



(86) Date de dépôt PCT/PCT Filing Date: 2008/01/30
(87) Date publication PCT/PCT Publication Date: 2008/08/07
(45) Date de délivrance/Issue Date: 2016/08/30
(85) Entrée phase nationale/National Entry: 2009/07/28
(86) N° demande PCT/PCT Application No.: EP 2008/051137
(87) N° publication PCT/PCT Publication No.: 2008/092901
(30) Priorité/Priority: 2007/01/30 (EP07101395.7)

(51) Cl.Int./Int.Cl. *C12N 9/16* (2006.01),
A01H 5/00 (2006.01), *A23J 3/14* (2006.01),
C12N 15/55 (2006.01), *C12N 15/82* (2006.01),
C12P 19/00 (2006.01), *C12P 7/06* (2006.01)

(72) Inventeurs/Inventors:
SJOEHOLM, CARSTEN, DK;
LASSEN, SOEREN FLENSTED, DK;
SKOV, LARS KOBBEROEE, DK;
DE MARIA, LEONARDO, DK

(73) Propriétaire/Owner:
NOVOZYMES A/S, DK

(74) Agent: NORTON ROSE FULBRIGHT CANADA
LLP/S.E.N.C.R.L., S.R.L.

(54) Titre : POLYPEPTIDES AYANT UNE ACTIVITE PHYTASE ET POLYNUCLEOTIDES CODANT POUR CEUX-CI
(54) Title: POLYPEPTIDES HAVING PHYTASE ACTIVITY AND POLYNUCLEOTIDES ENCODING SAME

(57) **Abrégé/Abstract:**

The present invention relates to polypeptides having phytase activity. These polypeptides have an amino acid sequence which has at least 70% identity to either of three phytases derived from the bacterium *Buttiauxella*, and which comprises at least one of the following amino acids at the position indicated: 119N, 120L, and/or 121E. These phytases have an improved specific activity. Additional specific amino acid substitutions are also disclosed which characterize and distinguish additional phytases of the invention having improved properties such as temperature and/or pH stability, pH activity profile, temperature activity profile, substrate profile, improved performance in animal feed in vitro or in vivo. The invention also relates to isolated polynucleotides encoding the polypeptides, nucleic acid constructs, vectors, and host cells comprising the polynucleotides as well as methods for producing and using the polypeptides.

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

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 August 2008 (07.08.2008)

PCT

(10) International Publication Number
WO 2008/092901 A3

(51) International Patent Classification:

A01H 5/00 (2006.01) *C12N 15/82* (2006.01)
C12N 9/16 (2006.01) *C12N 15/55* (2006.01)

(21) International Application Number:

PCT/EP2008/051137

(22) International Filing Date: 30 January 2008 (30.01.2008)**(25) Filing Language:**

English

(26) Publication Language:

English

(30) Priority Data:

07101395.7 30 January 2007 (30.01.2007) EP

(71) Applicant (for all designated States except US):
NOVOZYMES A/S [DK/DK]; Krogshoejvej 36,
DK-2880 Bagsvaerd (DK).**(72) Inventors:** **SJOEHOLM, Carsten**; Alleroedvej 17,
DK-3450 Alleroed (DK). **LASSEN, Soeren Flensted**;
Gydebakken 3B, DK-3520 Farum (DK). **SKOV, Lars**
Kobberoe; Aage Berntsens Alle 63, DK-2750 Ballerup
(DK). **DE MARIA, Leonardo**; Kong Georgsvej 4, 4th,
DK-2000 Frederiksberg (DK).**(81) Designated States (unless otherwise indicated, for every**
kind of national protection available): AE, AG, AL, AM,

AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA,
CH, 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, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC,
LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN,
MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH,
PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV,
SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, ZA,
ZM, ZW.

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

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments
- with (an) indication(s) in relation to deposited biological material furnished under Rule 13bis separately from the description

(88) Date of publication of the international search report:
6 November 2008**(54) Title:** POLYPEPTIDES HAVING PHYTASE ACTIVITY AND POLYNUCLEOTIDES ENCODING SAME

Fig. 2

```
Program: needle
Matrix: BLOSUM62
Gap initiation penalty: 10.0
Gap extension penalty: 0.5
Number of identical residues: 315
Length of shortest sequence: 413
% Identity: 315/413 = 76.3%
SEQ2_mature      1                               NTPASGYQVEKVVILS      17
                               :|..|...|...|...|...|
UNIPROTQ6U677    1  MTISLFTHSPTRLLKCMPLAFIAASMLTTASVASETEPSGYLEKVVILS      50
SEQ2_mature     18  RHGVRAITKMTQIMRDVTPATWPEWPKLGYLTTPRGEHLISLMGGFYREK      67
                               |...|...|...|...|...|...|...|...|...|...|...|...|...|...|...|
UNIPROTQ6U677    51  RHGVRAITKMTQIMRDVTPATWPEWPKLGYLTTPRGEHLISLMGGFYREK      100
SEQ2_mature     68  FQQCGILSQSGCPTPNSIYVADVDQRTLKTEAFAGLAGLAPQCGLTIEHQ      117
                               |...|...|...|...|...|...|...|...|...|...|...|...|...|...|...|
UNIPROTQ6U677   101  FQQLGILSRGRCPTANDVYVWADVQRTKRTGEAFAGLAGLAPQCGLTIEHQ      150
SEQ2_mature    118  QNLEKADPLFHFVKAATCSKDKTRLQQAWEKEAQTPIENLNQHYIPSLAL      167
                               |...|...|...|...|...|...|...|...|...|...|...|...|...|...|...|
UNIPROTQ6U677   151  QDIKQADPLFHFVKAATCSKDKTRLQQAWEKEAQTPIENLNQHYIPSLAL      200
SEQ2_mature    168  MNTILNFTSAWCKHSAKSCDLAQSMPSKLSIKONGNKVALDGAVGLS      217
                               |...|...|...|...|...|...|...|...|...|...|...|...|...|...|...|
UNIPROTQ6U677   201  MSSVLNFPKSTYCCQHSADQTCDLAQAMPKSLSEKDNKNKVALDGAVGLS      250
SEQ2_mature    218  STLAEIFLLEYAQGMPQAAWGIHSEQDWAEELKLENAQFDLMARIPYIA      267
                               |...|...|...|...|...|...|...|...|...|...|...|...|...|...|...|
UNIPROTQ6U677   251  STIARTPLIERYAQGMPDAWGRITNSQDQWNAITLITINNAQTDIMSRTPYTA      300
SEQ2_mature    268  RHNCTPLLOAISNALDPNATASKLPDISPNKILFIAGHDNTIANISCMIL      317
                               |...|...|...|...|...|...|...|...|...|...|...|...|...|...|...|
UNIPROTQ6U677   301  KHNGTPLLQTIYSAINSPSSRELPELSADNKILFPAGEDTHIANIAGMF      350
SEQ2_mature    318  NMRWTLPGQPDNTPPGGALIFERLADKAGKYVSVSMVYQTLQRLAQTP      367
                               |...|...|...|...|...|...|...|...|...|...|...|...|...|...|...|
UNIPROTQ6U677   351  GMSWALPGQPDNTPPGGALIFERLADKAGKYVSVSMVYQTLQRLAQTP      400
SEQ2_mature    368  LSLKEFAGSVQLKIPGQNEQTAEGYCLPCTFKRVSVSQSEEPGQQLQ      413
                               |...|...|...|...|...|...|...|...|...|...|...|...|...|...|...|
UNIPROTQ6U677   401  LTLDPKPGSVALKLIPGQNEQTAEGYCLPCTFKRVSVSQSEEPGQQLQ      444
```

(57) Abstract: The present invention relates to polypeptides having phytase activity. These polypeptides have an amino acid sequence which has at least 70% identity to either of three phytases derived from the bacterium *Buttiauxella*, and which comprises at least one of the following amino acids at the position indicated: 119N, 120L, and/or 121E. These phytases have an improved specific activity. Additional specific amino acid substitutions are also disclosed which characterize and distinguish additional phytases of the invention having improved properties such as temperature and/or pH stability, pH activity profile, temperature activity profile, substrate profile, improved performance in animal feed *in vitro* or *in vivo*. The invention also relates to isolated polynucleotides encoding the polypeptides, nucleic acid constructs, vectors, and host cells comprising the polynucleotides as well as methods for producing and using the polypeptides.

WO 2008/092901 A3

phytase activity and polynucleotides encoding the polypeptides. The phytases of the invention have an improved specific activity, an improved stability such as an improved temperature and/or pH stability, an improved pH activity profile, an improved temperature activity profile, an improved substrate profile, an improved performance in animal feed in vitro, and/or an improved performance in animal feed in vivo.

Summary of the Invention

In a first aspect, the present invention relates to a polypeptide having phytase activity and having an amino acid sequence which a) has at least 70% identity to amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and/or amino acids 1-413 of SEQ ID NO:6 when aligned to the respective amino acid sequence using the Needle program with the BLOSUM62 substitution matrix, a gap opening penalty of 10.0, and a gap extension penalty of 0.5; and b) comprises at least one of the following amino acids at the position indicated: 119N, 120L, and/or 121E, when aligned as described in a) to amino acids 1-413 of SEQ ID NO:2 and using an amino acid residue numbering corresponding to amino acids 1-413 of SEQ ID NO:2.

In a particular embodiment, the polypeptide comprises at least one of the following amino acids at the position indicated: 109Q, 111G, 119N, 120L, and/or 121E.

In another particular embodiment, the polypeptide a) has at least 78% identity, such as, at least 80% identity, at least 85% identity, at least 90% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity to amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and/or amino acids 1-413 of SEQ ID NO:6; and b) comprises at least one of the following amino acids at the position indicated: 109Q, 111G, 119N, 120L, 121E, and/or 193Q.

In another aspect, the invention relates to a polypeptide having phytase activity and having an amino acid sequence which a) has at least 78% identity, such as, at least 80% identity, at least 85% identity, at least 90% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity to amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and/or amino acids 1-413 of SEQ ID NO:6; and b) comprises at least one of the following amino acids at the position indicated: 1S, 10I, 38S, 66E, 71K, 81A, 109Q, 111G, 119N, 120L, 121E, 141R, 142L, 152M, 155E, 193Q, 214V, 239K, 245D, 248E, 255A,T, 268A,T, 277T, 283D,E, 285K, 287D, 288A,V, 293G, 296S, 303L, 314A, 337I, 345A, 350I, 364A, 371K, 372E, 396P, 399K, 406E, and/or 413P.

In another aspect, the invention relates to a polypeptide having phytase activity, selected from the group consisting of: (i) a polypeptide comprising a mature part of SEQ ID NO:2 (such as the sequence of amino acids 1-413 of SEQ ID NO:2); (ii) a variant of (i)

WO 2008/092901

PCT/EP2008/051137

comprising at least one of the following substitutions: K26E, N37Y, A89T, D92E, T134I,V, H160R, S164F, T171I, T176K, A178P, S188N, D190E, A192G, K207E,T, A209S, D211C, A235V, E248L, Q256H,Y, A261E, N270K, D283N, A288E, I303F, and/or N318D; and (iii) a variant of (i) or (ii) comprising at least one of the following substitutions: N1S, V9I, T38S, E66Q, Q71K, T81A, R141Q, L142V, T152M, E155D, V214I, K239N, D245E, E248S, A255T, R268A,T, A277T, D283E, N285K, T287D, A288V, D293G, P296S, I303L, S314A, I337V, A345S, V350I, A364S, K371N, E372Q, P396S, K399T, E406V, and/or Q413P.

In another aspect, the invention relates to a polypeptide having phytase activity, selected from the group consisting of: (i) a polypeptide comprising a mature part of SEQ ID NO:4 (such as the sequence of amino acids 1-413 of SEQ ID NO:4); (ii) a variant of (i) comprising at least one of the following substitutions: K26E, N37Y, A89T, D92E, T134I,V, H160R, S164F, T171I, T176K, A178P, S188N, D190E, A192G, K207E,T, A209S, D211C, A235V, S248L, Q256H,Y, A261E, N270K, E283N, V288E, I303F, and/or N318D; and (iii) a variant of (i) or (ii) comprising at least one of the following substitutions: S1N, I9V, S38T, Q66E, K71Q, T81A, Q141R, V142L, M152T, E155D, I214V, N239K, E245D, S248E, T255A, A268R,T, T277A, E283D, K285N, D287T, V288A, D293G, S296P, I303L, A314S, V337I, S345A, I350V, A364S, N371K, E372Q, S396P, T399K, V406E, and/or P413Q.

Yet another aspect of the invention relates to a polypeptide having phytase activity, selected from the group consisting of: (i) a polypeptide comprising a mature part of SEQ ID NO:6 (such as the sequence of amino acids 1-413 of SEQ ID NO:6); (ii) a variant of (i) comprising at least one of the following substitutions: K26E, N37Y, A89T, D92E, T134I,V, H160R, S164F, T171I, T176K, A178P, S188N, D190E, A192G, K207E,T, A209S, D211C, A235V, S248L, Q256H,Y, A261E, N270K, E283N, V288E, L303F, and/or N318D; and (iii) a variant of (i) or (ii) comprising at least one of the following substitutions: N1S, V9I, T38S, E66Q, Q71K, A81T, Q141R, V142L, T152M, D155E, I214V, N239K, E245D, S248E, A255T, T268A,R, T277A, E283D, K285N, D287T, V288A, G293D, P296S, L303I, A314S, V337I, S345A, V350I, S364A, N371K, Q372E, S396P, T399K, V406E, and/or Q413P.

Another aspect of the invention is directed to a phytase which has an amino acid sequence which has at least 80% identity, such as, at least 85% identity, at least 90% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity, to amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, or amino acids 1-413 of SEQ ID NO:6.

In another aspect, the present invention is directed to a phytase variant which has an amino acid sequence which has at least 70% identity, at least 75% identity, at least 80% identity, such as, at least 85% identity, at least 90% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity to a polypeptide comprising a mature part (such as the sequence of amino acids 34-446) of any

WO 2008/092901

PCT/EP2008/051137

one of the following GENESEQP sequences: AEH25057, AEH25059, AEH25058, AEH25067, AEH25071, AEH25072, AEH25074, AEH25076, AEH25073, AEH25070, AEH25069, AEH25066, AEH25060, AEH25068, AEH25063, AEH25065, AEH25062, AEH25061, AEH25056, AEH25051, or AEH25064; which variant has phytase activity and comprises at least one of the following substitutions: N1S, V10I, T38S, Q66E, Q71K, T81A, E109Q, H111G, D119N, I120L, K121E, Q141R, V142L, T152M, D155E, L193Q, I214V, N239K, E245D, S248E, V255A,T, R268A,T, A277T, N283D,E, N285K, T287D, E288A,V, D293G, P296S, I303L, A314S, V337I, S345A, V350I, S364A, N371K, Q372E, S396P, T399K, V406E, and/or Q413P.

In a particular embodiment, the invention relates to a variant of a polypeptide comprising a mature part (such as the sequence of amino acids 34-446) of GENESEQP:AEH25075, which variant has phytase activity and comprises at least one of the following substitutions: N1S, V10I, T38S, Q66E, Q71K, T81A, E109Q, H111G, D119N, I120L, K121E, Q141R, V142L, T152M, D155E, L193Q, I214V, N239K, E245D, S248E, V255A,T, R268A,T, A277T, N283D,E, N285K, T287D, E288A,V, D293G, P296S, F303L, A314S, V337I, S345A, V350I, S364A, N371K, Q372E, S396P, T399K, V406E, and/or Q413P.

In each of these aspects, for calculating identity and determining amino acid residue positions, an alignment is produced as described under the first aspect above.

The invention also relates to polynucleotides encoding these polypeptides, as well as nucleic acid constructs, recombinant expression vectors, and recombinant host cells comprising the polynucleotides, and to methods for producing and using these polypeptides.

Another aspect of the present invention relates to animal feed compositions comprising a phytase of the present invention. The present invention also provides methods for improving the nutritional value of an animal feed using a phytase of the present invention.

Another aspect of the present invention relates to methods for treating proteins, including vegetable proteins, with a phytase of the present invention.

Yet another aspect of the present invention relates to the methods for producing a fermentation product, such as, e.g., ethanol, beer, wine, wherein the fermentation is carried out in the presence of a phytase of the present invention.

Brief Description of the Figures

Figure 1 shows a multiple alignment of the expected mature parts of SEQ ID NO:2, SEQ ID NO:4 and SEQ ID NO:6 of the invention together with the sequences with the following GENESEQP accession numbers: AEH25051, AEH25056, AEH25057, AEH25058, AEH25059, AEH25060, AEH25061, AEH25062, AEH25063, AEH25064, AEH25065,

WO 2008/092901

PCT/EP2008/051137

AEH25066, AEH25067, AEH25068, AEH25069, AEH25070, AEH25071, AEH25072, AEH25073, AEH25074, AEH25075, and AEH25076. The alignment was made using the Clustal method (Higgins, 1989, CABIOS 5: 151-153) using the LASERGENE™ MEGALIGN™ software (DNASTAR, Inc., Madison, WI) with an identity table and the following multiple alignment parameters: Gap penalty of 10 and gap length penalty of 10. Pairwise alignment parameters are Ktuple=1, gap penalty=3, windows=5, and diagonals=5.

Figure 2 shows an alignment of UNIPROT accession no. Q6U677 with amino acids 1-413 of SEQ ID NO:2. The alignment was made using the program needle with the matrix BLOSUM62, a gap initiation penalty of 10.0 and a gap extension penalty of 0.5.

Detailed Description of the Invention

Structural Considerations

The structures of three phytases of the invention (mature part of SEQ ID NOs:2, 4 and 6) were built by homology modelling, using as a template the structure of the E. coli AppA phytase (Protein Data Bank id.: 1DKN; Lim et al, Nat. Struct. Biol. (2000), vol. 2, pp. 108-113).

Positions 119 to 121 are facing the active site cleft and therefore the specific amino acids in these positions are expected to have an effect on specific activity. For increasing the specific activity, 119N is estimated as the better choice, followed by 121E, and finally 120L.

Positions 109 and 111 are close to a disulfide in a region which is a bit stressed, so the specific amino acids in these positions are expected to have an effect on stability, on thermostability in particular. 111G may be better than 109Q.

Position 193 is another region which may influence the stability of the enzyme. 193Q may provide a good stability.

Corresponding phytase variants, and other phytase variants (e.g., a variant comprising a conservative substitution, deletion, and/or insertion of one or more amino acids of amino acids), may be prepared by methods known in the art and tested as described in the experimental part.

