcggtgctctggaactgcagtggaccctgccgggtcaaccggataac
acgccgccgggcggtgaactggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatc
caagttagcctggtctttcagaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacg
ccgccgggcgaagtgaagctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttct
ctggcaggttttacccagattgttaatgaagcacgcatcccggcttgtagtctg*ggggggcgcagaa*
*gcagctgccaaagaggcggccgcaaaggtcaatctg*TGCCTTTCTTTCGGAACTGAGATCCTTACC
GTTGAGTACGGACCACTTCCTATTGGTAAGATCGTTTCTGAGGAAATTAACTGCTCAGTGTACTCT
GTTGATCCAGAAGGAAGAGTTTACACTCAGGCTATCGCACAATGGCACGATAGGGGTGAACAAGAG
GTTCTGGAGTACGAGCTTGAAGATGGATCCGTTATTCGTGCTACCTCTGACCATAGATTCTTGACT
ACAGATTATCAGCTTCTCGCTATCGAGGAAATCTTTGCTAGGCAACTTGATCTCCTTACTTTGGAG
AACATCAAGCAGACAGAAGAGGCTCTTGACAACCACAGACTTCCATTCCCTTTGCTCGATGCTGGA
ACCATCAAGTAA (SEQ ID NO: 100)

MVKVIGRRSLGVQRIFDIGLPQDHNFLLANGAIAANSAAEAAAKEAAAKE 50
AAAKEAAAKALNTPQSAFAAQSEPELKLESVVIVSRHGVRAPTKFTQLMQ 100
DVTPDAFYTWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKKGCPQS 150
GQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKT 200
GVCQLDVAQVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNLALK 250
REKQDESASLTQALPSELKVSADNVSLTGAWSLASMLTEIFLLQQAQGMP 300
EPGWGRITDSHQWNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTALT 350
PHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELQWTLPGQPDNTPPG 400
GELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLFLNTPPGEVKLTLA 450
GCEERNAQGMCSLAGFTQIVNEARIPACSLGGAEAAAKEAAAKVNLCLSF 500
GTEILTVEYGPLPIGKIVSEEINCSVYSVDPEGRVYTQAIAQWHDRGEQE 550
VLEYELEDGSVIRATSDHRFLTTDYQLLAIEEIFARQLDLLTLENIKQTE 600
EALDNHRLPFPLLDAGTIK* (SEQ ID NO: 101)

[00113] These engineered phytases can be evaluated the same as other molecules for thermal stability, heterologous expression levels from any desirable host (microbial, plant, or otherwise), specific activity, gastric stability or gastric digestion using known techniques (Thomas, K., et al., 2004, A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. *Regulatory Toxicology and Pharmacology*, *39*(2), 87-98.; FU, T. J. (2002). Digestion stability as a criterion for protein allergenicity assessment. *Annals of the New York Academy of Sciences*, *964*(1), 99-110, all of which are incorporated herein by reference as if fully set forth).

[00114] ***Example 4. Creating an engineered phytase using coiled coil domains***

[00115] The following molecules were design based on the engineered phytases from Example 3. These molecules contain linkers but the trans-splicing C- and N-inteins are substituted with N- and C-terminal coils, respectively. The four prototype designs differ in the linker length and composition.

[00116] Nucleotide and amino acid sequences of the four prototype coiled coil stabilized phytase are below. Coil sequences at the N-and C-terminus are capitalized, linker sequences are lower case italics, phytase sequences are lower case.

The nucleotide sequence encoding Phy02-33:cc17 (cc17-N: L33-1-Phy02-L33-2 : cc17-C) [AA_SEQ ID NO: 103] is as follows:

ATGAGGGCCAAGCAGCTGGAGGACAAGATTGAGGAGCTGCTGAGCAAGATCTACCACCTGGAGAAC
GAGATAGCCCGCCTGAAGAAGCTGATTGGCGAGCGC*agcgggggtggcagtggaggcggttcgacc*
*ccgcagtccgcatttgcc*gcccaatcggaaccggaactgaaactggaaagtgtggttattgtgtct
cgtcatggcgttcgcgctccgaccaaatttacgcagctgatgcaagatgtcaccccggacgccttc
tatacgtggccggtgaagctgggtgaactgaccccgcgtggcggtgaactgatcgcctatctgggt
cactactggcgtcagcgcctggtggcagatggtctgctgccgaaaaagggctgcccgcagagcggt
caagttgcaattatcgctgatgtcgacgaacgtacccgcaaaacgggtgaagcatttgcggccggt
ctggcaccggattgcgccattaccgttcatacgcaggcagataccagctctccggacccgctgttc
aacccgctgaaaaccggcgtctgtcagctggatgtcgcgcaagtgacggacgccattctggaacgt
gcaggcggttccatcgctgattttaccggtcactaccagacggcattccgtgaactggaacgcgtt
ctgaactttccgcagtcaaatctggcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgacc
caagccctgccgagtgaactgaaagtctccgccgacaatgtgtcactgaccggcgcatggtcactg
gcttcgatgctgacggaaatttttctgctgcagcaagcacagggtatgccggaaccgggttggggt
cgtatcaccgattcgcatcagtggaacacgctgctgagcctgcacaatgcgcagttcgacctgctg
caacgtaccccggaagtggcacgttcgcgcgccacgccgctgctggatctgattaaaaccgctctg
acgccgcatccgccgcagaagcaagcgtatggcgtgaccctgccgacgagcgttctgtttatcgcg
ggtcacgacaccaacctggcaaatctgggcggtgctctggaactgcagtggaccctgccgggtcaa
ccggataacacgccgccgggcggtgaactggttttcgaacgttggcgtcgcctgagcgacaattct
cagtggatccaagttagcctggtctttcagaccctgcagcaaatgcgcgataaaaccccgctgttc
ctgaacacgccgccgggcgaagtgaagctgaccctggcgggttgcgaagaacgtaacgcccagggc
atgtgttctctggcaggttttacccagattgttaatgaagcacgcatcccggcttgtagtctg*ggt*
*ggcgggagcggtggagggagtgggggcggt*CAGCTGGAGGACAAGATTGAGGAGCTGCTGAGCAAG
ATCTACCACCTGGAGAACGAGATAGCGAGGCTGAAGAAGCTGATTGGCTAA (SEQ ID NO: 102)

The nucleotide sequence encoding Phy02-38: cc17 (cc17-N: L38-1-Phy02-L38-2 : cc17-C) [AA_SEQ ID NO: 105] is as follows:

ATGAGGGCCAAGCAGCTGGAGGACAAGATTGAGGAGCTGCTGAGCAAGATCTACCACCTGGAGAAC
GAGATAGCCCGCCTGAAGAAGCTGATTGGCGAGCGC*agcggtggctcgtcagggagtacgacaacc*
*acgcgtatcaccccgcaatctgcgttcgct*gcccaatcggaaccggaactgaaactggaaagtgtg
gttattgtgtctcgtcatggcgttcgcgctccgaccaaatttacgcagctgatgcaagatgtcacc
ccggacgccttctatacgtggccggtgaagctgggtgaactgaccccgcgtggcggtgaactgatc
gcctatctgggtcactactggcgtcagcgcctggtggcagatggtctgctgccgaaaaagggctgc
ccgcagagcggtcaagttgcaattatcgctgatgtcgacgaacgtacccgcaaaacgggtgaagca
tttgcggccggtctggcaccggattgcgccattaccgttcatacgcaggcagataccagctctccg

gacccgctgttcaacccgctgaaaaccggcgtctgtcagctggatgtcgcgcaagtgacggacgcc
attctggaacgtgcaggcggttccatcgctgattttaccggtcactaccagacggcattccgtgaa
ctggaacgcgttctgaactttccgcagtcaaatctggcgctgaaacgcgaaaagcaggatgaaagt
gcgtccctgacccaagccctgccgagtgaactgaaagtctccgccgacaatgtgtcactgaccggc
gcatggtcactggcttcgatgctgacggaaatttttctgctgcagcaagcacagggtatgccggaa
ccgggttggggtcgtatcaccgattcgcatcagtggaacacgctgctgagcctgcacaatgcgcag
ttcgacctgctgcaacgtaccccggaagtggcacgttcgcgcgccacgccgctgctggatctgatt
aaaaccgctctgacgccgcatccgccgcagaagcaagcgtatggcgtgaccctgccgacgagcgtt
ctgtttatcgcgggtcacgacaccaacctggcaaatctgggcggtgctctggaactgcagtggacc
ctgccgggtcaaccggataacacgccgccgggcggtgaactggttttcgaacgttggcgtcgcctg
agcgacaattctcagtggatccaagttagcctggtctttcagaccctgcagcaaatgcgcgataaa
accccgctgttcctgaacacgccgccgggcgaagtgaagctgaccctggcgggttgcgaagaacgt
aacgcccagggcatgtgttctctggcaggttttacccagattgttaatgaagcacgcatcccggct
tgtagtctg*caaaacacgtttagccaggggagtagctcgggatcc*CAGCTGGAGGACAAGATTGAG
GAGCTGCTGAGCAAGATCTACCACCTGGAGAACGAGATAGCGAGGCTGAAGAAGCTGATTGGCTAA
(SEQ ID NO: 104)

The nucleotide sequence encoding Phy02-46: cc17 (cc17-N: L46-1 -Phy02- L46-2 : cc17-C) [AA_SEQ ID NO: 107] is as follows:

ATGAGGGCCAAGCAGCTGGAGGACAAGATTGAGGAGCTGCTGAGCAAGATCTACCACCTGGAGAAC
GAGATAGCCCGCCTGAAGAAGCTGATTGGCGAGCGC*agcgcctttgca*gcccaatcggaaccggaa
ctgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgaccaaatttacgcag
ctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctgggtgaactgaccccg
cgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggtggcagatggtctg
ctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtcgacgaacgtacc
cgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattaccgttcatacgcag
gcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctgtcagctggatgtc
gcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattttaccggtcactac
cagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatctggcgctgaaacgc
gaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagtctccgccgac
aatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttttctgctgcagcaa
gcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtggaacacgctgctg
agcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacgttcgcgcgccacg
ccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagcaagcgtatggcgtg
accctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaatctgggcggtgct
ctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggtgaactggttttc
gaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggtctttcagaccctg
cagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagtgaagctgaccctg
gcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttacccagattgttaat
gaagcacgcatcccggcttgtagtctg*ggtgcagctccagcggccgcaccggctaaacaggaagcg*
gcagctccggctcctgcagcgaaggcggaagcaccggccgcagctcctgcggcaaaagcgaccccg
*cag*CAGCTGGAGGACAAGATTGAGGAGCTGCTGAGCAAGATCTACCACCTGGAGAACGAGATAGCG
AGGCTGAAGAAGCTGATTGGCTAA (SEQ ID NO: 106)

The nucleotide sequence encoding Phy02-55: cc17 (cc17-N: L55 -1-Phy02- L55-2 :cc17-C)[AA_SEQ ID NO: 109] is as follows:

ATGAGGGCCAAGCAGCTGGAGGACAAGATTGAGGAGCTGCTGAGCAAGATCTACCACCTGGAGAAC
GAGATAGCCCGCCTGAAGAAGCTGATTGGCGAGCGC*agcgcagccgaagcccgctgcgaaggaggca*
gctgcgaaagaagcgggctgcaaaagaagcggcagctaaggctttgaataccccgcaatcggctttc
*gc*tgccaatcggaaccggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgc
gctccgaccaaatttacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtg

aagctgggtgaactgaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcag
cgcctggtggcagatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatc
gctgatgtcgacgaacgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgc
gccattaccgttcatacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaacc
ggcgtctgtcagctggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatc
gctgattttaccggtcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcag
tcaaatctggcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagt
gaactgaaagtctccgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacg
gaaatttttctgctgcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcg
catcagtggaacacgctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaa
gtggcacgttcgcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccg
cagaagcaagcgtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaac
ctggcaaatctgggcggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccg
ccgggcggtgaactggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagtt
agcctggtctttcagaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccg
ggcgaagtgaagctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggca
ggttttacccagattgttaatgaagcacgcatcccggcttgtagtct*ggggggcgcagaagcagct*
*gccaaagaggcggccgcaaaggtcaatctg*CAGCTGGAGGACAAGATTGAGGAGCTGCTGAGCAAG
ATCTACCACCTGGAGAACGAGATAGCGAGGCTGAAGAAGCTGATTGGCTAA (SEQ ID NO:
108)

Heat unstable coiled-coil modified phytase (controls; cc30 with the four prototype linkers).

The nucleotide sequence encoding Phy02-33:cc30 (cc30-N: L33 -1-Phy02- L33-2 : cc30-C) [AA_SEQ ID NO: 111] is as follows:

ATGAGGGCCAAGCAGCTGGAGGACAAGGTCGAGGAGCTGCTGAGCAAGAACTACCACCTGGAGAAC
GAGGTCGCCCGCCTGAAGAAGCTGGTGGGCACCCGC*agcgggggtggcagtggaggcggttcgacc*
*ccgcagtccgcatttgcc*gcccaatcggaaccggaactgaaactggaaagtgtggttattgtgtct
cgtcatggcgttcgcgctccgaccaaatttacgcagctgatgcaagatgtcaccccggacgccttc
tatacgtggccggtgaagctgggtgaactgaccccgcgtggcggtgaactgatcgcctatctgggt
cactactggcgtcagcgcctggtggcagatggtctgctgccgaaaaagggctgcccgcagagcggt
caagttgcaattatcgctgatgtcgacgaacgtacccgcaaaacgggtgaagcatttgcggccggt
ctggcaccggattgcgccattaccgttcatacgcaggcagataccagctctccggacccgctgttc
aacccgctgaaaaccggcgtctgtcagctggatgtcgcgcaagtgacggacgccattctggaacgt
gcaggcggttccatcgctgattttaccggtcactaccagacggcattccgtgaactggaacgcgtt
ctgaactttccgcagtcaaatctggcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgacc
caagccctgccgagtgaactgaaagtctccgccgacaatgtgtcactgaccggcgcatggtcactg
gcttcgatgctgacggaaatttttctgctgcagcaagcacagggtatgccggaaccgggttggggt
cgtatcaccgattcgcatcagtggaacacgctgctgagcctgcacaatgcgcagttcgacctgctg
caacgtaccccggaagtggcacgttcgcgcgccacgccgctgctggatctgattaaaaccgctctg
acgccgcatccgccgcagaagcaagcgtatggcgtgaccctgccgacgagcgttctgtttatcgcg
ggtcacgacaccaacctggcaaatctgggcggtgctctggaactgcagtggaccctgccgggtcaa
ccggataacacgccgccgggcggtgaactggttttcgaacgttggcgtcgcctgagcgacaattct
cagtggatccaagttagcctggtctttcagaccctgcagcaaatgcgcgataaaaccccgctgttc
ctgaacacgccgccgggcgaagtgaagctgaccctggcgggttgcgaagaacgtaacgcccagggc
atgtgttctctggcaggttttacccagattgttaatgaagcacgcatcccggcttgtagtct*gggt*
*ggcgggagcggtggagggagtgggggcggt*CAATTGGAAGATAAAGTGGAAGAGCTCCTGTCCAAA
AATTATCATCTGGAAAATGAGGTGGCCCGCTTGAAGAAACTCGTGGGATAA (SEQ ID NO:
110)

The nucleotide sequence encoding Phy02-38: cc30 (cc30-N: L38 -1 -Phy02- L38-2 :cc30-C) [AA_SEQ ID NO: 113] is as follows:

ATGAGGGCCAAGCAGCTGGAGGACAAGGTCGAGGAGCTGCTGAGCAAGAACTACCACCTGGAGAAC
GAGGTCGCCCGCCTGAAGAAGCTGGTGGGCACCCGC*agcggtggctcgtcagggagtacgacaacc*
*acgcgtatcaccccgcaatctgcgttcgct*gcccaatcggaaccggaactgaaactggaaagtgtg
gttattgtgtctcgtcatggcgttcgcgctccgaccaaatttacgcagctgatgcaagatgtcacc
ccggacgccttctatacgtggccggtgaagctgggtgaactgaccccgcgtggcggtgaactgatc
gcctatctgggtcactactggcgtcagcgcctggtggcagatggtctgctgccgaaaaagggctgc
ccgcagagcggtcaagttgcaattatcgctgatgtcgacgaacgtacccgcaaaacgggtgaagca
tttgcggccggtctggcaccggattgcgccattaccgttcatacgcaggcagataccagctctccg
gacccgctgttcaacccgctgaaaaccggcgtctgtcagctggatgtcgcgcaagtgacggacgcc
attctggaacgtgcaggcggttccatcgctgattttaccggtcactaccagacggcattccgtgaa
ctggaacgcgttctgaactttccgcagtcaaatctggcgctgaaacgcgaaaagcaggatgaaagt
gcgtccctgacccaagccctgccgagtgaactgaaagtctccgccgacaatgtgtcactgaccggc
gcatggtcactggcttcgatgctgacggaaatttttctgctgcagcaagcacagggtatgccggaa
ccgggttggggtcgtatcaccgattcgcatcagtggaacacgctgctgagcctgcacaatgcgcag
ttcgacctgctgcaacgtaccccggaagtggcacgttcgcgcgccacgccgctgctggatctgatt
aaaaccgctctgacgccgcatccgccgcagaagcaagcgtatggcgtgaccctgccgacgagcgtt
ctgtttatcgcgggtcacgacaccaacctggcaaatctgggcggtgctctggaactgcagtggacc
ctgccgggtcaaccggataacacgccgccgggcggtgaactggttttcgaacgttggcgtcgcctg
agcgacaattctcagtggatccaagttagcctggtctttcagaccctgcagcaaatgcgcgataaa
accccgctgttcctgaacacgccgccgggcgaagtgaagctgaccctggcgggttgcgaagaacgt
aacgcccagggcatgtgttctctggcaggttttacccagattgttaatgaagcacgcatcccggct
tgtagtctg*caaaacacgtttagccaggggagtagctcgggatcc*CAATTGGAAGATAAAGTGGAA
GAGCTCCTGTCCAAAAATTATCATCTGGAAAATGAGGTGGCCCGCTTGAAGAAACTCGTGGGATAA
(SEQ ID NO: 112)

The nucleotide sequence encoding Phy02-46: cc30 (cc30-N: L46-1 -Phy02- L46-2 :cc30-C) [AA_SEQ ID NO: 115] is as follows:

ATGAGGGCCAAGCAGCTGGAGGACAAGGTCGAGGAGCTGCTGAGCAAGAACTACCACCTGGAGAAC
GAGGTCGCCCGCCTGAAGAAGCTGGTGGGCACCCGC*agcgcctttgca*gcccaatcggaaccggaa
ctgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgaccaaatttacgcag
ctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctgggtgaactgaccccg
cgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggtggcagatggtctg
ctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtcgacgaacgtacc
cgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattaccgttcatacgcag
gcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctgtcagctggatgtc
gcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgattttaccggtcactac
cagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatctggcgctgaaacgc
gaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaagtctccgccgac
aatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttttctgctgcagcaa
gcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtggaacacgctgctg
agcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacgttcgcgcgccacg
ccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagcaagcgtatggcgtg
accctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaatctgggcggtgct
ctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggtgaactggttttc
gaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggtctttcagaccctg
cagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagtgaagctgaccctg
gcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttacccagattgttaat
gaagcacgcatcccggcttgtagtctg*ggtgcagctccagcggccgcaccggctaaacaggaagcg*
gcagctccggctcctgcagcgaaggcggaagcaccggccgcagctcctgcggcaaaagcgaccccg
*cag*CAATTGGAAGATAAAGTGGAAGAGCTCCTGTCCAAAAATTATCATCTGGAAAATGAGGTGGCC
CGCTTGAAGAAACTCGTGGGATAA (SEQ ID NO: 114)

The nucleotide sequence encoding Phy02-55: cc30 (cc30-N: L55-1 -Phy02- L55-2 :cc30-C) [AA_SEQ ID NO: 117] is as follows:

ATGAGGGCCAAGCAGCTGGAGGACAAGGTCGAGGAGCTGCTGAGCAAGAACTACCACCTGGAGAAC
GAGGTCGCCCGCCTGAAGAAGCTGGTGGGCACCCGC*agcgcagccgaagccgctgcgaaggaggca*
gctgcgaaagaagcggctgcaaaagaagcggcagctaaggctttgaatacccccgcaatcggctttc
*gctgc*ccaatcggaaccggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgc
gctccgaccaaatttacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtg
aagctgggtgaactgaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcag
cgcctggtggcagatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatc
gctgatgtcgacgaacgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgc
gccattaccgttcatacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaacc
ggcgtctgtcagctggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatc
gctgattttaccggtcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcag
tcaaatctggcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagt
gaactgaaagtctccgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacg
gaaatttttctgctgcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcg
catcagtggaacacgctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaa
gtggcacgttcgcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccg
cagaagcaagcgtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaac
ctggcaaatctgggcggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccg
ccgggcggtgaactggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagtt
agcctggtctttcagaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccg
ggcgaagtgaagctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggca
ggttttacccagattgttaatgaagcacgcatcccggcttgtagtctg*ggggggcgcagaagcagct*
*gccaaagaggcggccgcaaaggtcaatctg*CAATTGGAAGATAAAGTGGAAGAGCTCCTGTCCAAA
AATTATCATCTGGAAAATGAGGTGGCCCGCTTGAAGAAACTCGTGGGATAA (SEQ ID NO: 116)

[00117] ***Example 5. Creating an engineered phytase using a tag-catcher domain set***

[00118] Using the methods described in Example 1, engineered phytases can be constructed using tag- and catcher-domains as described in FIGS. 4 – 7. FIG. 4 illustrates an engineered phytase with a tag- and catcher-domain attached to the amino- and carboxy-termini, respectively, of the phytase coding sequence (A) and binding of the tag- and catcher- domains to cyclize the phytase using non-covalent binding (B), and the form of the cyclized phytase that results following reaction of the tag-catcher domains to form a covalent bond (C). FIG. 5 illustrates an engineered phytase with a tag- and catcher-domain attached to the carboxy- and amino-termini, respectively, of the phytase coding sequence (A) and binding of the tag- and catcher-domains to cyclize the phytase using non-covalent binding (B), and the form of the cyclized phytase that results following reaction of the tag-catcher domains to

form a covalent bond (C). FIG. 6 illustrates an engineered phytase with a tag- and catcher-domains attached to linkers that connect to the amino- and carboxy-termini, respectively, of the phytase coding sequence (A), and binding of the tag- and catcher domains to cyclize the phytase using non-covalent binding (B), and the form of the cyclized phytase that results following reaction of the tag-catcher domains to form a covalent bond (C). FIG. 7 illustrates an engineered phytase with a tag- and catcher- domains attached to linkers that connect to the carboxy- and amino-termini, respectively, of the phytase coding sequence (A), and binding of the tag-and catcher-domains to cyclize the phytase using non-covalent binding (B), and the form of the cyclized phytase that results following reaction of the tag-catcher domains to form a covalent bond (C).

[00119] The tag- and catcher-domains can be directly connected to the phytase's termini, or connected to the termini using linkers. Unlike split inteins, which generally have a preferred termini to which each part of the intein attaches, tag- and catcher-domains can be used at either termini. For example, one engineered phytase may have the tag-domain connected to the target phytase's amino terminus without a linker (FIG. 4), or with a linker (FIG. 6), and have the catcher-domain connected to the target phytase's carboxy terminus without a linker (FIG. 4), or with a linker (FIG. 6). Similarly, one engineered phytase may have the tag-domain connected to the target phytase's carboxy terminus without a linker (FIG. 5), or with a linker (FIG. 7), and have the catcher-domain connected to the target phytase's amino terminus without a linker (FIG. 5), or with a linker (FIG. 7). The tag- and catcher-domains are capable of binding the termini of the target phytase in both configurations and forming a cyclic phytase through formation of a covalent bond. The following sequences illustrate how an engineered Phy02 phytase is constructed:

Tag-Domain:Tlinker1:Phy02:Clinker1:Catcher (linker is in bold and underlined):

atggcccacatcgtgatggtggacgcctacaagccgacgaag**ggttcaggggttccggt**gcccaa
tcggaaccggaactgaaactggaaagtgtggttattgtgtctcgtcatggcgttcgcgctccgacc
aaatttacgcagctgatgcaagatgtcaccccggacgccttctatacgtggccggtgaagctgggt

gaactgaccccgcgtggcggtgaactgatcgcctatctgggtcactactggcgtcagcgcctggtg
gcagatggtctgctgccgaaaaagggctgcccgcagagcggtcaagttgcaattatcgctgatgtc
gacgaacgtacccgcaaaacgggtgaagcatttgcggccggtctggcaccggattgcgccattacc
gttcatacgcaggcagataccagctctccggacccgctgttcaacccgctgaaaaccggcgtctgt
cagctggatgtcgcgcaagtgacggacgccattctggaacgtgcaggcggttccatcgctgatttt
accggtcactaccagacggcattccgtgaactggaacgcgttctgaactttccgcagtcaaatctg
gcgctgaaacgcgaaaagcaggatgaaagtgcgtccctgacccaagccctgccgagtgaactgaaa
gtctccgccgacaatgtgtcactgaccggcgcatggtcactggcttcgatgctgacggaaatttt
ctgctgcagcaagcacagggtatgccggaaccgggttggggtcgtatcaccgattcgcatcagtgg
aacacgctgctgagcctgcacaatgcgcagttcgacctgctgcaacgtaccccggaagtggcacgt
tcgcgcgccacgccgctgctggatctgattaaaaccgctctgacgccgcatccgccgcagaagcaa
gcgtatggcgtgaccctgccgacgagcgttctgtttatcgcgggtcacgacaccaacctggcaaat
ctgggcggtgctctggaactgcagtggaccctgccgggtcaaccggataacacgccgccgggcggt
gaactggttttcgaacgttggcgtcgcctgagcgacaattctcagtggatccaagttagcctggtc
tttcagaccctgcagcaaatgcgcgataaaaccccgctgttcctgaacacgccgccgggcgaagtg
aagctgaccctggcgggttgcgaagaacgtaacgcccagggcatgtgttctctggcaggttttacc
cagattgttaatgaagcacgcatcccggcttgtagtctg**gggagtggtggcagcgga**ggcgctatg
gttgataccttatcaggtttatcaagtgagcaaggtcagtccggtgatatgacaattgaagaagat
agtgctacccatattaaattctcaaaacgtgatgaggacggcaaagagttagctggtgcaactatg
gagttgcgtgattcatctggtaaaactattagtacatggatttcagatggacaagtgaaagatttc
tacctgtatccaggaaaatatacatttgtcgaaaccgcagcaccagacggttatgaggtagcaact
gctattacctttacagttaatgagcaaggtcaggttactgtaaatggcaaagcaactaaaggtgac
gctcatatt (SEQ ID NO: 118)

MAHIVMVDAYKPTK**GSGGSG**AQSEPELKLESVVIVSRHGVRAPTKFTQLMQDVTPDAFYTWPVKLG
ELTPRGGELIAYLGHYWRQRLVADGLLPKKGCPQSGQVAIIADVDERTRKTGEAFAAGLAPDCAIT
VHTQADTSSPDPLFNPLKTGVCQLDVAQVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNL
ALKREKQDESASLTQALPSELKVSADNVSLTGAWSLASMLTEIFLLQQAQGMPEPGWGRITDSHQW
NTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSVLFIAGHDTNLAN
LGGALELQWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSLVFQTLQQMRDKTPLFLNTPPGEV
KLTLAGCEERNAQGMCSLAGFTQIVNEARIPACSL**GSGGSG**GAMVDTLSGLSSEQGQSGDMTIEED
SATHIKFSKRDEDGKELAGATMELRDSSGKTISTWISDGQVKDFYLYPGKYTFVETAAPDGYEVAT
AITFTVNEQGQVTVNGKATKGDAHI(SEQ ID NO: 119)

[00120] As with the other engineered molecules described herein, optimization of the molecules and variants of the molecules and processes described herein can be used. Many different methods of optimization and mutagenesis may be employed, as described in Examples 2 and 3, and elsewhere in this specification.

