

(19)



**Евразийское
патентное
ведомство**

(21) **201692439** (13) **A1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ**

(43) Дата публикации заявки
2017.04.28

(51) Int. Cl. *C12N 9/62* (2006.01)

(22) Дата подачи заявки
2015.06.03

(54) **ПРОЛИН-СПЕЦИФИЧНАЯ ЭНДОПРОТЕАЗА И ЕЕ ПРИМЕНЕНИЕ**

(31) **14170879.2; 14172644.8; 14172645.5**

(32) **2014.06.03; 2014.06.17; 2014.06.17**

(33) **EP**

(86) **PCT/EP2015/062328**

(87) **WO 2015/185593 2015.12.10**

(71) Заявитель:
ДСМ АйПи АССТЕС Б.В. (NL)

(72) Изобретатель:
**Лан Ван Дер Ян Метске, Брйне-
Паулус Де Ангела, Кристис Шанталь,
Спанс Маргине, Вондерворт Ван Де
Петер Йозеф Ида (NL)**

(74) Представитель:
Бадаева Т.Н., Фелицына С.Б. (RU)

(57) Настоящее изобретение относится к полипептиду, имеющему активность пролин-специфичной эндопротеазы, где указанный полипептид имеет менее 70% остаточной активности после выдержки полипептида при температуре 65°C в течение 15 мин. Кроме того, изобретение относится к полипептиду, имеющему активность пролин-специфичной эндопротеазы, включающей аминокислотную последовательность по SEQ ID NO:1, где последовательность SEQ ID NO:1 содержит, по меньшей мере, одну аминокислотную замену, выбранную из группы, состоящей из P469A, P469C, P469D, P469E, P469F, P469G, P469H, P469I, P469K, P469L, P469M, P469N, P469Q, P469R, P469S, P469T, P469V, P469W, P469Y, нуклеиновой кислоте, кодирующей полипептид, обладающий активностью пролин-специфичной эндопротеазы, способу изготовления вариантного полипептида, имеющего активность пролин-специфичной эндопротеазы, рекомбинантной клетке-хозяину, к способу получения полипептида и к способу приготовления пищевого или кормового продукта, в котором используется указанный полипептид.

A1

201692439

201692439

A1

ПРОЛИН-СПЕЦИФИЧНАЯ ЭНДОПРОТЕАЗА И ЕЕ ПРИМЕНЕНИЕ

Настоящее изобретение относится к полипептиду, имеющему активность пролин-специфичной эндопротеазы, к композиции, включающей данный полипептид, к нуклеиновой кислоте, кодирующей пролин-специфичную эндопротеазу, к экспрессирующему вектору, содержащему нуклеиновую кислоту, кодирующую пролин-специфичную эндопротеазу, к рекомбинантной клетке-хозяину, к способу получения пролин-специфичной эндопротеазы и к способу приготовления пищевого или кормового продукта, в котором используется пролин-специфичная эндопротеаза.

Предшествующий уровень техники

Пролин-специфичные эндопротеазы представляют собой ферменты, которые гидролизуют белок или пептид, в положении, в котором в белке или пептиде находится пролин.

Пролин-специфичная эндопротеаза, например, может быть получена из *Aspergillus niger* или *Penicillium chrysogenum*, например, как описано в WO2002/046381 и WO2009/144269, соответственно.

Другие пролин-специфичную эндопротеазы известны из WO2012/174127. WO2012/174127 раскрывает пролин-специфические протеазы из *Botryotinia fuckeliana*, *Aspergillus clavatus*, *Sclerotinia sclerotiotum*, *Mycosphaerella graminicola*, *Neurospora crasse*, *Talaromyces stipitatus* и *Gibberella zeae*.

Пролин-специфичная эндопротеаза может быть использована в различных приложениях, например, при деградации глютена (см., например, WO2005/027953 или WO2003/068170). Глютен является нерастворимой белковой фракцией зерновых, таких как пшеница, рожь, овес и ячмень. Глютен представляет собой сложную смесь молекул глютеина и проламина, которые, как полагают, вызывают токсические эффекты, например у пациентов, страдающих от целиакии. Целиакия-спру или целиакия считается аутоиммунным заболеванием. Пациенты, страдающие от целиакии-спру должны следовать строгой безглютеновой диете, которую очень трудно соблюдать, поскольку глютен широко используется. Применение пролин-специфичной эндопротеазы в качестве лекарственного средства или диетической добавки может облегчить необходимость в строгой безглютеновой диете (WO2003/068170).