Phytase Polypeptides, Percentage of Identity

In the present context a phytase is a polypeptide having phytase activity, i.e. an enzyme which catalyzes the hydrolysis of phytate (myo-inositol hexakisphosphate) to (1) myo-inositol and/or (2) mono-, di-, tri-, tetra- and/or penta-phosphates thereof and (3) inorganic phosphate.

In the present context the term a phytase substrate encompasses, i.a., phytic acid

and any phytate (salt of phytic acid), as well as the phosphates listed under (2) above.

The ENZYME site at the internet (www.expasy.ch/enzyme/) is a repository of information relative to the nomenclature of enzymes. It is primarily based on the recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUB-MB) and it describes each type of characterized enzyme for which an EC (Enzyme Commission) number has been provided (Bairoch A. The ENZYME database, 2000, Nucleic Acids Res 28:304-305). See also the handbook Enzyme Nomenclature from NC-IUBMB, 1992).

According to the ENZYME site, three different types of phytases are known: A so-called 3-phytase (alternative name 1-phytase; a myo-inositol hexaphosphate 3-phosphohydrolase, EC 3.1.3.8), a so-called 4-phytase (alternative name 6-phytase, name based on 1L-numbering system and not 1D-numbering, EC 3.1.3.26), and a so-called 5-phytase (EC 3.1.3.72). For the purposes of the present invention, all three types are included in the definition of phytase.

In a particular embodiment, the phytases of the invention belong to the family of acid histidine phosphatases, which includes the Escherichia coli pH 2.5 acid phosphatase (gene appA) as well as fungal phytases such as Aspergillus awamorii phytases A and B (EC: 3.1.3.8) (gene phyA and phyB). The histidine acid phosphatases share two regions of sequence similarity, each centered around a conserved histidine residue. These two histidines seem to be involved in the enzymes' catalytic mechanism. The first histidine is located in the N-terminal section and forms a phosphor-histidine intermediate while the second is located in the C-terminal section and possibly acts as proton donor.

In a further particular embodiment, the phytases of the invention have a conserved active site motif, viz. R-H-G-V-R-A-P, (see amino acids 18 to 24 of SEQ ID NOs:2, 4, and 6).

For the purposes of the present invention the phytase activity is determined in the unit of FYT, one FYT being the amount of enzyme that liberates 1 micro-mol inorganic orthophosphate per min. under the following conditions: pH5.5; temperature 37°C; substrate: sodium phytate ($C_6H_6O_{24}P_6Na_{12}$) in a concentration of 10.8mmol/l. Phytase activity is preferably determined using the assay of Example 3 herein. Other suitable phytase assays are the FYT and FTU assays described in Example 1 of WO 00/20569. FTU is for determining phytase activity in feed and premix.

In a particular embodiment the phytase of the invention is isolated. The term "isolated" as used herein refers to a polypeptide which is at least 20% pure, preferably at least 40% pure, more preferably at least 60% pure, even more preferably at least 80% pure, most preferably at least 90% pure, and even most preferably at least 95% pure, such as, at least 96%, 97%, 98%, 99% and higher purity, as determined by SDS-PAGE. In particular, it is preferred that the polypeptides are in "essentially pure form", i.e., that the polypeptide

preparation is essentially free of other polypeptide material with which it is natively associated. This can be accomplished, for example, by preparing the polypeptide by means of well-known recombinant methods and/or by classical purification methods.

When used herein the term “mature part” refers to that part of the polypeptide which is secreted by a cell which contains, as part of its genetic equipment, a polynucleotide encoding the polypeptide. Generally, the mature polypeptide part refers to that part of the polypeptide which remains after the N-terminal signal peptide part is cleaved off once it has fulfilled its function of directing the encoded polypeptide into the cell’s secretory pathway. However, experience shows that sometimes also minor C-terminal truncations occur during the secretion process. The term mature part as used herein also takes into account such C-terminal truncations, if any. The expected signal peptide parts of SEQ ID NOs:2, 4 and 6 are amino acids -33 to -1 of SEQ ID NO:2, -9 to -1 of SEQ ID NO:4 (this is a partial signal peptide), and amino acids -33 to -1 of SEQ ID NO:6 (see the sequence listing included herewith). SEQ ID NO:8, which is encoded by SEQ ID NO:7, is also a signal peptide. As we are not presently aware of any C-terminal truncations having occurred during secretion, the expected mature parts of SEQ ID NOs:2, 4, and 6 are amino acids 1-413 thereof.

In various aspects of the present invention, the relatedness between two amino acid sequences is described by the parameter “identity”. For purposes of the present invention, the alignment of two amino acid sequences is determined by using the Needle program from the EMBOSS package (<http://emboss.org>), preferably in version 2.8.0. The Needle program implements the global alignment algorithm described in Needleman, S. B. and Wunsch, C. D. (1970) J. Mol. Biol. 48, 443-453. The substitution matrix used is BLOSUM62, the gap opening (initiation) penalty is 10.0, and the gap extension penalty is 0.5.

The degree of identity between an amino acid sequence of the present invention (“invention sequence”) and an amino acid sequence referred to in the claims (e.g. amino acids 1-413 of SEQ ID NO:2) is calculated as the number of exact matches in an alignment of the two sequences, divided by the length of the “invention sequence,” or the length of amino acids 1-413 of SEQ ID NO:2, whichever is the shortest. The result is expressed in percent identity.

An exact match occurs when the “invention sequence” and SEQ ID NO:2 have identical amino acid residues in the same positions of the overlap (in the alignment example below this is represented by “[”). The length of a sequence is the number of amino acid residues in the sequence (e.g. the length of amino acids 1-413 of SEQ ID NO:2 is 413).

In the purely hypothetical alignment example below, the overlap is the amino acid sequence “HTWGER-NL” of Sequence 1; or the amino acid sequence “HGWGEDANL” of Sequence 2. In the example a gap is indicated by a “-”.

Hypothetical alignment example:

Sequence 1: ACMSHTWGER-NL

| | | | |

Sequence 2: HGWGEDANLAMNPS

In a particular embodiment, the percentage of identity of an amino acid sequence of
 5 a polypeptide with, or to, e.g., amino acids 1-413 of SEQ ID NO:2 is determined by i)
 aligning the two amino acid sequences using the Needle program, with the BLOSUM62
 substitution matrix, a gap opening penalty of 10.0, and a gap extension penalty of 0.5; ii)
 counting the number of exact matches in the alignment; iii) dividing the number of exact
 matches by the length of the shortest of the two amino acid sequences, and iv) converting
 10 the result of the division of iii) into percentage.

In the above hypothetical example, the number of exact matches is 6, the length of
 the shortest one of the two amino acid sequences is 12; accordingly the percentage of
 identity is 50%.

Another example of an alignment is shown in Fig. 2, where the phytase from
 15 Obesumbacterium proteus (UNIPROT:Q6U677) is aligned with amino acids 1-413 of SEQ ID
 NO:2. From this alignment it appears that the percentage of identity of UNIPROT:Q6U677 to
 amino acids 1-413 of SEQ ID NO:2 is 76.3% (315 exact matches, the length of the shortest
 sequence is 413).

In particular embodiments of the phytase of the invention (i.e. according to any one
 20 of the various aspects thereof), the degree of identity to any one of amino acids 1-413 of
 SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6
 is at least 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%,
 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least
 99%. In still further particular embodiments, the degree of identity is at least 90.0%, 90.2%,
 25 90.4%, 90.6%, 90.8%, 91.0%, 91.2%, 91.4%, 91.6%, 91.8%, 92.0%, 92.2%, 92.4%, 92.6%,
 92.8%, 93.0%, 93.2%, 93.4%, 93.6%, 93.8%, 94.0%, 94.2%, 94.4%, 94.6%, 94.8%, 95.0%,
 95.2%, 95.4%, 95.6%, 95.8%, 96.0%, 96.2%, 96.4%, 96.6%, 96.8%, 97.0%, 97.2%, 97.4%,
 97.6%, 97.8%, 98.0%, 98.2%, 98.4%, 98.6%, 98.8%, 99.0%, 99.2%, 99.4%, 99.6%, or at
 least 99.8%.

In still further particular embodiments, the phytase of the invention has (or has an
 30 amino acid sequence which differs by) no more than 1, 2, 3, 4, 5, 6, 7, 8, 9, or no more than
 10 alterations as compared to any one of amino acids 1-413 of SEQ ID NO:2, amino acids
 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6; no more than 11, 12, 13,
 14, 15, 16, 17, 18, 19, or no more than 20 alterations as compared to any one of amino
 35 acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of
 SEQ ID NO:6; no more than 21, 22, 23, 24, 25, 26, 27, 28, 29, or no more than 30
 alterations as compared to any one of amino acids 1-413 of SEQ ID NO:2, amino acids 1-

413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6; no more than 31, 32, 33, 34, 35, 36, 37, 38, 39, or no more than 40 alterations as compared to any one of amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6; no more than 41, 42, 43, 44, 45, 46, 47, 48, 49, or no more than 50 alterations as compared to any one of amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6; no more than 51, 52, 53, 54, 55, 56, 57, 58, 59, or no more than 60 alterations as compared to any one of amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6; no more than 61, 62, 63, 64, 65, 66, 67, 68, 69, or no more than 70 alterations as compared to any one of amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6; no more than 71, 72, 73, 74, 75, 76, 77, 78, 79, or no more than 80 alterations as compared to any one of amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6; no more than 81, 82, 83, 84, 85, 86, 87, 88, 89, or no more than 90 alterations as compared to any one of amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6; no more than 91, 92, 93, 94, 95, 96, 97, 98, 99, or no more than 100 alterations as compared to any one of amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6; no more than 101, 102, 103, 104, 105, 106, 107, 108, 109, or no more than 110 alterations as compared to any one of amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6; no more than 111, 112, 113, 114, 115, 116, 117, 118, 119, or no more than 120 alterations as compared to any one of amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6; or no more than 121, 122, 123, or 124 alterations as compared to any one of amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6.

In alternative embodiments of the phytase of the invention, the degree of identity to any one of amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6 is at least 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, or at least 69%.

Position Numbering

The nomenclature used herein (i.e., according to any one of the various aspects of the invention) for defining amino acid positions is based on the amino acid sequence of the phytase derived from *Buttiauxella gaviniae* DSM 18930, viz. the expected mature sequence thereof which is amino acids 1-413 of SEQ ID NO:2. Accordingly, in the present context, the basis for numbering positions is SEQ ID NO:2 starting with N1 and ending with Q413. This position numbering is shown as the first line in the alignment of Fig. 1 and this numbering is

also shown in the alignment of Fig. 2 (the upper row of the alignment).

Alterations, such as Substitutions, Deletions, Insertions

A phytase of the invention (i.e., according to any one of the various aspects of the invention), be it a wildtype or a variant, can comprise various types of alterations relative to a template (i.e. a reference or comparative amino acid sequence such as amino acids 1-413 of SEQ ID NO:2): An amino acid can be substituted with another amino acid; an amino acid can be deleted; an amino acid can be inserted; as well as any combination of any number of such alterations. In the present context the term "insertion" is intended to cover also N- and/or C-terminal extensions, and the term "deletion" is intended to cover also N- and/or C-terminal truncations.

The general nomenclature used herein for a single alteration is the following: XDcY, where "X" and "Y" independently designate a one-letter amino acid code, or a "*", "D" designates a number, and "c" designates an alphabetical counter (a, b, c, and so forth), which is only present in insertions. Reference is made to Table 1 below which describes purely hypothetical examples of applying this nomenclature to various types of alterations.

Table 1

Type	Description	Example
Substitution	X=Amino acid in template D=Position in template c empty Y=Amino acid in variant	G80A 80 AALNNSIGVLGVAPSAELYAVKVLGASGSG : AALNNSIAVLGVAPSAELYAVKVLGASGSG
Insertion	X="*" D=Position in template before the insertion c="a" for first insertion at this position, "b" for next, etc	*80aT *80bY *85aS 80 85 AALNNSIG.VLGVA.PSAELYAVKVLGASG AALNNSIGTYVLGVASPSAELYAVKVLGASG
Deletion	X=Amino acid in template D=Position in template c empty Y="*"	V81* 80 AALNNSIGVLGVAPSAELYAVKVLGASGSG AALNNSIG.LGVAPSAELYAVKVLGASGSG
N-terminal	Insertions at position "0".	*0aA *0bT *0cG

extension		1 ...AQSVPWGISRVO ATGAQSVPWGISRVO
C-terminal extension	Insertions after the N-terminal amino acid.	*275aS *275bT 270 275 ATSLGSTNLYGSGLVNAEAATR.. ATSLGSTNLYGSGLVNAEAATRST

As explained above, the position number ("D") is counted from the first amino acid residue of amino acids 1-413 of SEQ ID NO:2.

Several alterations in the same sequence are separated by "/" (slash), e.g. the designation "1*/2*/3*" means that the amino acids in position number 1, 2, and 3 are all deleted, and the designation "104A/105F" means that the amino acid in position number 104 is substituted by A, and the amino acid in position number 105 is substituted by F.

Alternative alterations are separated by "," (comma), e.g., the designation "255A,T" means that the amino acid in position 255 is substituted with A or T.

The commas used herein in various other enumerations of possibilities mean what they usually do grammatically, viz. often and/or. E.g., the first comma in the listing "255A,T, 406E, and/or 413P" denotes an alternative (A or T), whereas the two next commas should be interpreted as and/or options: 255A or 255T, and/or 406E, and/or 413P.

In the present context, "at least one" (e.g. amino acids at the position indicated, or substitution) means one or more, e.g. 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids (or substitutions); or 12, 14, 15, 16, 18, 20, 22, 24, 25, 28, or 30 amino acids (or substitutions); and so on, up to a maximum number of alterations of 125, 130, 140, 150, 160, 170, 180, 190, or of 200.

A substitution or extension without any indication of what to substitute or extend with refers to the insertion of any natural, or non-natural, amino acid, except the one that occupies this position in the template.

Identifying Corresponding Position Numbers

As explained above, the mature phytase of *Buttiauxella gaviniae* DSM 18930 (SEQ ID NO:2) is used herein as the standard for position numbering and, thereby, also for the nomenclature.

For another phytase, be it a wildtype or a variant, the position corresponding to position D in SEQ ID NO:2 is found by aligning the two sequences as specified above in the

section entitled "Phytase polypeptides, percentage of identity". From the alignment, the position in the sequence of the invention corresponding to position D of SEQ ID NO:2 can be clearly and unambiguously identified (the two positions on top of each other in the alignment).

5 Here are some purely hypothetical examples derived from Table 1 above which in the third column includes a number of alignments of two sequences:

Consider the third cell in the first row of Table 1: The upper sequence is the template, the lower the variant. Position number 80 refers to amino acid residue G in the template. Amino acid A occupies the corresponding position in the variant. Accordingly, this
10 substitution is designated G80A.

Consider now the third cell in the second row of Table 1: The upper sequence is again the template and the lower the variant. Position number 80 again refers to amino acid residue G in the template. The variant has two insertions, viz. TY, after G80 and before V81 in the template. Whereas the T and Y of course would have their own "real" position number
15 in the variant amino acid sequence, for the present purposes we always refer to the template position numbers, and accordingly the T and the Y are said to be in position number 80a and 80b, respectively.

Finally, consider the third cell in the last row of Table 1: Position number 275 refers to the last amino acid of the template. A C-terminal extension of ST are said to be in position
20 number 275a and 275b, respectively, although, again, of course they have their own "real" position number in the variant amino acid sequence.

A real example of such alignment is shown in Fig. 2, where the phytase from *Obesumbacterium proteus* (UNIPROT:Q6U677) is aligned with amino acids 1-413 of SEQ ID NO:2. From this alignment it is inferred that, i.e., amino acids 119N, 120L, and 121E of SEQ
25 ID NO:2 correspond to 119D, 120I and 120K of UNIPROT:Q6U677, respectively.

Various embodiments

A polypeptide according to the second aspect of the invention has an amino acid sequence which a) has at least 78% identity to amino acids 1-413 of SEQ ID NO:2, amino
30 acids 1-413 of SEQ ID NO:4, and/or amino acids 1-413 of SEQ ID NO:6; and b) comprises at least one of the following amino acids at the position indicated: 1S, 10I, 38S, 66E, 71K, 81A, 109Q, 111G, 119N, 120L, 121E, 141R, 142L, 152M, 155E, 193Q, 214V, 239K, 245D, 248E, 255A,T, 268A,T, 277T, 283D,E, 285K, 287D, 288A,V, 293G, 296S, 303L, 314A, 337I, 345A, 350I, 364A, 371K, 372E, 396P, 399K, 406E, and/or 413P.

35 Each of the amino acids at the positions indicated under b) above make a difference to the *Buttiauxella* NCIMB 41248 wildtype phytase and its variants with the following GENESEQP accession numbers: AEH25051, AEH25056, AEH25057, AEH25058,

WO 2008/092901

PCT/EP2008/051137

AEH25059, AEH25060, AEH25061, AEH25062, AEH25063, AEH25064, AEH25065, AEH25066, AEH25067, AEH25068, AEH25069, AEH25070, AEH25071, AEH25072, AEH25073, AEH25074, AEH25075, and AEH25076.

In particular embodiments thereof, the amino acid sequence comprises a) at least one of 1S, 10I, 38S, 71K, 109Q, 111G, 119N, 120L, 121E, 152M, 155E, 193Q, 255T, 268A, 277T, 283E, 285K, 287D, 288V, 296S, 364A, 350I, 372E, and/or 413P (as represented by SEQ ID NO:4); b) at least one of 66E, 81A, 109Q, 111G, 119N, 120L, 121E, 193Q, 255A, 268T, 277T, 283E, 285K, 287D, 288V, 293G, and/or 303L (as represented by SEQ ID NO:6); and c) at least one of 66E, 109Q, 111G, 119N, 120L, 121E, 141R, 142L, 155E, 193Q, 214V, 239K, 245D, 248E, 255A, 283D, 288A, 314S, 337I, 345A, 364A, 371K, 372E, 396P, 399K, and/or 406E (as represented by SEQ ID NO:2).

In another particular embodiment the polypeptide of the invention has been altered inspired by the *Buttiauxella* wildtype and mutant phytases disclosed in WO 2006/043178, viz. it has an amino acid sequence which comprises at least one of the following amino acids at the position indicated: 26E, 37Y, 89T, 92E, 134I,V, 160R, 164F, 171I, 176K, 178P, 188N, 190E, 192G, 207E,T, 209S, 211C, 235V, 248L, 256H,Y, 261E, 270K, 303F, and/or 318D.