[00121] One skilled in the art would also recognize that any of the target phytases could be used in any of the examples described above with different molecular structures and binding domains. For example, the tag- and catcher-domains can be attached to the CQBscks phytase, with or without linkers, to create a version of the phytase with improved thermal stability. Likewise any other structures, including inteins and coiled coils, could be used with

CQBscks or any other target phytase to improve the target phytase's thermal stability.

[00122] ***Example 6. Assaying for phytase activity***

[00123] Phytase assays are necessary for engineering phytases for improved thermal stability as described herein. See Engelen et al., 2001, Determination of phytase activity in feed by a colorimetric enzymatic method: collaborative interlaboratory study. *Journal of AOAC International, 84*(3), 629-633; and US 7,629,139, issued December 8, 2009, all of which are incorporated herein by reference as if fully set forth. These assays often rely on comparing the amount of phosphate released from sodium phytate over time with a phosphate standard curve and adjusting for background phosphate levels and enzyme levels. Measurements are commonly reported in phytase units (FTUs), which are defined as a mass of phosphate (commonly a micromole of inorganic phosphate) released per unit time (commonly one minute) under a given set of assay conditions (commonly 37°C, pH 5.5 under an excess of sodium phosphate, but other conditions are also reported and used in research and industry). These methods can be used with microbially produced phytases and engineered phytases, as well as those produced from other host expression systems, including plant expression systems.

[00124] To conduct the assay, enzyme extracts must be prepared from the expression host. Many different protein preparation methods exist and are known in the art. In each, case cells are disrupted using a method such as mechanical disruption (*e.g.*, using a French press), liquid homogenization, sonication, repetitive freezing and thawing cycles, a detergent and chemical lysis, or manual grinding. Following lysis of the cells, the lysate may be used directly, or may be further fractionated to enrich for the desired protein, or even purified to a nearly pure protein substance (see "Current Protocols in Molecular Biology," 10.0.1-10.0.23, April, 2010, John Wiley & Sons, Inc., which is incorporated herein by reference as if fully set forth). Cellular lysis and protein extraction can even be automated to a large extent, facilitating the processing of many samples simultaneously. For protein extraction from

plants, or seeds, generally larger tissue samples must first be disrupted, often through milling or grinding, and sometimes including freezing of the sample or repetitive freezing and thawing cycles, and then the protein can be extracted in a method similar to those described and referred to above.

[00125] Phytase activity was measured starting with up to 1 mL of cellular lysate, protein extracts are diluted 100-fold in assay buffer (250 mM sodium acetate, pH 5.5, 1mM calcium chloride, 0.01% Tween 20). Seventy-five (75) microliters of the diluted extracts or 75 µl of buffer-only controls were dispensed into individual wells of a round-bottom 96-well plate. One-hundred fifty (150) microliters of freshly-prepared phytic acid (9.1 mM dodecasodium salt from Biosynth International, Staad, Switzerland, prepared in assay buffer) were added to each well. Plates were sealed and incubated for 60 minutes at 37°C. One-hundred fifty (150) microliters of stop solution (20mM ammonium molybdate, 5mM ammonium vanadate, 4% nitric acid) was added to each well, mixed thoroughly via pipetting, and allowed to incubate at room temperature for 10 minutes. Plates were centrifuged at 3000×G for 10 minutes, and 100 µL of the clarified supernatants were transferred to the wells of a flat-bottom 96-well plate. Absorbance at 415nm from each sample was compared to that of negative controls (buffer-only, no enzyme) and potassium phosphate standards. The standard curve was prepared by mixing 50 µl of potassium phosphate standards (0-1.44mM, prepared in assay buffer) with 100 µL of freshly-prepared phytic acid, followed by 100 µL of stop solution.

[00126] ***Example 7. Testing the thermal stability of cyclized phytases***

[00127] In order to determine the thermal stability of an engineered phytase, the activity of the engineered phytase must be measured following different temperature treatments. Measurement of phytase activity can be conducted using a phytase assay as known in the art. Phytase assays that may be used to measure phytase activity are also described in Example 6 herein. While many different procedures could be used to investigate the thermal stability of an engineered phytase, one method was used herein as an example,

recognizing that other procedures, experimental designs, and assay methods may be used in this analysis. Furthermore, the exact experimental conditions may vary dramatically depending on the breadth and depth of the analysis. Preferred procedures use a microbial expression system to rapidly produce the engineered phytase to be tested, and other control molecules that may be included in the evaluation, regardless of the final production system used to produce the engineered phytase at a greater scale. Microbial expression systems that may be used in this evaluation include *E. coli, Saccharomyces cerevisiae, Pichia pastoris, Bacillus, Aspergillus niger,* and *Trichoderma reesei* expression systems, although other systems may also be used. Following evaluation from a microbial expression system, it would be beneficial to repeat the evaluation using materials produced by the final production system whenever those materials are available.

[00128] To evaluate the thermal stability of an engineered phytase, it is desirable to test the engineered phytase and corresponding target phytase (without any molecular structures attached to the target phytase), at different temperatures, and for different lengths of time, under desirable conditions. Ideally, the experimental design for these tests would use a known molar quantity of engineered phytase and target phytase, incubating the molecules separately in a desired buffer for a length of time ranging from zero seconds (an untreated negative control) up to 30 minutes or more. Measurements can be taken at any desired time interval, but shorter time intervals will be necessary if activity values above the background of the assay are to be measured at higher temperatures. A constant temperature and pH of the buffer are used in each incubation. Temperatures in the range of 60°C up to 90°C or more would be of interest in determining the thermal stability of the engineered phytase relative to its corresponding target phytase. Likewise, pH values in the range from 2 up to 7 or more would be relevant for determining the thermal stability of the phytases at physiologically relevant levels of acidity. Following incubation, a sample of the incubation mixture is taken and the enzymatic activity is measured at a standard temperature (preferably between 25°C and 37°C) and pH (preferably between 5 and 7). The measured

activities of the engineered phytase can then be compared against the target phytase and the improvement in thermal stability can be determined. Target phytases Phy02, Nov9X, and CQBscks were incubated individually along with the engineered Phy02 phytases described herein. Incubations were conducted at pH 5.5 in a water bath set at 65°C, 70°C, 75°C, 80°C, 85°C, and 90°C. For each incubation, samples were removed at 15 seconds, 40 seconds, 1 minute, 1.5 minutes, 2 minutes, 3 minutes, 5 minutes, 10 minutes, and 15 minutes. Prior to each incubation, a sample was taken to represent the zero time point, where no elevated temperature exposure occurred. The activity measured at the zero time point was within the experimental variation of the maximum activity observed in the experiment. From the zero time point and each incubation sample, the activity was measured in triplicate as described in Example 6, at 37°C and pH 5.5. The activity of the engineered Phy02 phytases were then compared with the activity of the target phytases Phy02, Nov9X, and CQBscks. Nov9X showed the lowest activity across the treatments, with Phy02 and CQBscks showing greater activity at the different treatments. Engineered Phy02 phytases were selected that had elevated activity relative to the target enzymes in the different treatments.

[00129] Often times, experimental conditions are less than ideal and variations on the procedures described in this example are used. It is desirable to make activity measurements in at least triplicate, to be able to determine the variation in the activity measurement under a given set of conditions, but in some case only duplicate or single measurements may be feasible. In many cases, it's not feasible to purify each engineered enzyme or target enzyme in order to use equimolar concentrations. Often times, this is also not necessary given that expression levels for the different phytase enzymes from a given expression system may be similar. In these cases, enzyme loading into the incubations may be based upon culture volume, lysate volume, amount of total protein, or a similar variable. It's also not necessary to use purified enzyme in these evaluations, as the relative change in thermal stability can be used to compare enzymes and evaluate improvements in thermal stability. To evaluate the relative changes in thermal stability, the

activity levels measured across time points at a given temperature are normalized to the zero time point by dividing the activity measured at all subsequent time points by the activity measured at the zero time point and multiplying by 100 percent. Thus, if for example an engineered Phy02 enzyme was measured to have 1000 FTU at the zero time point, and the following measurements were made at a given temperature (for example 90°C) 950 FTU at 15 seconds, 902 FTU at 40 seconds, 857 FTU at one minute, 797 FTU at 1.5 minutes, 741 FTU at two minutes, 669 FTU at three minutes, 545 FTU at five minutes, 400 FTU at 10 minutes, and 238 FTU at 15 minutes, then the percent activity measurements would be calculated to give 100% (0 s), 95% (15s), 90.2% (40s), 85.7% (1m), 79.7% (1.5m), 74.1% (2m), 66.9% (3m), 54.5% (5m), 40.0% (10m), and 23.8% (15m). If the corresponding values for the target enzyme were determined to be 100% (0 s), 85% (15s), 60.2% (40s), 25.7% (1m), 5.1% (1.5m), 1.3% (2m), 1.5% (3m), 0.9% (5m), 0.0% (10m), and 0.0% (15m), then it would be clear to one skilled in the art that the engineered Phy02 phytase had improved thermal stability relative to the target phytase. This procedure may be repeated at multiple temperatures and other pH values to define the differences in thermal stability between the engineered phytase and target phytase in greater detail and more precision. Using relative measurements and readily available automation, many engineered phytase variants can be readily screened and evaluated, and the most improved enzymes selected for commercial use.

[00130] Furthermore, other methods exist to determine thermal stability. Differential scanning calorimetry is a method known in the art, which can provide very accurate measurements of thermal stability.

[00131] ***Example 8. Thermal stability optimization of engineered phytases***

[00132] Any of the molecules or procedures described in the previous examples can be continued to develop further improvements in the engineered phytase's thermal stability or other properties. Properties of particular commercial and scientific interest include the specific activity of the engineered phytase,

expression level of the engineered phytase in a variety of heterologous expression systems (including microbial expression systems, plant expression systems, and mammalian expression systems), gastric and pepsin stability of the engineered phytase, and pepsin digestibility of the engineered phytase. Many methods exist for further optimizing the engineered phytase to have improved thermal stability or other properties. These methods include site directed mutagenesis, saturation mutagenesis, random mutagenesis, sequence shuffling, modeling, and others. In addition, these methods can easily be employed using automated screening systems, enabling the evaluation of millions of variants within reasonable time frames.

[00133] For optimization of engineered phytases whose coding sequences comprise an intein sequence, several methods can be particularly useful, including saturating mutagenesis and site directed mutagenesis. It is known in the art that mutations which occur near the intein-extein junction can have a significant impact on intein splicing, thus enabling the development of molecules that bind but don't splice, bind and create an isopeptide, bind and selectively cleave one portion of the split intein, or bind and fully splice to form a covalent bond at the insertion site (Xu, M. Q., & Perler, F. B. (1996). The mechanism of protein splicing and its modulation by mutation. *The EMBO journal, 15*(19), 5146, which is incorporated herein by reference as if fully set forth). Thus mutations at the -3 to -1 position in the target phytase at the intein junction, as well as mutations at the +1 to +3 positions relative to the intein insertion site commonly have a significant effect on the extent of the binding and splicing reactions, as well as the rate of reaction under different conditions. Mutations at these sites may improve the rate of splicing, thereby improving the rate of cyclization of the phytase and in some cases the observed thermal stability of the enzyme (as evaluated in Example 7). Because preferred insertion cassettes have been identified for many inteins, these cassettes may be successfully used in a target phytase backbone to improve intein splicing and therefore the thermal stability of the resulting engineered phytase or in linkers for the same purpose and effect. Similarly, other mutations in the protein coding sequence, including the molecular

structures, may be used to improve thermal stability. For insertion cassettes for inteins, see Apgar et al. 2012, which is incorporated herein by reference as if fully set forth.

[00134] Specific activity, heterologous expression levels, gastric stability, and pepsin digestion may also be improved by further mutagenesis studies on an engineered phytase constructed in this study. The procedures used to optimize these properties would be carried out in an analogous way to thermal stability optimization, but in each case a different property would be considered in the evaluations program.

[00135] ***Example 9. Descriptions of expression cassettes for engineered phytases***

[00136] *Cyclic phytase sequences and maps for plant expression.* Sequences containing different variants of cyclic phytases for plant expression have been assembled as expression cassettes with KpnI restriction site at 5' and EcoRI restriction site at 3' ends. All sequences for individual genetic elements were codon optimized for expression in maize. Two cassettes per each individual sequence were designed with one for cytoplasmic and the other for endoplasmic reticulum (ER) targeted protein expression. To generate final plant expression constructs, each expression cassette can be cloned into KpnI-EcoRI digested vector such as pAG4500. A representative map of resulting construct pAG4918 which contains expression cassette ZmZ27:Gp41-1C:Phy02opt:Gp41-1N:NosT (the Phy02opt cassette) cloned in this way is illustrated on FIG. 8. As shown in FIG. 8, the Phy02opt expression cassette including polynucleotides encoding the ZmZ27 promoter, Gp41-1C intein, Phy02opt phytase, Gp41-1N intein, and NosT terminator can be introduced into pAG4918 at the KpnI site (position 10227) and the EcoR1 site (position 283). pAG4918 also carries a plant selectable marker comprised of a *Zea mays* ubiquitin (ZmUbi1) promoter, a *Zea mays* ubiquitin (ZmUbi1) intron, a *Zea mays* (Zm) Kozak, the phosphomannose isomerase coding sequence, and NosT terminator a phosphomannose isomerase (PMI) gene, and the NosT terminator. Both the Phy02 opt and the plant selectable marker cassettes are

integrated into pAG4918 between the right border (RT) and the left border (LB). pAG4918 includes the spectinomycin adenylyltransferase gene (aadA), the streptothricin adenyltransferase gene, the cohesive site (*cos*) of bacteriophage λ and the Ori origin of replication. pAG4918 or similar vectors can be transferred from *E. coli* to *Agrobacterium tumefaciens* LBA4404 via conjugal transfer, during which the plasmid will integrate into pSB1 (a resident Ti plasmid) via homologous recombination. Co-culture of the resulting recombinant *Agrobacterium* strain with plant cells can result in the transfer of the pAG4918-derived DNA to the plant genome. Embodiments herein include a transformation vector having any one of engineered phytases.

[00137] Plant transformation vectors were assembled by inserting the expression cassettes or constructs described herein between the *Agrobacterium* T-DNA right border (RB) and left border (LB) sequences of pAG4500 or any suitable plasmid.

[00138] FIGS. 9A - 9C illustrate examples of expression cassettes for selected engineered phytases with split inteins attached to the ends of the phytase coding sequences. FIG. 9A illustrates the Phy02opt expression cassette the ZmZ27P : xGZein27ss : Gp41-1C : Phy02opt : Gp41-1N : DPNGSEKDEL : NosT including polynucleotides encoding the ZmZ27 promoter, GZein27ss signal sequence, Gp41-1C intein, Phy02opt phytase, Gp41-1N intein, DPNG linker, SEKDEL terminal extension sequence, and NosT terminator that can be introduced into pAG4918 at the KpnI site (position 10227) and the EcoR1 site (position 283). pAG4918 also carries a plant selectable marker comprised of a *Zea mays* ubiquitin (ZmUbi1) promoter, a *Zea mays* ubiquitin (ZmUbi1) intron, a *Zea mays* (Zm) Kozak, the phosphomannose isomerase (PMI) coding sequence, and NosT terminator. FIG. 9B illustrates the ZmZ27P:Ssp DnaE-C:Phy02opt:Ssp DnaE-N:NosT expression cassette. Referring to FIG. 9B, the expression cassette includes the ZmZ27 promoter, Ssp DnaE-C intein, Phy02opt phytase, Ssp DnaE-N intein, and NosT terminator. FIG. 9C illustrates the ZmZ27P:xGZein27ss:Ssp DnaE-C:Phy02opt:Ssp DnaE-N: DPNGSEKDEL: NosT expression cassette. Referring to FIG. 9C, the expression cassette includes the ZmZ27 promoter,

GZein27ss signal sequence, Ssp DnaE-C intein, Phy02opt phytase, Ssp DnaE-N intein, DPNG linker, SEKDEL terminal extension sequence, and NosT terminator.

[00139] FIGS. 10A - 10H are schematic diagrams illustrating expression cassettes for selected engineered phytases with split intein attached to linkers that connect to the ends of the phytase coding sequences.

[00140] FIG. 10A illustrates the ZmZ27P:Ssp DnaE-C:L33-1:Phy02opt:L33-2:Ssp DnaE-N:NosT expression cassette. Referring to FIG. 10A the expression cassette includes the ZmZ27 promoter, Ssp DnaE-C intein, L33-1 linker (L33-1), Phy02opt phytase, L33-2 linker (L33-2), Ssp DnaE-N intein, and NosT terminator. FIG. 10B illustrates the ZmZ27P:xGZein27ss: Ssp DnaE-C : L33-1 : Phy02opt : L33-2 : Ssp DnaE-N : DPNGSEKDEL : NosT expression cassette. Referring to FIG. 10B the expression cassette includes the ZmZ27 promoter, GZein27ss signal sequence, Ssp DnaE-C intein, L33-1 linker (L33-1), Phy02opt phytase, L33-2 linker (L33-2), Ssp DnaE-N intein, DPNG linker, SEKDEL terminal extension sequence, and NosT terminator. FIG. 10C illustrates the ZmZ27P:Ssp DnaE-C:L38-1:Phy02opt:L38-2:Ssp DnaE-N:NosT expression cassette. Referring to FIG. 10C the expression cassette includes the ZmZ27 promoter, Ssp DnaE-C intein, L38-1 linker (L38-1), Phy02opt phytase, L38-2 linker (L38-2), Ssp DnaE-N intein, and NosT terminator. FIG. 10D illustrates the ZmZ27P:xGZein27ss: Ssp DnaE-C:L38-1:Phy02opt:L38-2:Ssp DnaE-N:DPNGSEKDEL:NosT expression cassette. Referring to FIG. 10D the expression cassette includes the ZmZ27 promoter, GZein27ss signal sequence, Ssp DnaE-C intein, L38-1 linker (L38-1), Phy02opt phytase, L38-2 linker (L38-2), Ssp DnaE-N intein, DPNG linker, SEKDEL terminal extension sequence and NosT terminator. FIG. 10E illustrates the ZmZ27P:Ssp DnaE-C:L46-1:Phy02opt:L46-2:Ssp DnaE-N:NosT expression cassette. Referring to FIG. 10E the expression cassette includes the ZmZ27 promoter, Ssp DnaE-C intein, L46-1 linker (L46-1), Phy02opt phytase, L46-2 linker (L46-2), Ssp DnaE-N intein, and NosT terminator. FIG.10F illustrates the ZmZ27P :xGZein27ss : Ssp DnaE-C : L46-1 : Phy02opt : L46-2 : Ssp DnaE-N : DPNGSEKDEL : NosT expression cassette. Referring to FIG. 10F the

expression cassette includes the ZmZ27 promoter, GZein27ss signal sequence, Ssp DnaE-C intein, L46-1 linker (L46-1), Phy02opt phytase, L46-2 linker (L46-2), Ssp DnaE-N intein, DPNG linker, SEKDEL terminal extension sequence and NosT terminator. FIG. 10G illustrates the ZmZ27P:Ssp DnaE-C:L55-1:Phy02opt:L55-2:Ssp DnaE-N:NosT expression cassette. Referring to FIG. 10G the expression cassette includes the ZmZ27 promoter, GZein27ss signal sequence, Ssp DnaE-C intein, L55-1 linker (L55-1), Phy02opt phytase, L55-2 linker (L55-2), Ssp DnaE-N intein, DPNG linker, SEKDEL terminal extension sequence and NosT terminator. FIG. 10H illustrates the ZmZ27P : xGZein27ss : Ssp DnaE-C : L55-1 : Phy02opt : L55-2 : Ssp DnaE-N : DPNGSEKDEL : NosT expression cassette. Referring to FIG. 10H the expression cassette includes the ZmZ27 promoter, GZein27ss signal sequence, Ssp DnaE-C intein, L55-1 linker (L55-1), Phy02opt phytase, L55-2 linker (L55-2), Ssp DnaE-N intein, DPNG linker, SEKDEL terminal extension sequence and NosT terminator. Each one of the cassettes shown in FIGS. 10A - 10H has KpnI, EcoRI, and BamHI restriction sites, and can be cloned as a KpnI-EcoRI fragment into the T-DNA of the transformation vector.

[00141] The list of expression constructs is compiled in Table 6.

Table 6. Construct list

Vector	Expression cassette
pAG4918	ZmZ27P:Gp41-1C:Phy02opt:Gp41-1N:NosT
pAG4919	ZmZ27P:xGZein27ss:Gp41-1C:Phy02opt:Gp41-1N:DPNGSEKDEL:NosT
pAG4920	ZmZ27P:Ssp_DnaE–C:Phy02opt:Ssp_DnaE–N:NosT
pAG4921	ZmZ27P:xGZein27ss:Ssp_DnaE–C:Phy02opt:Ssp_DnaE–N:DPNGSEKDEL:NosT
pAG4922	ZmZ27P:Ssp_DnaE:L33-1:Phy02opt:L33-2:NosT
pAG4923	ZmZ27P:xGZein27ss:Ssp_DnaE:L33-1:Phy02opt:L33-2:DPNGSEKDEL:NosT
pAG4924	ZmZ27P:Ssp_DnaE:L38-1:Phy02opt:L38-2:NosT

pAG4925	ZmZ27P:xGZein27ss:Ssp_DnaE:L38-1:Phy02opt:L38-2:DPNGSEKDEL:NosT
pAG4926	ZmZ27P:Ssp_DnaE:L46-1:Phy02opt:L46-2:NosT
pAG4927	ZmZ27P:xGZein27ss:Ssp_DnaE:L46-1:Phy02opt:L46-2:DPNGSEKDEL:NosT
pAG4928	ZmZ27P:Ssp_DnaE:L55-1:Phy02opt:L55-2:NosT
pAG4929	ZmZ27P:xGZein27ss:Ssp_DnaE:L55-1:Phy02opt:L55-2:DPNGSEKDEL:NosT

[00142] Nucleotide sequences in vectors pAG4924, pAG4926, and pAG4928 are identical to those in pAG4922 with the exception of two linker sequences. Similarly, all nucleotide sequences in constructs pAG4925, pAG4927 and pAG4929 are the same as in pAG4923 except for two linker sequences. The linker sequences that are specified on provided maps of expression cassettes pAG4918-pAG4929 include L33-1, L33-2, L38-1, L38-2, L46-1, L46-2, L55-1 and L55-2 and are shown in Table 4.