Пролин-специфичные эндопротеазы также используются для уменьшения мутности в пиве, в которое пролин-специфические протеазы могут быть добавлены в ходе

нескольких стадий производства (WO 2002/046381).

Желательно, чтобы ферменты в пищевых и кормовых применениях имели соответствующий оптимум pH и предпочтительно не являлись активными в конечном пищевом продукте или напитке.

Целью настоящего изобретения является альтернативная пролин-специфичная эндопротеаза с улучшенными характеристиками.

Сущность изобретения

В одном из аспектов настоящее изобретение относится к полипептиду, имеющему активность пролин-специфичной эндопротеазы, где указанный полипептид содержит менее 70% остаточной активности на ацетил-AlaAlaPro-паранитроанилине (Ac-AAP-PNA), используемом в качестве субстрата после того, как полипептид был выдержан при температуре 65°C в течение 15 мин. Остаточная активность полипептида, обладающего активностью пролин-специфичной эндопротеазы, преимущественно определяется при использовании ацетил-AlaAlaPro-паранитроанилина (Ac-AAP-PNA) в качестве субстрата при температуре 20°C и при pH 4,5, например, в буфере при pH 4,5, например, в натрий-ацетатном (NaAc) буфере, который может включать еще одну соль, такую как NaCl. Остаточная активность может быть определена путем инкубации полипептида, как описано в данном документе, при температуре 20°C, и при pH 4,5 в течение 60 мин.

В другом своем аспекте настоящее изобретение относится к полипептиду, имеющему активность пролин-специфичной эндопротеазы, где указанный полипептид выбран из группы, состоящей из полипептида, имеющего активность пролин-специфичной эндопротеазы, необязательно имеющего менее 70% остаточной активности после того, как полипептид был выдержан при температуре 65°C в течение 15 мин., где полипептид выбран из группы, состоящей из:

i. полипептида, который, когда выровнен относительно аминокислотной последовательности по SEQ ID NO: 1, содержит аминокислотную замену в положении, соответствующем положению 469, где положение определяется относительно SEQ ID NO: 1;

ii. полипептида, который, когда выровнен относительно аминокислотной последовательности по SEQ ID NO: 1, содержит аминокислоту, выбранную из группы, состоящей из Ala (A), Cys (C), Asp (D), Glu (E), Phe (F), Gly (G), His (H), Ile (I), Lys (K), Leu (L), Met (M), Asn (N), Gln (Q), Arg (R), Ser (S), Thr (T), Val (V), Trp (W) и Tyr (Y), в положении, соответствующем положению 469, где положение определяется относительно SEQ ID NO: 1;

iii. полипептида, содержащего аминокислотную последовательность по SEQ ID NO: 1, где последовательность SEQ ID NO: 1 содержит, по меньшей мере, одну аминокислотную замену, выбранную из группы, состоящей из P469A, P469C, P469D, P469E, P469F, P469G, P469H, P469I, P469K, P469L, P469M, P469N, P469Q, P469R, P469S, P469T, P469V, P469W и P469Y, где аминокислотные замены определяются относительно SEQ ID NO: 1;

iv. полипептида по i)-iii), но утратившего сигнальную последовательность и/или последовательность пробелка;

v. полипептида по i)-iv), где полипептид идентичен, по меньшей мере, на 60%, 70%, 75%, 80, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% аминокислотной последовательности по SEQ ID NO: 1;

vi. полипептида, кодируемого нуклеиновой кислотой, которая идентична, по меньшей мере, на 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, или 99% или 100% SEQ ID NO: 2, или зрелой кодирующей последовательности SEQ ID NO: 2, где последовательность SEQ ID NO: 2 включает, по меньшей мере, одну мутацию, кодирующую, по меньшей мере, одну аминокислотную замену, выбранную из группы P469A, P469C, P469D, P469E, P469F, P469G, P469H, P469I, P469K, P469L, P469M, P469N, P469Q, P469R, P469S, P469T, P469V, P469W и P469Y, где аминокислотные замены определяются относительно SEQ ID NO: 1.