For example, the mature part of SEQ ID NO:2 is mutated to comprise at least one of the following substitutions: K26E, N37Y, A89T, D92E, T134I,V, H160R, S164F, T171I, T176K, A178P, S188N, D190E, A192G, K207E,T, A209S, D211C, A235V, E248L, Q256H,Y, A261E, N270K, D283N, A288E, I303F, and/or N318D. Other preferred substitutions in the mature part of SEQ ID NO:2 are inspired by SEQ ID NOs:4 and 6: N1S, V9I, T38S, E66Q, Q71K, T81A, R141Q, L142V, T152M, E155D, V214I, K239N, D245E, E248S, A255T, R268A,T, A277T, D283E, N285K, T287D, A288V, D293G, P296S, I303L, S314A, I337V, A345S, V350I, A364S, K371N, E372Q, P396S, K399T, E406V, and/or Q413P.

As another example, the mature part of SEQ ID NO:4 is mutated to comprise at least one of the following substitutions: K26E, N37Y, A89T, D92E, T134I,V, H160R, S164F, T171I, T176K, A178P, S188N, D190E, A192G, K207E,T, A209S, D211C, A235V, S248L, Q256H,Y, A261E, N270K, E283N, V288E, I303F, and/or N318D. Other preferred substitutions in the mature part of SEQ ID NO:2 are inspired by SEQ ID NOs:2 and 6: S1N, I9V, S38T, Q66E, K71Q, T81A, Q141R, V142L, M152T, E155D, I214V, N239K, E245D, S248E, T255A, A268R,T, T277A, E283D, K285N, D287T, V288A, D293G, S296P, I303L, A314S, V337I, S345A, I350V, A364S, N371K, E372Q, S396P, T399K, V406E, and/or P413Q.

As a still further example, the mature part of SEQ ID NO:6 is mutated to comprise at least one of the following substitutions: K26E, N37Y, A89T, D92E, T134I,V, H160R, S164F,

WO 2008/092901

PCT/EP2008/051137

T171I, T176K, A178P, S188N, D190E, A192G, K207E,T, A209S, D211C, A235V, S248L, Q256H,Y, A261E, N270K, E283N, V288E, L303F, and/or N318D. Other preferred substitutions in the mature part of SEQ ID NO:2 are inspired by SEQ ID NOs:2 and 4: N1S, V9I, T38S, E66Q, Q71K, A81T, Q141R, V142L, T152M, D155E, I214V, N239K, E245D, S248E, A255T, T268A,R, T277A, E283D, K285N, D287T, V288A, G293D, P296S, L303I, A314S, V337I, S345A, V350I, S364A, N371K, Q372E, S396P, T399K, V406E, and/or Q413P.

As explained above, various aspects of the present invention relate to polypeptides "having" an amino acid sequence which has a specified degree of identity to amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and/or amino acids 1-413 of SEQ ID NO:6, which have phytase activity (hereinafter "homologous polypeptides"); or to polypeptides "having" an amino acid sequence which differs by a maximum number of amino acids amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6.

A polypeptide of the present invention preferably comprises any one of amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, or amino acids 1-413 of SEQ ID NO:6; or an allelic variant thereof; or a fragment thereof that has phytase activity.

In another preferred embodiment, a polypeptide of the present invention consists of the amino acid sequence of amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, and amino acids 1-413 of SEQ ID NO:6; or an allelic variant thereof; or a fragment thereof that has phytase activity.

Anyone of the nucleotide sequences of SEQ ID NO:1, 3 and 5, or a subsequence thereof, as well as the amino acid sequences of SEQ ID NO:2, 4, and 6, or a fragment thereof, may be used to design a nucleic acid probe to identify and clone DNA encoding polypeptides having phytase activity from strains of different genera or species according to methods well known in the art. In particular, such probes can be used for hybridization with the genomic or cDNA of the genus or species of interest, following standard Southern blotting procedures, in order to identify and isolate the corresponding gene therein. Both DNA and RNA probes can be used. The probes are typically labelled for detecting the corresponding gene (for example, with ^{32}P , ^3H , ^{35}S , biotin, or avidin).

A genomic DNA library prepared from such other organisms may, therefore, be screened for DNA which hybridizes with the probes described above and which encodes a polypeptide having phytase activity. Genomic or other DNA from such other organisms may be separated by agarose or polyacrylamide gel electrophoresis, or other separation techniques. DNA from the libraries or the separated DNA may be transferred to and immobilized on nitrocellulose or other suitable carrier material. In order to identify a clone or DNA which is homologous with SEQ ID NO:1, 3 or 5 or a subsequence thereof, the carrier

material is used in a Southern blot.

In a preferred embodiment, the nucleic acid probe comprises nucleotides 454-462 of SEQ ID NO:1, nucleotides 384-392 of SEQ ID NO:3, or nucleotides 454-462 of SEQ ID NO:5 (all corresponding to the motif 119N, 120L, and 121E). In another preferred aspect, the nucleic acid probe is a polynucleotide sequence which encodes the polypeptides of any one of SEQ ID NOs:2, 4, or 6, or a subsequence thereof. In another preferred aspect, the nucleic acid probe is SEQ ID NO:1, 3 or 5.

The present invention also relates to variants comprising a conservative substitution, deletion, and/or insertion of one or more amino acids in the sequences of amino acids 1-413 of SEQ ID NO:2, amino acids 1-413 of SEQ ID NO:4, or amino acids 1-413 of SEQ ID NO:6, or the mature polypeptides thereof. Preferably, amino acid changes are of a minor nature, that is conservative amino acid substitutions or insertions that do not significantly affect the folding and/or activity of the protein; small deletions, typically of one to about 30 amino acids; small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue; a small linker peptide of up to about 20-25 residues; or a small extension that facilitates purification by changing net charge or another function, such as a poly-histidine tract, an antigenic epitope or a binding domain. A fragment is a preferred example of a variant comprising a deletion, as described above, and a fragment preferably retains phytase activity.

Examples of conservative substitutions are within the group of basic amino acids (arginine, lysine and histidine), acidic amino acids (glutamic acid and aspartic acid), polar amino acids (glutamine and asparagine), hydrophobic amino acids (leucine, isoleucine and valine), aromatic amino acids (phenylalanine, tryptophan and tyrosine), and small amino acids (glycine, alanine, serine, threonine and methionine). Amino acid substitutions which do not generally alter specific activity are known in the art and are described, for example, by H. Neurath and R.L. Hill, 1979, In, The Proteins, Academic Press, New York. The most commonly occurring exchanges are Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, and Asp/Gly.

In addition to the 20 standard amino acids, non-standard amino acids (such as 4-hydroxyproline, 6-N-methyl lysine, 2-aminoisobutyric acid, isovaline, and alpha-methyl serine) may be substituted for amino acid residues of a wild-type polypeptide. A limited number of non-conservative amino acids, amino acids that are not encoded by the genetic code, and unnatural amino acids may be substituted for amino acid residues. "Unnatural amino acids" have been modified after protein synthesis, and/or have a chemical structure in their side chain(s) different from that of the standard amino acids. Unnatural amino acids can be chemically synthesized, and preferably, are commercially available, and include

pipecolic acid, thiazolidine carboxylic acid, dehydroproline, 3- and 4-methylproline, and 3,3-dimethylproline.

Essential amino acids in the parent polypeptide can be identified according to procedures known in the art, such as site-directed mutagenesis or alanine-scanning
5 mutagenesis (Cunningham and Wells, 1989, Science 244: 1081-1085). In the latter technique, single alanine mutations are introduced at every residue in the molecule, and the resultant mutant molecules are tested for biological activity (i.e., phytase activity) to identify amino acid residues that are critical to the activity of the molecule. See also, Hilton et al., 1996, J. Biol. Chem. 271: 4699-4708. The active site of the enzyme or other biological
10 interaction can also be determined by physical analysis of structure, as determined by such techniques as nuclear magnetic resonance, crystallography, electron diffraction, or photoaffinity labelling, in conjunction with mutation of putative contact site amino acids. See, for example, de Vos et al., 1992, Science 255: 306-312; Smith et al., 1992, J. Mol. Biol. 224: 899-904; Wlodaver et al., 1992, FEBS Lett. 309:59-64. The identities of essential amino
15 acids can also be inferred from analysis of identities with polypeptides which are related to a polypeptide according to the invention.

Single or multiple amino acid substitutions can be made and tested using known methods of mutagenesis, recombination, and/or shuffling, followed by a relevant screening procedure, such as those disclosed by Reidhaar-Olson and Sauer, 1988, Science 241: 53-
20 57; Bowie and Sauer, 1989, Proc. Natl. Acad. Sci. USA 86: 2152-2156; WO 95/17413; or WO 95/22625. Other methods that can be used include error-prone PCR, phage display (e.g., Lowman et al., 1991, Biochem. 30:10832-10837; U.S. Patent No. 5,223,409; WO 92/06204), and region-directed mutagenesis (Derbyshire et al., 1986, Gene 46:145; Ner et al., 1988, DNA 7:127).

25 Mutagenesis/shuffling methods can be combined with high-throughput, automated screening methods to detect activity of cloned, mutagenized polypeptides expressed by host cells. Mutagenized DNA molecules that encode active polypeptides can be recovered from the host cells and rapidly sequenced using standard methods in the art. These methods allow the rapid determination of the importance of individual amino acid residues in a
30 polypeptide of interest, and can be applied to polypeptides of unknown structure.

A polypeptide of the present invention may be obtained from microorganisms of any genus. For purposes of the present invention, the term "obtained from" as used herein in connection with a given source shall mean that the polypeptide encoded by a nucleotide
35 sequence is produced by the source or by a strain in which the nucleotide sequence from the source has been inserted. In a preferred aspect, the polypeptide obtained from a given source is secreted extracellularly.

A polypeptide of the present invention may be a bacterial polypeptide. In particular

embodiments, the polypeptide is obtainable from the Proteobacteria; Gammaproteobacteria; Enterobacteriales; family of Enterobacteriaceae; preferably from the genus of *Buttiauxella*, e.g. selected from amongst the species of *Buttiauxella agrestis*, *Buttiauxella brennerae*, *Buttiauxella ferragutiae*, *Buttiauxella gaviniae*, *Buttiauxella izardii*, *Buttiauxella noackiae*, *Buttiauxella warmboldiae*, *Buttiauxella* sp. B22, *Buttiauxella* sp. BTN01, *Buttiauxella* sp. LBV 449, *Buttiauxella* sp. P5, *Buttiauxella* sp. PNBS, *Buttiauxella* sp. S212, *Buttiauxella* sp. S215, *Buttiauxella* sp. S218. <http://www.ncbi.nlm.nih.gov/Taxonomy/Browser>.

In a more preferred aspect, the polypeptide is a *Buttiauxella gaviniae* or *Buttiauxella agrestis* polypeptide, most preferably a *Buttiauxella gaviniae* DSM 18930, a *Buttiauxella agrestis* DSM 18931, or a *Buttiauxella agrestis* DSM 18932 polypeptide.

It will be understood that for the aforementioned species, the invention encompasses both the perfect and imperfect states, and other taxonomic equivalents, e.g., anamorphs, regardless of the species name by which they are known. Those skilled in the art will readily recognize the identity of appropriate equivalents.

Strains of these species are readily accessible to the public in a number of culture collections, such as the American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Centraalbureau Voor Schimmelcultures (CBS), and Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

Furthermore, such polypeptides may be identified and obtained from other sources including microorganisms isolated from nature (e.g., soil, composts, water, etc.) using the above-mentioned probes. Techniques for isolating microorganisms from natural habitats are well known in the art. The polynucleotide may then be obtained by similarly screening a genomic or cDNA library of another microorganism. Once a polynucleotide sequence encoding a polypeptide has been detected with the probe(s), the polynucleotide can be isolated or cloned by utilizing techniques which are well known to those of ordinary skill in the art (see, e.g., Sambrook et al., 1989, *supra*).

Polypeptides of the present invention also include fused polypeptides or cleavable fusion polypeptides in which another polypeptide is fused at the N-terminus or the C-terminus of the polypeptide or fragment thereof. A fused polypeptide is produced by fusing a nucleotide sequence (or a portion thereof) encoding another polypeptide to a nucleotide sequence (or a portion thereof) of the present invention. Techniques for producing fusion polypeptides are known in the art, and include ligating the coding sequences encoding the polypeptides so that they are in frame and that expression of the fused polypeptide is under control of the same promoter(s) and terminator.

Amended Properties

In a particular embodiment of any aspect of the invention the phytase has amended, preferably improved, properties. Examples of amended, preferably improved, properties are pl, temperature and/or pH stability, pH activity profile, temperature activity profile, substrate profile, and performance in animal feed in vitro or in vivo.

5 The terms "amended" and "improved" imply a comparison with another phytase. Examples of such other, reference, or comparative, phytases are the *Buttiauxella* NCIMB 41248 wildtype phytase and its variants with the following GENESEQP accession numbers: AEH25051, AEH25056, AEH25057, AEH25058, AEH25059, AEH25060, AEH25061, AEH25062, AEH25063, AEH25064, AEH25065, AEH25066, AEH25067, AEH25068,
10 AEH25069, AEH25070, AEH25071, AEH25072, AEH25073, AEH25074, AEH25075, and AEH25076.

 The phytases of the invention which comprise at least one of the following amino acids at the position indicated: 119N, 120L, and/or 121E have an improved specific activity as shown in the experimental part. In particular embodiments, the phytase of the invention
15 comprises i) 119N; ii) 121E; iii) 120L; iv) 119N and 121E; v) 119N and 120L; vi) 121E and 120L; or vii) 119N, 120L, and 121E. In still further particular embodiments the phytases of i)-vii) have an improved specific activity. Specific activity is determined as described below.

 The phytases of the invention which comprise at least one of the following amino acids at the position indicated: 109Q, and/or 111G are expected to have an improved
20 stability, in particular an improved thermostability. In particular embodiments, the phytase of the invention comprises i) 111G; ii) 109Q; or iii) 111G and 109Q. In still further particular embodiments the phytases of i)-iii) have an improved stability, preferably an improved thermostability. Thermostability is determined as described below.

 The phytases of the invention which comprise 193Q are also expected to have an
25 improved stability, preferably an improved thermostability. Thermostability is determined as described below.

 Additional specific amino acids at specified positions, as well as additional specific amino acid substitutions are also disclosed herein, which characterize and distinguish additional phytases of the invention which may have improved properties such as pl,
30 temperature stability, pH stability, pH activity profile, temperature activity profile, substrate profile, and/or improved performance in animal feed in vitro or in vivo.

Specific activity

 In a particular embodiment, the phytase of the invention has an improved specific
35 activity relative to a reference phytase. More in particular, the specific activity of a phytase of the invention is at least 105%, relative to the specific activity of a reference phytase determined by the same procedure. In still further particular embodiments, the relative

specific activity is at least 110, 115, 120, 125, 130, 140, 145, 150, 160, 170, 180, 190, 200, 220, 240, 260, 280, 300, 350 or even 400%, still relative to the specific activity of the reference phytase as determined by the same procedure.

Examples of reference phytases are listed above. Preferred reference phytases for this embodiment are GENESEQP:AEH25072 (the variant disclosed in Table 3 of WO 2006/043178), and GENESEQP:AEH25051 (the wildtype phytase from *Buttiauxella* P1-29 disclosed in Example 10 of WO 2006/043178).

In the alternative, the term high specific activity refers to a specific activity of at least 240 FYT/mg Enzyme Protein (EP). In particular embodiments, the specific activity is at least 300, 350, 400, 450, 500, 550, 600, or at least 650 FYT/mg EP.

Specific activity is measured on highly purified samples (an SDS poly acryl amide gel should show the presence of only one component). Preferably, the sample of the enzyme is at least 95% pure, as determined by SDS-PAGE. The enzyme protein concentration may be determined by amino acid analysis, and the phytase activity in the units of FYT, determined as described in Example 3, i.e. on the substrate of sodium phytate at pH5.5 and 37°C. Specific activity is a characteristic of the specific phytase in question, and it is calculated as the phytase activity measured in FYT units per mg phytase variant enzyme protein. See Example 4 for further details.

Isoelectric point (pI)

In particular embodiments, the pI for the phytase of the invention is in the range of i) 7.0-8.0; ii) 7.2-7.8; or iii) 7.4-7.6. The pI is determined as described in Example 5.

pH profile

In a particular embodiment, the phytase of the invention has an amended pH profile as compared to a reference phytase. Examples of reference phytases are listed above.

In particular embodiments, the phytase of the invention has a relative activity at pH 2.0 which is at least 30% of the activity at pH 4.5 (optimum pH), preferably at least 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, or at least 41%.

In additional particular embodiments, the phytase of the invention has a relative activity at pH 6.0 which is at least 10% of the activity at pH 4.5 (optimum pH), preferably at least 15%, 20%, 25%, 30%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, or at least 43%.

The pH profile (phytase activity as a function of pH) is determined on highly purified samples (an SDS poly acryl amide gel should show the presence of only one component). Preferably, the sample of the enzyme is at least 95% pure, as determined by SDS-PAGE. The phytase activity is determined in the pH range of 2.0 to 7.5 using a buffer cocktail (50mM glycine, 50mM acetic acid and 50mM Bis-Tris (Bis-(2-hydroxyethyl) imino-tris

(hydroxymethyl) methan), and on the substrate of sodium phytate at pH5.5 and 37°C. A preferred assay is the assay of Example 3, except that the buffer is replaced with the above indicated buffer. More details are found in Example 6.

5 pH stability

In a particular embodiment, the phytase of the invention has an amended pH stability as compared to a reference phytase. Examples of reference phytases are listed above.

In particular embodiments, the phytase of the invention has a residual phytase activity after incubation for 1½ hours at 40°C and at pH selected from 2.0, 3.0, 4.0, 5.0, 6.0, 10 7.0, and 8.0 of at least 50%, preferably at least 55%, 60%, 65%, 70%, 75%, 80%, 82%, 84%, 86%, 90%, 92%, or at least 94%, relative to the phytase activity at 0 hours (before start of the incubation). In preferred embodiments, the incubation pH is i) 2.0, ii) 3.0, iii) 4.0, iv) 5.0, v) 6.0, vi) 7.0, and vii) 8.0. In even more preferred embodiments, the incubation pH is 2.0, and the residual activity is at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, or at least 15 82%. The phytase is incubated in 0.1M glycine, 0.1M acetic acid, 0.1M Bis-Tris, adjusted to the desired pH. The phytase activity is determined on the substrate of sodium phytate at pH5.5 and 37°C. The activity assay of Example 3 where the buffer is replaced with the above indicated buffer is a preferred assay. More details are found in Example 7.