[00143] Relevant sequences of plant expression cassettes for cyclic phytases

[00144] >***ZmZ27P***:Gp411C:Phy02opt:Gp411N:*NosT*

[00145] ZmZ27P is shown in bold upper case font and italicized, gp411 is underlined, NosT is italicized.

ggtacc***AAAGTAATCATATTATTTTATGTGTGAATCTTCTTTACTTTTTCATTTGATTATGATTAT***
GAAGGTATGACCTTCATAACCTTCGTCCGAAATCCATTATATCCAAAGGAAAATAATGCTTCGAAG
GACGAAGGATTTTGATATTTAACATTTTATGTTGCCTTGTTCTTAATTCATAGCATTTGAGAACAA
GTCCCCAACACCAATCTTTATCTTTACTATATTAAAGCACCAGTTCAACGATCGTCTCGTGTCAAT
ATTATTAAAAAACTCCTACATTTCTTTATAATCAACCCGCACTCTTATAATCTCTTCTCTTACTAC
TATAATAAGAGAGTTTATGTACAAAATAAGGTGAAATTATGTATAAGTGTTCTGGACCTTGGTTGT
TGGCTCATATTCACACAACCTAATCAATAGAAAACATATGTTTTATTAAAACAAAATTTATCATAT
ATATATATATATATATATATATATATATATATATATATATATATAATATAAACCGTAGCAATGCACAGGCAT
ATGACTAGTGGCAACTTAATACCATGTGTGTATTAAGATGAATAAGAGGTATCCAAATAAATAACT
TGTTCGCTTACGTCTGGATCGAAAGGGGTTGGAAACGATTAAATCTCTTCCTAGTCAAAATTAAAT
AGAAGGAGATTTAATCGATTTCTCCCAATCCCCTTCGATCCAGGTGCAACCGAATAAGTCCTTAAA
TGTTGAGGAACACGAAACAACCATGCATTGGCATGTAAAGCTCCAAGAATTCGTTGTATCCTTAAC
AACTCACAGAACATCAACCAAAATTGCACGTCAAGGGTATTGGGTAAGAAACAATCAAACAAATCC
TCTCTGTGTGCAAAGAAACACGGTGAGTCATGCCGAGATCATACTCATCTGATATACATGCTTACA
GCTCACAAGACATTACAAACAACTCATATTGCATTACAAAGATCGTTTCATGAAAAATAAAATAGG
CCGGAACAGGACAAAAATCCTTGACGTGTAAAGTAAATTTACAACAAAAAAAAAAGCCATATGTCAA
GCTAAATCTAATTCGTTTTACGTAGATCAACAACCTGTAGAAGGCAACAAAACTGAGCCACGCAGA
AGTACAGAATGATTCCAGATGAACCATCGACGTGCTACGTAAAGAGAGTGACGAGTCATATACATT
TGGCAAGAAACCATGAAGCTGCCTACAGCCGTCTCGGTGGCATAAGAACACAAGAAATTGTGTTAA
TTAATCAAAGCTATAAATAACGCTCGCATGCCTGTGCACTTCTCCATCACCACCACTGGGTCTTCA
GACCATTAGCTTTATCTACTCCAGAGCGCAGAAGAACCCGATCGACACCggatccacc<u>ATGATGCT</u>
<u>GAAGAAGATCCTGAAGATCGAGGAGCTGGACGAGAGGGAGCTGATCGACATCGAGGTGAGCGGCAA</u>

CCACCTGTTCTACGCCAACGACATCCTGACCCACAACAGCGCCCAGTCCGAGCCGGAGCTGAAGCT
GGAGTCCGTGGTGATCGTGTCGCGCCACGGGGTGCGCGCCCCGACCAAGTTCACGCAGCTCATGCA
GGACGTGACCCCGGACGCCTTCTACACCTGGCCGGTGAAGCTCGGCGAGCTGACCCCGCGCGGCGG
CGAGCTGATCGCCTACCTCGGCCACTACTGGCGCCAGCGCCTCGTGGCCGACGGCCTCCTCCCGAA
GAAGGGCTGCCCGCAGTCCGGCCAGGTGGCGATCATCGCCGACGTGGACGAGCGCACCCGCAAGAC
GGGCGAGGCCTTCGCCGCCGGCCTCGCCCCGGACTGCGCCATCACCGTGCACACCCAGGCCGACAC
CTCCTCCCCGGACCCGCTCTTCAACCCGCTCAAGACCGGCGTGTGCCAGCTCGACGTGGCCCAGGT
GACCGACGCCATCCTGGAGCGCGCCGGCGGCTCCATCGCCGACTTCACCGGCCACTACCAGACCGC
CTTCCGCGAGCTGGAGCGCGTGCTCAACTTCCCGCAGTCGAACCTCGCCCTCAAGCGCGAGAAGCA
GGACGAGTCCGCCTCCCTCACCCAGGCCCTCCCGTCCGAGCTGAAGGTGTCCGCCGACAACGTGTC
CCTCACCGGCGCCTGGTCCCTCGCCTCCATGCTCACCGAAATCTTCCTCCTCCAGCAGGCCCAGGG
CATGCCGGAGCCGGGCTGGGGCCGCATCACCGACTCCCACCAGTGGAACACCCTCCTCTCCCTCCA
CAACGCCCAGTTCGACCTCCTCCAGCGCACCCCGGAGGTGGCCCGCTCCCGCGCCACCCCGCTCCT
CGACCTCATCAAGACCGCCCTCACCCCGCACCCGCCGCAGAAGCAGGCCTACGGCGTGACCCTCCC
GACCTCGGTGCTCTTCATCGCCGGCCACGACACCAACCTCGCCAACCTCGGCGGCGCCCTGGAGCT
GCAGTGGACCCTCCCGGGCCAGCCGGACAACACCCCGCCGGGCGGCGAGCTGGTGTTCGAGCGCTG
GCGCCGCCTCTCCGACAACTCCCAGTGGATTCAGGTGTCCCTCGTGTTCCAGACCCTCCAGCAGAT
GCGCGACAAGACCCCGCTCTTCCTCAACACCCCGCCGGGCGAGGTGAAGCTCACCCTGGCCGGCTG
CGAGGAGCGCAACGCGCAGGGCATGTGCTCCCTCGCCGGCTTCACCCAGATCGTGAACGAGGCCCG
CATCCCGGCCTGCTCCCTCTGCCTGGACCTGAAGACCCAGGTGCAGACCCCGCAGGGCATGAAGGA
GATCAGCAACATCCAGGTGGGCGACCTGGTGCTGAGCAACACCGGCTACAACGAGGTGCTGAACGT
GTTCCCGAAGAGCAAGAAGAAGAGCTACAAGATCACCCTGGAGGACGGCAAGGAGATCATCTGCAG
CGAGGAGCACCTGTTCCCGACCCAGACCGGCGAGATGAACATCAGCGGCGGCCTGAAGGAGGGCAT
GTGCCTGTACGTGAAGGAGTGAacctaggtccccgaatttccccgatcgttcaaacatttggcaata
aagtttcttaagattgaatcctgttgccggtcttgcgatgattatcatataatttctgttgaatta
cgttaagcatgtaataattaacatgtaatgcatgacgttatttatgagatgggttttatgattag
agtcccgcaattatacatttaatacgcgatagaaaacaaaatatagcgcgcaaactaggataaatt
atcgcgcgcggtgtcatctatgttactagatcgggaattg (SEQ ID NO: 128)

[00146] >***ZmZ27P***:xGZein27ss:Gp411-C:Phy02opt:Gp411-N:*DPNG***SEKDEL**:*NosT*

[00147] ZmZ27P is shown in bold upper case font and italicized, gp411 is underlined, DPNG is in upper case and italicized, SEKDEL is in bold upper case, NosT is italicized.

ggtacc***AAAGTAATCATATTATTTTATGTGTGAATCTTCTTTACTTTTTCATTTGATTATGATTAT***
GAAGGTATGACCTTCATAACCTTCGTCCGAAATCCATTATATCCAAAGGAAAATAATGCTTCGAAG
GACGAAGGATTTTGATATTTAACATTTTATGTTGCCTTGTTCTTAATTCATAGCATTTGAGAACAA
GTCCCCAACACCAATCTTTATCTTTACTATATTAAAGCACCAGTTCAACGATCGTCTCGTGTCAAT
ATTATTAAAAAACTCCTACATTTCTTTATAATCAACCCGCACTCTTATAATCTCTTCTCTTACTAC
TATAATAAGAGAGTTTATGTACAAAATAAGGTGAAATTATGTATAAGTGTTCTGGACCTTGGTTGT
TGGCTCATATTCACACAACCTAATCAATAGAAAACATATGTTTTATTAAAACAAAATTTATCATAT
ATATATATATATATATATATATATATATATATATATATATATAATATAAACCGTAGCAATGCACAGGCAT
ATGACTAGTGGCAACTTAATACCATGTGTGTATTAAGATGAATAAGAGGTATCCAAATAAATAACT
TGTTCGCTTACGTCTGGATCGAAAGGGGTTGGAAACGATTAAATCTCTTCCTAGTCAAAATTAAAT
AGAAGGAGATTTAATCGATTTCTCCCAATCCCCTTCGATCCAGGTGCAACCGAATAAGTCCTTAAA
TGTTGAGGAACACGAAACAACCATGCATTGGCATGTAAAGCTCCAAGAATTCGTTGTATCCTTAAC
AACTCACAGAACATCAACCAAAATTGCACGTCAAGGGTATTGGGTAAGAAACAATCAAACAAATCC
TCTCTGTGTGCAAAGAAACACGGTGAGTCATGCCGAGATCATACTCATCTGATATACATGCTTACA
GCTCACAAGACATTACAAACAACTCATATTGCATTACAAAGATCGTTTCATGAAAAATAAAATAGG
CCGGAACAGGACAAAAATCCTTGACGTGTAAAGTAAATTTACAACAAAAAAAAAGCCATATGTCAA

GCTAAATCTAATTCGTTTTACGTAGATCAACAACCTGTAGAAGGCAACAAAACTGAGCCACGCAGA
AGTACAGAATGATTCCAGATGAACCATCGACGTGCTACGTAAAGAGAGTGACGAGTCATATACATT
TGGCAAGAAACCATGAAGCTGCCTACAGCCGTCTCGGTGGCATAAGAACACAAGAAATTGTGTTAA
TTAATCAAAGCTATAAATAACGCTCGCATGCCTGTGCACTTCTCCATCACCACCACTGGGTCTTCA
GACCATTAGCTTTATCTACTCCAGAGCGCAGAAGAACCCGATCGACACCggatccaccATGAGGGT
GTTGCTCGTTGCCCTCGCTCTCCTGGCTCTCGCTGCGAGCGCCACCAGCATGATGCTGAAGAAGAT
CCTGAAGATCGAGGAGCTGGACGAGAGGGAGCTGATCGACATCGAGGTGAGCGGCAACCACCTGTT
CTACGCCAACGACATCCTGACCCACAACAGCGCTGCGCAGTCCGAGCCGGAGCTGAAGCTGGAGTC
CGTGGTGATCGTGTCGCGCCACGGGGTGCGCGCCCCGACCAAGTTCACGCAGCTCATGCAGGACGT
GACCCCGGACGCCTTCTACACCTGGCCGGTGAAGCTCGGCGAGCTGACCCCGCGCGGCGGCGAGCT
GATCGCCTACCTCGGCCACTACTGGCGCCAGCGCCTCGTGGCCGACGGCCTCCTCCCGAAGAAGGG
CTGCCCGCAGTCCGGCCAGGTGGCGATCATCGCCGACGTGGACGAGCGCACCCGCAAGACGGGCGA
GGCCTTCGCCGCCGGCCTCGCCCCGGACTGCGCCATCACCGTGCACACCCAGGCCGACACCTCCTC
CCCGGACCCGCTCTTCAACCCGCTCAAGACCGGCGTGTGCCAGCTCGACGTGGCCCAGGTGACCGA
CGCCATCCTGGAGCGCGCCGGCGGCTCCATCGCCGACTTCACCGGCCACTACCAGACCGCCTTCCG
CGAGCTGGAGCGCGTGCTCAACTTCCCGCAGTCGAACCTCGCCCTCAAGCGCGAGAAGCAGGACGA
GTCCGCCTCCCTCACCCAGGCCCTCCCGTCCGAGCTGAAGGTGTCCGCCGACAACGTGTCCCTCAC
CGGCGCCTGGTCCCTCGCCTCCATGCTCACCGAAATCTTCCTCCTCCAGCAGGCCCAGGGCATGCC
GGAGCCGGGCTGGGGCCGCATCACCGACTCCCACCAGTGGAACACCCTCCTCTCCCTCCACAACGC
CCAGTTCGACCTCCTCCAGCGCACCCCGGAGGTGGCCCGCTCCCGCGCCACCCCGCTCCTCGACCT
CATCAAGACCGCCCTCACCCCGCACCCGCCGCAGAAGCAGGCCTACGGCGTGACCCTCCCGACCTC
GGTGCTCTTCATCGCCGGCCACGACACCAACCTCGCCAACCTCGGCGGCGCCCTGGAGCTGCAGTG
GACCCTCCCGGGCCAGCCGGACAACACCCCGCCGGGCGGCGAGCTGGTGTTCGAGCGCTGGCGCCG
CCTCTCCGACAACTCCCAGTGGATTCAGGTGTCCCTCGTGTTCCAGACCCTCCAGCAGATGCGCGA
CAAGACCCCGCTCTTCCTCAACACCCCGCCGGGCGAGGTGAAGCTCACCCTGGCCGGCTGCGAGGA
GCGCAACGCGCAGGGCATGTGCTCCCTCGCCGGCTTCACCCAGATCGTGAACGAGGCCCGCATCCC
GGCCTGCTCCCTCTGCCTGGACCTGAAGACCCAGGTGCAGACCCCGCAGGGCATGAAGGAGATCAG
CAACATCCAGGTGGGCGACCTGGTGCTGAGCAACACCGGCTACAACGAGGTGCTGAACGTGTTCCC
GAAGAGCAAGAAGAAGAGCTACAAGATCACCCTGGAGGACGGCAAGGAGATCATCTGCAGCGAGGA
GCACCTGTTCCCGACCCAGACCGGCGAGATGAACATCAGCGGCGGCCTGAAGGAGGGCATGTGCCT
GTACGTGAAGGAG*GACCCGAACGGC***TCCGAGAAGGACGAGCTG**TGAcctagg*tccccgaatttccc*
cgatcgttcaaacatttggcaataaagtttcttaagattgaatcctgttgccggtcttgcgatgat
tatcatataatttctgttgaattacgttaagcatgtaataattaacatgtaatgcatgacgttatt
tatgagatgggttttatgattagagtcccgcaattatacatttaatacgcgatagaaaacaaaat
atagcgcgcaaactaggataaattatcgcgcgcggtgtcatctatgttactagatcgggaattg
(SEQ ID NO: 129)

[00148] >***ZmZ27P***:Ssp_DnaE–C:Phy02opt:Ssp_DnaE–N:*NosT*

[00149] ZmZ27P is shown in bold upper case font and italicized, SSp_DnaE is underlined, NosT is italicized.

ggtacc***AAAGTAATCATATTATTTTATGTGTGAATCTTCTTTACTTTTTCATTTGATTATGATTAT***
GAAGGTATGACCTTCATAACCTTCGTCCGAAATCCATTATATCCAAAGGAAAATAATGCTTCGAAG
GACGAAGGATTTTGATATTTAACATTTTATGTTGCCTTGTTCTTAATTCATAGCATTTGAGAACAA
GTCCCCAACACCAATCTTTATCTTTACTATATTAAAGCACCAGTTCAACGATCGTCTCGTGTCAAT
ATTATTAAAAAACTCCTACATTTCTTTATAATCAACCCGCACTCTTATAATCTCTTCTCTTACTAC
TATAATAAGAGAGTTTATGTACAAAATAAGGTGAAATTATGTATAAGTGTTCTGGACCTTGGTTGT
TGGCTCATATTCACACAACCTAATCAATAGAAAACATATGTTTTATTAAAACAAAATTTATCATAT
ATATATATATATATATATATATATATATATATATATATATAATATAAACCGTAGCAATGCACAGGCAT
ATGACTAGTGGCAACTTAATACCATGTGTGTATTAAGATGAATAAGAGGTATCCAAATAAATAACT
TGTTCGCTTACGTCTGGATCGAAAGGGGTTGGAAACGATTAAATCTCTTCCTAGTCAAAATTAAAT
AGAAGGAGATTTAATCGATTTCTCCCAATCCCCTTCGATCCAGGTGCAACCGAATAAGTCCTTAAA
TGTTGAGGAACACGAAACAACCATGCATTGGCATGTAAAGCTCCAAGAATTCGTTGTATCCTTAAC

AACTCACAGAACATCAACCAAAATTGCACGTCAAGGGTATTGGGTAAGAAACAATCAAACAAATCC
TCTCTGTGTGCAAAGAAACACGGTGAGTCATGCCGAGATCATACTCATCTGATATACATGCTTACA
GCTCACAAGACATTACAAACAACTCATATTGCATTACAAAGATCGTTTCATGAAAAATAAAATAGG
CCGGAACAGGACAAAAATCCTTGACGTGTAAAGTAAATTTACAACAAAAAAAAAGCCATATGTCAA
GCTAAATCTAATTCGTTTTACGTAGATCAACAACCTGTAGAAGGCAACAAAACTGAGCCACGCAGA
AGTACAGAATGATTCCAGATGAACCATCGACGTGCTACGTAAAGAGAGTGACGAGTCATATACATT
TGGCAAGAAACCATGAAGCTGCCTACAGCCGTCTCGGTGGCATAAGAACACAAGAAATTGTGTTAA
TTAATCAAAGCTATAAATAACGCTCGCATGCCTGTGCACTTCTCCATCACCACCACTGGGTCTTCA
GACCATTAGCTTTATCTACTCCAGAGCGCAGAAGAACCCGATCGACACCggatccaccATGGTTAA
GGTGATTGGAAGACGTTCTCTTGGTGTTCAAAGGATCTTCGATATCGGATTGCCACAAGACCACAA
CTTTCTTCTCGCTAATGGTGCCATCGCTGCCAATAGCGCTGCGCAGTCCGAGCCGGAGCTGAAGCT
GGAGTCCGTGGTGATCGTGTCGCGCCACGGGGTGCGCGCCCCGACCAAGTTCACGCAGCTCATGCA
GGACGTGACCCCGGACGCCTTCTACACCTGGCCGGTGAAGCTCGGCGAGCTGACCCCGCGCGGCGG
CGAGCTGATCGCCTACCTCGGCCACTACTGGCGCCAGCGCCTCGTGGCCGACGGCCTCCTCCCGAA
GAAGGGCTGCCCGCAGTCCGGCCAGGTGGCGATCATCGCCGACGTGGACGAGCGCACCCGCAAGAC
GGGCGAGGCCTTCGCCGCCGGCCTCGCCCCGGACTGCGCCATCACCGTGCACACCCAGGCCGACAC
CTCCTCCCCGGACCCGCTCTTCAACCCGCTCAAGACCGGCGTGTGCCAGCTCGACGTGGCCCAGGT
GACCGACGCCATCCTGGAGCGCGCCGGCGGCTCCATCGCCGACTTCACCGGCCACTACCAGACCGC
CTTCCGCGAGCTGGAGCGCGTGCTCAACTTCCCGCAGTCGAACCTCGCCCTCAAGCGCGAGAAGCA
GGACGAGTCCGCCTCCCTCACCCAGGCCCTCCCGTCCGAGCTGAAGGTGTCCGCCGACAACGTGTC
CCTCACCGGCGCCTGGTCCCTCGCCTCCATGCTCACCGAAATCTTCCTCCTCCAGCAGGCCCAGGG
CATGCCGGAGCCGGGCTGGGGCCGCATCACCGACTCCCACCAGTGGAACACCCTCCTCTCCCTCCA
CAACGCCCAGTTCGACCTCCTCCAGCGCACCCCGGAGGTGGCCCGCTCCCGCGCCACCCCGCTCCT
CGACCTCATCAAGACCGCCCTCACCCCGCACCCGCCGCAGAAGCAGGCCTACGGCGTGACCCTCCC
GACCTCGGTGCTCTTCATCGCCGGCCACGACACCAACCTCGCCAACCTCGGCGGCGCCCTGGAGCT
GCAGTGGACCCTCCCGGGCCAGCCGGACAACACCCCGCCGGGCGGCGAGCTGGTGTTCGAGCGCTG
GCGCCGCCTCTCCGACAACTCCCAGTGGATTCAGGTGTCCCTCGTGTTCCAGACCCTCCAGCAGAT
GCGCGACAAGACCCCGCTCTTCCTCAACACCCCGCCGGGCGAGGTGAAGCTCACCCTGGCCGGCTG
CGAGGAGCGCAACGCGCAGGGCATGTGCTCCCTCGCCGGCTTCACCCAGATCGTGAACGAGGCCCG
CATCCCGGCCTGCTCCCTCTGCCTTTCTTTCGGAACTGAGATCCTTACCGTTGAGTACGGACCACT
TCCTATTGGTAAGATCGTTTCTGAGGAAATTAACTGCTCAGTGTACTCTGTTGATCCAGAAGGAAG
AGTTTACACTCAGGCTATCGCACAATGGCACGATAGGGGTGAACAAGAGGTTCTGGAGTACGAGCT
TGAAGATGGATCCGTTATTCGTGCTACCTCTGACCATAGATTCTTGACTACAGATTATCAGCTTCT
CGCTATCGAGGAAATCTTTGCTAGGCAACTTGATCTCCTTACTTTGGAGAACATCAAGCAGACAGA
AGAGGCTCTTGACAACCACAGACTTCCATTCCCTTTGCTCGATGCTGGAACCATCAAGTAAcctag
g*tccccgaatttccccgatcgttcaaacatttggcaataaagtttcttaagattgaatcctgttgc*
cggtcttgcgatgattatcatataatttctgttgaattacgttaagcatgtaataattaacatgta
atgcatgacgttatttatgagatgggttttatgattagagtcccgcaattatacatttaatacgc
gatagaaaacaaaatatagcgcgcaaactaggataaattatcgcgcgcggtgtcatctatgttact
agatcgggaattg (SEQ ID NO: 130)

[00150] >***ZmZ27P***:xGZein27ss:Ssp_DnaE–C:Phy02opt:Ssp_DnaE–N:*DPNG***SEKDEL**:*NosT*

[00151] ZmZ27P is shown in bold upper case font and italicized, Ssp_DnaE is underlined, DPNG is in upper case and italicized, SEKDEL is in bold upper case, NosT is italicized.

ggtacc***AAAGTAATCATATTATTTTATGTGTGAATCTTCTTTACTTTTTCATTTGATTATGATTAT***
GAAGGTATGACCTTCATAACCTTCGTCCGAAATCCATTATATCCAAAGGAAAATAATGCTTCGAAG
GACGAAGGATTTTGATATTTAACATTTTATGTTGCCTTGTTCTTAATTCATAGCATTTGAGAACAA
GTCCCCAACACCAATCTTTATCTTTACTATATTAAAGCACCAGTTCAACGATCGTCTCGTGTCAAT
ATTATTAAAAAACTCCTACATTTCTTTATAATCAACCCGCACTCTTATAATCTCTTCTCTTACTAC

TATAATAAGAGAGTTTATGTACAAAATAAGGTGAAATTATGTATAAGTGTTCTGGACCTTGGTTGT
TGGCTCATATTCACACAACCTAATCAATAGAAAACATATGTTTTATTAAAACAAAATTTATCATAT
ATATATATATATATATATATATATATATATATATATATATATAATATAAACCGTAGCAATGCACAGGCAT
ATGACTAGTGGCAACTTAATACCATGTGTGTATTAAGATGAATAAGAGGTATCCAAATAAATAACT
TGTTCGCTTACGTCTGGATCGAAAGGGGTTGGAAACGATTAAATCTCTTCCTAGTCAAAATTAAAT
AGAAGGAGATTTAATCGATTTCTCCCAATCCCCTTCGATCCAGGTGCAACCGAATAAGTCCTTAAA
TGTTGAGGAACACGAAACAACCATGCATTGGCATGTAAAGCTCCAAGAATTCGTTGTATCCTTAAC
AACTCACAGAACATCAACCAAAATTGCACGTCAAGGGTATTGGGTAAGAAACAATCAAACAAATCC
TCTCTGTGTGCAAAGAAACACGGTGAGTCATGCCGAGATCATACTCATCTGATATACATGCTTACA
GCTCACAAGACATTACAAACAACTCATATTGCATTACAAAGATCGTTTCATGAAAAATAAAATAGG
CCGGAACAGGACAAAAATCCTTGACGTGTAAAGTAAATTTACAACAAAAAAAAAGCCATATGTCAA
GCTAAATCTAATTCGTTTTACGTAGATCAACAACCTGTAGAAGGCAACAAAACTGAGCCACGCAGA
AGTACAGAATGATTCCAGATGAACCATCGACGTGCTACGTAAAGAGAGTGACGAGTCATATACATT
TGGCAAGAAACCATGAAGCTGCCTACAGCCGTCTCGGTGGCATAAGAACACAAGAAATTGTGTTAA
TTAATCAAAGCTATAAATAACGCTCGCATGCCTGTGCACTTCTCCATCACCACCACTGGGTCTTCA
*GACCATTAGCTTTATCTACTCCAGAGCGCAGAAGAACCCGATCGACACC*ggatccaccATGAGGGT
GTTGCTCGTTGCCCTCGCTCTCCTGGCTCTCGCTGCGAGCGCCACCAGCATGGTTAAGGTGATTGG
AAGACGTTCTCTTGGTGTTCAAAGGATCTTCGATATCGGATTGCCACAAGACCACAACTTTCTTCT
CGCTAATGGTGCCATCGCTGCCAATAGCGCTGCGCAGTCCGAGCCGGAGCTGAAGCTGGAGTCCGT
GGTGATCGTGTCGCGCCACGGGGTGCGCGCCCCGACCAAGTTCACGCAGCTCATGCAGGACGTGAC
CCCGGACGCCTTCTACACCTGGCCGGTGAAGCTCGGCGAGCTGACCCCGCGCGGCGGCGAGCTGAT
CGCCTACCTCGGCCACTACTGGCGCCAGCGCCTCGTGGCCGACGGCCTCCTCCCGAAGAAGGGCTG
CCCGCAGTCCGGCCAGGTGGCGATCATCGCCGACGTGGACGAGCGCACCCGCAAGACGGGCGAGGC
CTTCGCCGCCGGCCTCGCCCCGGACTGCGCCATCACCGTGCACACCCAGGCCGACACCTCCTCCCC
GGACCCGCTCTTCAACCCGCTCAAGACCGGCGTGTGCCAGCTCGACGTGGCCCAGGTGACCGACGC
CATCCTGGAGCGCGCCGGCGGCTCCATCGCCGACTTCACCGGCCACTACCAGACCGCCTTCCGCGA
GCTGGAGCGCGTGCTCAACTTCCCGCAGTCGAACCTCGCCCTCAAGCGCGAGAAGCAGGACGAGTC
CGCCTCCCTCACCCAGGCCCTCCCGTCCGAGCTGAAGGTGTCCGCCGACAACGTGTCCCTCACCGG
CGCCTGGTCCCTCGCCTCCATGCTCACCGAAATCTTCCTCCTCCAGCAGGCCCAGGGCATGCCGGA
GCCGGGCTGGGGCCGCATCACCGACTCCCACCAGTGGAACACCCTCCTCTCCCTCCACAACGCCCA
GTTCGACCTCCTCCAGCGCACCCCGGAGGTGGCCCGCTCCCGCGCCACCCCGCTCCTCGACCTCAT
CAAGACCGCCCTCACCCCGCACCCGCCGCAGAAGCAGGCCTACGGCGTGACCCTCCCGACCTCGGT
GCTCTTCATCGCCGGCCACGACACCAACCTCGCCAACCTCGGCGGCGCCCTGGAGCTGCAGTGGAC
CCTCCCGGGCCAGCCGGACAACACCCCGCCGGGCGGCGAGCTGGTGTTCGAGCGCTGGCGCCGCCT
CTCCGACAACTCCCAGTGGATTCAGGTGTCCCTCGTGTTCCAGACCCTCCAGCAGATGCGCGACAA
GACCCCGCTCTTCCTCAACACCCCGCCGGGCGAGGTGAAGCTCACCCTGGCCGGCTGCGAGGAGCG
CAACGCGCAGGGCATGTGCTCCCTCGCCGGCTTCACCCAGATCGTGAACGAGGCCCGCATCCCGGC
CTGCTCCCTCTGCCTTTCTTTCGGAACTGAGATCCTTACCGTTGAGTACGGACCACTTCCTATTGG
TAAGATCGTTTCTGAGGAAATTAACTGCTCAGTGTACTCTGTTGATCCAGAAGGAAGAGTTTACAC
TCAGGCTATCGCACAATGGCACGATAGGGGTGAACAAGAGGTTCTGGAGTACGAGCTTGAAGATGG
ATCCGTTATTCGTGCTACCTCTGACCATAGATTCTTGACTACAGATTATCAGCTTCTCGCTATCGA
GGAAATCTTTGCTAGGCAACTTGATCTCCTTACTTTGGAGAACATCAAGCAGACAGAAGAGGCTCT
TGACAACCACAGACTTCCATTCCCTTTGCTCGATGCTGGAACCATCAAG*GACCCGAACGGC***TCCGA**
GAAGGACGAGCTGTAAcctagg*tccccgaatttccccgatcgttcaaacatttggcaataaagttt*
cttaagattgaatcctgttgccggtcttgcgatgattatcatataatttctgttgaattacgttaa
gcatgtaataattaacatgtaatgcatgacgttatttatgagatgggttttatgattagagtccc
gcaattatacatttaatacgcgatagaaaacaaaatatagcgcgcaaactaggataaattatcgcg
cgcggtgtcatctatgttactagatcgggaattg (SEQ ID NO: 131)

[00152] >***ZmZ27P***:Ssp_DnaE:**L33-1**:Phy02opt:**L33-3**:*NosT*

(SSp_DnaE-C:**L33-1**:Phy02opt:**L33-2**:Ssp_DnaE-N)

[00153] ZmZ27P is shown in bold upper case font and italicized, Ssp_DnaE is underlined, linker is in bold, DPNG is in upper case and italicized, SEKDEL is in bold upper case, and NosT is italicized.