Настоящее изобретение также относится к композиции, содержащей полипептид, имеющий пролин-специфичную эндопротеазу, как описано в данном документе.

В другом аспекте настоящее изобретение относится к способу получения вариантного полипептида, имеющего активность пролин-специфичной эндопротеазы, как описано в данном документе.

Изобретение также относится к нуклеиновой кислоте, кодирующей пролин-специфичную эндопротеазу, последовательность которой идентична, по меньшей мере, на 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, или 100% SEQ ID NO: 2, или зрелой кодирующей последовательности SEQ ID NO: 2, где последовательность SEQ ID NO: 2 содержит, по меньшей мере, одну мутацию, кодирующую, по меньшей мере, одну аминокислотную замену, выбранную из группы, состоящей из P469A, P469C, P469D, P469E, P469F, P469G, P469H, P469I, P469K, P469L, P469M, P469N, P469Q, P469R, P469S, P469T, P469V, P469W и P469Y, где аминокислотные замены определяются относительно SEQ ID NO: 1.

В другом аспекте настоящее изобретение относится к экспрессирующему вектору, содержащему полинуклеотид, кодирующий полипептид, описанный в данном документе.

В другом своем аспекте настоящее изобретение относится к рекомбинантной клетке-хозяину, содержащей полинуклеотидную последовательность или экспрессирующий вектор, как описано в данном документе.

В еще одном аспекте настоящее изобретение относится к способу получения полипептида, включающему культивирование клетки-хозяина, как описано в данном документе в условиях, обеспечивающих экспрессию полипептида, и получение полипептида.

В другом своем аспекте настоящее изобретение относится к способу приготовления пищевого или кормового продукта, включающему инкубирование промежуточной формы пищевого или кормового продукта с полипептидом, или композицией, содержащей полипептид, описанный в данном документе, и приготовления пищевого продукта.

Настоящее изобретение также относится к пищевому или кормовому продукту, получаемому способом, как описано в настоящем документе.

Определения

Термин «хлебобулочные изделия» в данном документе определен как любой продукт, изготовленный из теста или жидкого теста. Продукт может иметь мягкий или хрустящий характер и может быть белого, светлого или темного типа. Хлебобулочные изделия включают, без ограничения перечисленным, хлеб, например, белый хлеб, хлеб из непросеянной муки или ржаной хлеб, хлеб типа французский багет, продукты слоеного теста, такие как (датская) сдоба, круассаны или слоеное тесто, лаваш, лепешки, тако, пирожные, торты, печенье, бисквиты, пончики, рогалики, пироги, кексы, хлеб, приготовленный на пару, и хрустящий хлеб. Типы хлебобулочных изделий, способы их описания и производства известны специалистам в данной области, см., например, «Baking Science and Technology», by E.J. Pyler, L.A. Gorton, 2008, (2 volumes) Sosland Publishing Company, Kansas, USA, or «Baked Products: Science, Technology and Practice» by S.P. Cauvain, L.S. Young, 2006, Blackwell Publishing Ltd, Oxford, UK.

Термин «комплементарная цепь» может быть использован взаимозаменяемо с термином «комплементарность». Комплементарность цепи нуклеиновой кислоты может быть комплементарностью к кодирующей цепи или комплементарностью к некодирующей цепи. Когда речь идет о двухцепочечных нуклеиновых кислотах, комплементарность к нуклеиновой кислоте, кодирующей полипептид, относится к нити, которая комплементарна нити, кодирующей аминокислотную последовательность, или к любой молекуле нуклеиновой кислоты, содержащей то же самое.