In still further particular embodiments, the phytase of the invention has a residual 20 phytase activity after incubation for 24 hours at 40°C and at pH selected from 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, and 8.0 of at least 50%, preferably at least 51%, 52%, 54%, 56%, 58%, 60%, 62%, 64%, 66%, 68%, 70%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 82%, 84%, 86%, or at least 88%, relative to the phytase activity at 0 hours (before start of the incubation). In preferred embodiments, the incubation pH is i) 2.0, ii) 3.0, iii) 4.0, iv) 5.0, v) 25 6.0, vi) 7.0, and vii) 8.0. In even more preferred embodiments, the incubation pH is 2.0, and the residual activity is at least 50%, 51%, 52%, 54%, or at least 56%. The phytase is incubated in 0.1M glycine, 0.1M acetic acid, 0.1M Bis-Tris, adjusted to the desired pH. The phytase activity is determined on the substrate of sodium phytate at pH5.5 and 37°C. The activity assay of Example 3 where the buffer is replaced with the above indicated buffer is a 30 preferred assay. More details are found in Example 7.

Temperature profile

In a particular embodiment, the phytase of the invention has an amended pH stability as compared to a reference phytase. Examples of reference phytases are listed above.

35 In particular embodiments, the phytase of the invention has a relative activity at 30°C which is at least 20% of the activity at 60°C (optimum temperature), preferably at least 22%,

24%, 26%, 28%, 30%, 32%, 34%, 36%, 38%, or at least 40%.

In additional particular embodiments, the phytase of the invention has a relative activity at 70°C which is at least 10% of the activity at 60°C (optimum temperature), preferably at least 12%, 14%, 16%, 18%, 20%, 22%, or at least 23%.

5 The temperature profile (phytase activity as a function of temperature) is determined on highly purified samples (an SDS poly acryl amide gel should show the presence of only one component). Preferably, the sample of the enzyme is at least 95% pure, as determined by SDS-PAGE. The phytase activity is determined in the temperature range of 20-90°C at pH4.0, in the alternative at pH5.5. In both cases a 0.25M sodium acetate buffer is used. The
10 activity is determined on the substrate of sodium phytate. A preferred assay is the assay of Example 3, except of course that the temperature differs, and if desired also the pH, cf. above. More details are found in Example 8.

Thermostability

15 In a particular embodiment, the phytase of the invention has an amended, preferably improved, thermostability as compared to a reference phytase. Examples of reference phytases are listed above.

In particular embodiments, the phytase of the invention has a residual phytase activity after incubation for a desired time (such as 15 minutes, 30 minutes, 1 hour, 1½
20 hours, or 2 hours) at an elevated temperature (such as 65, 70, 75, 80, 85, 90, or 95°C) at pH 4.5 of at least 50%, preferably at least 55%, 60%, 65%, 70%, 75%, 80%, 82%, 84%, 86%, 90%, 92%, or at least 94%, relative to the phytase activity at 0 hours (before start of the incubation). In a preferred embodiment, the incubation time is 1 hour, the pH is 4.5 and the temperature 80°C. The phytase is incubated in 0.25 M sodium acetate. The phytase activity
25 is determined on the substrate of sodium phytate at pH5.5 and 37°C. The activity assay of Example 3 is a preferred assay.

In the alternative, Differential Scanning Calorimetry (DSC) measurements may be used to determine the denaturation temperature, Td, of the purified phytase protein. The Td is indicative of the thermostability of the protein: The higher the Td, the higher the
30 thermostability. DSC measurements may be performed at various pH values, e.g. using the VP-DSC from Micro Cal. Scans are performed at a constant scan rate of 1.5°C/min from 20-90°C. Preferred pH values are 4.0 and 5.5, preferably 4.0. Before running the DSC, the phytases are desalted, e.g. using NAP-5 columns (Pharmacia) equilibrated in appropriate buffers (e.g. 25mM sodium acetate pH 4.0; 0.1M sodium acetate, pH5.5). Data-handling is
35 performed using the MicroCal Origin software (version 4.10), and the denaturation temperature, Td (also called the melting temperature, Tm) is defined as the temperature at

the apex of the peak in the thermogram.

In particular embodiments, the phytase of the invention has a Td, which may be determined as described above, of at least 60°C. In still further particular embodiments, the Td is at least 61, 62, 64, 66, 68, 70, 72, 74, 76, 78, or at least 80°C.

5

Performance in animal feed

In a particular embodiment the phytase of the invention has an improved performance in animal feed as compared to a reference phytase. Examples of reference phytases are listed above.

10 The performance in animal feed may be determined using an in vitro model simulating the gastro-intestinal conditions in a monogastric animal, e.g. as follows:

Feed samples composed of 30% soybean meal and 70% maize meal with added CaCl₂ to a concentration of 5 g calcium per kg feed are prepared and pre-incubated at 40°C and pH 3.0 for 30 minutes followed by addition of pepsin (3000 U/g feed) and suitable dosages of the phytases (identical dosages are used for all phytases to be tested to allow comparison), for example between 0.25 to 0.75 phytase units FYT/g feed. A blank with no phytase activity is also included as reference. The samples are then incubated at 40°C and pH 3.0 for 60 minutes followed by pH 4.0 for 30 minutes.

20 The reactions are stopped and phytic acid and inositol-phosphates extracted by addition of HCl to a final concentration of 0.5 M and incubation at 40°C for 2 hours, followed by one freeze-thaw cycle and 1 hour incubation at 40°C.

Phytic acid and inositol-phosphates are separated by high performance ion chromatography as described by Chen et al in Journal of Chromatography A (2003) vol. 1018, pp. 41-52 and quantified as described by Skoglund et al in J. Agric. Food Chem. (1997), vol. 45, pp. 431-436.

Released phosphorous is then calculated as the difference in inositol-phosphate bound phosphorous (IP-P) between phytase-treated and non-treated samples. The relative performance of a specified phytase is calculated as the percentage of the phosphorous released relative to the desired reference phytase.

30 In particular embodiments the relative performance in vitro of the phytase of the invention is at least 105%, preferably at least 110, 120, 130, 140, 150, 160, 170, 180, 190, or at least 200%.

Polynucleotides, Nucleic Acid Constructs, and Expression Vectors

35 The present invention also relates to polynucleotides comprising a nucleotide sequence which encode a polypeptide of the present invention. The polynucleotides are

preferably substantially pure, or isolated, which refers to a polynucleotide preparation free of other extraneous or unwanted nucleotides and in a form suitable for use within genetically engineered protein production systems. Thus, a substantially pure polynucleotide contains at most 10%, preferably at most 8%, more preferably at most 6%, more preferably at most 5%, more preferably at most 4%, more preferably at most 3%, even more preferably at most 2%, most preferably at most 1%, and even most preferably at most 0.5% by weight of other polynucleotide material with which it is natively associated. A substantially pure polynucleotide may, however, include naturally occurring 5' and 3' untranslated regions, such as promoters and terminators. It is preferred that the substantially pure polynucleotide is at least 90% pure, preferably at least 92% pure, more preferably at least 94% pure, more preferably at least 95% pure, more preferably at least 96% pure, more preferably at least 97% pure, even more preferably at least 98% pure, most preferably at least 99%, and even most preferably at least 99.5% pure by weight. The polynucleotides of the present invention are preferably in a substantially pure form. In particular, it is preferred that the polynucleotides disclosed herein are in "essentially pure form", i.e., that the polynucleotide preparation is essentially free of other polynucleotide material with which it is natively associated. Herein, the term "substantially pure polynucleotide" is synonymous with the terms "isolated polynucleotide" and "polynucleotide in isolated form." The polynucleotides may be of genomic, cDNA, RNA, semisynthetic, synthetic origin, or any combinations thereof.

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

The term "nucleic acid construct" as used herein refers to a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or which is modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature. The term nucleic acid construct is synonymous with the term "expression cassette" when the nucleic acid construct contains the control sequences required for expression of a coding sequence of the present invention.

The term "control sequences" is defined herein to include all components, which are necessary or advantageous for the expression of a polynucleotide encoding a polypeptide of the present invention. Each control sequence may be native or foreign to the nucleotide sequence encoding the polypeptide. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be

provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the nucleotide sequence encoding a polypeptide.

5 The control sequence may be an appropriate promoter sequence, a nucleotide sequence which is recognized by a host cell for expression of a polynucleotide encoding a polypeptide of the present invention.

10 Examples of suitable promoters for directing the transcription of the nucleic acid constructs of the present invention, especially in a bacterial host cell, are the promoters obtained from the *E. coli* lac operon, *Streptomyces coelicolor* agarase gene (*dagA*), *Bacillus subtilis* levansucrase gene (*sacB*), *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus stearothermophilus* maltogenic amylase gene (*amyM*), *Bacillus amyloliquefaciens* alpha-amylase gene (*amyQ*), *Bacillus licheniformis* penicillinase gene (*penP*), *Bacillus subtilis* *xylA* and *xylB* genes, and prokaryotic beta-lactamase gene (Villa-Kamaroff et al., 1978, Proceedings of the National Academy of Sciences USA 75: 3727-3731), as well as
15 the *tac* promoter (DeBoer et al., 1983, Proceedings of the National Academy of Sciences USA 80: 21-25). Further promoters are described in "Useful proteins from recombinant bacteria" in Scientific American, 1980, 242: 74-94; and in Sambrook et al., 1989, supra.

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

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

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

The control sequence may also be a signal peptide coding region that codes for an amino acid sequence linked to the amino terminus of a polypeptide and directs the encoded polypeptide into the cell's secretory pathway. The 5' end of the coding sequence of the
35 nucleotide sequence may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region which encodes the secreted polypeptide. Alternatively, the 5' end of the coding sequence may contain a signal

peptide coding region which is foreign to the coding sequence. The foreign signal peptide coding region may be required where the coding sequence does not naturally contain a signal peptide coding region. Alternatively, the foreign signal peptide coding region may simply replace the natural signal peptide coding region in order to enhance secretion of the polypeptide. However, any signal peptide coding region which directs the expressed polypeptide into the secretory pathway of a host cell of choice may be used in the present invention.

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

In a preferred aspect, the signal peptide coding region is nucleotides 1 to 81 of SEQ ID NO:7 which encode amino acids 1 to 27 of SEQ ID NO:8.

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

Where both signal peptide and propeptide regions are present at the amino terminus of a polypeptide, the propeptide region is positioned next to the amino terminus of a polypeptide and the signal peptide region is positioned next to the amino terminus of the propeptide region.

It may also be desirable to add regulatory sequences which allow the regulation of the expression of the polypeptide relative to the growth of the host cell. Examples of regulatory systems are those which cause the expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory systems in prokaryotic systems include the lac, tac, and trp operator systems. Other examples of regulatory sequences are those which allow for gene amplification. In these cases, the nucleotide sequence encoding the polypeptide would be operably linked with the regulatory sequence.

The present invention also relates to recombinant expression vectors comprising a

polynucleotide of the present invention, a promoter, and transcriptional and translational stop signals. The term "expression vector" is defined herein as a linear or circular DNA molecule that comprises a polynucleotide encoding a polypeptide of the invention, and which is operably linked to additional nucleotides that provide for its expression. The various
5 nucleic acids and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleotide sequence encoding the polypeptide at such sites. Alternatively, a nucleotide sequence of the present invention may be expressed by inserting the nucleotide sequence or a nucleic acid construct comprising the sequence
10 into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

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

The recombinant expression vector may be any vector (e.g., a plasmid or virus) which can be conveniently subjected to recombinant DNA procedures and can bring about expression of the nucleotide sequence. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The
20 vectors may be linear or closed circular plasmids.

A vector comprising a polynucleotide of the present invention is introduced into a host cell so that the vector is maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector as described earlier. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur
25 during replication. The choice of a host cell will to a large extent depend upon the gene encoding the polypeptide and its source.

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

The vectors of the present invention preferably contain one or more selectable markers which permit easy selection of transformed cells. A selectable marker is a gene the

product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like.

A conditionally essential gene may function as a non-antibiotic selectable marker. Non-limiting examples of bacterial conditionally essential non-antibiotic selectable markers are the *dal* genes from *Bacillus subtilis*, *Bacillus licheniformis*, or other *Bacilli*, that are only essential when the bacterium is cultivated in the absence of D-alanine. Also the genes encoding enzymes involved in the turnover of UDP-galactose can function as conditionally essential markers in a cell when the cell is grown in the presence of galactose or grown in a medium which gives rise to the presence of galactose. Non-limiting examples of such genes are those from *B. subtilis* or *B. licheniformis* encoding UTP-dependent phosphorylase (EC 2.7.7.10), UDP-glucose-dependent uridylyltransferase (EC 2.7.7.12), or UDP-galactose epimerase (EC 5.1.3.2). Also a xylose isomerase gene such as *xylA*, of *Bacilli* can be used as selectable markers in cells grown in minimal medium with xylose as sole carbon source. The genes necessary for utilizing gluconate, *gntK*, and *gntP* can also be used as selectable markers in cells grown in minimal medium with gluconate as sole carbon source. Other examples of conditionally essential genes are known in the art. Antibiotic selectable markers confer antibiotic resistance to such antibiotics as ampicillin, kanamycin, chloramphenicol, erythromycin, tetracycline, neomycin, hygromycin or methotrexate.

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

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

For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. The origin of

replication may be any plasmid replicator mediating autonomous replication which functions in a cell. The term "origin of replication" or "plasmid replicator" is defined herein as a nucleotide sequence that enables a plasmid or vector to replicate in vivo.

5 Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, and pACYC184 permitting replication in *E. coli*, and pUB110, pE194, pTA1060, and pAM β 1 permitting replication in *Bacillus*.

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

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

Host cells

20 The present invention also relates to recombinant host cells, comprising a polynucleotide of the present invention, which are advantageously used in the recombinant production of the polypeptides. The term "host cell", as used herein, includes any cell type which is susceptible to transformation, transfection, transduction, and the like with a nucleic acid construct comprising a polynucleotide of the present invention.

25 The host cell may be a unicellular microorganism, e.g., a prokaryote, or a non-unicellular microorganism, e.g., a eukaryote.

Useful unicellular microorganisms are bacterial cells such as gram positive bacteria including, but not limited to, a *Bacillus* cell, e.g., *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus*
30 *stearothermophilus*, *Bacillus subtilis*, and *Bacillus thuringiensis*; or a *Streptomyces* cell, e.g., *Streptomyces lividans* and *Streptomyces murinus*, or gram negative bacteria such as *E. coli* and *Pseudomonas* sp. In a preferred aspect, the bacterial host cell is a *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus stearothermophilus*, or *Bacillus subtilis* cell. In another preferred aspect, the *Bacillus* cell is an alkalophilic *Bacillus*.

35 The introduction of a vector into a bacterial host cell may, for instance, be effected by protoplast transformation (see, e.g., Chang and Cohen, 1979, *Molecular General Genetics* 168: 111-115), using competent cells (see, e.g., Young and Spizizin, 1961, *Journal of*

Bacteriology 81: 823-829, or Dubnau and Davidoff-Abelson, 1971, Journal of Molecular Biology 56: 209-221), electroporation (see, e.g., Shigekawa and Dower, 1988, Biotechniques 6: 742-751), or conjugation (see, e.g., Koehler and Thorne, 1987, Journal of Bacteriology 169: 5771-5278).

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

Methods of Production

10 The present invention also relates to methods for producing a polypeptide of the present invention, comprising (a) cultivating a cell, which in its wild-type form is capable of producing the polypeptide, under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide. Preferably, the cell is of the genus *Buttiauxella*, and more preferably *Buttiauxella agrestis* or *Buttiauxella gaviniae*.

15 The present invention also relates to methods for producing a polypeptide of the present invention, comprising (a) cultivating a host cell under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide.

20 In the production methods of the present invention, the cells are cultivated in a nutrient medium suitable for production of the polypeptide using methods well known in the art. For example, the cell may be cultivated by shake flask cultivation, and small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the polypeptide to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (e.g., in catalogues of the American Type Culture Collection). If the polypeptide is secreted into the nutrient medium, the polypeptide can be recovered directly from the medium. If the polypeptide is not secreted, it can be recovered from cell lysates.

25 The polypeptides may be detected using methods known in the art that are specific for the polypeptides. These detection methods may include use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. For example, an enzyme assay may be used to determine the activity of the polypeptide as described herein.

30 The resulting polypeptide may be recovered using methods known in the art. For example, the polypeptide may be recovered from the nutrient medium by conventional procedures including, but not limited to, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation.

 The polypeptides of the present invention may be purified by a variety of procedures

known in the art including, but not limited to, chromatography (e.g., ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (e.g., preparative isoelectric focusing), differential solubility (e.g., ammonium sulfate precipitation), SDS-PAGE, or extraction (see, e.g., Protein Purification, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989). A typical purification scheme may include centrifugation, germ filtration, ammonium sulphate precipitation (using e.g. 80% ammonium sulphate saturation), centrifugation, re-suspension of pellets in 50mM sodium acetate buffer pH4.5, filtration, dialysis against 50mM sodium acetate buffer and cation exchange chromatography (S-sepharose, loading with 50mM sodium acetate pH4.5, eluting with a linear salt gradient (50mM sodium acetate pH4.5 + 1M NaCl)).

Transgenic Plants

The present invention also relates to a transgenic plant, plant part, or plant cell which has been transformed with a nucleotide sequence encoding a polypeptide having phytase activity of the present invention so as to express and produce the polypeptide in recoverable quantities. The polypeptide may be recovered from the plant or plant part. Alternatively, the plant or plant part containing the recombinant polypeptide may be used as such for improving the quality of a food or feed, e.g., improving nutritional value, palatability, and rheological properties, or to destroy an antinutritive factor.

In a particular embodiment, the polypeptide is targeted to the endosperm storage vacuoles in seeds. This can be obtained by synthesizing it as a precursor with a suitable signal peptide, see Horvath et al in PNAS, Feb. 15, 2000, vol. 97, no. 4, p. 1914-1919.

The transgenic plant can be dicotyledonous (a dicot) or monocotyledonous (a monocot) or engineered variants thereof. Examples of monocot plants are grasses, such as meadow grass (blue grass, *Poa*), forage grass such as *Festuca*, *Lolium*, temperate grass, such as *Agrostis*, and cereals, e.g., wheat, oats, rye, barley, rice, sorghum, triticale (stabilized hybrid of wheat (*Triticum*) and rye (*Secale*), and maize (corn). Examples of dicot plants are tobacco, legumes, such as sunflower (*Helianthus*), cotton (*Gossypium*), lupins, potato, sugar beet, pea, bean and soybean, and cruciferous plants (family *Brassicaceae*), such as cauliflower, rape seed, and the closely related model organism *Arabidopsis thaliana*. Low-phytate plants as described e.g. in US patent no. 5,689,054 and US patent no. 6,111,168 are examples of engineered plants.