ggtacc***AAAGTAATCATATTATTTTATGTGTGAATCTTCTTTACTTTTTCATTTGATTATGATTAT***
GAAGGTATGACCTTCATAACCTTCGTCCGAAATCCATTATATCCAAAGGAAAATAATGCTTCGAAG
GACGAAGGATTTTGATATTTAACATTTTATGTTGCCTTGTTCTTAATTCATAGCATTTGAGAACAA
GTCCCCAACACCAATCTTTATCTTTACTATATTAAAGCACCAGTTCAACGATCGTCTCGTGTCAAT
ATTATTAAAAAACTCCTACATTTCTTTATAATCAACCCGCACTCTTATAATCTCTTCTCTTACTAC
TATAATAAGAGAGTTTATGTACAAAATAAGGTGAAATTATGTATAAGTGTTCTGGACCTTGGTTGT
TGGCTCATATTCACACAACCTAATCAATAGAAAACATATGTTTTATTAAAACAAAATTTATCATAT
ATATATATATATATATATATATATATATATATATATATATATATAATATAAACCGTAGCAATGCACAGGCAT
ATGACTAGTGGCAACTTAATACCATGTGTGTATTAAGATGAATAAGAGGTATCCAAATAAATAACT
TGTTCGCTTACGTCTGGATCGAAAGGGGTTGGAAACGATTAAATCTCTTCCTAGTCAAAATTAAAT
AGAAGGAGATTTAATCGATTTCTCCCAATCCCCTTCGATCCAGGTGCAACCGAATAAGTCCTTAAA
TGTTGAGGAACACGAAACAACCATGCATTGGCATGTAAAGCTCCAAGAATTCGTTGTATCCTTAAC
AACTCACAGAACATCAACCAAAATTGCACGTCAAGGGTATTGGGTAAGAAACAATCAAACAAATCC
TCTCTGTGTGCAAAGAAACACGGTGAGTCATGCCGAGATCATACTCATCTGATATACATGCTTACA
GCTCACAAGACATTACAAACAACTCATATTGCATTACAAAGATCGTTTCATGAAAAATAAAATAGG
CCGGAACAGGACAAAAATCCTTGACGTGTAAAGTAAATTTACAACAAAAAAAAAAGCCATATGTCAA
GCTAAATCTAATTCGTTTTACGTAGATCAACAACCTGTAGAAGGCAACAAAACTGAGCCACGCAGA
AGTACAGAATGATTCCAGATGAACCATCGACGTGCTACGTAAAGAGAGTGACGAGTCATATACATT
TGGCAAGAAACCATGAAGCTGCCTACAGCCGTCTCGGTGGCATAAGAACACAAGAAATTGTGTTAA
TTAATCAAAGCTATAAATAACGCTCGCATGCCTGTGCACTTCTCCATCACCACCACTGGGTCTTCA
GACCATTAGCTTTATCTACTCCAGAGCGCAGAAGAACCCGATCGACACCggatccaccATGGTTAA
GGTGATTGGAAGACGTTCTCTTGGTGTTCAAAGGATCTTCGATATCGGATTGCCACAAGACCACAA
CTTTCTTCTCGCTAATGGTGCCATCGCTGCCAAT**agcggcggcggcagcggcggcggcagcacccc**
gcagagcgccttcgccGCTGCGCAGTCCGAGCCGGAGCTGAAGCTGGAGTCCGTGGTGATCGTGTC
GCGCCACGGGGTGCGCGCCCCGACCAAGTTCACGCAGCTCATGCAGGACGTGACCCCGGACGCCTT
CTACACCTGGCCGGTGAAGCTCGGCGAGCTGACCCCGCGCGGCGGCGAGCTGATCGCCTACCTCGG
CCACTACTGGCGCCAGCGCCTCGTGGCCGACGGCCTCCTCCCGAAGAAGGGCTGCCCGCAGTCCGG
CCAGGTGGCGATCATCGCCGACGTGGACGAGCGCACCCGCAAGACGGGCGAGGCCTTCGCCGCCGG
CCTCGCCCCGGACTGCGCCATCACCGTGCACACCCAGGCCGACACCTCCTCCCCGGACCCGCTCTT
CAACCCGCTCAAGACCGGCGTGTGCCAGCTCGACGTGGCCCAGGTGACCGACGCCATCCTGGAGCG
CGCCGGCGGCTCCATCGCCGACTTCACCGGCCACTACCAGACCGCCTTCCGCGAGCTGGAGCGCGT
GCTCAACTTCCCGCAGTCGAACCTCGCCCTCAAGCGCGAGAAGCAGGACGAGTCCGCCTCCCTCAC
CCAGGCCCTCCCGTCCGAGCTGAAGGTGTCCGCCGACAACGTGTCCCTCACCGGCGCCTGGTCCCT
CGCCTCCATGCTCACCGAAATCTTCCTCCTCCAGCAGGCCCAGGGCATGCCGGAGCCGGGCTGGGG
CCGCATCACCGACTCCCACCAGTGGAACACCCTCCTCTCCCTCCACAACGCCCAGTTCGACCTCCT
CCAGCGCACCCCGGAGGTGGCCCGCTCCCGCGCCACCCCGCTCCTCGACCTCATCAAGACCGCCCT
CACCCCGCACCCGCCGCAGAAGCAGGCCTACGGCGTGACCCTCCCGACCTCGGTGCTCTTCATCGC
CGGCCACGACACCAACCTCGCCAACCTCGGCGGCGCCCTGGAGCTGCAGTGGACCCTCCCGGGCCA
GCCGGACAACACCCCGCCGGGCGGCGAGCTGGTGTTCGAGCGCTGGCGCCGCCTCTCCGACAACTC
CCAGTGGATTCAGGTGTCCCTCGTGTTCCAGACCCTCCAGCAGATGCGCGACAAGACCCCGCTCTT
CCTCAACACCCCGCCGGGCGAGGTGAAGCTCACCCTGGCCGGCTGCGAGGAGCGCAACGCGCAGGG
CATGTGCTCCCTCGCCGGCTTCACCCAGATCGTGAACGAGGCCCGCATCCCGGCCTGCTCCCTC**gg**
cggcggcagcggcggcggcagcggcggcggcTGCCTTTCTTTCGGAACTGAGATCCTTACCGTTGA
GTACGGACCACTTCCTATTGGTAAGATCGTTTCTGAGGAAATTAACTGCTCAGTGTACTCTGTTGA
TCCAGAAGGAAGAGTTTACACTCAGGCTATCGCACAATGGCACGATAGGGGTGAACAAGAGGTTCT
GGAGTACGAGCTTGAAGATGGATCCGTTATTCGTGCTACCTCTGACCATAGATTCTTGACTACAGA
TTATCAGCTTCTCGCTATCGAGGAAATCTTTGCTAGGCAACTTGATCTCCTTACTTTGGAGAACAT
CAAGCAGACAGAAGAGGCTCTTGACAACCACAGACTTCCATTCCCTTTGCTCGATGCTGGAACCAT
CAAGTAAcctagg*tccccgaatttccccgatcgttcaaacatttggcaataaagtttcttaagatt*
gaatcctgttgccggtcttgcgatgattatcataatttctgttgaattacgttaagcatgtaat

aattaacatgtaatgcatgacgttatttatgagatgggtttttatgattagagtcccgcaattata
catttaatacgcgatagaaaacaaaatatagcgcgcaaactaggataaattatcgcgcgcggtgtc
atctatgttactagatcgggaattg (SEQ ID NO: 132)

[00154] >***ZmZ27P***:xGZein27ss:Ssp_DnaE:**L33-1**:Phy02opt:L33-2:*DPNG***SEKDEL**:*NosT*

[00155] ZmZ27P is shown in bold upper case font and italicized, Ssp_DnaE is underlined, L33 linker is in bold upper case, DPNG is in upper case and italicized, SEKDEL is in bold upper case, and NosT is italicized.

ggtacc***AAAGTAATCATATTATTTTATGTGTGAATCTTCTTTACTTTTTCATTTGATTATGATTAT***
GAAGGTATGACCTTCATAACCTTCGTCCGAAATCCATTATATCCAAAGGAAAATAATGCTTCGAAG
GACGAAGGATTTTGATATTTAACATTTTATGTTGCCTTGTTCTTAATTCATAGCATTTGAGAACAA
GTCCCCAACACCAATCTTTATCTTTACTATATTAAAGCACCAGTTCAACGATCGTCTCGTGTCAAT
ATTATTAAAAAACTCCTACATTTCTTTATAATCAACCCGCACTCTTATAATCTCTTCTCTTACTAC
TATAATAAGAGAGTTTATGTACAAAATAAGGTGAAATTATGTATAAGTGTTCTGGACCTTGGTTGT
TGGCTCATATTCACACAACCTAATCAATAGAAAACATATGTTTTATTAAAACAAAATTTATCATAT
ATATATATATATATATATATATATATATATATATATATATATAATATAAACCGTAGCAATGCACAGGCAT
ATGACTAGTGGCAACTTAATACCATGTGTGTATTAAGATGAATAAGAGGTATCCAAATAAATAACT
TGTTCGCTTACGTCTGGATCGAAAGGGGTTGGAAACGATTAAATCTCTTCCTAGTCAAAATTAAAT
AGAAGGAGATTTAATCGATTTCTCCCAATCCCCTTCGATCCAGGTGCAACCGAATAAGTCCTTAAA
TGTTGAGGAACACGAAACAACCATGCATTGGCATGTAAAGCTCCAAGAATTCGTTGTATCCTTAAC
AACTCACAGAACATCAACCAAAATTGCACGTCAAGGGTATTGGGTAAGAAACAATCAAACAAATCC
TCTCTGTGTGCAAAGAAACACGGTGAGTCATGCCGAGATCATACTCATCTGATATACATGCTTACA
GCTCACAAGACATTACAAACAACTCATATTGCATTACAAAGATCGTTTCATGAAAAATAAAATAGG
CCGGAACAGGACAAAAATCCTTGACGTGTAAAGTAAATTTACAACAAAAAAAAAAGCCATATGTCAA
GCTAAATCTAATTCGTTTTACGTAGATCAACAACCTGTAGAAGGCAACAAAACTGAGCCACGCAGA
AGTACAGAATGATTCCAGATGAACCATCGACGTGCTACGTAAAGAGAGTGACGAGTCATATACATT
TGGCAAGAAACCATGAAGCTGCCTACAGCCGTCTCGGTGGCATAAGAACACAAGAAATTGTGTTAA
TTAATCAAAGCTATAAATAACGCTCGCATGCCTGTGCACTTCTCCATCACCACCACTGGGTCTTCA
GACCATTAGCTTTATCTACTCCAGAGCGCAGAAGAACCCGATCGACACCggatccaccATGAGGGT
GTTGCTCGTTGCCCTCGCTCTCCTGGCTCTCGCTGCGAGCGCCACCAGCATGGTTAAGGTGATTGG
AAGACGTTCTCTTGGTGTTCAAAGGATCTTCGATATCGGATTGCCACAAGACCACAACTTTCTTCT
CGCTAATGGTGCCATCGCTGCCAAT**agcggcggcggcagcggcggcggcagcaccccgcagagcgc**
cttcgccGCTGCGCAGTCCGAGCCGGAGCTGAAGCTGGAGTCCGTGGTGATCGTGTCGCGCCACGG
GGTGCGCGCCCCGACCAAGTTCACGCAGCTCATGCAGGACGTGACCCCGGACGCCTTCTACACCTG
GCCGGTGAAGCTCGGCGAGCTGACCCCGCGCGGCGGCGAGCTGATCGCCTACCTCGGCCACTACTG
GCGCCAGCGCCTCGTGGCCGACGGCCTCCTCCCGAAGAAGGGCTGCCCGCAGTCCGGCCAGGTGGC
GATCATCGCCGACGTGGACGAGCGCACCCGCAAGACGGGCGAGGCCTTCGCCGCCGGCCTCGCCCC
GGACTGCGCCATCACCGTGCACACCCAGGCCGACACCTCCTCCCCGGACCCGCTCTTCAACCCGCT
CAAGACCGGCGTGTGCCAGCTCGACGTGGCCCAGGTGACCGACGCCATCCTGGAGCGCGCCGGCGG
CTCCATCGCCGACTTCACCGGCCACTACCAGACCGCCTTCCGCGAGCTGGAGCGCGTGCTCAACTT
CCCGCAGTCGAACCTCGCCCTCAAGCGCGAGAAGCAGGACGAGTCCGCCTCCCTCACCCAGGCCCT
CCCGTCCGAGCTGAAGGTGTCCGCCGACAACGTGTCCCTCACCGGCGCCTGGTCCCTCGCCTCCAT
GCTCACCGAAATCTTCCTCCTCCAGCAGGCCCAGGGCATGCCGGAGCCGGGCTGGGGCCGCATCAC
CGACTCCCACCAGTGGAACACCCTCCTCTCCCTCCACAACGCCCAGTTCGACCTCCTCCAGCGCAC
CCCGGAGGTGGCCCGCTCCCGCGCCACCCCGCTCCTCGACCTCATCAAGACCGCCCTCACCCCGCA
CCCGCCGCAGAAGCAGGCCTACGGCGTGACCCTCCCGACCTCGGTGCTCTTCATCGCCGGCCACGA
CACCAACCTCGCCAACCTCGGCGGCGCCCTGGAGCTGCAGTGGACCCTCCCGGGCCAGCCGGACAA
CACCCCGCCGGGCGGCGAGCTGGTGTTCGAGCGCTGGCGCCGCCTCTCCGACAACTCCCAGTGGAT
TCAGGTGTCCCTCGTGTTCCAGACCCTCCAGCAGATGCGCGACAAGACCCCGCTCTTCCTCAACAC

CCCGCCGGGCGAGGTGAAGCTCACCCTGGCCGGCTGCGAGGAGCGCAACGCGCAGGGCATGTGCTC
CCTCGCCGGCTTCACCCAGATCGTGAACGAGGCCCGCATCCCGGCCTGCTCCCTC**ggcggcggcag**
cggcggcggcagcggcggcggcTGCCTTTCTTTCGGAACTGAGATCCTTACCGTTGAGTACGGACC
ACTTCCTATTGGTAAGATCGTTTCTGAGGAAATTAACTGCTCAGTGTACTCTGTTGATCCAGAAGG
AAGAGTTTACACTCAGGCTATCGCACAATGGCACGATAGGGGTGAACAAGAGGTTCTGGAGTACGA
GCTTGAAGATGGATCCGTTATTCGTGCTACCTCTGACCATAGATTCTTGACTACAGATTATCAGCT
TCTCGCTATCGAGGAAATCTTTGCTAGGCAACTTGATCTCCTTACTTTGGAGAACATCAAGCAGAC
AGAAGAGGCTCTTGACAACCACAGACTTCCATTCCCTTTGCTCGATGCTGGAACCATCAAG*GACCC*
*GAACGGC***TCCGAGAAGGACGAGCTG**TAA*cctaggtccccgaatttccccgatcgttcaaacatttg*
gcaataaagtttcttaagattgaatcctgttgccggtcttgcgatgattatcatataatttctgtt
gaattacgttaagcatgtaataattaacatgtaatgcatgacgttatttatgagatgggttttat
gattagagtcccgcaattatacatttaatacgcgatagaaaacaaaatatagcgcgcaaactagga
taaattatcgcgcgcggtgtcatctatgttactagatcgggaattg *(SEQ ID NO: 133)*

[00156] ***Example 11. Expression of cyclized phytases in transgenic plants***

[00157] Independently transgenic maize plants that had been transformed with vectors as described above were grown to maturity, and cross-pollinated with wild-type (untransformed) maize plants. Approximately 20 seeds were harvested from each of these plants. Seed was milled through a 0.5mm screen to produce a fine powder. Enzyme was then extracted and assayed for phytase activity as described below.

[00158] *Phytase assay from seed, brief description of the protocol.* Enzyme extracts were prepared by incubating 15mg milled seed flour for 1 hour at room temperature in 1.5ml of 25mM sodium borate, pH10, 0.01% Tween 20. Extracts were then diluted 100-fold in assay buffer (250mM sodium acetate, pH5.5, 1mM calcium chloride, 0.01% Tween 20). Seventy-five (75) microliters of the diluted extracts or 75µl of buffer-only controls were dispensed into individual wells of a round-bottom 96-well plate. One-hundred fifty (150) microliters of freshly-prepared phytic acid (9.1mM dodecasodium salt from Biosynth International, Staad, Switzerland, prepared in assay buffer) were added to each well. Plates were sealed and incubated for 60min at 37°C. 150 µL of stop solution (20mM ammonium molybdate, 5mM ammonium vanadate, 4% nitric acid) was added to each well, mixed thoroughly via pipetting, and allowed to incubate at room temperature for 10 min. Plates were centrifuged at 3000×G for 10 minutes, and 100 µL of the clarified supernatants were transferred to the wells of a flat-bottom 96-well plate. Absorbance at 415nm from each sample was compared to that of negative controls (buffer-only, no

enzyme) and potassium phosphate standards. The standard curve was prepared by mixing 50 µl of potassium phosphate standards (0-1.44mM, prepared in assay buffer) with 100 µL of freshly-prepared phytic acid, followed by 100 µL of stop solution.

[00159] Phytase activity varied significantly in seed from independent transgenic plants, as expected.

[00160] ***Example 12. Thermal stability of cyclic phytases in pelleting processes***

[00161] To determine the thermal stability of an engineered phytase, feed must be mixed containing a specified level of the engineered phytase, the corresponding target phytase, and any control phytases that it is desired to compare the thermal stability with and include in the evaluation. For testing thermal stability in feed, it is beneficial to mix several diets at a few different dosing levels, and then evaluate each in a series of pelleting processes conducted at different temperatures. Doses used in the evaluation may include 500 FTU/ kg, 1000FTU/kg, or 3000 FTU/kg. Temperatures used in the evaluation may include 60°C, 65°C, 70°C, 75°C, 80°C, 85°C, 90°C, and 95°C, or any other desired temperatures. The residence time in the pelleting process may range from 15 seconds or less, up to one minute or more. For each formulated diet, for each enzyme (and the negative control diets containing no enzyme), a pre-pelleting sample is taken in addition to samples taken after pelleting. From these samples, the activity is measured and compared. Pelleted samples are compared with the corresponding mash samples in each treatment, and also compared with the identical treatments with other enzymes included in the trial. Engineered enzymes that maintain the highest percentage of activity post-pelleting at the highest temperatures demonstrate the greatest degree of thermal stability. Engineered phytases that demonstrate higher thermal stability than the corresponding target phytase have improved thermal performance and are candidates for commercial development.

[00162] ***Example 13. Performance of cyclic phytases in broilers and pigs***

[00163] A basal corn-soy diet was prepared with a low content of inorganic phosphate. Replicate diets were prepared from this basal diet by adding enzyme in the form of Quantum Blue (AB Enzymes) or milled corn grain expressing either Phy02, Nov9X, engineered cyclic Phy02, or engineered cyclic Nov9X, varying the total amount of enzyme incorporated into each diet. For Phy02 and Nov9X, a small amount of corn was omitted from the basal diet to account for the transgenic grain that was being added back to supply the enzyme. Control diets were prepared in which the amount of inorganic phosphate was increased relative to the basal diet.

[00164] Male broiler chicks were distributed among various feed treatments in pens with about 12 birds per pen, and 6 replicate pens per treatment. The feed was provided to one set of birds in mash form, and pelleted feeds was provided to another set of birds. After 14, 21, 28, 35, and 42 days, birds are weighed and compared to determine the effect of the various enzyme treatments on broiler production.

[00165] Similarly, pigs were distributed among various feed treatments in pens with about 7 pigs per pen, and 5 replicate pens per treatment. The pelleted feed was provided the pigs. After 21, 35, and 49 days, pigs are weighed and compared to determine the effect of the various enzyme treatments on broiler production.

[00166] ***Example 14. Heat stability of modified phytases***

[00167] FIG. 11 illustrates expression profiles of SspDnaE-C:Phy02:SspDnaE-N constructs. Referring to this figure, “C” represents the crude extract, “S” represents the soluble fraction, “*” marks the position of the target protein in the crude extract and “o” marks the position of cyclic Phy02 in the crude extract. Coomassie gel of IPTG induced expression cultures. Constructs were cloned between the EcoRI and XhoI sites of pETDuetI (Novagen) and transformed into Shuffle T7 (NEB) *E.coli* expression host. To analyze expression profiles, overnight starter cultures in LB+ Carbenicillin

(100 mg/L) were 40-fold diluted to fresh medium and grown at 30°C, 250 rpm to OD_{600}= 0.6, then IPTG was added to 0.5 mM final concentration and the cultures were grown for another 3 hours. Cells were harvested at 3000 g for 10 minutes, washed with one culture volume of phytase wash buffer (250 mM NaOAc pH=5.5 and 1mM $CaCl_2$) and cells were pelleted as before. Cell pellet was lysed (30°C, 250 rpm, 1 hr) in phytase lysis buffer that contains 1X Fastbreak (Promega) with Benzonase (50U/mL, Novagen). Sample preparation for the Coomassie gel was as follows: Crude extract (C) was made by mixing equal volumes of lysate with 2 X Laemmli sample buffer (Bio-Rad) containing 5% beta-mercaptoethanol. To prepare the soluble fraction (S) lysates were centrifuged at 5000 g for 10 min and the supernatants were mixed with equal volumes of loading dye as before. The heat soluble fraction (H) was made by incubating the lysates at 55 °C for 15 min followed by centrifugation at 5000 g for 10 minutes and supernatants were mixed with equal volumes of loading dye. Before loading, SDS/PAGE samples were heated at 95 °C for 5min and 5 µL aliquots were loaded to Criterion XT 12% Bis-Tris gels together with 10 µL of the Mw marker (Precision Plus Protein Kaleidoscope, Bio-Rad). After separation of the proteins, gel was stained with SimplyBlue Safe Stain (Novex by Life Technologies).

[00168] Referring to FIG. 11, it was observed that Phy02 represented comparably in the crude (C), soluble (S) and heat soluble (H) fraction. Expression levels of SspDnaE-C:phy02:SspDnaE-N fusion proteins were comparable but showed significant difference in solubility: without linker (-) the protein was primarily non-soluble, while the linker containing constructs primarily expressed to the soluble fraction and were well represented in the heat soluble fractions as well. Phy02 and its intein-modified versions were resolved at the expected size of the linear molecules (marked "*", around 58KD), except two constructs with the longest linkers (linker 46 and 55), that in addition to the linear proteins showed faster moving new protein species (marked "o") at comparable levels in the crude (C), soluble (S) and heat soluble (H) fractions. Higher mobility is a hallmark of cyclic Phy02 as established by

comparing mobility of cyclization competent SpyTag : Phy02: SpyCatcher with the cyclization deficient mutant (see FIG. 12).

[00169] FIG. 12 illustrates the heat stability assay of Phy02. Referring to FIG. 12, the crude extract was prepared as described in FIG. 11 and diluted 50X in phytase wash buffer. 150 µL aliquots in PCR tubes were heat treated in a PCR block programmed to for identical block a lid temperature. Tubes were withdrawn at specified time points and incubated at room temp for 1 hour to allow for refolding. Each sample was diluted to 250-, 1000-, 5000- and 20000 –fold and phytase activity was assayed based on established protocol.

[00170] The graph illustrates heat stability of the unmodified Phy02 in crude cell lysates pretreated at 70°C, 75 °C and 80°C over 4 min in samples taken in 30 sec intervals. Full activity was retained only in the 70°C/ 30 sec sample. Increasing either heat exposure time and/or temperature quickly diminished phytase activity. One minute exposure to 75°C or 80°C reduced the unmodified Phy02 phytase activity to levels borderline detectable or undetectable, respectively.

[00171] FIGS. 13A - 13B illustrate heat stability of SspDnaE-C:Phy02:SspDnaE-N constructs. Expression culture and preparation of crude extract was as in FIG. 11. Heat pretreatment was performed at 75 °C for 60 sec and phytase activity was assayed as in FIG.12. FIG. 13A shows enzyme activity of untreated (37 °C) and heat treated (75 °C/ 60 sec) samples. FIG. 13B shows residual phytase activity in heat pretreated samples as percentage of activity of their respective untreated control (37 °C).

[00172] Each linker modified trans-splicing Phy02 retained some activity after a heat pretreatment that completely abolished phytase activity of the unmodified Phy02 control. The two clones with the longest linkers (linker 46 and 55) showed the highest heat tolerance at retained ~ 10% activity in the heat pretreated samples. Intein fusion without linker (DnaE-sPhy02_DnaE) did not improve heat stability.

[00173] FIG. 14 illustrates expression profiles of SpyTag:Phy02:SpyCatcher wild type and mutated forms. Coomassie gel of IPTG induced expression cultures.

[00174] Constructs were cloned between the NcoI and XhoI sites of pETDuetI (Novagen) and transformed into Shuffle T7 (NEB) *E. coli* expression host. The cyclization deficient mutant carried an alanine mutation in the SpyTag (AHIVMVDAYKPTK [SEQ ID NO: 216] for wild type and AHIVMVAAYKPTK [SEQ ID NO: 217] for mutant). Induction cultures, preparation of the crude (C), soluble (S) and heat soluble (H) fraction and SDS/PAGE were the same as in FIG. 11. Position of the target proteins are marked by asterisk in the crude extracts.

[00175] Both the wild type and the mutated SpyTag:Phy02:SpyCatcher expressed to the soluble fraction and were equally represented in the heat soluble (H), soluble (S) fractions as well as in the crude (C). While the cyclization competent version (wt) separated at the expected size for the linear molecule at 63 kD (552 amino acids), the cyclization deficient mutant (mut) moved fast on the gel. This observation is consistent with the interpretation that intramolecular interaction between SpyTag and SpyCatcher leads to intramolecular cyclization of the cyclization competent molecule. Mutation in the SpyTag prevented cyclization. Cyclic Phy02 has higher mobility than the cyclization deficient linear molecule. The cyclization competent wild type SpyTag:Phy02:SpyCatcher dominantly express the high mobility Phy02 form indicating that cyclization is highly efficient.

[00176] FIG. 15A illustrates SpyTag:Phy02:SpyCatcher improves heat tolerance of phytase. Phytase activity of heat pretreated samples. Expression of recombinant proteins was as described in FIG. 14, heat pretreatment and enzyme assay was performed as in FIG. 12 at 75°C and 80 °C and aliquots were taken at 30 sec intervals over 120 sec. Left panel show enzyme activity after heat treatment at 75 °C, right panel after heat treatment at 80 °C, respectively. The cyclization competent wild type SpyTag:Phy02:SpyCatcher (wt) showed dramatically improved heat stability and remained stable at 80 °C over the entire length of heat pretreatment tested. The cyclization deficient mutant SpyTag:Phy02:SpyCatcher (mut) also displayed improved heat stability compared to the unmodified Phy02.

[00177] FIG. 15B illustrates SpyTag:Phy02:SpyCatcher improves heat tolerance of phytase. Retention of phytase activity of heat pretreated samples. Phytase activities of heat pretreated samples of FIG. 17A are graphed as percentage of their respective untreated control. The cyclization competent phytase (wt) retained more than 35 % activity at 80 ^{o}C that remained stable over the entire heat treatment period of 2 minutes. In contrast, the cyclization disabled linear form (mut) quickly lost activity at 80 ^{o}C, but thermo-tolerance exceeded heat stability of the unmodified Phy02. This beneficial effect possibly due to retention of the refolding functionality of the SpyCatcher in the cyclization disabled mutated SpyTag construct. Possibly, the differences between heat tolerance of phytase activity of the cyclic and linear molecules could indicate the extent to which cyclization and refolding impact on heat stability.