Термин «регуляторная последовательность» может быть использован

взаимозаменяемо с термином «нуклеотидная последовательность, регулирующая экспрессию». Термин, при использовании в данном описании, относится к нуклеотидной последовательности, необходимой для и/или влияющий на экспрессию функционально связанной кодирующей последовательности в конкретном организме-хозяине или *in vitro*. Когда две нуклеотидной последовательности, функционально связаны, они, как правило, будет находиться в одной и той же ориентации, а также в одной и той же рамке считывания. Они, как правило, будут по существу непрерывными, хотя это может не потребоваться. Нуклеотидные последовательности, регулирующие экспрессию, такие как, в числе прочего, соответствующие последовательности инициации транскрипции, терминации, промотора, лидера, сигнального пептида, пропептида, препропептида или энхансера; последовательность Шайна-Дельгарно, репрессорные или активаторные последовательности; сигналы эффективного процессинга РНК, такие как сигналы сплайсинга и сигналы полиаденилирования; последовательности, которые стабилизируют цитоплазматическую мРНК; последовательности, которые повышают эффективность трансляции (например, сайты связывания рибосом); последовательности, которые повышают стабильность белка; и, при необходимости, последовательности, которые усиливают секрецию белка, могут быть любой последовательностью нуклеиновой кислоты, обнаруживающей активность в выбранном организме-хозяине и могут быть получены из генов, кодирующих белки, которые являются либо эндогенными, либо гетерологичными клетке-хозяину. Каждая регуляторная последовательность может быть нативной или чужеродной для нуклеотидной последовательности, кодирующей полипептид. При необходимости, регуляторная последовательность может быть снабжена линкерами с целью введения специфических сайтов рестрикции, облегчающих лигирование регуляторных последовательностей с кодирующей областью нуклеотидной последовательности, кодирующей полипептид. Регуляторные последовательности могут быть оптимизированы для конкретных целей.

А «молочный продукт» относится к любому виду продукта на основе молока, предназначенному для применения в качестве продукта питания, корма или напитка, включая, без ограничения перечисленным, сыр, молоко, обезжиренное молоко, подкисленное молоко, пахту, сгущенное молоко, спреды, маргарины, йогурт, мороженое, молоко, сливочное масло, ЕМС (фермент-модифицированный сыр), дульче де лече, кофейные сливки, сливки, топленое масло, молочный аналог, и так далее. Сыр может быть любым видом сыра, например, свежим сыром, твердым сыром, творогом, сливочным сыром, сыром с белой плесенью, сыром с голубой плесенью и плавленым сыром. Примерами свежего сыра являются рикотта, сливочный сыр, сыр невшатель или сыр

«коттедж». Примерами твердых сыров являются чеддер, данбо, манчего, сент-паулин, чеддер, монтерей, колби, эдам, гауда, мюнстера, швейцарский сыр, грюйер, эмментальский сыр, пармиджано реджано, грана падано, пармезан, пекорино, проволоне и романо. Примеры творожного сыра, представляют собой сыр фета, сыр котиха, сыр паста филата, такой как моцарелла, сыр квесо фреско. Примером сливочного сыра является сыр филадельфия. Примером сыра с белой плесенью является бри и камамбер. Примерами сыра с голубой плесенью являются горгонзола и данаблю.

При использовании в данном описании термин «эндогенный» относится к нуклеотидной или аминокислотной последовательности встречающиеся в хозяине в природе.

Эндопептидазы или эндопротеиназы способны разрушать пептидные связи нетерминальных аминокислот (т.е. в белке), в отличие от экзопептидаз, которые разрушают пептидные связи либо с амино- либо с карбоксильного конца. Эндопептидазы не склонны к разрушению пептидов на мономеры, но приводят к появлению относительно больших пептидных фрагментов. Специфическая генерация относительно больших фрагментов весьма предпочтительна во многих приложениях, связанных с пищевыми продуктами и кормами. Частным случаем эндопептидазы является олигопептидаза, чьими субстратами являются олигопептиды, а не белки.

Термин «экспрессия» включает любую стадию, приводящую к получению полипептида, включая, без ограничения перечисленным, транскрипцию, процессинг РНК, трансляцию, пост-трансляционную модификацию и секрецию.

Полинуклеотиды по настоящему изобретению, как описано в настоящем документе, могут быть сверхэкспрессированы в клетке-хозяине по изобретению по сравнению с родительскими клетками, в котором указанный ген не сверхэкспрессирован. Сверхэкспрессия полинуклеотидной последовательности определяется в данном документе как экспрессия последовательности указанного гена, которая приводит к активности полипептида, кодируемого указанной последовательностью в клетке-хозяине, по меньшей мере, в 1,1, по меньшей мере, в 1,25 или, по меньшей мере, в 1,5 больше активности полипептида в клетке-хозяине; предпочтительно, если активность указанного полипептида является, по меньшей мере, 2-кратной, более предпочтительно, по меньшей мере, 3-кратной, более предпочтительно, по меньшей мере, 4-кратной, более предпочтительно, по меньшей мере, 5-кратной, еще более предпочтительно, по меньшей мере, 10-кратной и наиболее предпочтительно, по меньшей мере, 20-кратной по отношению к активности полипептида в родительской клетке.