Examples of plant parts are stem, callus, leaves, root, fruits, seeds, and tubers, as well as the individual tissues comprising these parts, e.g. epidermis, mesophyll, parenchyma, vascular tissues, meristems. Also specific plant cell compartments, such as chloroplast, apoplast, mitochondria, vacuole, peroxisomes, and cytoplasm are considered to be a plant part. Furthermore, any plant cell, whatever the tissue origin, is considered to be a

plant part. Likewise, plant parts such as specific tissues and cells isolated to facilitate the utilisation of the invention are also considered plant parts, e.g. embryos, endosperms, aleurone and seed coats.

Also included within the scope of the present invention are the progeny of such plants, plant parts and plant cells.

The transgenic plant or plant cell expressing a polypeptide of the present invention may be constructed in accordance with methods known in the art. Briefly, the plant or plant cell is constructed by incorporating one or more expression constructs encoding a polypeptide of the present invention into the plant host genome and propagating the resulting modified plant or plant cell into a transgenic plant or plant cell.

Conveniently, the expression construct is a nucleic acid construct which comprises a nucleic acid sequence encoding a polypeptide of the present invention operably linked with appropriate regulatory sequences required for expression of the nucleic acid sequence in the plant or plant part of choice. Furthermore, the expression construct may comprise a selectable marker useful for identifying host cells into which the expression construct has been integrated and DNA sequences necessary for introduction of the construct into the plant in question (the latter depends on the DNA introduction method to be used).

The choice of regulatory sequences, such as promoter and terminator sequences and optionally signal or transit sequences are determined, for example, on the basis of when, where, and how the polypeptide is desired to be expressed. For instance, the expression of the gene encoding a polypeptide of the present invention may be constitutive or inducible, or may be developmental, stage or tissue specific, and the gene product may be targeted to a specific cell compartment, tissue or plant part such as seeds or leaves. Regulatory sequences are, for example, described by Tague et al., 1988, Plant Physiology 86: 506.

For constitutive expression, the following promoters may be used: The 35S-CaMV promoter (Franck et al., 1980, Cell 21: 285-294), the maize ubiquitin 1 (Christensen AH, Sharrock RA and Quail 1992. Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation), or the rice actin 1 promoter (Plant Mo. Biol. 18, 675-689.; Zhang W, McElroy D. and Wu R 1991, Analysis of rice Act1 5' region activity in transgenic rice plants. Plant Cell 3, 1155-1165). Organ-specific promoters may be, for example, a promoter from storage sink tissues such as seeds, potato tubers, and fruits (Edwards & Coruzzi, 1990, Ann. Rev. Genet. 24: 275-303), or from metabolic sink tissues such as meristems (Ito et al., 1994, Plant Mol. Biol. 24: 863-878), a seed specific promoter such as the glutelin, prolamin, globulin, or albumin promoter from rice (Wu et al., 1998, Plant and Cell Physiology 39: 885-889), a Vicia faba promoter from the legumin B4 and the unknown seed protein gene from

Vicia faba (Conrad et al., 1998, Journal of Plant Physiology 152: 708-711), a promoter from a seed oil body protein (Chen et al., 1998, Plant and Cell Physiology 39: 935-941), the storage protein napA promoter from Brassica napus, or any other seed specific promoter known in the art, e.g., as described in WO 91/14772. Furthermore, the promoter may be a
5 leaf specific promoter such as the rbcS promoter from rice or tomato (Kyoizuka et al., 1993, Plant Physiology 102: 991-1000, the chlorella virus adenine methyltransferase gene promoter (Mitra and Higgins, 1994, Plant Molecular Biology 26: 85-93), or the aldP gene promoter from rice (Kagaya et al., 1995, Molecular and General Genetics 248: 668-674), or a wound inducible promoter such as the potato pin2 promoter (Xu et al., 1993, Plant
10 Molecular Biology 22: 573-588). Likewise, the promoter may be inducible by abiotic treatments such as temperature, drought or alterations in salinity or inducible by exogenously applied substances that activate the promoter, e.g. ethanol, oestrogens, plant hormones like ethylene, abscisic acid, gibberellic acid, and/or heavy metals.

A promoter enhancer element may also be used to achieve higher expression of
15 the polypeptide in the plant. For instance, the promoter enhancer element may be an intron which is placed between the promoter and the nucleotide sequence encoding a polypeptide of the present invention. For instance, Xu et al., 1993, supra disclose the use of the first intron of the rice actin 1 gene to enhance expression.

Still further, the codon usage may be optimized for the plant species in question to
20 improve expression (see Horvath et al referred to above).

The selectable marker gene and any other parts of the expression construct may be chosen from those available in the art.

The nucleic acid construct is incorporated into the plant genome according to conventional techniques known in the art, including Agrobacterium-mediated transformation,
25 virus-mediated transformation, microinjection, particle bombardment, biolistic transformation, and electroporation (Gasser et al., 1990, Science 244: 1293; Potrykus, 1990, Bio/Technology 8: 535; Shimamoto et al., 1989, Nature 338: 274).

Presently, Agrobacterium tumefaciens-mediated gene transfer is the method of choice for generating transgenic dicots (for a review, see Hooykas and Schilperoort, 1992,
30 Plant Molecular Biology 19: 15-38), and it can also be used for transforming monocots, although other transformation methods are more often used for these plants. Presently, the method of choice for generating transgenic monocots, supplementing the Agrobacterium approach, is particle bombardment (microscopic gold or tungsten particles coated with the transforming DNA) of embryonic calli or developing embryos (Christou, 1992, Plant Journal
35 2: 275-281; Shimamoto, 1994, Current Opinion Biotechnology 5: 158-162; Vasil et al., 1992, Bio/Technology 10: 667-674). An alternative method for transformation of monocots is

based on protoplast transformation as described by Omirulleh et al., 1993, Plant Molecular Biology 21: 415-428.

Following transformation, the transformants having incorporated therein the expression construct are selected and regenerated into whole plants according to methods well-known in the art. Often the transformation procedure is designed for the selective elimination of selection genes either during regeneration or in the following generations by using e.g. co-transformation with two separate T-DNA constructs or site specific excision of the selection gene by a specific recombinase.

The present invention also relates to methods for producing a polypeptide of the present invention comprising (a) cultivating a transgenic plant or a plant cell comprising a nucleic acid sequence encoding a polypeptide having phytase activity of the present invention under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide.

Transgenic Animals

The present invention also relates to a transgenic, non-human animal and products or elements thereof, examples of which are body fluids such as milk and blood, organs, flesh, and animal cells. Techniques for expressing proteins, e.g. in mammalian cells, are known in the art, see e.g. the handbook Protein Expression: A Practical Approach, Higgins and Hames (eds), Oxford University Press (1999), and the three other handbooks in this series relating to Gene Transcription, RNA processing, and Post-translational Processing. Generally speaking, to prepare a transgenic animal, selected cells of a selected animal are transformed with a nucleic acid sequence encoding a polypeptide having phytase activity of the present invention so as to express and produce the polypeptide. The polypeptide may be recovered from the animal, e.g. from the milk of female animals, or the polypeptide may be expressed to the benefit of the animal itself, e.g. to assist the animal's digestion. Examples of animals are mentioned below in the section headed Animal Feed.

To produce a transgenic animal with a view to recovering the polypeptide from the milk of the animal, a gene encoding the polypeptide may be inserted into the fertilized eggs of an animal in question, e.g. by use of a transgene expression vector which comprises a suitable milk protein promoter, and the gene encoding the polypeptide. The transgene expression vector is microinjected into fertilized eggs, and preferably permanently integrated into the chromosome. Once the egg begins to grow and divide, the potential embryo is implanted into a surrogate mother, and animals carrying the transgene are identified. The resulting animal can then be multiplied by conventional breeding. The polypeptide may be purified from the animal's milk, see e.g. Meade, H.M. et al (1999): Expression of

recombinant proteins in the milk of transgenic animals, Gene expression systems: Using nature for the art of expression. J. M. Fernandez and J. P. Hoeffler (eds.), Academic Press.

5 In the alternative, in order to produce a transgenic non-human animal that carries in the genome of its somatic and/or germ cells a nucleic acid sequence including a heterologous transgene construct including a transgene encoding the polypeptide, the transgene may be operably linked to a first regulatory sequence for salivary gland specific expression of the polypeptide, as disclosed in WO 00/064247.

Compositions and Uses

10 In still further aspects, the present invention relates to compositions comprising a polypeptide of the present invention, as well as methods of using these.

The polypeptide compositions may be prepared in accordance with methods known in the art and may be in the form of a liquid or a dry composition. For instance, the polypeptide composition may be in the form of granulates or microgranulates. The
15 polypeptide to be included in the composition may be stabilized in accordance with methods known in the art.

The phytase of the invention can be used for degradation, in any industrial context, of, for example, phytate, phytic acid, and/or the mono-, di-, tri-, tetra- and/or penta-phosphates of myo-inositol. It is well known that the phosphate moieties of these
20 compounds chelates divalent and trivalent cations such as metal ions, i.a. the nutritionally essential ions of calcium, iron, zinc and magnesium as well as the trace minerals manganese, copper and molybdenum. Besides, the phytic acid also to a certain extent binds proteins by electrostatic interaction.

Accordingly, preferred uses of the polypeptides of the invention are in animal feed
25 preparations (including human food) or in additives for such preparations.

In a particular embodiment, the polypeptide of the invention can be used for improving the nutritional value of an animal feed. Non-limiting examples of improving the nutritional value of animal feed (including human food), are: Improving feed digestibility; promoting growth of the animal; improving feed utilization; improving bio-availability of
30 proteins; increasing the level of digestible phosphate; improving the release and/or degradation of phytate; improving bio-availability of trace minerals; improving bio-availability of macro minerals; eliminating the need for adding supplemental phosphate, trace minerals, and/or macro minerals; and/or improving egg shell quality. The nutritional value of the feed is therefore increased, and the growth rate and/or weight gain and/or feed conversion (i.e.
35 the weight of ingested feed relative to weight gain) of the animal may be improved.

Furthermore, the polypeptide of the invention can be used for reducing phytate level of manure.

Animals, Animal Feed, and Animal Feed Additives

The term animal includes all animals, including human beings. Examples of animals are non-ruminants, and ruminants. Ruminant animals include, for example, animals such as
5 sheep, goat, and cattle, e.g. cow such as beef cattle and dairy cows. In a particular embodiment, the animal is a non-ruminant animal. Non-ruminant animals include mono-gastric animals, e.g. pig or swine (including, but not limited to, piglets, growing pigs, and sows); poultry such as turkeys, ducks and chickens (including but not limited to broiler chicks, layers); fish (including but not limited to salmon, trout, tilapia, catfish and carp); and
10 crustaceans (including but not limited to shrimp and prawn).

The term feed or feed composition means any compound, preparation, mixture, or composition suitable for, or intended for intake by an animal.

In the use according to the invention the polypeptide can be fed to the animal before, after, or simultaneously with the diet. The latter is preferred.

15 In a particular embodiment, the polypeptide, in the form in which it is added to the feed, or when being included in a feed additive, is substantially pure. In a particular embodiment it is well-defined. The term "well-defined" means that the phytase preparation is at least 50% pure as determined by Size-exclusion chromatography (see Example 12 of WO 01/58275). In other particular embodiments the phytase preparation is at least 60, 70, 80,
20 85, 88, 90, 92, 94, or at least 95% pure as determined by this method.

A substantially pure, and/or well-defined polypeptide preparation is advantageous. For instance, it is much easier to dose correctly to the feed a polypeptide that is essentially free from interfering or contaminating other polypeptides. The term dose correctly refers in particular to the objective of obtaining consistent and constant results, and the capability of
25 optimising dosage based upon the desired effect.

For the use in animal feed, however, the phytase polypeptide of the invention need not be that pure; it may e.g. include other polypeptides, in which case it could be termed a phytase preparation.

The phytase preparation can be (a) added directly to the feed (or used directly in a
30 treatment process of proteins), or (b) it can be used in the production of one or more intermediate compositions such as feed additives or premixes that is subsequently added to the feed (or used in a treatment process). The degree of purity described above refers to the purity of the original polypeptide preparation, whether used according to (a) or (b) above.

Polypeptide preparations with purities of this order of magnitude are in particular
35 obtainable using recombinant methods of production, whereas they are not so easily obtained and also subject to a much higher batch-to-batch variation when the polypeptide is produced by traditional fermentation methods.

Such polypeptide preparation may of course be mixed with other polypeptides.

The polypeptide can be added to the feed in any form, be it as a relatively pure polypeptide, or in admixture with other components intended for addition to animal feed, i.e. in the form of animal feed additives, such as the so-called pre-mixes for animal feed.

5 In a further aspect the present invention relates to compositions for use in animal feed, such as animal feed, and animal feed additives, e.g. premixes.

Apart from the polypeptide of the invention, the animal feed additives of the invention contain at least one fat-soluble vitamin, and/or at least one water soluble vitamin, and/or at least one trace mineral. The feed additive may also contain at least one macro mineral.

10 Further, optional, feed-additive ingredients are colouring agents, e.g. carotenoids such as beta-carotene, astaxanthin, and lutein; aroma compounds; stabilisers; antimicrobial peptides; polyunsaturated fatty acids; reactive oxygen generating species; and/or at least one other polypeptide selected from amongst phytase (EC 3.1.3.8 or 3.1.3.26); phosphatase (EC 3.1.3.1; EC 3.1.3.2; EC 3.1.3.39); xylanase (EC 3.2.1.8); galactanase (EC 3.2.1.89);
15 alpha-galactosidase (EC 3.2.1.22); protease (EC 3.4.-.-), phospholipase A1 (EC 3.1.1.32); phospholipase A2 (EC 3.1.1.4); lysophospholipase (EC 3.1.1.5); phospholipase C (3.1.4.3); phospholipase D (EC 3.1.4.4); amylase such as, for example, alpha-amylase (EC 3.2.1.1); and/or beta-glucanase (EC 3.2.1.4 or EC 3.2.1.6).

20 In a particular embodiment these other polypeptides are well-defined (as defined above for phytase preparations).

The phytase of the invention may also be combined with other phytases, for example ascomycete phytases such as *Aspergillus* phytases, for example derived from *Aspergillus ficuum*, *Aspergillus niger*, or *Aspergillus awamori*; or basidiomycete phytases, for example derived from *Peniophora lycii*, *Agrocybe pediades*, *Trametes pubescens*, or *Paxillus involutus*; or derivatives, fragments or variants thereof which have phytase activity.

Thus, in preferred embodiments of the use in animal feed of the invention, and in preferred embodiments of the animal feed additive and the animal feed of the invention, the phytase of the invention is combined with such phytases.

30 Examples of antimicrobial peptides (AMP's) are CAP18, Leucocin A, Tritrpticin, Protegrin-1, Thanatin, Defensin, Lactoferrin, Lactoferricin, and Ovispirin such as Novispirin (Robert Lehrer, 2000), Plectasins, and Statins, including the compounds and polypeptides disclosed in WO 03/044049 and WO 03/048148, as well as variants or fragments of the above that retain antimicrobial activity.

35 Examples of antifungal polypeptides (AFP's) are the *Aspergillus giganteus*, and *Aspergillus niger* peptides, as well as variants and fragments thereof which retain antifungal activity, as disclosed in WO 94/01459 and WO 02/090384.

Examples of polyunsaturated fatty acids are C18, C20 and C22 polyunsaturated fatty

acids, such as arachidonic acid, docosohexaenoic acid, eicosapentaenoic acid and gamma-linoleic acid.

Examples of reactive oxygen generating species are chemicals such as perborate, persulphate, or percarbonate; and polypeptides such as an oxidase, an oxygenase or a syntethase.

Usually fat- and water-soluble vitamins, as well as trace minerals form part of a so-called premix intended for addition to the feed, whereas macro minerals are usually separately added to the feed. Either of these composition types, when enriched with a polypeptide of the invention, is an animal feed additive of the invention.

In a particular embodiment, the animal feed additive of the invention is intended for being included (or prescribed as having to be included) in animal diets or feed at levels of 0.01 to 10.0%; more particularly 0.05 to 5.0%; or 0.2 to 1.0% (% meaning g additive per 100 g feed). This is so in particular for premixes.

The following are non-exclusive lists of examples of these components:

Examples of fat-soluble vitamins are vitamin A, vitamin D3, vitamin E, and vitamin K, e.g. vitamin K3.

Examples of water-soluble vitamins are vitamin B12, biotin and choline, vitamin B1, vitamin B2, vitamin B6, niacin, folic acid and panthothenate, e.g. Ca-D-panthothenate.

Examples of trace minerals are manganese, zinc, iron, copper, iodine, selenium, and cobalt.

Examples of macro minerals are calcium, phosphorus and sodium.

The nutritional requirements of these components (exemplified with poultry and piglets/pigs) are listed in Table A of WO 01/58275. Nutritional requirement means that these components should be provided in the diet in the concentrations indicated.

In the alternative, the animal feed additive of the invention comprises at least one of the individual components specified in Table A of WO 01/58275. At least one means either of, one or more of, one, or two, or three, or four and so forth up to all thirteen, or up to all fifteen individual components. More specifically, this at least one individual component is included in the additive of the invention in such an amount as to provide an in-feed-concentration within the range indicated in column four, or column five, or column six of Table A.

The present invention also relates to animal feed compositions. Animal feed compositions or diets have a relatively high content of protein. Poultry and pig diets can be characterised as indicated in Table B of WO 01/58275, columns 2-3. Fish diets can be characterised as indicated in column 4 of this Table B. Furthermore such fish diets usually have a crude fat content of 200-310 g/kg.

WO 01/58275 corresponds to US 09/779334.

An animal feed composition according to the invention has a crude protein content of 50-800 g/kg, and furthermore comprises at least one polypeptide as claimed herein.

5 Furthermore, or in the alternative (to the crude protein content indicated above), the animal feed composition of the invention has a content of metabolisable energy of 10-30 MJ/kg; and/or a content of calcium of 0.1-200 g/kg; and/or a content of available phosphorus of 0.1-200 g/kg; and/or a content of methionine of 0.1-100 g/kg; and/or a content of methionine plus cysteine of 0.1-150 g/kg; and/or a content of lysine of 0.5-50 g/kg.

10 In particular embodiments, the content of metabolisable energy, crude protein, calcium, phosphorus, methionine, methionine plus cysteine, and/or lysine is within any one of ranges 2, 3, 4 or 5 in Table B of WO 01/58275 (R. 2-5).

Crude protein is calculated as nitrogen (N) multiplied by a factor 6.25, i.e. Crude protein (g/kg)= N (g/kg) x 6.25. The nitrogen content is determined by the Kjeldahl method
15 (A.O.A.C., 1984, Official Methods of Analysis 14th ed., Association of Official Analytical Chemists, Washington DC).