[00178] ***Example 15. Intein splicing is required for attaining elevated heat tolerance of the cyclic phytase constructs***

[00179] The prototype cyclic phytase was constructed by using the rigid linker 55-1 and 55-2 and the trans-splicing intein gp41-1 and created the gp41-1C:L55-1:Phy02:L55-2:gp41-1N [Amino acid (AA)_SEQ ID NO: 201 and nucleic acid(NA) _SEQ ID NO: 200]. In addition, a solubility optimized version of the construct that have a solubility enhancer thioredoxin domain (TrxH) [AA_SEQ ID NO: 197 and NA_Seq ID NO: 196] at the N-terminus attached with an Asp-Pro-Asn-Gly linker (DPNG) [AA_SEQ ID NO: 199 and NA_SEQ ID NO: 198] to a mutated version of the gp41-1C (MTT) encoding the construct of TrxH:DPNG:gp41-1C[MTT]:L55-1:Phy02:L55-2:gp41-1N [AA_205 and NA_204] was created.

[00180] Constructs were cloned between the EcoRI and XhoI sites of pETDuetI, expressed from the Shuffle T7 *E.coli* host and were tested for phytase heat stability. Induction cultures and preparation of crude lysates were as described for FIG. 11. For heat treatments, 150 µL of the crude lysates in PCR tubes were heated for 1 min at the specified temperatures in PCR blocks, then tubes were incubated at room temp for 1 hr to allow for refolding. Each sample was diluted to 250-, 1000-, 5000- and 2000-fold and

phytase activity was assayed as in FIG. 12. FIG. 16 illustrates heat pretreatment of cyclic phytases gp41-L55-1:Phy02:L55-2:gp41-1N (closed circle) and TrxH:DPNG:gp41-1C[MTT]:L55-1:Phy02:L55-2:gp41-1N (closed square) compared to the wild type enzyme Phy02 (vertical line) and an empty vector (horizontal line). Referring to FIG. 16, it was observed the wild type phytase (Phy02) quickly lost activity above 75 ^{o}C while both cyclic phytase constructs retained activity at 85 ^{o}C, showing 16% versus 8% activity in the prototype vs. solubility optimized phytase, respectively.

[00181] To evaluate whether protein cyclization is required for acquisition of heat tolerance, splicing was disabled by mutating splicing essential amino acid residues in two cyclic phytase constructs with different linkers, in the TrxH:DPNG:gp41-1C [MTT]:L46-1:Phy02:L46-2:gp41-1N [AA_SEQ ID NO: 207 and NA_SEQ ID NO: 206] and the gp41-1C[MTT]:L55-1:Phy02:L55-2:gp41-1N [AA_SEQ ID NO: 205 and NA_SEQ ID NO: 204]. Splicing disabling mutations were either the gp41-1C intein C-terminal Asn residue to Ala [N125A] or the gp41-1C C-terminal flanking Ser residue to Ala [S1A] in +1 position of the linkers. The following mutants were created: [N125A-1] splicing disabled TrxH:DPNG:gp41-1C[MTT]:L46-1:Phy02:L46-2:gp41-1N [AA_SEQ ID NO: 209 and NA_SEQ ID NO: 208], [N125A-2]splicing disabled gp41-1C[MTT]:L55-1:Phy02:L55-2:gp41-1N [AA_SEQ ID NO: 213 and NA_SEQ ID NO: 212], [S1A-1] splicing disabled TrxH:DPNG:gp41-1C[MTT]:L46-1:Phy02:L46-2:gp41-1N [AA_SEQ ID NO: 211 and NA_SEQ ID NO: 210], and [S1A-2] splicing disabled gp41-1C[MTT]:L55-1:Phy02:L55-2:gp41-1N [AA_SEQ ID NO: 215 and NA_SEQ ID NO: 214].

[00182] Constructs were cloned between the EcoRI and XhoI sites of pETDuetI, expressed from Shuffle T7 *E.coli* host and heat tolerance of splicing enabled and disabled constructs were tested after heat pretreatment at 85 ^{o}C/ 1 min. FIG. 17 illustrates phytase activity of the splicing enabled and the splicing disabled (intein N125A and linker S1A) cyclic phytases gp41-1C:L55-1:Phy02:L55-2:gp41-1N and TrxH:DPNG:gp41-1C[MTT]:L55-1:Phy02:L55-2:gp41-1N and wild type Phy02 phytase following pretreatment at 85^{o}C for 1 minute. Referring to FIG. 17, it was observed that at 37 ^{o}C, all constructs

showed phytase activity. After 1 min exposure to 85 °C however, only the splicing enabled constructs retained activity. The splicing disabled mutants all displayed heat sensitivity similar to the intein unmodified wild type phytase. These results are consistent with the interpretation that acquisition of heat tolerance depends from intein splicing mediated protein cyclization.

References:

Apgar, J., Ross, M., Zuo, X., Dohle, S., Sturtevant, D., Shen, B., ... & Raab, R. M. (2012). A predictive model of intein insertion site for use in the engineering of molecular switches. *PloS one*, *7*(5), e37355.

Arakawa, T., Chong, D. K., & Langridge, W. H. (1998). Efficacy of a food plant-based oral cholera toxin B subunit vaccine. *Nature Biotechnology*, *16*(3), 292–297. doi:10.1038/nbt0398-292.

Basu, S. S., Winslow, S., Nelson, A., Ono, M., & Betts, S. (2009). EXTRACTION METHODS AND ASSAYS FOR FEED ENZYMES.

Cervelli, M., Di Caro, O., Di Penta, A., Angelini, R., Federico, R., Vitale, A., & Mariottini, P. (2004). A novel C-terminal sequence from barley polyamine oxidase is a vacuolar sorting signal. *Plant Journal*, *40*(3), 410–418. doi:10.1111/j.1365-313X.2004.02221.X.

"Current Protocols in Molecular Biology," 10.0.1-10.0.23, April, 2010, John Wiley & Sons, Inc.

Engelen, A. J., Heeft, F. C., Randsdorp, P. H., Somers, W. A., Schaefer, J., & van der Vat, B. J. (2001). Determination of phytase activity in feed by a colorimetric enzymatic method: collaborative interlaboratory study. *Journal of AOAC International*, *84*(3), 629-633 .

FU, T. J. (2002). Digestion stability as a criterion for protein allergenicity assessment. *Annals of the New York Academy of Sciences*, *964*(1), 99-110.

Gogarten, J. P., Senejani, A. G., Zhaxybayeva, O., Olendzenski, L., & Hilario, E. (2002). Inteins: structure, function, and evolution. *Annual Reviews in Microbiology*, *56*(1), 263-287.

Haq, T. a, Mason, H. S., Clements, J. D., & Arntzen, C. J. (1995). Oral immunization with a recombinant bacterial antigen produced in transgenic plants. *Science (New York, N.Y.)*, *268*(5211), 714–716. doi:10.1126/science.7732379.

Lau SYM, Taneja AK and Hodges RS (1984) Synthesis of a model protein of defined secondary and quaternary structure. Effect of chain length on the stabilization and formation of two-stranded α-helical coiled-coils. J. Biol. Chem. 259 (21), 13253-61.

Korban, S. S. (2002). Targeting and expression of antigenic proteins in transgenic plants for production of edible oral vaccines. *In Vitro Cellular & Developmental Biology - Plant*, *38*(3), 231–236. doi:10.1079/IVP2002292.

Munro, S., & Pelham, H. R. (1987). A C-terminal signal prevents secretion of luminal ER proteins. *Cell*, *48*(5), 899–907. doi:10.1016/0092-8674(87)90086-9.

Parry DA, Fraser RD and Squire JM (2008) Fifty years of coiled-coils and alpha-helical bundles: a close relationship between sequence and structure. J Struct Biol. 163(3), 258-69.

Perler, F. B. (2002). InBase: the intein database. *Nucleic acids research*, *30*(1), 383-384.

Perler, F. B., Davis, E. O., Dean, G. E., Gimble, F. S., Jack, W. E., Neff, N., ... & Belfort, M. (1994). Protein splicing elements: inteins and exteins--a definition of terms and recommended nomenclature. *Nucleic acids research*, *22*(7), 1125.

Schoene, C., Fierer, J. O., Bennett, S. P., & Howarth, M. (2014). SpyTag/SpyCatcher cyclization confers resilience to boiling on a mesophilic enzyme. *Angewandte Chemie International Edition*, *53*(24), 6101-6104.

Thomas, K., Aalbers, M., Bannon, G. A., Bartels, M., Dearman, R. J., Esdaile, D. J., ... & Zawodny, J. (2004). A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. *Regulatory Toxicology and Pharmacology*, *39*(2), 87-98.

US 7,629,139, issued December 8, 2009.

Woolfson DN (2005) The design of coiled-coil structures and assemblies. Adv. Protein Chem.70, 79-f112

Xu, M. Q., & Perler, F. B. (1996). The mechanism of protein splicing and its modulation by mutation. *The EMBO journal*, *15*(19), 5146.

Zakeri, B., Fierer, J. O., Celik, E., Chittock, E. C., Schwarz-Linek, U., Moy, V. T., & Howarth, M. (2012). Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proceedings of the National Academy of Sciences*, *109*(12), E690-E697.

[00183] The references cited throughout this application, are incorporated for all purposes apparent herein and in the references themselves as if each reference was fully set forth. For the sake of presentation, specific ones of these references are cited at particular locations herein. A citation of a reference at a particular location indicates a manner(s) in which the teachings of the reference are incorporated. However, a citation of a reference at a particular location does not limit the manner in which all of the teachings of the cited reference are incorporated for all purposes.

[00184] It is understood, therefore, that this invention is not limited to the particular embodiments disclosed, but is intended to cover all modifications which are within the spirit and scope of the invention as defined by the appended claims; the above description; and/or shown in the attached drawings.

* * *

CLAIMS

What is claimed is:

1. An engineered phytase comprising a target phytase, a first binding element and a second binding element, wherein each of the first binding element and the second binding is fused to the target phytase, the first binding element interacts with the second binding element to cause cyclization of the engineered phytase, and enhance thermal stability of the target phytase, and wherein the first binding element or the second binding element is selected from the group consisting of: a tag domain, a catcher domain, an intein or part thereof, and a coiled-coil dimerization domain or part thereof.

2. The engineered phytase of claim 1, wherein upon the interaction, each of the first binding element and the second binding element is capable of being released from the engineered phytase spontaneously.

3. The engineered phytase of claim 1, wherein the first binding element or the second binding element is fused to the N-terminus or the C-terminus of the target phytase.

4. The engineered phytase of claim 3, wherein the C-terminus of the first binding element is linked to and contiguous with the N-terminus of the target phytase.

5. The engineered phytase of claim 3, wherein the N-terminus of the second binding element is linked to and contiguous with the C-terminus of the target phytase.

6. The engineered phytase of claim 1, wherein the target phytase is selected from the group consisting of phytases derived from *Escherichia coli,*

Aspergillus niger, *Peniophora lycii*, *Neurospora crassa*, and *Schwaniomyces accidentalis*.

7. The engineered phytase of claim 1, wherein the target phytase comprises an amino acid sequence with at least 90% identity to a reference sequence selected from the group consisting of SEQ ID NOS: 53 - 54, 56, and 219.

8. The engineered phytase of claim 1, wherein the first binding element is a C-intein of an intein.

9. The engineered phytase of claim 8, wherein the C-intein comprises an amino acid sequence with at least 90% identity to a reference sequence selected from the group consisting of: SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 189, 191, and 195.

10. The engineered phytase of claim 1, wherein the second binding element is an N-intein of an intein.

11. The engineered phytase of claim 10, wherein the N-intein comprises an amino acid sequence with at least 90% identity to a reference sequence selected from the group consisting of: SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 187, and 193.

12. The engineered phytase of claim 1, wherein the first binding element is a C-coil of the coiled-coil dimerization domain.

13. The engineered phytase of claim 12, wherein the C-coil comprises an amino acid sequence with at least 90% identity to a reference sequence of SEQ ID NOS: 38 or 40.

14. The engineered phytase of claim 1, wherein the second binding element is an N-coil of a coiled-coil dimerization domain.

15. The engineered phytase of claim 14, wherein the N-coil comprises an amino acid sequence with at least 90% identity to a reference sequence of SEQ ID NOS: 37 or 39.

16. The engineered phytase of claim 1, wherein the first binding element is a tag domain.

17. The engineered phytase of claim 16, wherein the tag domain comprises an amino acid sequence with at least 90% identity to a reference sequence of SEQ ID NOS: 33 or 34.

18. The engineered phytase of claim 1, wherein the second binding element is a catcher domain.

19. The engineered phytase of claim 18, wherein the catcher domain comprises an amino acid sequence with at least 90% identity to a reference sequence of SEQ ID NOS: 35 or 36.

20. The engineered phytase of claim 1 further comprising a first linker, wherein the first linker is contiguous with and between the first binding element and the target phytase.

21. The engineered phytase of claim 20, wherein the first linker comprises a sequence with at least 90% identity to a sequence selected from the group consisting of: SEQ ID NOS: 41, 43, 45, 47, 48, 50, and 51.

22. The engineered phytase of claim 1 further comprising a second linker, wherein the second linker is contiguous with and between the target phytase and the second binding element.

23. The engineered phytase of claim 22, wherein the second linker comprises a sequence with at least 90% identity to a sequence selected from the group consisting of: SEQ ID NOS: 42, 44, 46, 49, 50, and 51.

24. The engineered phytase of claim 1 comprising an amino acid sequence having at least 90% identity to a reference sequence selected from the group consisting of: SEQ ID NOS: 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 201, 203, 205, and 207.

25. The engineered phytase of claim 1, wherein the phytase activity is stable at a temperature in a range from 70°C to 90°C.

26. The engineered phytase of claim 1 included in a host, wherein the host is selected form the group consisting of a microorganism, a plant cell, a phage, a virus, a mammalian cell, and an insect cell.

27. An engineered nucleic acid encoding the engineered phytase of any one of claims 1 - 26.

28. An engineered nucleic acid encoding an engineered phytase comprising a target phytase, a first binding element and a second binding element, wherein each of the first binding element and the second binding is fused to the target phytase, the first binding element interacts with the second binding element to cause cyclization of the engineered phytase, and enhance thermal stability of the target phytase, and wherein the first binding element or the second binding element is selected from the group consisting of: a tag

domain, a catcher domain, an intein or part thereof, and a coiled-coil dimerization domain or part thereof.

29. The engineered nucleic acid of claim 28, wherein upon the interaction, each of the first binding element and the second binding element is capable of being released from the engineered phytase spontaneously.

30. The engineered nucleic acid of claim 28 comprising a sequence encoding the target phytase selected from the group consisting of phytases derived from *Escherichia coli, Aspergillus niger, Peniophora lycii, Neurospora crassa,* and *Schwaniomyces accidentalis.*

31. The engineered nucleic acid of 28 comprising a sequence encoding the target phytase and having at least 90% identity to a reference sequence selected from the group consisting of SEQ ID NOS: 52, 55, 185, and 219.

32. The engineered nucleic acid of claim 28, wherein the first binding element is fused to the N-terminus of the target phytase.

33. The engineered nucleic acid of claim 28, wherein the second binding element is fused to the C-terminus of the target phytase.

34. The engineered nucleic acid of claim 28 comprising the sequence encoding the first binding element, wherein the first binding element is a C-intein of an intein.

35. The engineered nucleic acid of claim 34 comprising the sequence with at least 90% identity to a reference sequence selected from the group consisting of: SEQ ID NOS: 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 188, 190, and 194.

36. The engineered nucleic acid of claim 28 comprising the sequence encoding the second binding element, wherein the second binding element is an N-intein of an intein.

37. The engineered nucleic acid of claim 36 comprising the sequence with at least 90% identity to a reference sequence selected from the group consisting of: SEQ ID NOS: 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 186, and 192.

38. The engineered nucleic acid of claim 28 comprising the sequence encoding the first binding element, wherein the first binding element is a C-coil of the coiled-coil dimerization domain.

39. The engineered nucleic acid of claim 38 comprising the sequence with at least 90% identity to a reference sequence of SEQ ID NOS: 179 or 181.

40. The engineered nucleic acid of claim 28 comprising the sequence encoding the second binding element, wherein the second binding element is an N-coil of a coiled-coil dimerization domain.

41. The engineered nucleic acid of claim 40 comprising the sequence with at least 90% identity to a reference sequence of SEQ ID NOS: 178 or 180.

42. The engineered nucleic acid of claim 28 comprising the sequence encoding the first binding element, wherein the first binding element is a tag domain.

43. The engineered nucleic acid of claim 42 comprising a sequence with at least 90% identity to a reference sequence of SEQ ID NOS: 174 or 175.

44. The engineered nucleic acid of claim 28 comprising the sequence encoding the second binding element, wherein the second binding element is a catcher domain.

45. The engineered nucleic acid of claim 44 comprising a sequence with at least 90% identity to a reference sequence of SEQ ID NOS: 176 or 177.

46. The engineered nucleic acid of claim 28 further comprising a sequence encoding a first linker, wherein the first linker is contiguous with and between the first binding element and the target phytase.

47. The engineered nucleic acid of claim 46 comprising a sequence with at least 90% identity to a sequence selected from the group consisting of: SEQ ID NOS: 120, 122, 124, 126, 182, 183, and 184.

48. The engineered nucleic acid of claim 28 further comprising a sequence encoding a second linker, wherein the second linker is contiguous with and between the target phytase and the second binding element.

49. The engineered nucleic acid of claim 48 comprising a sequence with at least 90% identity to a sequence selected from the group consisting of: SEQ ID NOS: 121, 123, 125, 127, 183 and 184.

50. The engineered nucleic acid of claim 28 comprising a sequence having at least 90% identity to a reference sequence selected from the group consisting of: SEQ ID NOS: 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 128 - 133, 200, 202, 204 and 206.

51. The engineered nucleic acid of claim 28 comprising a sequence encoding the engineered phytase having stable phytase activity at a temperature in a range from 70°C to 90°C.

52. The engineered nucleic acid of claim 28 expressed in the host, wherein the host is selected form the group consisting of a microorganism, a plant cell, a phage, a virus, a mammalian cell, and an insect cell.

53. A vector comprising the engineered nucleic acid of any one of claims 28 - 52.

54. A host comprising the engineered phytase of any one of claims 1 - 26, wherein the host is selected from the group consisting of: a microorganism, a plant cell, a phage, a virus, a mammalian cell, and an insect cell.

55. A method of enhancing thermal stability of a target phytase comprising producing the engineered phytase of any one of claims 1 - 26.

56. An animal feed comprising an engineered phytase comprising a target phytase, a first binding element and a second binding element, wherein each of the first binding element and the second binding is fused to the target phytase, the first binding element interacts with the second binding element to cause cyclization of the engineered phytase, and enhance thermal stability of the target phytase, and wherein the first binding element or the second binding element is selected from the group consisting of: a tag domain, a catcher domain, an intein or part thereof, and a coiled-coil dimerization domain or part thereof.

57. An animal feed comprising an engineered phytase of the one of claims 1 - 26.

58. A method of preparing an animal feed comprising adding the engineered phytase of any one of claims 1 - 26 to the animal feed.

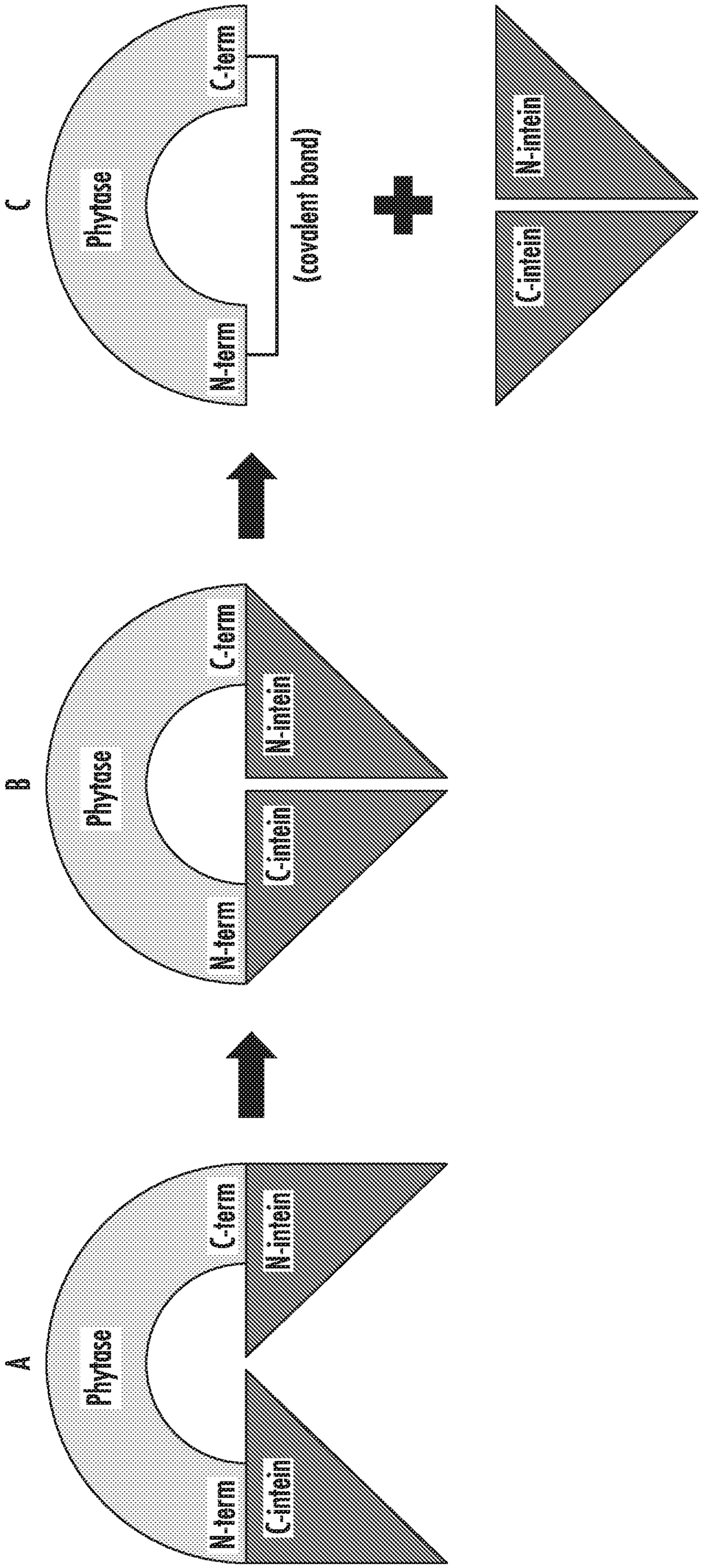
A
Phytase
C-term
N-term
N-intein
C-intein
(covalent bond)
B
Phytase
C-term
N-term
N-intein
C-intein

FIG. 1

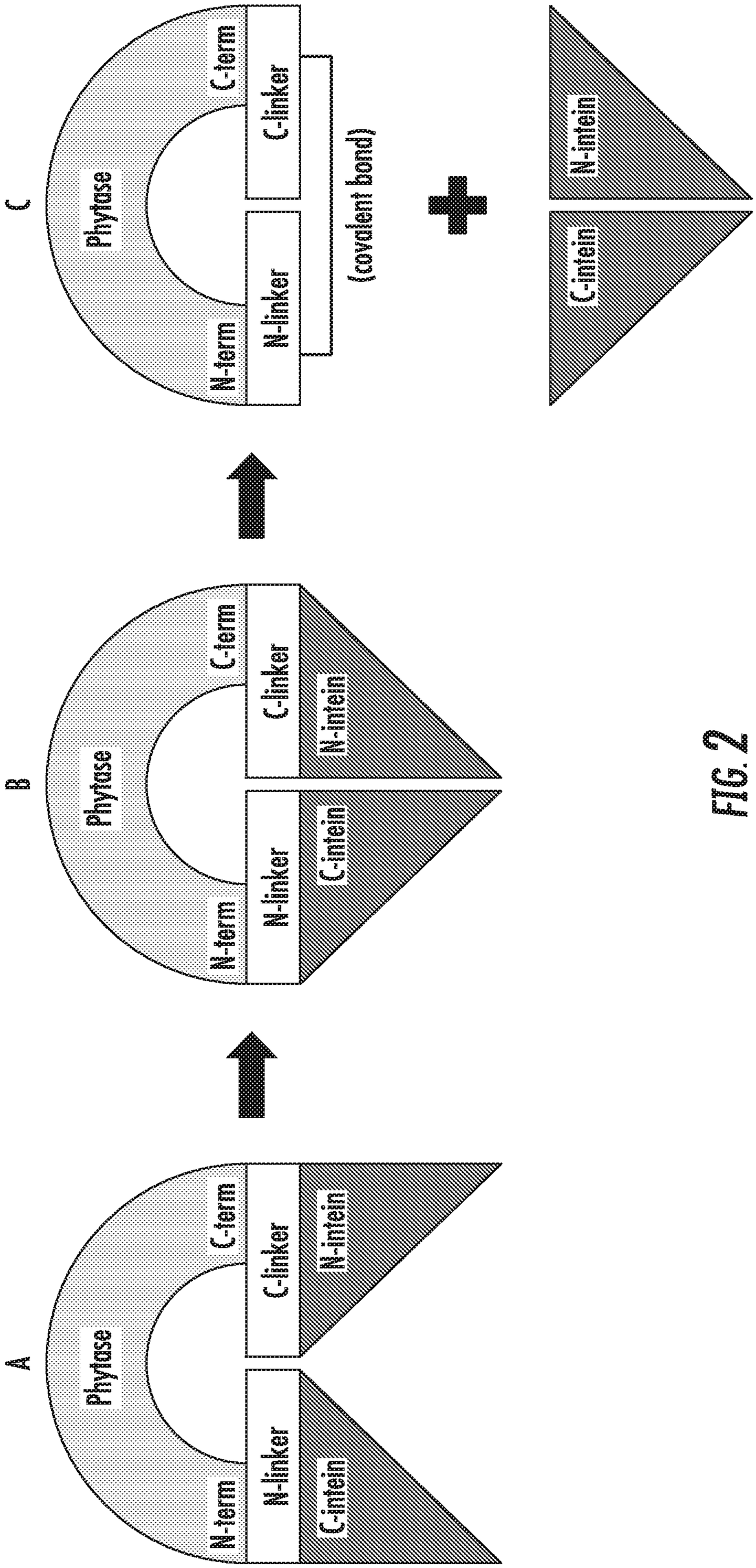
A
Phytase
C-term
C-linker
N-linker
N-term
N-interm
C-interm
(covalent bond)
B
Phytase
C-term
C-linker
N-linker
N-term
N-interm
C-interm

FIG. 2

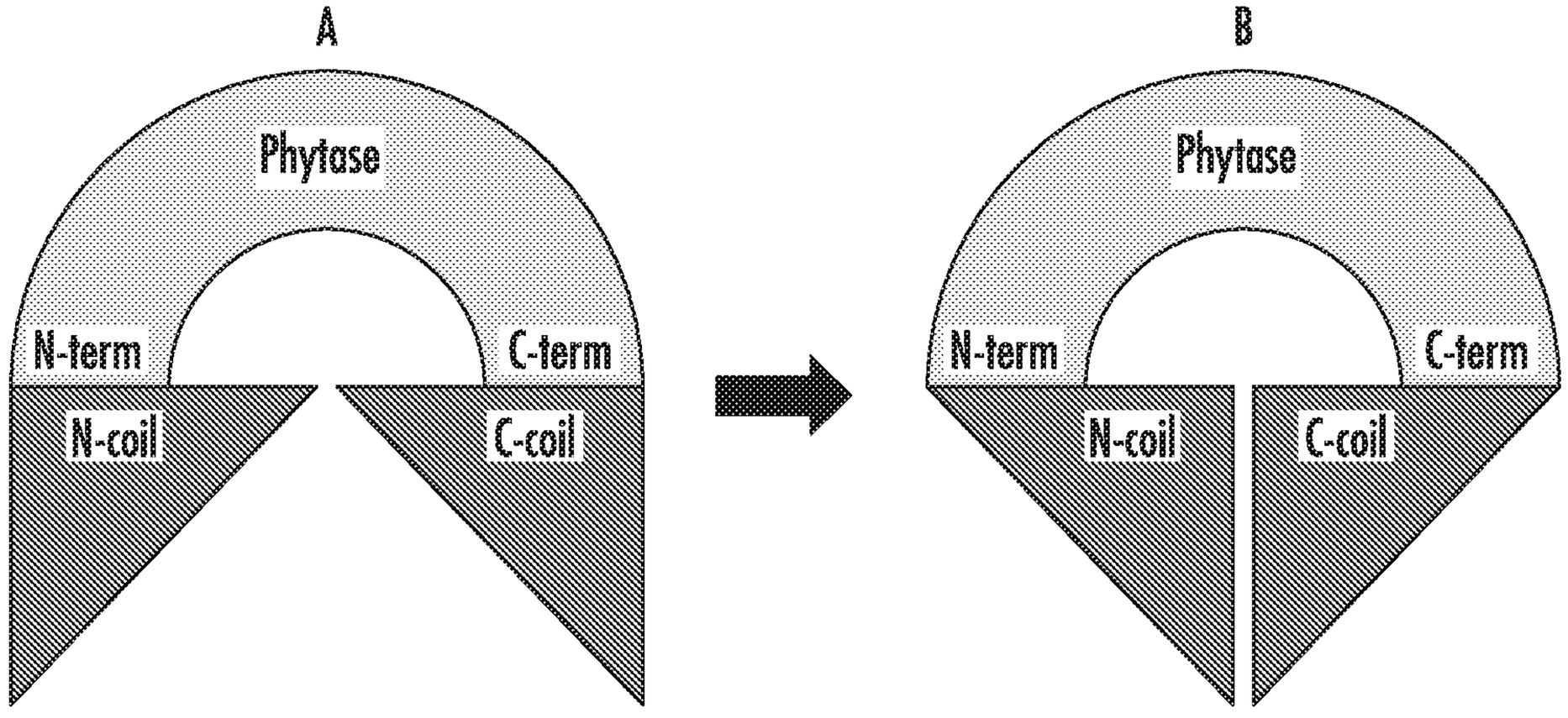
A
Phytase
N-term
C-term
N-coil
C-coil
B
Phytase
N-term
C-term
N-coil
C-coil

FIG. 3

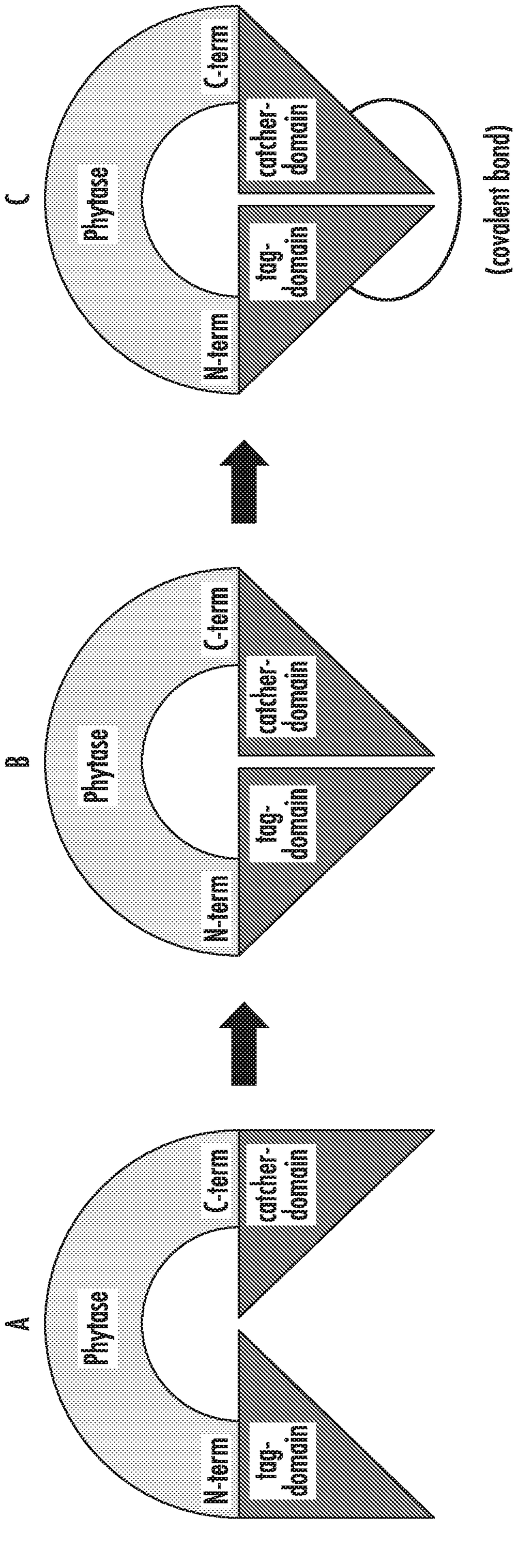
A
Phytase
C-term
catcher-domain
N-term
tag-domain
B
Phytase
C-term
catcher-domain
N-term
tag-domain
C
Phytase
C-term
catcher-domain
N-term
tag-domain
(covalent bond)

FIG. 4

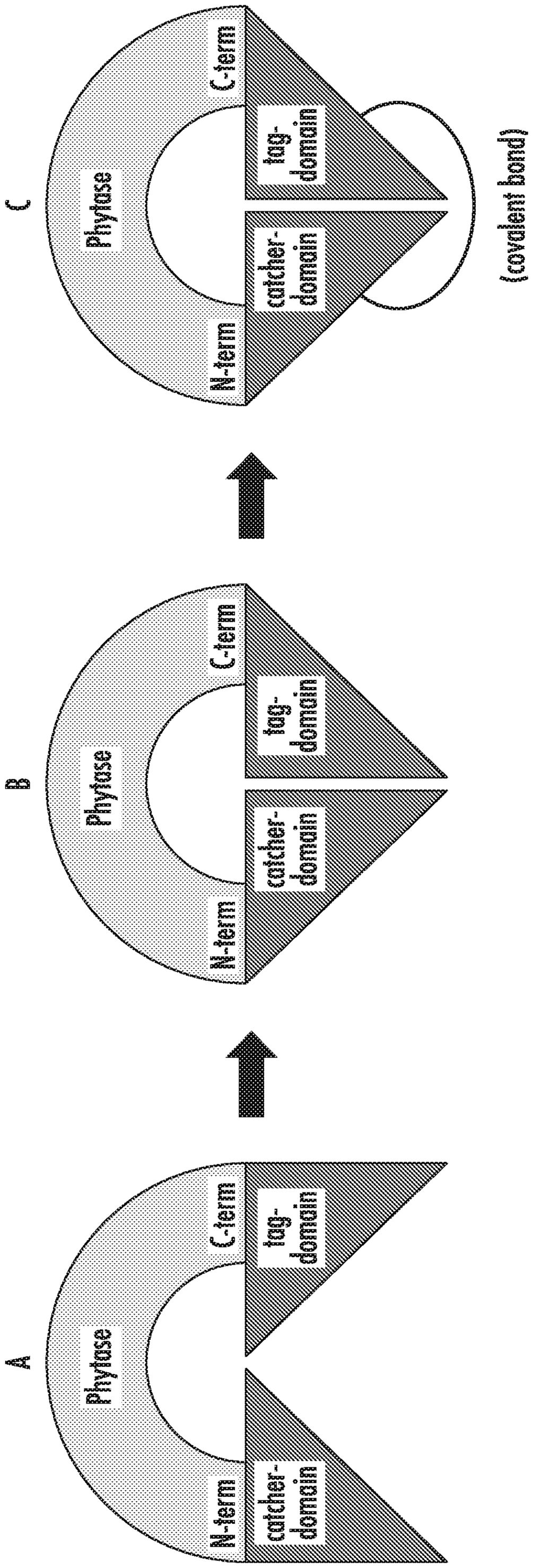
C
Phytase
C-term
tag-domain
catcher-domain
(covalent bond)
N-term
B
Phytase
C-term
tag-domain
catcher-domain
N-term
A
Phytase
C-term
tag-domain
N-term
catcher-domain

FIG. 5

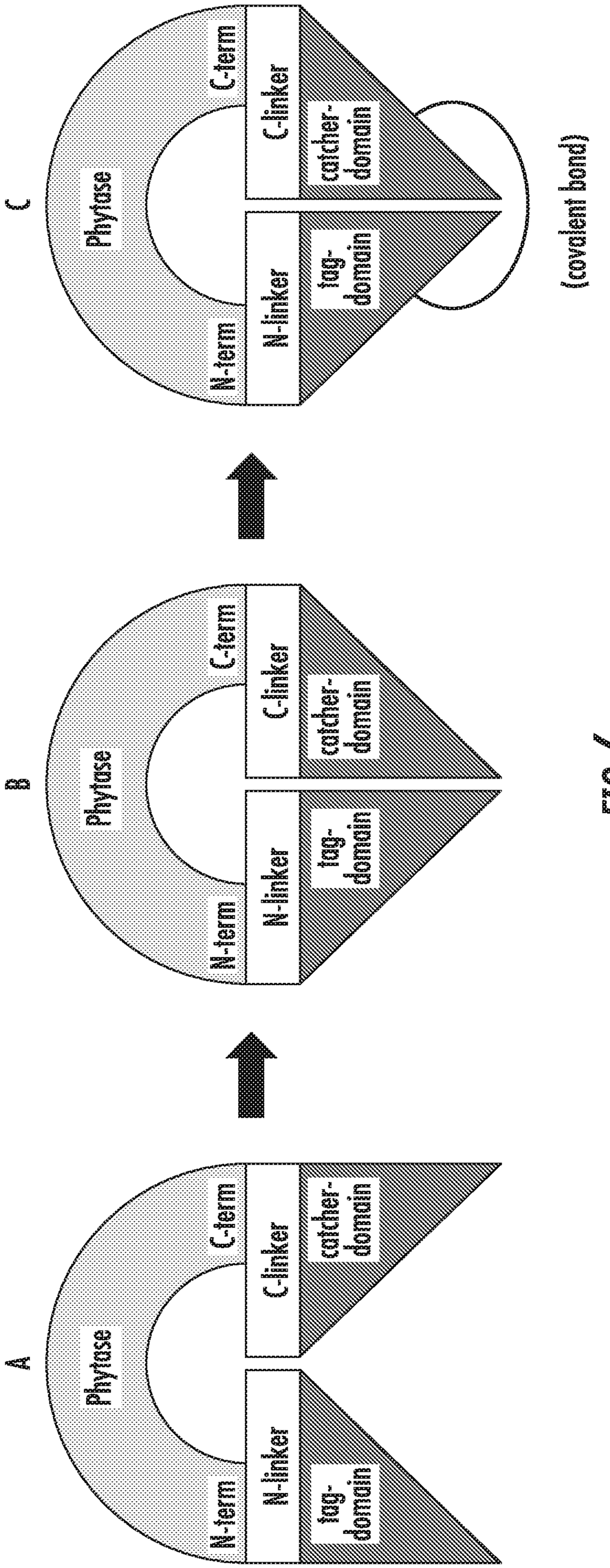
A
Phytase
N-term
C-term
N-linker
C-linker
tag-domain
catcher-domain
(covalent bond)
B
C

FIG. 6

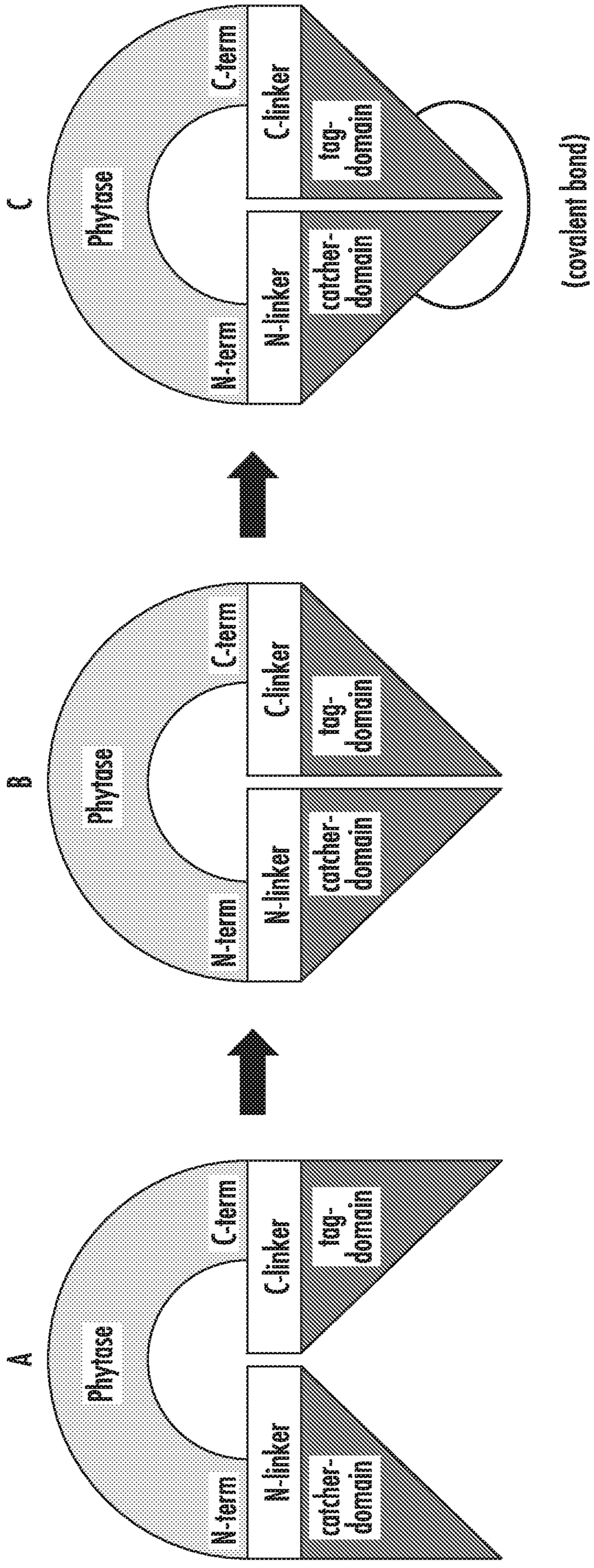
C
Phytase
C-term
C-linker
tag-domain
N-term
N-linker
catcher-domain
(covalent bond)
B
Phytase
C-term
C-linker
tag-domain
N-term
N-linker
catcher-domain
A
Phytase
C-term
C-linker
tag-domain
N-term
N-linker
catcher-domain

FIG. 7

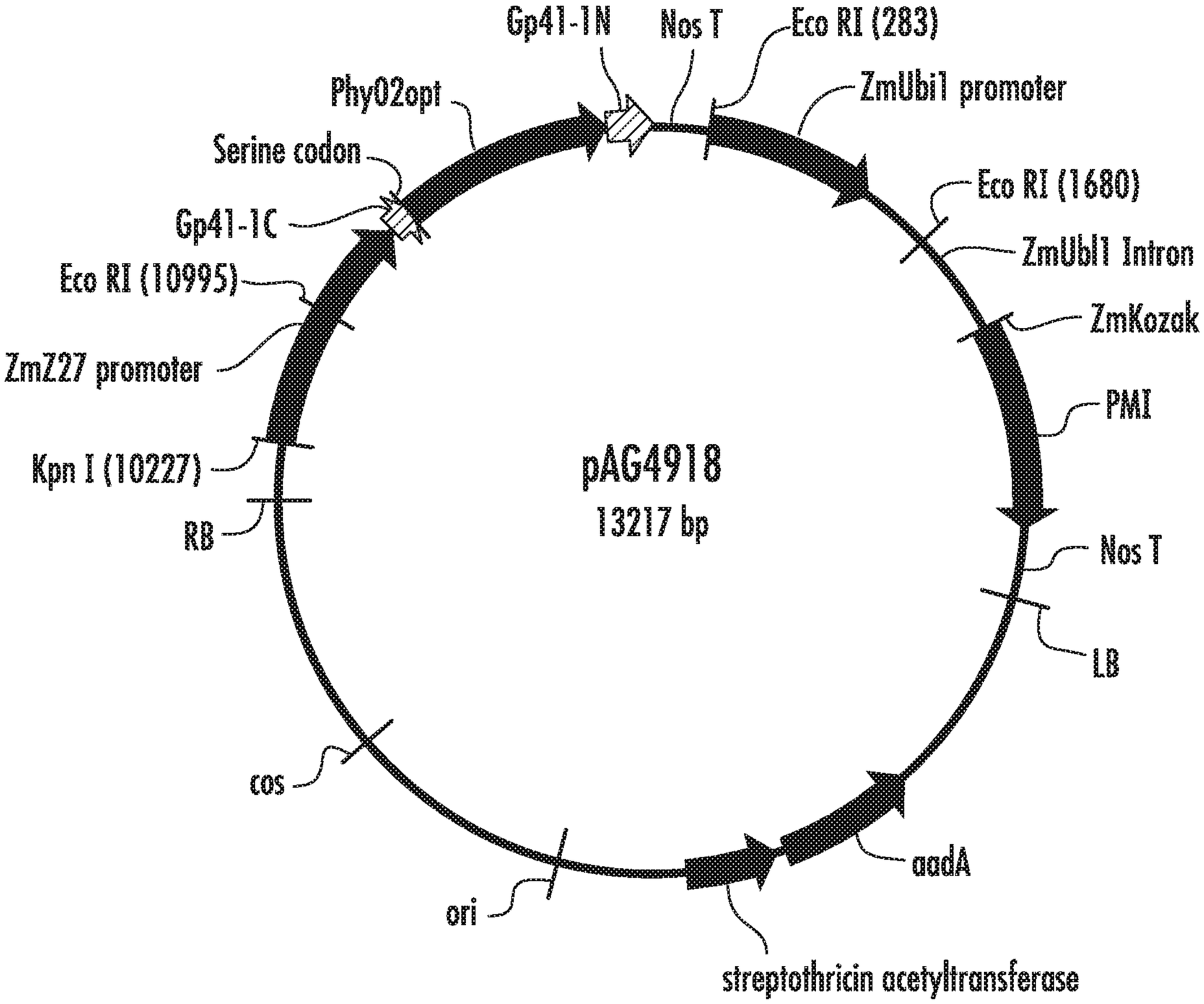
Gp41-1N
Nos T
Eco RI (283)
ZmUbi1 promoter
Eco RI (1680)
ZmUbi1 Intron
ZmKozak
PMI
Nos T
LB
aadA
streptothricin acetyltransferase
ori
cos
Kpn I (10227)
RB
ZmZ27 promoter
Eco RI (10995)
Gp41-1C
Serine codon
PhyO2opt
pAG4918
13217 bp

FIG. 8

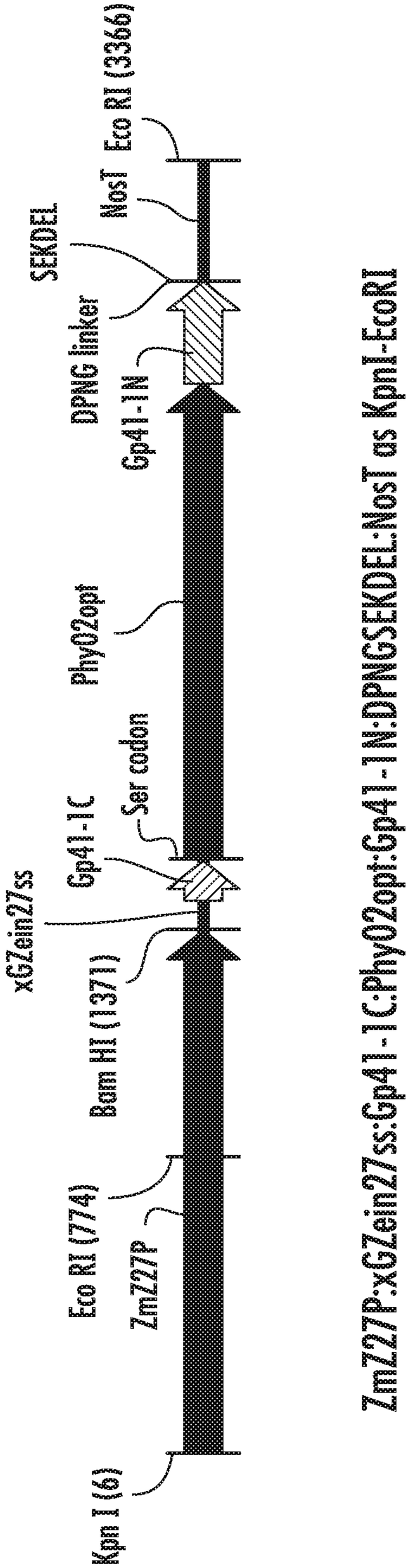
Kpn I (6)
Zm27P
Eco RI (774)
xGZein27ss
Bam HI (1371)
Ser codon
Gp41-1C
DPNG linker
SEKDEL
NosT
Eco RI (3366)
Phy02opt
Gp41-1N
Zm27P:xGZein27ss:Gp41-1C:Phy02opt:Gp41-1N:DPNGSEKDEL:NosT as KpnI-EcoRI
3370 bp

FIG. 9A

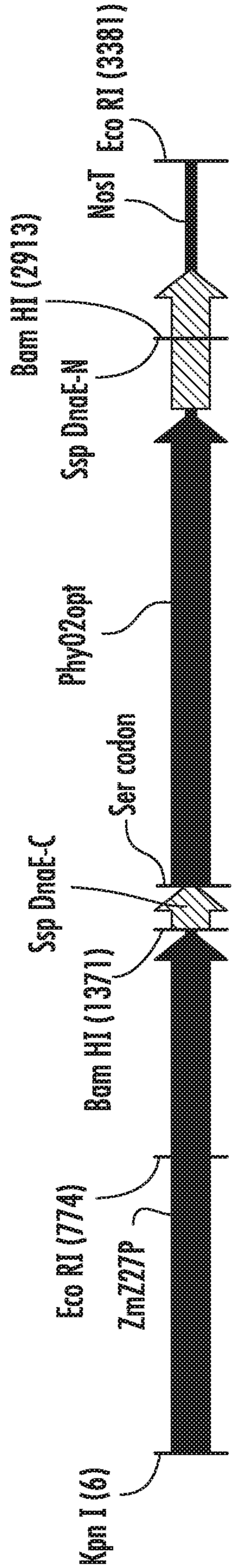
Kpn I (6)
Eco RI (774)
Zm27P
Bam HI (1371)
Ssp DnaE-C
Ser codon
Ssp DnaE-N
Bam HI (2913)
Phy02opt
Most Eco RI (3381)
Zm27P:Ssp DnaE-C:Phy02opt:Ssp DnaE-N:Most as Kpn-EcoRI
3385 bp

FIG. 9B

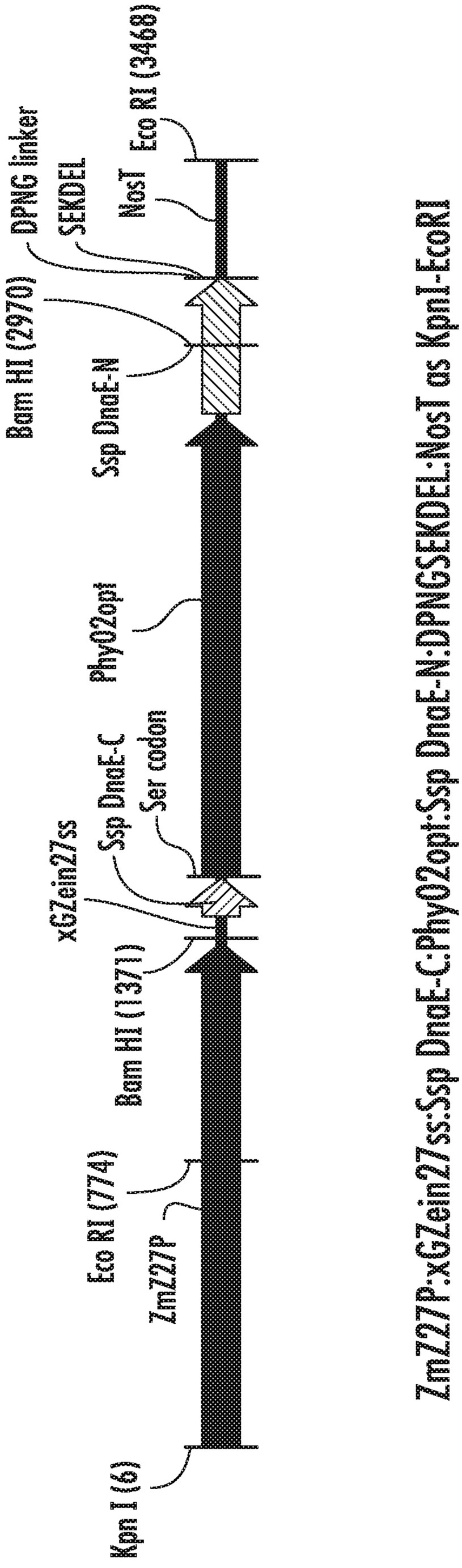
Kpn I (6)
Eco RI (774)
Zm27P
Bam HI (1371)
xGZein27ss
Ssp DnaE-C
Ser codon
Phy02opt
Ssp DnaE-N
Bam HI (2970)
DPNG linker
SEKDEL
Nost
Eco RI (3468)
Zm27P:xGZein27ss:Ssp DnaE-C:Phy02opt:Ssp DnaE-N:DPNGSEKDEL:Nost ds Kpn-EcoRI
3472 bp