Metabolisable energy can be calculated on the basis of the NRC publication Nutrient requirements in swine, ninth revised edition 1988, subcommittee on swine nutrition, committee on animal nutrition, board of agriculture, national research council. National
20 Academy Press, Washington, D.C., pp. 2-6, and the European Table of Energy Values for Poultry Feed-stuffs, Spelderholt centre for poultry research and extension, 7361 DA Beekbergen, The Netherlands. Grafisch bedrijf Ponsen & Iooijen bv, Wageningen. ISBN 90-71463-12-5.

The dietary content of calcium, available phosphorus and amino acids in complete
25 animal diets is calculated on the basis of feed tables such as Veevoedertabel 1997, gegevens over chemische samenstelling, verteerbaarheid en voederwaarde van voedermiddelen, Central Veevoederbureau, Runderweg 6, 8219 pk Lelystad. ISBN 90-72839-13-7.

In a particular embodiment, the animal feed composition of the invention contains at
30 least one protein. The protein may be an animal protein, such as meat and bone meal, and/or fish meal; or it may be a vegetable protein. The term vegetable proteins as used herein refers to any compound, composition, preparation or mixture that includes at least one protein derived from or originating from a vegetable, including modified proteins and protein-derivatives. In particular embodiments, the protein content of the vegetable proteins
35 is at least 10, 20, 30, 40, 50, or 60% (w/w).

Vegetable proteins may be derived from vegetable protein sources, such as legumes and cereals, for example materials from plants of the families Fabaceae (Leguminosae),

Cruciferaeae, Chenopodiaceae, and Poaceae, such as soy bean meal, lupin meal and rapeseed meal.

In a particular embodiment, the vegetable protein source is material from one or more plants of the family Fabaceae, e.g. soybean, lupine, pea, or bean.

5 In another particular embodiment, the vegetable protein source is material from one or more plants of the family Chenopodiaceae, e.g. beet, sugar beet, spinach or quinoa.

Other examples of vegetable protein sources are rapeseed, sunflower seed, cotton seed, and cabbage.

Soybean is a preferred vegetable protein source.

10 Other examples of vegetable protein sources are cereals such as barley, wheat, rye, oat, maize (corn), rice, triticale, and sorghum.

In still further particular embodiments, the animal feed composition of the invention contains 0-80% maize; and/or 0-80% sorghum; and/or 0-70% wheat; and/or 0-70% Barley; and/or 0-30% oats; and/or 0-40% soybean meal; and/or 0-25% fish meal; and/or 0-25%
15 meat and bone meal; and/or 0-20% whey.

Animal diets can e.g. be manufactured as mash feed (non pelleted) or pelleted feed. Typically, the milled feed-stuffs are mixed and sufficient amounts of essential vitamins and minerals are added according to the specifications for the species in question. Polypeptides can be added as solid or liquid polypeptide formulations. For example, a solid polypeptide
20 formulation is typically added before or during the mixing step; and a liquid polypeptide preparation is typically added after the pelleting step. The polypeptide may also be incorporated in a feed additive or premix.

The final polypeptide concentration in the diet is within the range of 0.01-200 mg polypeptide protein per kg diet, for example in the range of 5-30 mg polypeptide protein per
25 kg animal diet.

The phytase of the invention should of course be applied in an effective amount, i.e. in an amount adequate for improving solubilisation and/or improving nutritional value of feed. It is at present contemplated that the polypeptide is administered in one or more of the following amounts (dosage ranges): 0.01-200; 0.01-100; 0.5-100; 1-50; 5-100; 10-100; 0.05-
30 50; or 0.10-10 – all these ranges being in mg phytase polypeptide protein per kg feed (ppm).

For determining mg phytase polypeptide protein per kg feed, the phytase is purified from the feed composition, and the specific activity of the purified phytase is determined using a relevant assay. The phytase activity of the feed composition as such is also determined using the same assay, and on the basis of these two determinations, the dosage
35 in mg phytase protein per kg feed is calculated.

The same principles apply for determining mg phytase polypeptide protein in feed additives. Of course, if a sample is available of the phytase used for preparing the feed

additive or the feed, the specific activity is determined from this sample (no need to purify the phytase from the feed composition or the additive).

Methods For Producing Fermentation Products and Co-Products

Yet another aspect of the present invention relates to the methods for producing a fermentation product/co-product, such as, e.g., ethanol, beer, wine, distillers dried grains (DDG), wherein the fermentation is carried out in the presence of a phytase of the present invention. Examples of fermentation processes include, for example, the processes described in WO 01/62947. Fermentation is carried out using a fermenting microorganism, such as, yeast.

In a particular embodiment, the present invention provides methods for producing ethanol, comprising fermenting (using a fermenting microorganism, such as yeast) a carbohydrate containing material (e.g., starch) in the presence of a phytase of the present invention.

In another embodiment, the present invention provides methods for producing ethanol comprising hydrolyzing starch, e.g., by a liquefaction and/or saccharification process, a raw starch hydrolysis process, fermenting the resulting starch in the presence of a phytase of the present invention, and producing ethanol.

The phytase may be added to the fermentation process at any suitable stage and in any suitable composition, including alone or in combination with other enzymes, such as, one or more alpha-amylases, glucoamylases, proteases, cellulases.

In another embodiment, the present invention provides methods for producing ethanol comprising hydrolyzing biomass, and fermenting the resulting biomass in the presence of a phytase of the present invention.

Examples

Chemicals used as buffers and substrates were commercial products of at least reagent grade.

Example 1: Cloning and Expression of *Buttiauxella* Phytases

A multiple alignment was made of the following histidine acid phosphatases (HAP): appA *Escherichia coli* (SPTREMBL:Q8GN88), *Citrobacter gillenii* DSM 13694 phytase (geneseqp:aeH04533), *Citrobacter amalonaticus* ATCC 25407 phytase (geneseqp:aeH04535), *Citrobacter braakii* phytase (geneseqp:aeH04827), and ypo1648 *Yersinia pestis* CO92 (SPTREMBL:Q8ZFP6). Two degenerate oligonucleotide primers were designed on the basis of consensus sequences:

2123: 5'- CATGGTGTGCGNGCNCCNACNAA -3' (SEQ ID NO:9, forward primer)
 2065: 5'- CCCACCAGGNGGNGTRTTRTCNGGYTG -3' (SEQ ID NO:10, reverse primer),
 wherein Y designates T or C, R designates A or G, and N designates A, C, G or T.

The primers were used for PCR screening of a number of bacterial species at
 5 annealing temperatures between 40 and 50°C but typically as touch down program starting
 with 50°C and then reducing the annealing temperature with 1°C for each cycle over the
 next 10 cycles before conducting standard PCR.

A partial phytase gene in the form of an approximately 950 bp PCR fragment was
 identified in Buttiauxella agrestis DSM 18932.

10 The PCR fragment was isolated from agarose gel and the fragment was sequenced
 using the same PCR primers as those with which the fragment was generated. By
 translation of the nucleotide sequence, it was confirmed that the DNA fragment was part of a
 HAP phytase gene.

For obtaining the full length nucleotide sequence of the gene, the DNA Walking
 15 SpeedUp Kit (DWSK-V102 from Seegene, Inc., 2nd Fl., Myungji Bldg., 142-21, Samsung-
 dong, Kangnam-gu, Seoul, 135-090, Korea) was used, which is designed to capture
 unknown target sites. For this purpose, 4 specific oligonucleotides were designed and used
 with the kit:

2138 TSP1N: 5'- ATTGCGGAGCAAGCCCTG-3' (SEQ ID NO:11)

20 2139 TSP2N: 5'- TCGCCAGTTTTTAAGCGNGCG-3' (SEQ ID NO:12)

2136 TSP1C: 5'- TGAATATGCGCAAGGGATG -3' (SEQ ID NO:13)

2137 TSP2C: 5'- TGGGGGTCAGAGCAAGAGTGGG-3' (SEQ ID NO:14)

The full length nucleotide sequence encoding the phytase from *Buttiauxella agrestis*
 DSM 18932 is shown in the sequence listing as SEQ ID NO:5, and the corresponding
 25 encoded amino acid sequence has SEQ ID NO:6. The first 33 amino acids of SEQ ID NO:6
 (i.e. amino acids -33 to -1) are a signal peptide, as predicted by Signal P V3.0 (see
www.cbs.dtu.dk/services/SignalP/).

The same multiple alignment described above was used for design of two other
 degenerate oligonucleotide PCR primers:

30 315: 5'- CGCGTGGTGATTGTGTCCMGNCAYGGNGT -3' (SEQ ID NO:15, forward primer)

316: 5'- CCAGGTTGGTATCATGGCCNGCDATRAA -3' (SEQ ID NO:16, reverse primer),

wherein Y designates T or C, M designates A or C, N designates A, C, G or T, and R
 designates A or G.

The primers were used for PCR screening of a number of bacterial species at the
 35 same conditions as described above.

Partial phytase genes in the form of approximately 900 bp PCR fragments were
 identified in *Buttiauxella gaviniae* DSM 18930 as well as in *Buttiauxella agrestis* DSM 18931.

The two PCR fragments were isolated from agarose gel and sequenced using primers 1978 and 1979:

1978: 5'- CGCGTGGTGATTGTGTCC -3' (SEQ ID NO:17)

1979: 5'- CCAGGTTGGTATCATGGCC -3' (SEQ ID NO:18)

5 By translation of the nucleotide sequence, it was confirmed that both DNA fragments were PCR amplified from HAP phytase genes.

Four new primers were designed for each gene for obtaining the full length nucleotide sequence of the gene according to the instructions given in the DNA Walking SpeedUp Kit (DWSK-V102 from Seegene, Inc., 2nd Fl., Myungji Bldg., 142-21, Samsung-dong, Kangnam-gu, Seoul, 135-090, Korea).

The four specific oligonucleotide PCR primers designed for the *Buttiauxella gaviniae* DSM 18930 phytase gene were:

2029 TSP1N: 5'- AAGCTTCGCCAGTTTAAAGCG-3' (SEQ ID NO:19)

2030 TSP2N: 5'- TTGAGTTTGGTGTGGGGCAACTG-3' (SEQ ID NO:20)

15 2031 TSP1C: 5'- TGGCAACAAAGTCGCTCTCG-3' (SEQ ID NO:21)

2032 TSP2C: 5'- TCCTGCTGGAATATGCGCAAGG-3' (SEQ ID NO:22)

The four specific oligonucleotide PCR primers designed for the *Buttiauxella agrestis* DSM 18931 phytase gene were:

2017 TSP1N: 5'- TTCGCCCGTTTAAAGCGTG-3' (SEQ ID NO:23?)

20 2018 TSP1C: 5'- ACTGCCCTGCGATAAAATGCCC-3' (SEQ ID NO:24)

2019 TSP2N: 5'- TTTCCTGCTGGAATATGCGC-3' (SEQ ID NO:25)

2020 TSP2C: 5'- TTGATGGCGCGCACACCTTAC-3' (SEQ ID NO:26)

The full length nucleotide sequence encoding the phytase from *Buttiauxella gaviniae* DSM 18930 is shown in the sequence listing as SEQ ID NO:1, and the corresponding encoded amino acid sequence has SEQ ID NO:2. The first 33 amino acids of SEQ ID NO:2 (i.e. amino acids -33 to -1) are expected to be a signal peptide (predicted by Signal P V3.0, see www.cbs.dtu.dk/services/SignalP/).

The full length nucleotide sequence encoding the mature phytase, as well as a partial sequence coding for the signal peptide from *Buttiauxella agrestis* DSM 18931 is shown in the sequence listing as SEQ ID NO:3. The corresponding encoded amino acid sequence has SEQ ID NO:4. The first 9 amino acids of SEQ ID NO:4 (i.e. amino acids -9 to -1) are expected to be a part of the signal peptide.

The three phytase genes were expressed in *Bacillus subtilis* as follows:

The signal peptide encoding sequence of SEQ ID NO:7 (encoding the signal peptide of SEQ ID NO:8 and derived from a *Bacillus licheniformis* protease) was fused by PCR in frame to the gene encoding the mature phytase from each of the three phytases in turn.

The DNA coding for the resulting coding sequence was integrated by homologous

recombination on the *Bacillus subtilis* host cell genome. The gene constructs were expressed under the control of a triple promoter system (as described in WO 99/43835), consisting of the promoters from *Bacillus licheniformis* alpha-amylase gene (*amyL*), *Bacillus amyloliquefaciens* alpha-amylase gene (*amyQ*), and the *Bacillus thuringiensis* *cryIIIA* promoter including stabilizing sequence. The gene coding for Chloramphenicol acetyl-transferase was used as marker as described in, e.g., Diderichsen et al., A useful cloning vector for *Bacillus subtilis*. Plasmid, 30, p. 312, 1993.

For each of the three constructs, chloramphenicol resistant transformants were cultured in PS-1 medium (10% sucrose, 4% soybean flour, 1% $\text{Na}_3\text{PO}_4 \cdot 12\text{H}_2\text{O}$, 0.5% CaCO_3 , and 0.01% pluronic acid), while shaking at 250 RPM at 30°C. After 2-5 days of incubation the supernatants were removed and the phytase activity was identified by applying 20 μl of the supernatant into 4 mm diameter holes punched out in 1% LSB-agarose plates containing 0.1M Sodium acetate pH4.5 and 0.1% Inositol hexaphosphoric acid. The plates were left over night at 37°C and a buffer consisting of 0.25M CaCl_2 and 500mM MES (adjusted to pH 6.5 with 4N NaOH) were poured over the plates. The plates were left at room temperature for 1h and the inositolphosphate phosphatase, or phytase, activity was then identified as a clear zone.

Several phytase positive transformants for each of the three constructs were analyzed by DNA sequencing to ensure the correct DNA sequence of the constructs. One correct clone was selected for each of the three constructs.

Example 2: Fermentation and Purification of *Buttiauxella* Phytases

The strain of *Bacillus subtilis* harboring the *Buttiauxella agrestis* DSM 18932 phytase construct and capable of expressing the phytase having SEQ ID NO:6 (mature part) was cultivated at 30°C and with 250rpm for 6 days in SK-1M medium (Sodium Caseinate (Miprodan 30 from Arla) 40g, Maltodextrin 01 (Glucidex 6, catalogue no. 332203 from Roquette), 200g, Soybean Meal 50g, Dowfax 63N10 (a non-ionic surfactant from Dow) 0.1ml, tap water up to 1000 ml, CaCO_3 tablet 0.5g/100 ml).

The strain of *Bacillus subtilis* harboring the *Buttiauxella agrestis* DSM 18931 phytase construct and capable of expressing the phytase having SEQ ID NO:4 (mature part) was cultivated at 30°C and with 250rpm for 5 days in the PS-1 medium which is described in Example 1.

The strain of *Bacillus subtilis* harboring the *Buttiauxella gaviniae* DSM 18930 phytase construct and capable of expressing the phytase having SEQ ID NO:2 (mature part) was cultivated at 30°C and with 250 rpm for 5 days in the PS-1 medium which is described in Example 1.

The fermentation supernatant with the phytase of SEQ ID NO:2 was first centrifuged at 7200rpm and 5°C for 2 hours and filtered through a Fast PES Bottle top filter with a 0.22µm cut-off. Next, the filtered supernatant was pre-treated as follows:

5 The sample solution was washed with water and concentrated using an ultrafiltration unit (Filtron, from Filtron Technology Corporation) equipped with a 10kDa cut-off ultrafiltration membrane.

Then pH was adjusted to 5.0 with 6M HCl, which caused a minor precipitation. This was removed by centrifugation of the sample at 7200rpm for 20 minutes at 5°C. The supernatant, containing the phytase, was added 2.5ml 1M sodium acetate pH 5.0 and
10 filtered through a Fast PES bottle top filter with a 0.22µm cut-off. After this the pH of the solution was measured to 5.0 and the conductivity was found to be 1.8mS/cm. After pretreatment the phytase was purified by chromatography on SP SepharoseTM, approximately 85ml in a XK26 column, using as buffer A 20mM sodium acetate pH5.0, and as buffer B 20mM sodium acetate + 1M NaCl pH 5.0.

15 The fractions from the column were analyzed for phytase activity and fractions with activity were pooled.

The purification of the phytase of SEQ ID NO:4 was performed essentially as described above, except that 10% acetic acid was used to adjust the pH. Again some precipitation was observed and this was removed by centrifugation. The pH of the solution
20 was measured to 5.0, while the conductivity was found to be approximately 1.2mS/cm before the column chromatography step.

The fermentation supernatant with the phytase of SEQ ID NO:6 was first centrifuged at 7200rpm and 5°C for one hour and filtered through a sandwich of four WhatmanTM glass microfibre filters (2.7, 1.6, 1.2 and 0.7 micrometer). Following this the
25 solution was filtered through a Seitz-EKS depth filter using pressure. The solution was added solid ammonium sulfate giving a final concentration of 1.5M and the pH was adjusted to 6.0 using 6M HCl.

The phytase-containing solution was applied to a butyl-sepharose column, approximately 50ml in a XK26 column, using as buffer A 25mM bis-tris + 1.5M ammonium
30 sulfate pH 6.0, and as buffer B 25mM bis-tris pH 6.0. However, the enzyme did not bind to the column and almost all activity was found in the flow-through and wash fractions. The flow-through and wash fractions were combined and solid ammonium sulfate was added to about 80% saturation. The solution was left overnight at 5°C in order to complete precipitation. The precipitate was isolated (no activity was found in the supernatant) and
35 dissolved in Milli-Q water. This solution was dialyzed against Milli-Q water and pH was adjusted to 4.5.

Following this the phytase was purified by chromatography on S Sepharose,

approximately 150ml in a XK26 column, using as buffer A 50mM sodium acetate pH 4.5, and as buffer B 50mM sodium acetate + 1M NaCl pH 4.5.

The fractions from the column were analyzed for phytase activity and fractions with activity were pooled. Finally, the fractions with phytase activity were concentrated using an
5 Amicon ultra-15 filtering device with a 30kDa cut-off membrane.

The molecular weight, as estimated from SDS-PAGE, was approximately 40 kDa for all three phytases and the purity was in all cases > 95%.

Example 3: Phytase Activity Assay

10 Phytase activity may suitably be determined by the following assay:

75 microliter phytase-containing enzyme solution, appropriately diluted in 0.25M sodium acetate, 0.005% (w/v) TweenTM-20. pH5.5, is dispensed in a microtiter plate well, e. g. NUNC 269620, and 75 microliter substrate is added (prepared by dissolving 100mg sodium phytate from rice (Aldrich Cat.No. 274321) in 10ml 0.25M sodium acetate buffer, pH5.5).