FIG. 9C

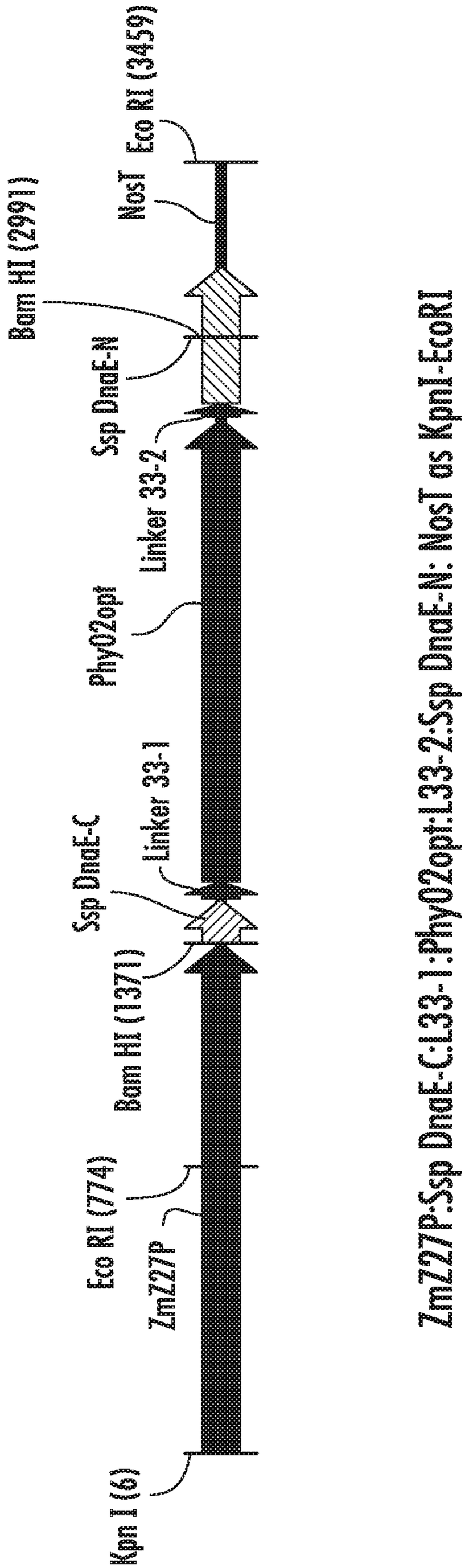
ZmZ27P:Ssp DnaE-C::L33-1::Phy02opt::L33-2:Ssp DnaE-N: NosT as KpnI-EcoRI
3463 bp
KpnI (6)
ZmZ27P
EcoRI (774)
Ssp DnaE-C
BamHI (1371)
Linker 33-1
Phy02opt
Linker 33-2
Ssp DnaE-N
BamHI (2971)
NosT EcoRI (3459)

FIG. 10A

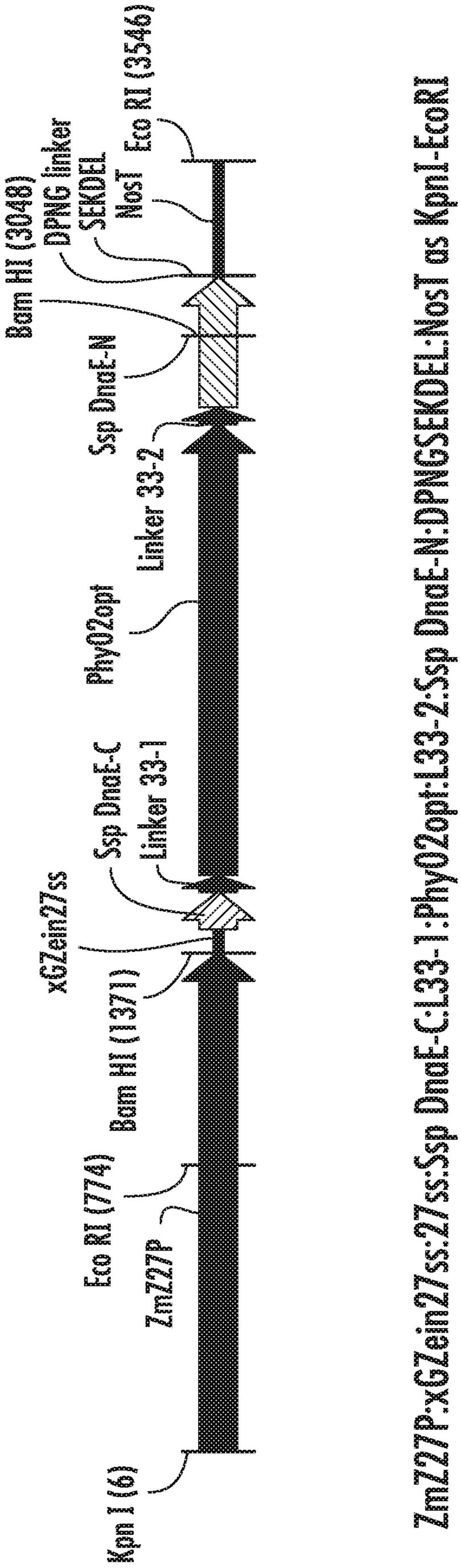
Kpn I (67)
Eco RI (774)
Zm27P
Bam HI (1371)
xGZein27ss
Ssp DnaE-C
Linker 33-1
Phy02opt
Linker 33-2
Ssp DnaE-N
Bam HI (3048)
DPNG linker
SEKDEL
NosT Eco RI (3546)
Zm27P:xGZein27ss:Ssp DnaE-C:L33-1:Phy02opt:L33-2:Ssp DnaE-N:DPNGSEKDEL:NosT Kpn I-EcoRI
3550 bp

FIG. 10B

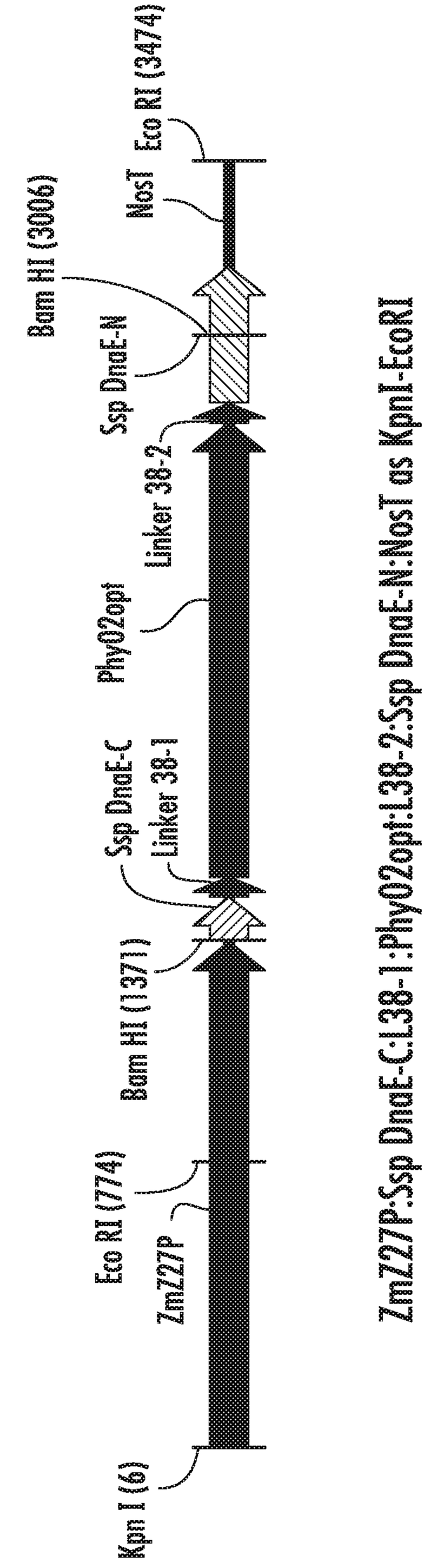
Kpn I (6)
Eco RI (774)
ZmZ27P
Phy02opt
Bam HI (1371)
Ssp DnaE-C
Linker 38-1
Ssp DnaE-N
Linker 38-2
NosT
Eco RI (3474)
Bam HI (3006)
ZmZ27P:Ssp DnaE-C:L38-1:Phy02opt:L38-2:Ssp DnaE-N:NosT as KpnI-EcoRI
3478 bp

FIG. 10C

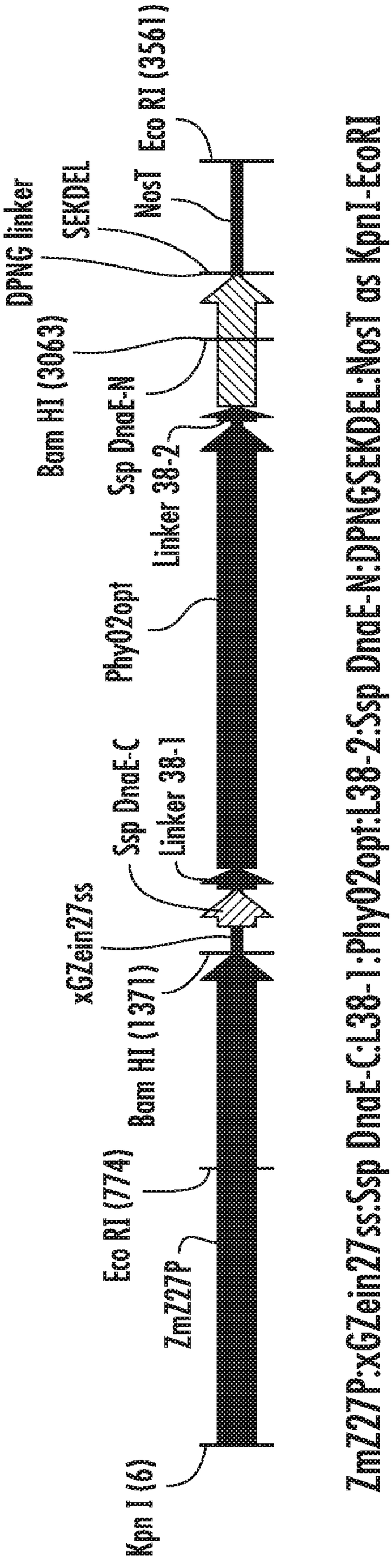
Kpn I (6)
Eco RI (774)
ZmZ27P
Bam HI (1371)
xGZein27ss
Linker 38-1
Ssp DmaE-C
Phy02opt
Linker 38-2
Ssp DmaE-N
Bam HI (3063)
DPNG linker
SEKDEL
Nost Eco RI (3561)
ZmZ27P:xGZein27ss:Ssp DmaE-C:L38-1:Phy02opt:L38-2:Ssp DmaE-N:DPNGSEKDEL:Nost Eco RI
3565 bp

FIG. 10D

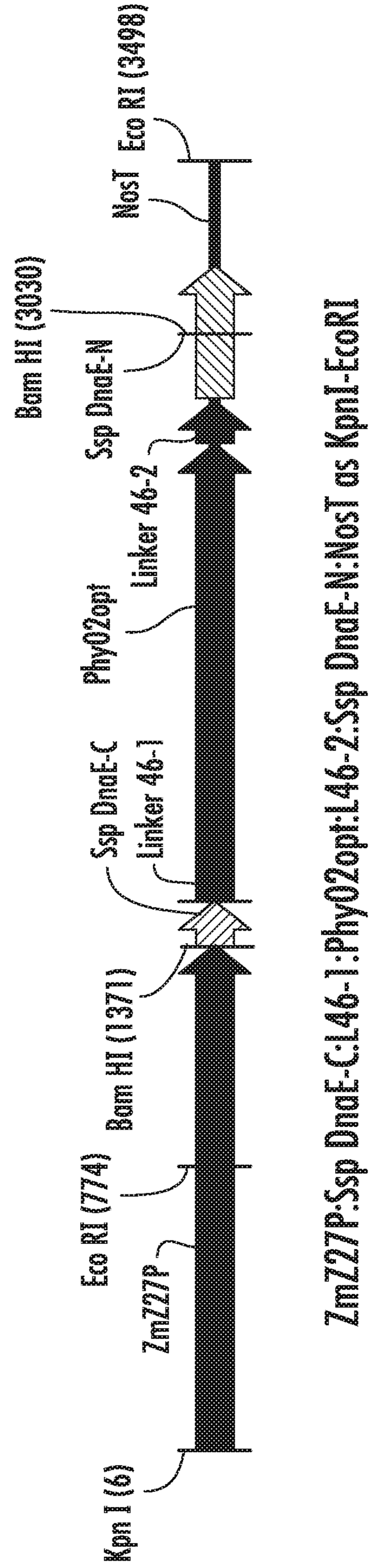
Eco RI (3498)
NosT
Bam HI (3030)
Ssp DnaE-N
Linker 46-2
Phy02opt
Ssp DnaE-C
Linker 46-1
Bam HI (1371)
Eco RI (774)
Zm27P
Kpn I (6)
Zm27P:Ssp DnaE-C:L46-1:Phy02opt:L46-2:Ssp DnaE-N:NosT ds Kpn I-Eco RI
3502 bp

FIG. 10E

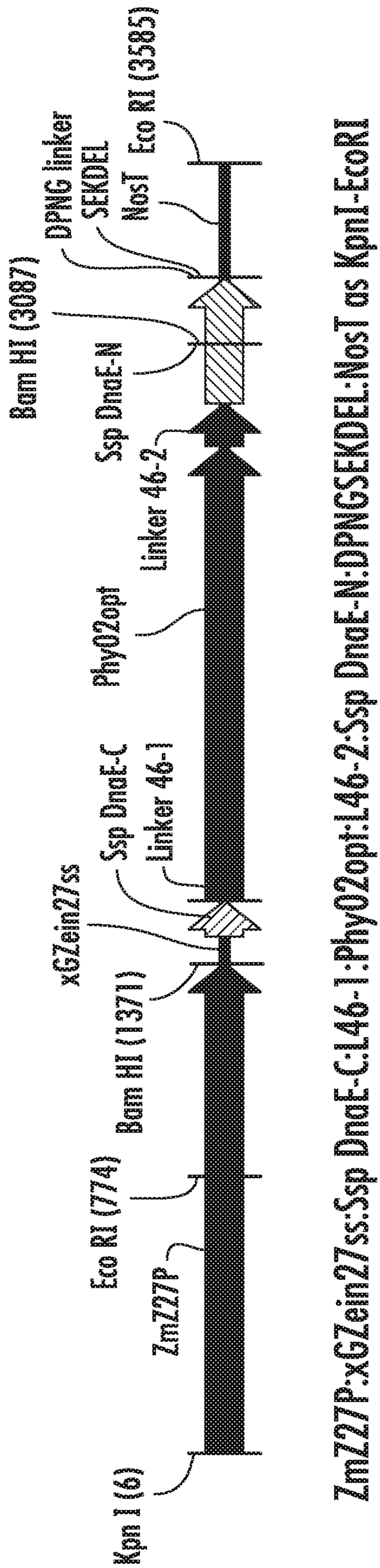
ZmZ27P:xGZein27ss:Ssp DnaE-C:L46-1:Phy02opt:L46-2:Ssp DnaE-N:DPNGSEKDEL:Nost ns KpnI-EcoRI
3589 bp
ZmZ27P
Eco RI (774)
Bam HI (1371)
xGZein27ss
Ssp DnaE-C
Linker 46-1
Kpn I (6)
Phy02opt
Linker 46-2
Ssp DnaE-N
Bam HI (3087)
DPNG
SEKDEL
Nost Eco RI (3585)

FIG. 10F

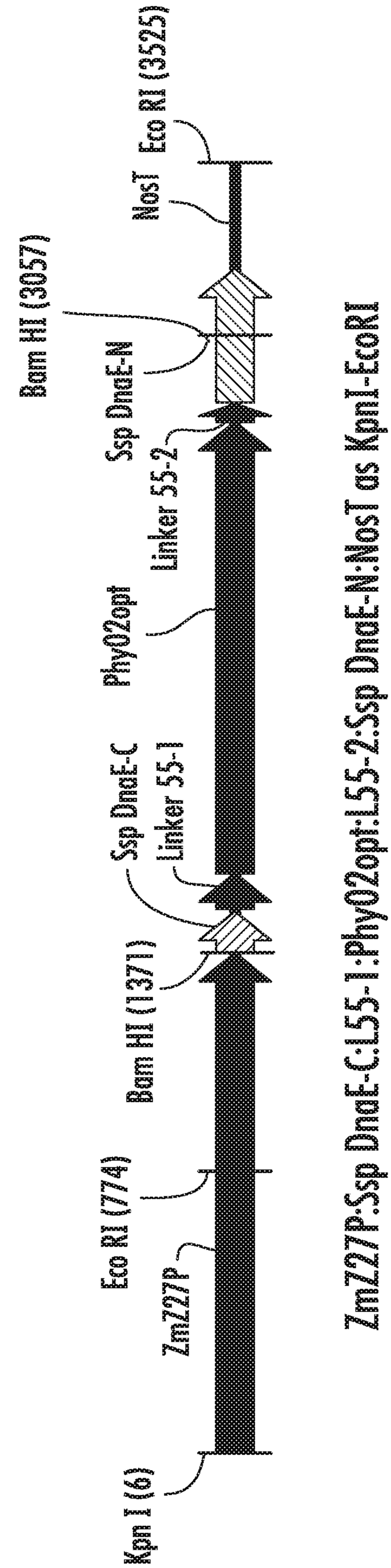
Xba I (6057)
Eco RI (3525)
NosT
Phy02opt
Linker 55-2
Ssp DnaE-N
Bam HI (3057)
Eco RI (774)
Zm727P
Bam HI (1371)
Ssp DnaE-C
Linker 55-1
Kpn I (6)
Zm727P:Ssp DnaE-C:L55-1:Phy02opt:L55-2:Ssp DnaE-N:NosT us KpnI-EcoRI
3529 bp

FIG. 10G

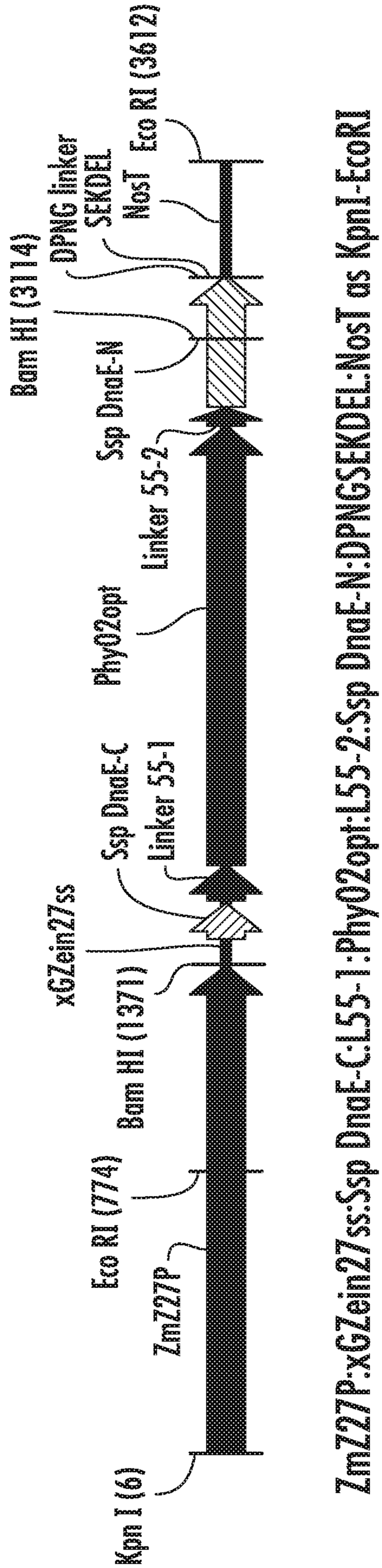
Kpn I (6)
Eco RI (774)
Zm27P
Bam HI (1371)
xGZein27ss
Ssp DnaE-C
Linker 55-1
Phy02opt
linker 55-2
Ssp DnaE-N
Bam HI (3127)
DPNG linker
SEKDEL
NosT Eco RI (3612)
Zm27P:xGZein27ss:Ssp DnaE-C:L55-1:Phy02opt:L55-2:Ssp DnaE-N:DPNGSEKDEL:NosT Eco RI
3616 bp

FIG. 10H

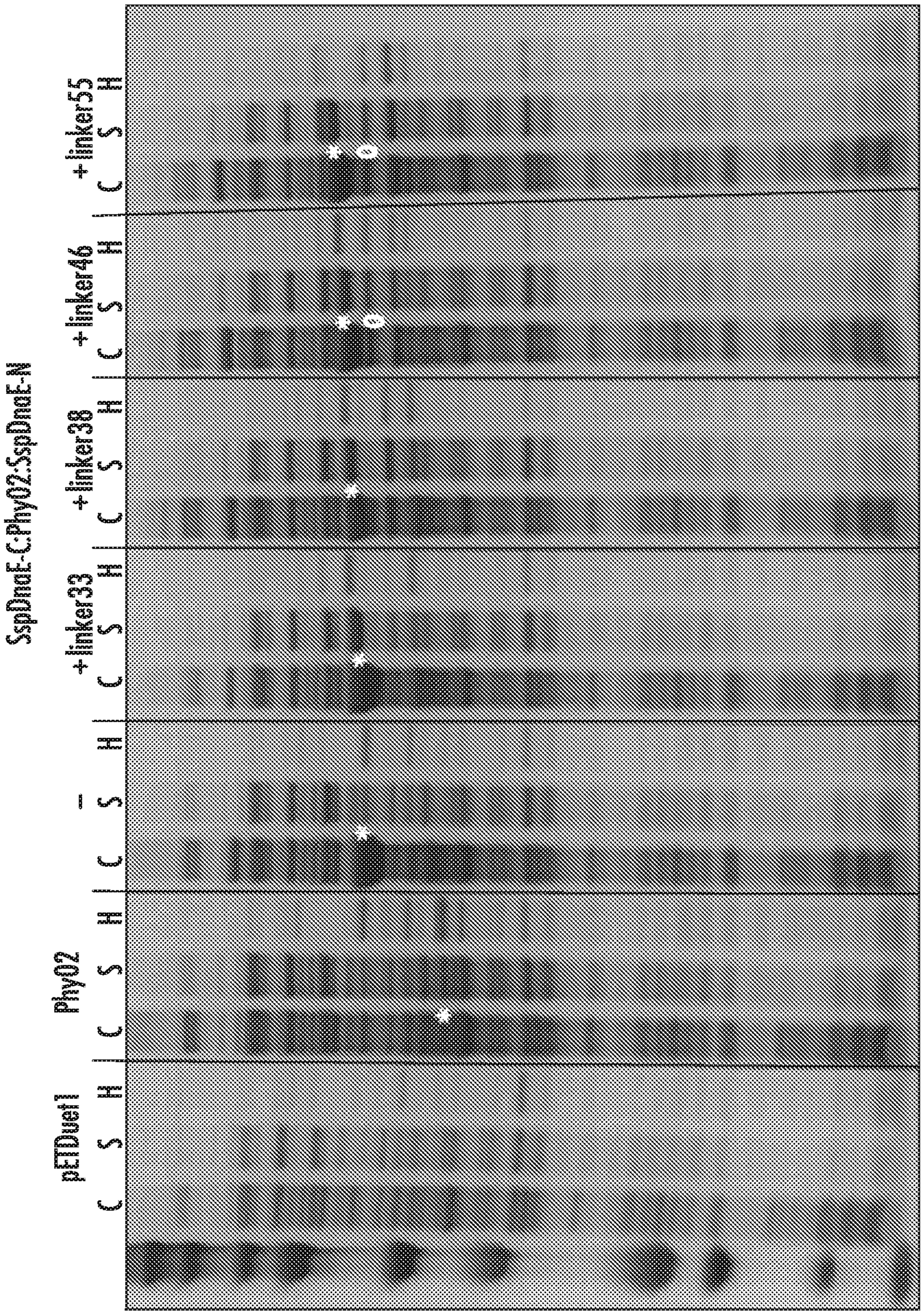
SspDnaE-C:Phy02:SspDnaE-N
pETDuet
C S H
Phy02
C S H
+linker33
C S H
+linker38
C S H
+linker46
C S H
+linker55
C S H
kDa
250
150
100
75
50
37
25
20
15
10

FIG. 17

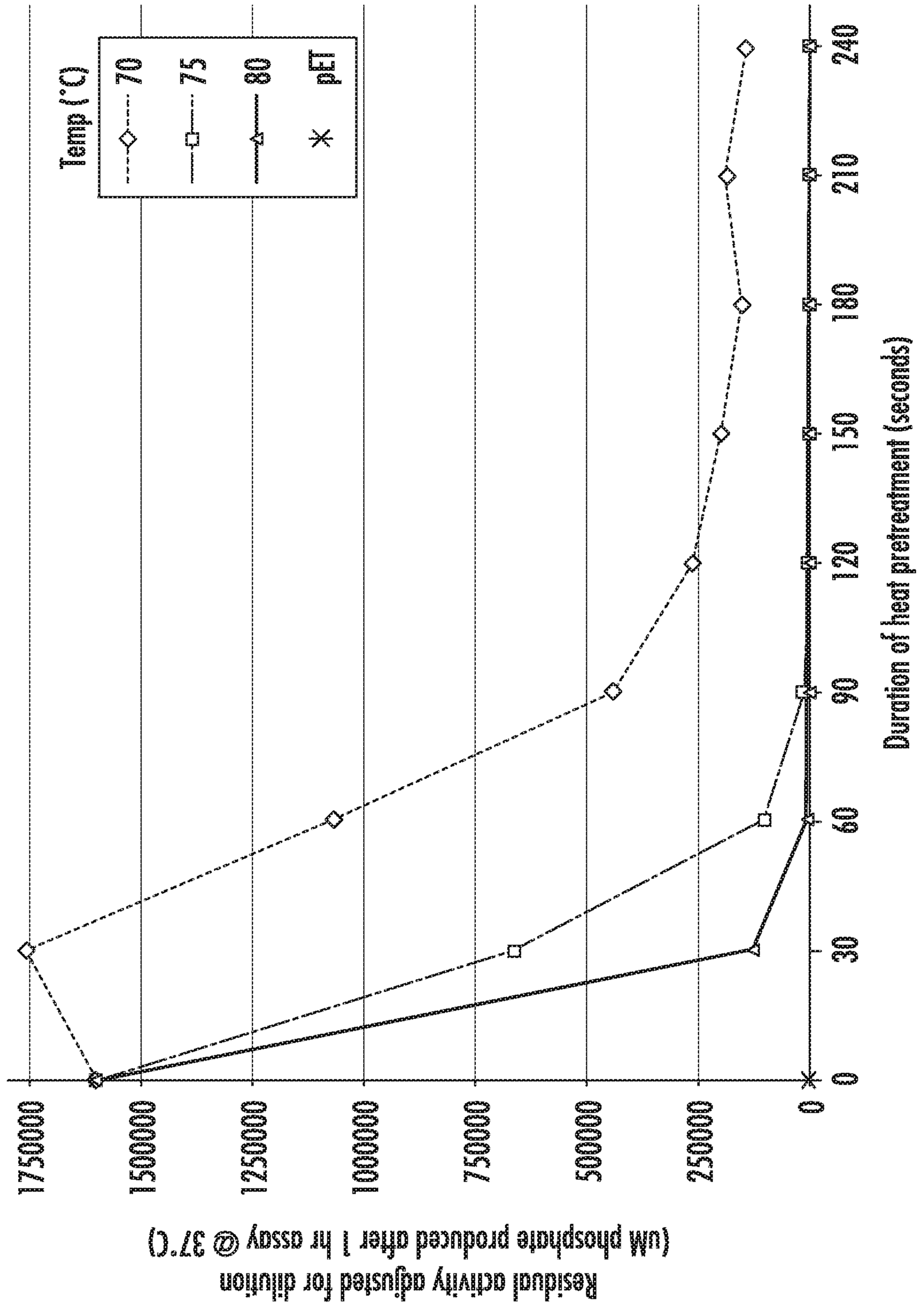
Temp (°C)
70
75
80
pET
Duration of heat pretreatment (seconds)
0
30
60
90
120
150
180
210
240
Residual activity adjusted for dilution
(um phosphate produced after 1 hr assay @ 37°C)
0
250000
500000
750000
1000000
1250000
1500000
1750000