15 The plate is sealed and incubated 15min. shaken with 750rpm at 37°C. After incubation, 75 microliter stop reagent is added (the stop reagent being prepared by mixing 10 ml molybdate solution (10% (w/v) ammonium hepta-molybdate in 0.25% (w/v) ammonia solution), 10ml ammonium vanadate (0.24% commercial product from Bie&Berntsen, Cat.No. LAB17650), and 20ml 21.7% (w/v) nitric acid), and the absorbance at 405nm is measured in a microtiter
20 plate spectrophotometer. The phytase activity is expressed in the unit of FYT, one FYT being the amount of enzyme that liberates 1 micromole inorganic ortho-phosphate per minute under the conditions above. An absolute value for the measured phytase activity may be obtained by reference to a standard curve prepared from appropriate dilutions of inorganic phosphate, or by reference to a standard curve made from dilutions of a phytase
25 enzyme preparation with known activity (such standard enzyme preparation with a known activity is available on request from Novozymes A/S, Krogshoejvej 36, DK-2880 Bagsvaerd).

Example 4: Specific activity

30 The specific activity of the phytases from *Buttiauxella gaviniae* DSM 18930, *Buttiauxella agrestis* DSM 18931, and *Buttiauxella agrestis* DSM 18932 (having the amino acid sequences of the mature parts of SEQ ID NO:2, 4 and 6, respectively) was determined in sodium acetate buffer, pH 5.5. The phytases were highly purified as described in Example 2, i.e. only one component was identified on an SDS poly acryl amide gel.

35 The protein concentration was determined by amino acid analysis as follows: An aliquot of the sample was hydrolyzed in 6M HCl, 0.1% phenol for 16h at 110°C in an evacuated glass tube. The resulting amino acids were quantified using an Applied

Biosystems 420A amino acid analysis system operated according to the manufacturer's instructions. From the amounts of the amino acids the total mass - and thus also the concentration - of protein in the hydrolyzed aliquot was calculated.

The phytase activity was determined in the units of FYT as described in Example 3, and the specific activity was calculated as the phytase activity measured in FYT units per mg phytase enzyme protein.

The resulting specific activities for the three phytases are shown in Table 2. The specific activity was determined on sodium phytate at pH5.5 and 37°C.

10 Table 2

Phytase	Specific activity (FYT/mg protein)
SEQ ID NO:2	700
SEQ ID NO:4	800
SEQ ID NO:6	650

The wildtype (wt) phytase from *Buttiauxella* P1-29 described in WO 2006/043178 has a specific activity at pH3.5 and 37°C of about 300U/mg (see Example 10 of this WO publication). According to Example 2 thereof, the ratio of activities at pH3.5 and pH5.5 is about 1.3, while according to Fig. 2 (pH activity profile of the purified phytase) this ratio is rather 1.5. In any event, when the specific activity at pH3.5 is about 300 U/mg, the specific activity at pH5.5 is not higher than $300/1.3 = 231\text{U/mg}$ (maybe it is rather $300/1.5 = 200\text{U/mg}$).

WO 2006/043178, also discloses a variant phytase with a specific activity at pH4.0 and 37°C which is 115% of the specific activity of the wt. Using again Fig. 2, the specific activity of the wt at pH 4.0 is not higher than $80/60 \times 300 \text{ U/mg} = 400\text{U/mg}$. The specific activity at pH4.0 of the variant is then $1.15 \times 400 \text{ U/mg} = 460\text{U/mg}$, while at pH5.5 the specific activity of this variant is approximately $40/80 \times 460 \text{ U/mg} = 230\text{U/mg}$ (in fact, 230 is precisely 115% of the 200 U/mg of the wildtype at pH 5.5, as calculated above).

Apart from the pH-difference, which has been taken into account above, the assay conditions of WO 2006/043178 are, for all practical purposes, identical to those of the present invention, except possibly for the presence in WO 2006/043178 of 0.8mM CaCl_2 in the reaction mixture. However, we have tested the activity of the phytases of SEQ ID NOs:2, 4 and 6 in the assay of Example 3 including even more Ca^{2+} (viz. 1mM CaCl_2) but found no difference.

Accordingly, the phytases of the present invention have a very much improved

specific activity at pH5.5 as compared to the phytases disclosed in WO 2006/043178. In fact, the specific activity of the phytases of the invention is improved in the whole range of pH3.5-5.5. This is apparent from Example 6 below, which shows that for the phytases of the invention the activity in this range is at least as high as the activity at pH5.5.

5

Example 5: pI

The isoelectric point, pI, for the three phytases was determined using isoelectric focusing gels (Novex pH 310 IEF gel from Invitrogen, catalog number EC6655A2) run as described by the manufacturer. The pI for the *Buttiauxella agrestis* DSM 18931 and DSM 18932 phytases (SEQ ID NOs:4 and 6) is 7.4, while the pI for the *Buttiauxella gaviniae* DSM 18930 (SEQ ID NO:2) phytase is 7.6.

10

Example 6: pH profile

The pH profile (phytase activity as a function of pH) of the three phytases was determined in the pH range of 2.0 to 7.5 as described in Example 3, except that a buffer cocktail (50mM glycine, 50mM acetic acid and 50mM Bis-Tris (Bis-(2-hydroxyethyl)imino-tris(hydroxymethyl)methan)) was used instead of the 0.25M sodium acetate pH5.5 buffer.

15

The results are shown in Table 3 below, relative to the value at the optimum pH for each phytase (pH4.5).

20

Table 3: Relative phytase activity versus pH

pH	<i>Buttiauxella gaviniae</i> DSM 18930 (SEQ ID NO:2)	<i>Buttiauxella agrestis</i> DSM 18931 (SEQ ID NO:4)	<i>Buttiauxella noackiea</i> DSM 18932 (SEQ ID NO:6)
2.0	37	41	40
2.5	46	49	49
3.0	66	65	66
3.5	83	86	78
4.0	96	99	99
4.5	100	100	100
5.0	92	91	83
5.5	70	70	69
6.0	38	43	39
6.5	11	14	14
7.0	-3	1	1
7.5	-5	-2	0

As compared to the pH activity profile of the wildtype (wt) phytase from *Buttiauxella* P1-29 described in WO 2006/043178 (see Fig. 2 thereof), the phytases of the invention appear to have a higher relative activity at low as well as high pH-values (at pH2.0 and pH6.0 the activity of the phytases of the present invention is twice as high, and eight times as high, respectively, as compared to the *Buttiauxella* P1-29 wt phytase).

Example 7: pH stability

The pH stability of the purified phytases of SEQ ID NOs:2 and 4 at 40°C was determined by measuring residual phytase activity after incubation at 40°C and at various pH values for 1.5 and 24 hours. The phytases were incubated in 0.1M glycine, 0.1M acetic acid, 0.1M Bis-Tris, adjusted to the desired pH. Samples of the respective incubation mixtures were withdrawn after 0, 1.5 and 24 hours, the pH of the samples was adjusted to 5.5 by dilution in 0.25M sodium acetate, 0.005% (w/v) Tween20, pH5.5), and the residual activity at pH5.5 was determined using the method described in Example 3. The results, normalized to the activity found at 0 hours, are shown in Table 4 below.

Table 4: pH stability at 40°C

pH	<i>Buttiauxella</i> <i>gaviniae</i> (SEQ ID NO:2)	<i>Buttiauxella</i> <i>agrestis</i> (SEQ ID NO:4)	pH	<i>Buttiauxella</i> <i>gaviniae</i> (SEQ ID NO:2)	<i>Buttiauxella</i> <i>agrestis</i> (SEQ ID NO:4)
1.5 hours			24 hours		
2.0	77	83	2.0	51	56
3.0	74	85	3.0	66	66
4.0	81	95	4.0	77	88
5.0	80	68	5.0	78	67
6.0	82	67	6.0	76	67
7.0	79	75	7.0	76	73
8.0	80	74	8.0	77	70

Both phytases are very stable for 1.5 hours in the entire range of pH2.0-8.0, whereas when incubated for 24 hours a certain loss of activity is observed at the lower pH values (pH2.0-3.0).

Example 8: Temperature profile

The temperature profile (phytase activity as a function of temperature) of the three

phytases was determined in the temperature range of 20-90°C essentially as described in Example 3, however, the enzymatic reactions (100 microliter phytase-containing enzyme solution + 100 microliter substrate) were performed in PCR tubes instead of microtiter plates. After a 15 minute reaction period at desired temperature the tubes were cooled to 20°C for 20 seconds and 150 microliter of the reaction mixture was transferred to a microtiter plate. 75 microliter stop reagent was added and the absorbance at 405 nm was measured in a microtiter plate spectrophotometer.

The temperature profiles for the phytases of SEQ ID NOs:2 and 4 were determined at pH4.0 (0.25M sodium acetate), whereas the temperature profile for the phytase of SEQ ID NO:6 was determined at pH5.5 (0.25M sodium acetate).

The results are shown in Table 5 below, for each phytase relative to the activity at the optimum temperature.

Table 5: Temperature profile at pH 4.0 / 5.5

Temperature (°C)	20	30	40	50	60	70	80	90
SEQ ID NO:2	21	36	43	75	100	23	9	6
SEQ ID NO:4	24	40	54	82	100	17	8	7
SEQ ID NO:6	22	32	50	74	100	14	8	3

All phytases appear to have an optimum temperature of about 60°C, are fairly active in the temperature range of 30-60°C, and they also show a decent activity at 20°C as well as at 70°C. In the range of 80-90°C the activity is insignificant.

Deposit of Biological Material

The following biological material has been deposited under the terms of the Budapest Treaty with DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH), Inhoffenstraße 7 B, D-38124 Braunschweig, Germany, and given the following accession numbers:

Deposit	Accession Number	Date of Deposit
Buttiauxella gaviniae	DSM 18930	15-JAN-2007
Buttiauxella agrestis	DSM 18931	15-JAN-2007
Buttiauxella agrestis	DSM 18932	15-JAN-2007

The strains have been deposited under conditions that assure that access to the culture will be available during the pendency of this patent application to one determined by the Commissioner of Patents. The deposits represent substantially pure cultures of the deposited

strains. The deposits are available as required by foreign patent laws in countries wherein counterparts of the subject application or its progeny are filed. However, it should be understood that the availability of a deposit does not constitute a license to practice the subject invention in derogation of patent rights granted by governmental action.

5 The above strains were isolated from samples collected in Denmark in 2005.

 The scope of the claims should not be limited by the preferred embodiments set forth in the examples, but should be given the broadest interpretation consistent with the description as a whole.

10

15

10991.204-WO

PCT

Print Out (Original in Electronic Form)

(This sheet is not part of and does not count as a sheet of the international application)

0-1	Form PCT/RO/134 (SAFE) Indications Relating to Deposited Microorganism(s) or Other Biological Material (PCT Rule 13bis)	
0-1-1	Prepared Using	PCT Online Filing Version 3.5.000.193 MT/FOP 20020701/0.20.5.9
0-2	International Application No.	
0-3	Applicant's or agent's file reference	10991.204-WO
1	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
1-1	page	49
1-2	line	8
1-3	Identification of deposit	
1-3-1	Name of depositary institution	DSMZ DSMZ-Deutsche Sammlung von Mikroor- ganismen und Zellkulturen GmbH
1-3-2	Address of depositary institution	Mascheroder Weg 1b, D-38124 Braunschweig, Germany
1-3-3	Date of deposit	15 January 2007 (15.01.2007)
1-3-4	Accession Number	DSMZ DSM 18930
1-5	Designated States for Which Indications are Made	all designations
2	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
2-1	page	49
2-2	line	9
2-3	Identification of deposit	
2-3-1	Name of depositary institution	DSMZ DSMZ-Deutsche Sammlung von Mikroor- ganismen und Zellkulturen GmbH
2-3-2	Address of depositary institution	Mascheroder Weg 1b, D-38124 Braunschweig, Germany
2-3-3	Date of deposit	15 January 2007 (15.01.2007)
2-3-4	Accession Number	DSMZ DSM 18931
2-5	Designated States for Which Indications are Made	all designations

10991.204-WO

PCT

Print Out (Original in Electronic Form)

(This sheet is not part of and does not count as a sheet of the international application)

3	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
3-1	page	49
3-2	line	10
3-3	Identification of deposit	
3-3-1	Name of depositary institution	DSMZ DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH
3-3-2	Address of depositary institution	Mascheroder Weg 1b, D-38124 Braunschweig, Germany
3-3-3	Date of deposit	15 January 2007 (15.01.2007)
3-3-4	Accession Number	DSMZ DSM 18932
3-5	Designated States for Which Indications are Made	all designations

FOR RECEIVING OFFICE USE ONLY

0-4	This form was received with the international application: (yes or no)	
0-4-1	Authorized officer	

FOR INTERNATIONAL BUREAU USE ONLY

0-5	This form was received by the international Bureau on:	
0-5-1	Authorized officer	

CLAIMS:

1. A polypeptide having phytase activity and having an amino acid sequence which
 - a) has at least 70% identity to amino acids 1-413 of SEQ ID NO:2, amino acids 1- 413 of SEQ ID NO:4, and/or amino acids 1-413 of SEQ ID NO:6 when aligned to the respective amino acid sequence using the Needle program with the BLOSUM62 substitution matrix, a gap opening penalty of 10.0, and a gap extension penalty of 0.5; and
 - b) comprises at least one of the following amino acids at the position indicated: 119N, 120L, and/or 121E, when aligned as described in a) to amino acids 1-413 of SEQ ID NO:2 and using an amino acid residue numbering corresponding to amino acids 1-413 of SEQ ID NO:2.

2. A polypeptide having phytase activity and having an amino acid sequence which
 - a) has at least 70% identity to amino acids 1-413 of SEQ ID NO:2, amino acids 1- 413 of SEQ ID NO:4, and/or amino acids 1-413 of SEQ ID NO:6 when aligned to the respective amino acid sequence using the Needle program with the BLOSUM62 substitution matrix, a gap opening penalty of 10.0, and a gap extension penalty of 0.5; and
 - b) comprises at least one of the following amino acids at the position indicated: 109Q, 111G, 119N, 120L, and/or 121E, when aligned as described in a) to amino acids 1-413 of SEQ ID NO:2 and using an amino acid residue numbering corresponding to amino acids 1-413 of SEQ ID NO:2.

3. A polypeptide having phytase activity and having an amino acid sequence which
 - a) has at least 78% identity to amino acids 1-413 of SEQ ID NO:2, amino acids 1- 413 of SEQ ID NO:4, and/or amino acids 1-413 of SEQ ID NO:6 when aligned to the respective amino acid sequence using the Needle program with the BLOSUM62 substitution matrix, a gap opening penalty of 10.0, and a gap extension penalty of 0.5; and
 - b) comprises at least one of the following amino acids at the position indicated: 109Q, 111G, 119N, 120L, 121E, and/or 193Q, when aligned as described in a) to amino acids 1-413 of SEQ ID NO:2 and using an amino acid residue numbering corresponding to amino acids 1-413 of SEQ ID NO:2.

4. The polypeptide of claim 3 further comprising at least one of the following amino acids at the position indicated: 1S, 10I, 38S, 66E, 71K, 81A, 141R, 142L, 152M, 155E,

214V, 239K, 245D, 248E, 255A,T, 268A,T, 277T, 283D,E, 285K, 287D, 288A,V, 293G, 296S, 303L, 314A, 337I, 345A, 350I, 364A, 371K, 372E, 396P, 399K, 406E, and/or 413P, or at least one of the following amino acids at the position indicated: 26E, 37Y, 89T, 92E, 134I,V, 160R, 164F, 171I, 176K, 178P, 188N, 190E, 192G, 207E,T, 209S, 211C, 235V, 248L, 256H,Y, 261E, 270K, 303F, and/or 318D when aligned to the respective amino acid sequence using the Needle program with the BLOSUM62 substitution matrix, a gap opening penalty of 10.0, and a gap extension penalty of 0.5; and using an amino acid residue numbering corresponding to amino acids 1-413 of SEQ ID NO:2.

5. A polypeptide having phytase activity which has an amino acid sequence which has at least 93.6% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to amino acids 1-413 of SEQ ID NO:2, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, or at least 99% identity to amino acids 1-413 of SEQ ID NO:4, or at least 95.8% identity, at least 97% identity, at least 98% identity, or at least 99% identity to amino acids 1-413 of SEQ ID NO:6.

6. A polypeptide having phytase activity, selected from the group consisting of:

- (i) a polypeptide comprising amino acids 1-413 of SEQ ID NO:2;
- (ii) a variant of (i) which has at least 93.6% identity to amino acids 1-413 of SEQ ID NO:2 and which comprises at least one of the following substitutions: K26E, N37Y, A89T, D92E, T134I,V, H160R, S164F, T171I, T176K, A178P, S188N, D190E, A192G, K207E,T, A209S, D211C, A235V, E248L, Q256H,Y, A261E, N270K, D283N, A288E, I303F, and/or N318D; and
- (iii) a variant of (i) or (ii) which has at least 93.6% identity to amino acids 1-413 of SEQ ID NO:2 and which comprises at least one of the following substitutions: N1S, V10I, T38S, E66Q, Q71K, T81A, R141Q, L142V, T152M, E155D, V214I, K239N, D245E, E248S, A255T, R268A,T, A277T, D283E, N285K, T287D, A288V, D293G, P296S, I303L, S314A, I337V, A345S, V350I, A364S, K371N, E372Q, P396S, K399T, E406V, and/or Q413P;

wherein for determining amino acid residue positions in ii) and iii) the amino acid sequence of the polypeptide is aligned to amino acids 1-413 of SEQ ID NO:2 using the Needle program with the BLOSUM62 substitution matrix, a gap opening penalty of 10.0, and a gap extension penalty of 0.5, and using an amino acid residue numbering corresponding to amino acids 1-413 of SEQ ID NO:2.

7. A polypeptide having phytase activity, selected from the group consisting of:

(i) a polypeptide comprising amino acids 1-413 of SEQ ID NO:4;

(ii) a variant of (i) which has at least 95% identity to amino acids 1-413 of SEQ ID NO:4 and which comprises at least one of the following substitutions: K26E, N37Y, A89T, D92E, T134I,V, H160R, S164F, T171I, T176K, A178P, S188N, D190E, A192G, K207E,T, A209S, D211C, A235V, S248L, Q256H,Y, A261E, N270K, E283N, V288E, I303F, and/or N318D; and

(iii) a variant of (i) or (ii) which has at least 95% identity to amino acids 1-413 of SEQ ID NO:4 and which comprises at least one of the following substitutions: S1N, I10V, S38T, Q66E, K71Q, T81A, Q141R, V142L, M152T, E155D, I214V, N239K, E245D, S248E, T255A, A268R,T, T277A, E283D, K285N, D287T, V288A, D293G, S296P, I303L, A314S, V337I, S345A, I350V, A364S, N371K, E372Q, S396P, T399K, V406E, and/or P413Q;

wherein for determining amino acid residue positions in ii) and iii) the amino acid sequence of the polypeptide is aligned to amino acids 1-413 of SEQ ID NO:4 using the Needle program with the BLOSUM62 substitution matrix, a gap opening penalty of 10.0, and a gap extension penalty of 0.5, and using an amino acid residue numbering corresponding to amino acids 1-413 of SEQ ID NO:4.