FIG. 12

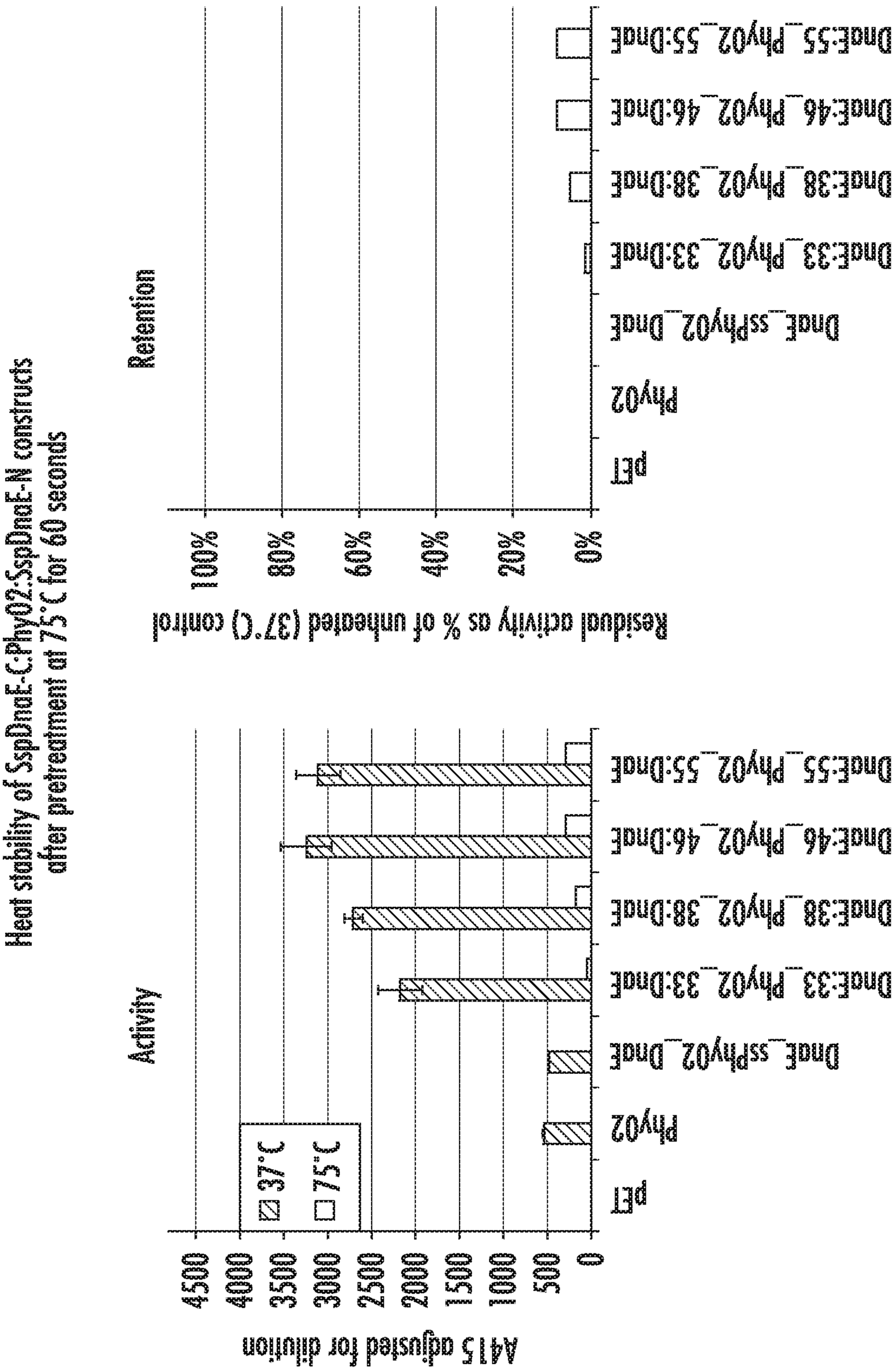

FIG. 13A

FIG. 13B

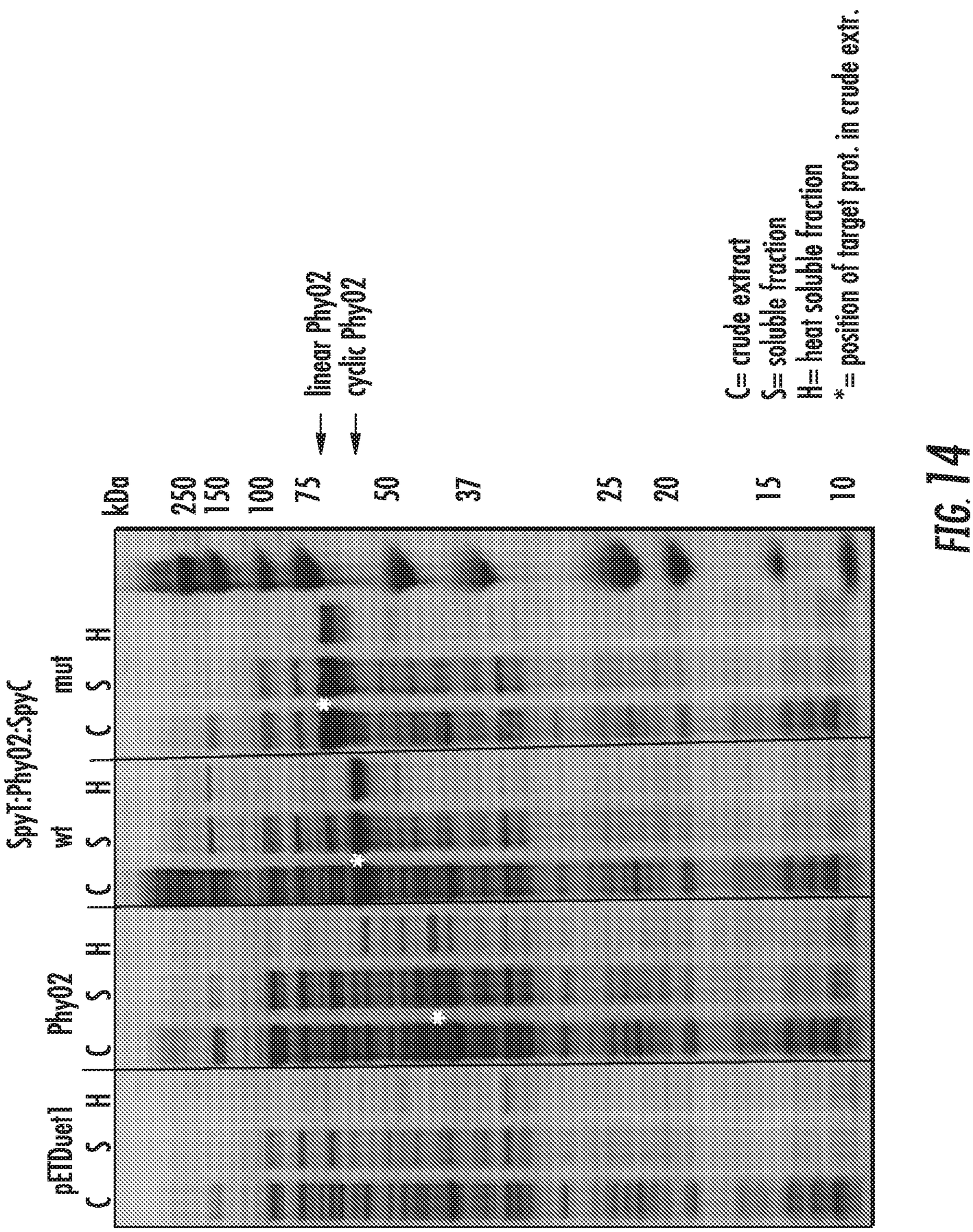
kDa
250
150
100
75
50
37
25
20
15
10
← linear Phy02
← cyclic Phy02
C= crude extract
S= soluble fraction
H= heat soluble f
*= position of tar
pETDuet1
Phy02
SpyT:Phy02:SpyC
wt
mut
C S H C S H C S H

FIG. 14

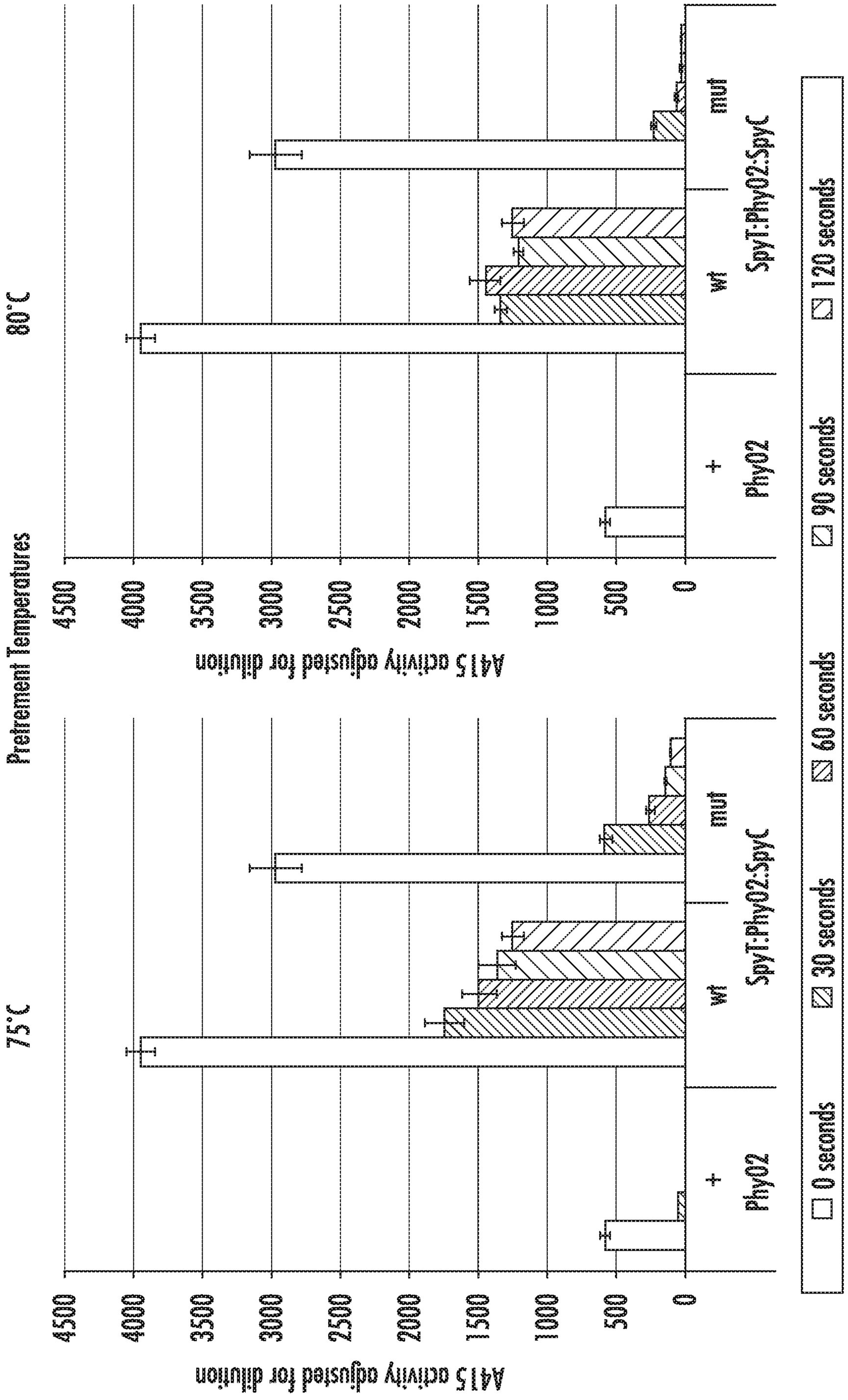
A415 activity adjusted for dilution
4500
4000
3500
3000
2500
2000
1500
1000
500
0
Pretreament Temperatures
80°C
75°C
Phy02
+
wt
mut
SpyT:Phy02:SpyC
90 seconds
120 seconds
30 seconds
60 seconds
0 seconds
Duration of Heat Pretreatment

FIG. 15A

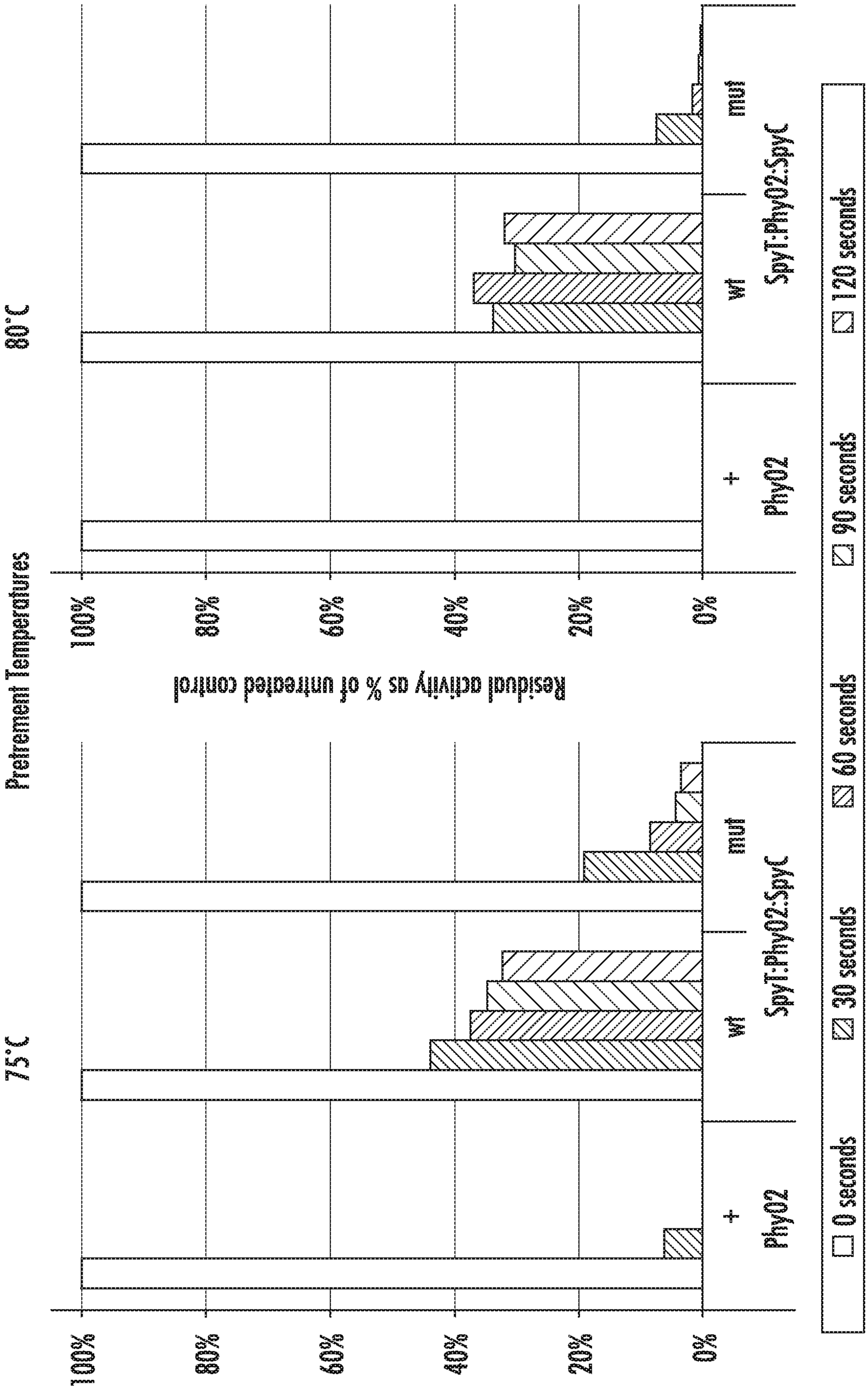
80°C
Pretreament Temperatures
100%
80%
60%
40%
20%
0%
Residual activity as % of untreated control
wt
mut
SpyT:Phy02:SpyC
+
Phy02
90 seconds
120 seconds
75°C
30 seconds
60 seconds
0 seconds
Duration of Heat Pretreatment

FIG. 15B

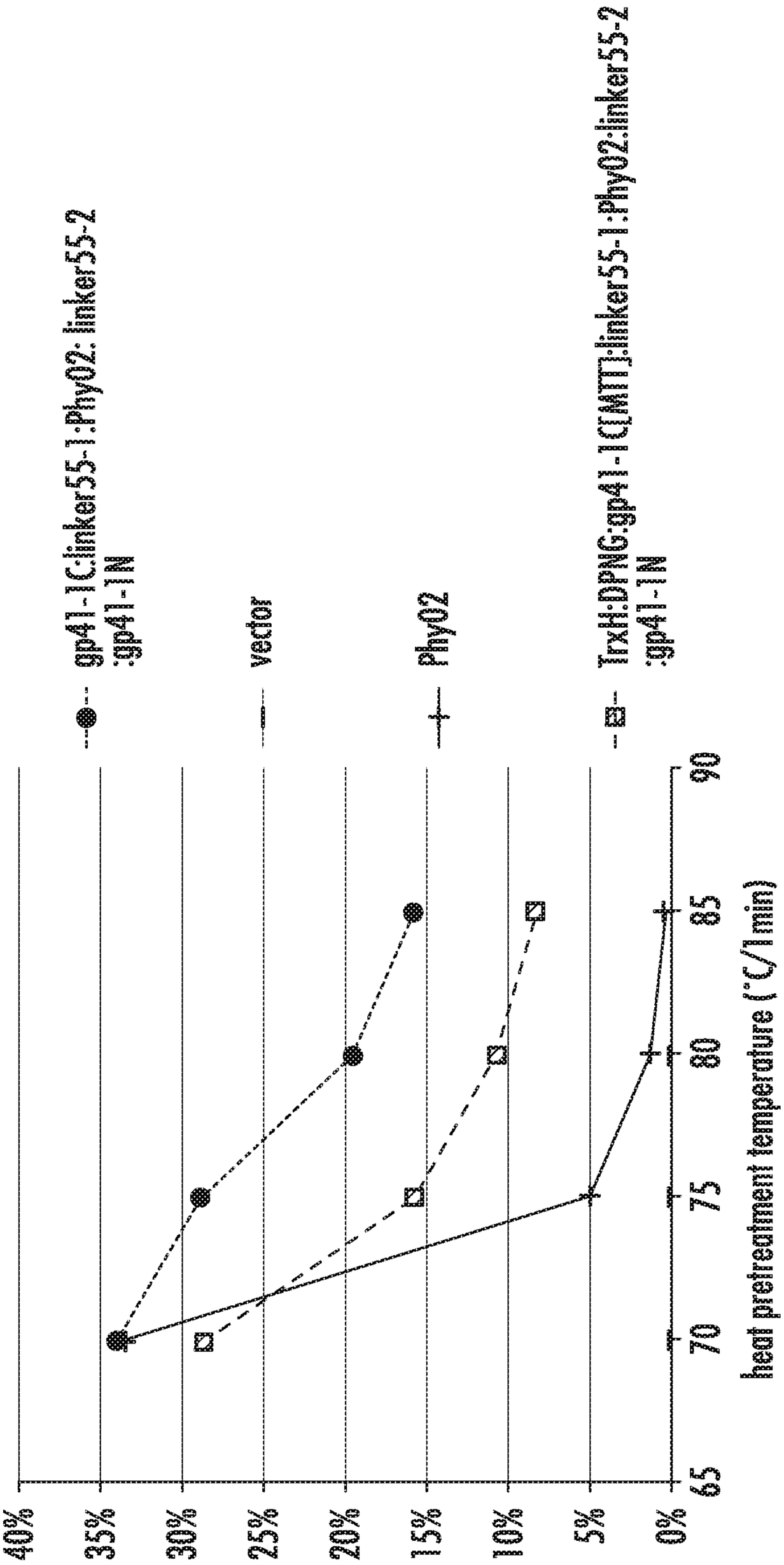
gp41-1C:linker55-1:Phy02:linker55-2
:gp41-1M
vector
Phy02
TrxH:DHFR:gp41-1CM1:linker55-1::Phy02:linker55-2
:gp41-1M
40%
35%
30%
25%
20%
15%
10%
5%
0%
65
70
75
80
85
heat pretreatment temperature (°C/1 min)
retention of activity after heat treatment
(% of unheated control)

FIG. 16

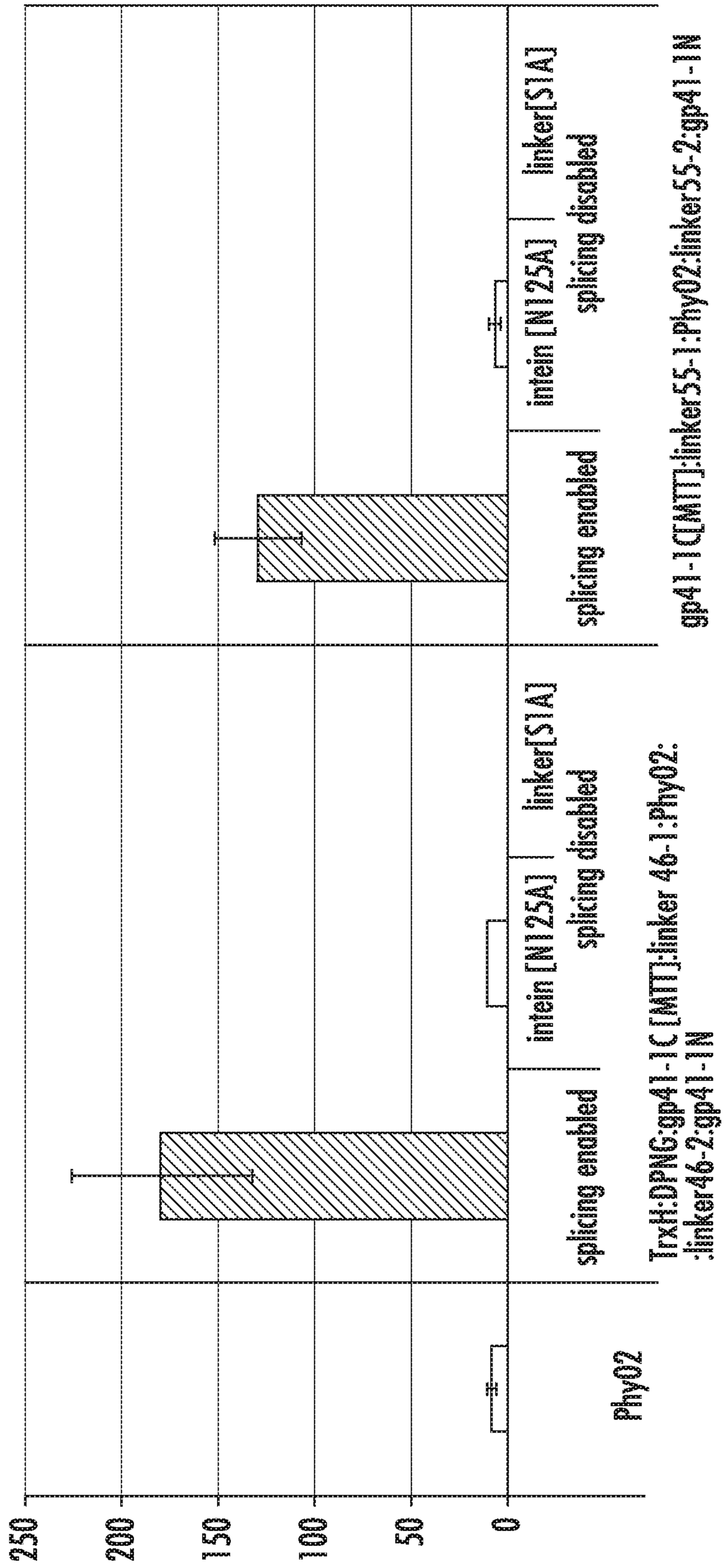
phytase activity (A415)
after 85°C/1min heat pretreatment
250
200
150
100
50
0
Phy02
TrxH:BPNG:gp41-1C EMTP:linker46-1:Phy02:linker46-2:gp41-1M
:linker46-2:gp41-1M
splicing enabled
splicing disabled
intein [NI25A]
TrxH:BPNG:gp41-1C EMTP:linker55-1:Phy02:linker55-2:gp41-1M
:linker55-2:gp41-1M
splicing enabled
splicing disabled
intein [NI25A]

FIG. 17

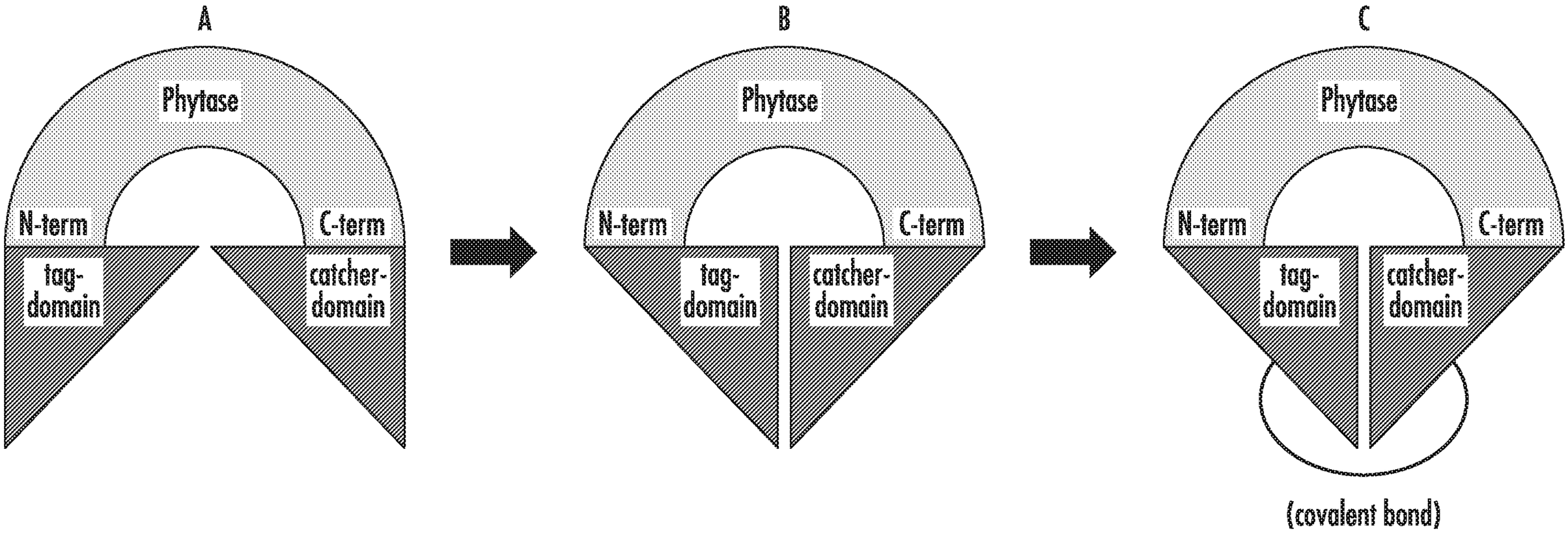
A
Phytase
N-term
C-term
tag-
domain
catcher-
domain
B
Phytase
N-term
C-term
tag-
domain
catcher-
domain
C
Phytase
N-term
C-term
tag-
domain
catcher-
domain
(covalent bond)

FIG. 4