8. A polypeptide having phytase activity, selected from the group consisting of:

(i) a polypeptide comprising amino acids 1-413 of SEQ ID NO:6;

(ii) a variant of (i) which has at least 95.8% identity to amino acids 1-413 of SEQ ID NO:6 and which comprises at least one of the following substitutions: K26E, N37Y, A89T, D92E, T134I,V, H160R, S164F, T171I, T176K, A178P, S188N, D190E, A192G, K207E,T, A209S, D211C, A235V, S248L, Q256H,Y, A261E, N270K, E283N, V288E, L303F, and/or N318D; and

(iii) a variant of (i) or (ii) which has at least 95.8% identity to amino acids 1-413 of SEQ ID NO:6 and which comprises at least one of the following substitutions: N1S, V10I, T38S, E66Q, Q71K, A81T, Q141R, V142L, T152M, D155E, I214V, N239K, E245D, S248E, A255T, T268A,R, T277A, E283D, K285N, D287T, V288A, G293D, P296S, L303I, A314S, V337I, S345A, V350I, S364A, N371K, Q372E, S396P, T399K, V406E, and/or Q413P;

wherein for determining amino acid residue positions in ii) and iii) the amino acid sequence

of the polypeptide is aligned to amino acids 1-413 of SEQ ID NO:6 using the Needle program with the BLOSUM62 substitution matrix, a gap opening penalty of 10.0, and a gap extension penalty of 0.5, and using an amino acid residue numbering corresponding to amino acids 1-413 of SEQ ID NO:6.

9. A variant polypeptide which has at least 80% identity, at least 85% identity, at least 90% identity, at least 95% identity, at least 96% identity, at least 97% identity, at least 98% identity, at least 99% identity to a polypeptide comprising amino acids 1-413 of any one of the sequences: AEH25057, AEH25059, AEH25058, AEH25067, AEH25071, AEH25072, AEH25074, AEH25075, AEH25076, AEH25073, AEH25070, AEH25069, AEH25066, AEH25060, AEH25068, AEH25063, AEH25065, AEH25062, AEH25061, AEH25056, AEH25051, and AEH25064 as represented in figure 1; which variant has phytase activity and comprises at least one of the following substitutions: E109Q, H111G, D119N, I120L, K121E and/or L193Q; wherein for determining amino acid residue positions the amino acid sequence of the variant is aligned to amino acids 1-413 of SEQ ID NO:2 using the Needle program with the BLOSUM62 substitution matrix, a gap opening penalty of 10.0, and a gap extension penalty of 0.5, and using an amino acid residue numbering corresponding to amino acids 1-413 of SEQ ID NO:2.

10. A polypeptide having phytase activity, selected from the group consisting of:

(a) a polypeptide having an amino acid sequence which has at least 93.6% identity with the sequence of amino acids 1-413 of SEQ ID NO:2; at least 95.0% identity with the sequence of amino acids 1-413 of SEQ ID NO:4, or at least 95.8% identity with the sequence of amino acids 1-413 of SEQ ID NO:6; and

(b) a variant comprising a conservative substitution, deletion, and/or insertion of one or more amino acids of amino acids 1-413 of SEQ ID NO:2; SEQ ID NO:4, or SEQ ID NO:6 wherein the amino acid sequence differs by no more than 26 point alterations as compared to amino acids 1-413 of SEQ ID NO:2, no more than 21 point alterations as compared to amino acids 1-413 of SEQ ID NO:4, and/or no more than 17 point alterations as compared to amino acids 1-413 of SEQ ID NO:6,

wherein for calculating identity in a) the amino acid sequence of the polypeptide is aligned to amino acids 1-413 of SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6 using the Needle program with the BLOSUM62 substitution matrix, a gap opening penalty of 10.0, and a gap extension penalty of 0.5.

11. A polynucleotide comprising a nucleotide sequence which encodes the polypeptide of any of claims 1-10.
12. A nucleic acid construct comprising the polynucleotide of claim 11 operably linked to one or more control sequences that direct the production of the polypeptide in an expression host.
13. A recombinant host cell comprising the nucleic acid construct of claim 12.
14. A method for producing the polypeptide of any one of claims 1-10 comprising (a) cultivating a cell, which in its wild-type form produces the polypeptide, under conditions conducive for production of the polypeptide; and (b) recovering the polypeptide, or (b) cultivating a host cell comprising a nucleic acid construct comprising a nucleotide sequence encoding the polypeptide of any one of claims 1-10 under conditions conducive for production of the polypeptide; and (c) recovering the polypeptide.
15. Use of the polypeptide of any one of claims 1-10 in animal feed; in the preparation of animal feed; for improving the nutritional value of animal feed; for reducing phytate levels in animal manure; for improving digestability of vegetable proteins; or for liberating phosphorous from a phytase substrate.
16. Use of the polypeptide of any one of claims 1-10 in the preparation of ethanol using a fermentation process.

1/9

Fig. 1

Numbering	1	10	20
AEH25057	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25059	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25058	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25067	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25071	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25072	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25074	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25076	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25075	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25073	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25070	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25069	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25066	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25060	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25068	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25063	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25065	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25062	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25061	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25056	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTEM	
AEH25051	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
AEH25064	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
SEQ4	-----FSLGLTAYAS	DTPASGYQ	VEKVVILSRHGVRAPTKM
SEQ6	MTISAFNRKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	
SEQ2	MTISAFNHKKLTLHPGLFVALSAIFSLGSTAYANDTPASGYQ	VEKVVILSRHGVRAPTKM	

.*****.******.*

2/9

Fig. 1 - continued

Numbering	30	40	50	60	70	80
AEH25057	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25059	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25058	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25067	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25071	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25072	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25074	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25076	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25075	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25073	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25070	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25069	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25066	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25060	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25068	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25063	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25065	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25062	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25061	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25056	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25051	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
AEH25064	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQQ
SEQ4	TQTMRDVT	PN S WPEW	PKLGYIT	PRGEHLI	SLMGGFY	RQKFQQ K GILS
SEQ6	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	R E KFQQQILS
SEQ2	TQTMRDVT	PNTWPEW	PKLGYIT	PRGEHLI	SLMGGFY	R E KFQQQILS
	*****	:	*****	:	*****	:

3/9

Fig. 1 - continued

Numbering	90	100	110	120	130	140
AEH25057	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGVCSMDKTQVQQAVE				
AEH25059	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGICSMDKTQVQQAVE				
AEH25058	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGTCSMDKTQVQQAVE				
AEH25067	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGICSMDKTQVQQAVE				
AEH25071	WTDVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGICSMDKTQVQQAVE				
AEH25072	WTDVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGICSMDKTQVQQAVE				
AEH25074	WADVEQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGICSMDKTQVQQAVE				
AEH25076	WADVEQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGICSMDKTQVQQAVE				
AEH25075	WTDVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGICSMDKTQVQQAVE				
AEH25073	WTDVEQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGTCSMDKTQVQQAVE				
AEH25070	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGICSMDKTQVQQAVE				
AEH25069	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGICSMDKTQVQQAVE				
AEH25066	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGTCSMDKTQVQQAVE				
AEH25060	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGTCSMDKTQVQQAVE				
AEH25068	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGTCSMDKTQVQQAVE				
AEH25063	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGTCSMDKTQVQQAVE				
AEH25065	WADVEQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGTCSMDKTQVQQAVE				
AEH25062	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGTCSMDKTQVQQAVE				
AEH25061	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGTCSMDKTQVQQAVE				
AEH25056	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGTCSMDKTQVQQAVE				
AEH25051	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGTCSMDKTQVQQAVE				
AEH25064	WADVDQRTLKTGEAFLAGLAPECHLTIHHQQD	IKKADPLFHPVKAGTCSMDKTQVQQAVE				
SEQ4	WADVDQRTLKTGEAFLAGLAP	QCGLTIHHQQNLEKADPLFHPVKAGTCSMDKTQVQQAVE				
SEQ6	WADVDQRTLKTGEAFLAGLAP	QCGLTIHHQQNLEKADPLFHPVKAGTCSMDKTQVQQAVE				
SEQ2	WADVDQRTLKTGEAFLAGLAP	QCGLTIHHQQNLEKADPLFHPVKAGTCSMDKT	RLQQAVE			
	* : ** : ***** : * ***** : : ***** ***** : : *****					

4/9

Fig. 1- continued

Numbering	150	160	170	180	190	200
AEH25057	KEAQTPIDN	LNQHYIPFL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNK
AEH25059	KEAQTPIDN	LNQHYIPFL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNK
AEH25058	KEAQTPIDN	LNQHYIPFL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNT
AEH25067	KEAQTPIDN	LNQHYIPFL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNT
AEH25071	KEAQTPIDN	LNQHYIPSL	ALMNTTTL	NFSKSPWC	QKHSADK	SCDLGLSMPSKLSIKDNGNE
AEH25072	KEAQTPIDN	LNQHYIPSL	ALMNTTTL	NFSKSPWC	QKHSADK	SCDLGLSMPSKLSIKDNGNE
AEH25074	KEAQTPIDN	LNQRYIPSL	ALMNTTTL	NFSKSPWC	QKHSADK	SCDLGLSMPSKLSIKDNGNE
AEH25076	KEAQTPIDN	LNQRYIPSL	ALMNTTTL	NFSKSPWC	QKHSADK	SCDLGLSMPSKLSIKDNGNE
AEH25075	KEAQTPIDN	LNQRYIPSL	ALMNTTTL	NFSKSPWC	QKHSADK	SCDLGLSMPSKLSIKDNGNE
AEH25073	KEAQTPIDN	LNQRYIPSL	ALMNTTTL	NFSKSPWC	QKHSADK	NCDLALSMPSKLSIKDNGNE
AEH25070	KEAQTPIDN	LNQHYIPFL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNE
AEH25069	KEAQTPIDN	LNQHYIPFL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNT
AEH25066	KEAQTPIDN	LNQHYIPFL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNK
AEH25060	KEAQTPIDN	LNQHYIPFL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNE
AEH25068	KEAQTPIDN	LNQHYIPFL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNE
AEH25063	KEAQTPIDN	LNQHYIPFL	ALMNTTTL	NFSKSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNK
AEH25065	KEAQTPIDN	LNQRYIPFL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNK
AEH25062	KEAQTPIDN	LNQHYIPFL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNK
AEH25061	KEAQTPIDN	LNQHYIPFL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNK
AEH25056	KEAQTPIDN	LNQHYIPFL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNK
AEH25051	KEAQTPIDN	LNQHYIPFL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNK
AEH25064	KEAQTPIDN	LNQHYIPSL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLGLSMPSKLSIKDNGNK
SEQ4	KEAQMPI	ENLNQHYIP	SLALMNTTTL	NFSTSAWC	QKHSADK	SCDLAQSMPSKLSIKDNGNK
SEQ6	KEAQTPIDN	LNQHYIPSL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLAQSMPSKLSIKDNGNK
SEQ2	KEAQTPIDN	LNQHYIPSL	ALMNTTTL	NFSTSAWC	QKHSADK	SCDLAQSMPSKLSIKDNGNK
	*****	**.*	*****.*	*****	*****.*	*****.*

5/9

Fig. 1 - continued

Numbering	210	220	230	240	250	260
AEH25057	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MARTPYIA		
AEH25059	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MARTPYIA		
AEH25058	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MARTPYIA		
AEH25067	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MARTPYIA		
AEH25071	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MERTPYIA		
AEH25072	VSLDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVYF	DL	MERTPYIA		
AEH25074	VALDGAIGLSSTLAEIFLLEYAQGMPQVWGN	IHSEQEWASLLKLHNVHF	DL	MERTPYIA		
AEH25076	VALDGAIGLSSTLAEIFLLEYAQGMPQVWGN	IHSEQEWASLLKLHNVHF	DL	MERTPYIA		
AEH25075	VALDGAIGLSSTLAEIFLLEYAQGMPQVWGN	IHSEQEWASLLKLHNVHF	DL	MERTPYIA		
AEH25073	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MERTPYIA		
AEH25070	VSLDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MERTPYIA		
AEH25069	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MERTPYIA		
AEH25066	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MERTPYIA		
AEH25060	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MARTPYIA		
AEH25068	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MARTPYIA		
AEH25063	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MARTPYIA		
AEH25065	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MARTPYIA		
AEH25062	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVYF	DL	MARTPYIA		
AEH25061	VALCGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MARTPYIA		
AEH25056	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MARTPYIA		
AEH25051	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MARTPYIA		
AEH25064	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNVQF	DL	MARTPYIA		
SEQ4	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHN	TQ	F	DL	MARTPYIA
SEQ6	VALDGAIGLSSTLAEIFLLEYAQGMPQAAWGN	IHSEQEWASLLKLHNAQ	F	DL	MARTPYIA	
SEQ2	VALDGA	VGLSSTLAEIFLLEYAQGMPQAAWGN	KIHSEQDWAELLKLHNAQ	F	DL	MARTPYIA
	: **:******.***:*****:* *****. **** *****					

6/9

Fig. 1- continued

Numbering	270	280	290	300	310	320
AEH25057	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25059	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25058	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25067	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25071	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25072	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25074	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25076	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25075	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25073	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25070	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25069	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25066	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25060	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25068	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25063	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25065	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25062	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25061	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25056	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25051	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
AEH25064	RHNGTPLLQAISNALNP	NATESKLPDISPDNKIL	FIAGHDTN	IANIAGMLNMRWTL	PGQP	
SEQ4	A HNGTPLLQ T ISNAL EPKADV SKLPDIS S DNKILFIAGHDTN	IANIAGMLNMRWTL	PGQP			
SEQ6	T HNGTPLLQ T ISNAL EPKADV SKLP G ISPDNKILF L AGHDTN	IANIAGMLNMRWTL	PGQP			
SEQ2	RHNGTPLLQAISNAL DPNAT ASKLPDISPDNKILFIAGHDTN	IANI S GMLNMRWTL	PGQP			
	*:*****:*****:*:*	*****.*.*.*****:*****:***:*****				

Fig. 1- continued

Numbering	330	340	350	360	370	380
AEH25057	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25059	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25058	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25067	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25071	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25072	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25074	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25076	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25075	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25073	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25070	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25069	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25066	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25060	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25068	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25063	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25065	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25062	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25061	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25056	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25051	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
AEH25064	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
SEQ4	DNTPPGGALVFERLADKSGKQY	ISVSMVYQTLEQLR	AQTPLSLN	EPAGSVQLKIPGCNDQ		
SEQ6	DNTPPGGALVFERLADKSGKQYVSVSMVYQTLEQLRSQTPLSLNQPAGSVQLKIPGCNDQ					
SEQ2	DNTPPGGAL	IFERLADK	AGKQYVSVSMVYQTLEQLR	AQTPLSL	KEPAGSVQLKIPGCNDQ	
	*****.*****.*****.*****.*****.*****.*****					

8/9

Fig. 1- continued

Numbering	390	400	410
AEH25057	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25059	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25058	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25067	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25071	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25072	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25074	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25076	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25075	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25073	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25070	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25069	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25066	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25060	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25068	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25063	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25065	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25062	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25061	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25056	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25051	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
AEH25064	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
SEQ4	TAEGYCPLSTFTRVVSQSVEPGCQL P		
SEQ6	TAEGYCPLSTFTRVVSQSVEPGCQLQ		
SEQ2	TAEGYCPL P T F K R V V S Q S E EPGCQLQ		
	***** . ** . ***** *****		

9/9

Fig. 2

Program: needle
 Matrix: BLOSUM62
 Gap initiation penalty: 10.0
 Gap extension penalty: 0.5
 Number of identical residues: 315
 Length of shortest sequence: 413
 % Identity: 315/413 = 76.3%

SEQ2_mature	1		NDTPASGYQVEKVVILS	17
			:: . . :	
UNIPROTQ6U677	1	MTISLFTHSPTRLLKCMPLAFIAASMLTTASYASETEPSGYQLEKVVILS	50	
SEQ2_mature	18	RHGVRAPTKMTQTMRDVTPTNTWPEWPKLGYITPRGEHLISLMGGFYREK	67	
UNIPROTQ6U677	51	RHGVRAPTKMTQTMRDVTPTNAWPEWPKLGYITPRGEHLVSLMGGFYRQK	100	
SEQ2_mature	68	FQQQGILSQGSCPTPNSIYVWADVVDQRTLKTGEAFLAGLAPQCGLTIHHQ	117	
		. : . . .:.		
UNIPROTQ6U677	101	FQQLGILSKGRCPTANDVYVWADVVDQTRKTGEAFLAGLAPECHLSIHHQ	150	
SEQ2_mature	118	QNLEKADPLFHPVKAGTCSMDKTRLQQAVEKEAQTPHENLNQHYIPSLAL	167	
		:::		
UNIPROTQ6U677	151	QDIKQADPLFHPVKAGVCTMEKTQVQQAVEQQAGMPIDQLNQHYRPALAL	200	
SEQ2_mature	168	MNTTLNLFSTSAWCQKHSADKSCDLAQSMPSKLSIKDNGNKVALDGAVGLS	217	
		::. . .:. : :		
UNIPROTQ6U677	201	MSSVLNFPKSTYCCQHSADQTCDLAQAMPKLSIKDNGNKVALDGAVGLS	250	
SEQ2_mature	218	STLAEIFLLEYAQGMPQAAWGKIHSEQDWAELLKLHNAQFDLMARTPYIA	267	
UNIPROTQ6U677	251	STLAEIFLLEYAQGMPDAAWGKIHSEQDWNALLTLHNAQFDLMSRTPYIA	300	
SEQ2_mature	268	RHNGTPLLQAISNALDPNATASKLPDISPDNKILFIAGHDTNIANISGML	317	
		: . .:. ::.....: ::: . .		
UNIPROTQ6U677	301	KHNGTPLLQTIIVSAINSQPSSRELPELSADNKILFPAGHDTNIANIAGMF	350	
SEQ2_mature	318	NMRWTLPGQPDNTPPGGALIFERLADKAGKQYVSVSMVYQTLEQLRAQTP	367	
		. . .		
UNIPROTQ6U677	351	GMSWALPGQPDNTPPGGALVFERWSDKTGKKYVSVQMMYQTLAQLRNQTP	400	
SEQ2_mature	368	LSLKEPAGSVQLKIPGCNDQTAEGYCPLPTFKRVVSQSEEPGCQLQ	413	
		: .:. . :		
UNIPROTQ6U677	401	LTLDKPAGSVALKIPGCDDQTAEGYCPLDTFTRLAKQNELVECQ	444	