Экспрессирующий вектор содержит полинуклеотид, кодирующий полипептид,

такой как полипептид в соответствии с настоящим изобретением, функционально связанный с соответствующими регуляторными последовательностями (например, промотором и стоп-сигналами транскрипции и трансляции) для экспрессии и/или трансляции полинуклеотида *in vitro*, или в клетке-хозяине.

Экспрессирующий вектор может быть любым вектором (например, плазмидой или вирусом), который может быть легко подвергнут процедурам рекомбинантных ДНК, и который может осуществлять экспрессию полинуклеотида. Выбор вектора, как правило, зависит от совместимости вектора с клеткой, в которую вектор должен быть введен. Векторы могут быть линейными или замкнутыми кольцевыми плазмидами. Вектор может быть автономно реплицирующимся вектором, т.е. вектором, который существует как экстра-хромосомная сущность, репликация которого не зависит от хромосомной репликации, например, может быть плазмидой, экстра-хромосомным элементом, мини-хромосомой или искусственной хромосомой. В ином случае, вектор может быть таким, который при введении в клетку-хозяина интегрируется в геном и реплицируется вместе с хромосомой (хромосомами), в которую он был интегрирован. Интегративный клонирующий вектор может интегрироваться в случайном порядке или в заданный целевой локус в хромосомах клетки-хозяина. Векторная система может быть одним вектором или плазмидой или двумя или несколькими векторами или плазмидами, которые вместе содержат полную ДНК для введения в геном клетки-хозяина, или транспозон. Вектор по изобретению может содержать один, два или более, например три, четыре или пять полинуклеотидов по изобретению, например, для избыточной экспрессии.

Термин «ген», при использовании в данном описании, относится к сегменту молекулы нуклеиновой кислоты, кодирующей полипептидную цепь, которая может включать или не включать регуляторные последовательности генов, предшествующие и последующие кодирующей последовательности, например, промоторы, энхансеры и т.д., а также интроны (интроны) между отдельными кодирующими сегментами (экзонами). Следует также иметь в виду, что определение гена может включать нуклеиновые кислоты, которые не кодируют полипептид, а скорее обеспечивают шаблоны для транскрипции функциональных молекул РНК, таких как тРНК, рРНК и т.д.

Клетка-хозяин, как определено в данном описании, представляет собой организм, подходящий для генетических манипуляций и организм, который можно культивировать при плотности клеток, пригодной для промышленного производства целевого продукта, такого как полипептид в соответствии с настоящим изобретением. Клетка-хозяин может быть клеткой-хозяином, обнаруживаемой в природе или клеткой-хозяином, полученной из родительской клетки-хозяина после генетической манипуляции или классического

мутагенеза. Предпочтительно, если клетка-хозяин представляет собой рекомбинантную клетку-хозяин. Клетка-хозяин может быть прокариотической, архебактериальной или эукариотической клеткой-хозяином. Прокариотическая клетка-хозяин, может быть, без ограничения указанным, бактериальной клеткой-хозяином. Эукариотическая клетка-хозяин, может быть, без ограничения указанным, клеткой-хозяином из дрожжей, грибов, амёб, водорослей, растений, животных или насекомых.

Термин «гетерологичный», при использовании в данном описании, относится к последовательностям нуклеиновых кислот или аминокислот, не встречающихся в природе в клетке-хозяине. Другими словами, нуклеотидная или аминокислотная последовательности не идентичны обнаруживаемым в естественных условиях в клетке-хозяине.

Термин «гибридизация» означает спаривание по существу комплементарных цепей олигомерных соединений, таких как соединения нуклеиновых кислот. Гибридизация может выполняться в условиях низкой, средней и высокой жесткости. Условия низкой степени жесткости, включают гибридизацию в $6\times$ хлориде натрия/цитрате натрия (SSC) при температуре около 45°C , с последующими двумя промывками в $0,2 \times$ SSC, 0,1% SDS, по меньшей мере, при 50°C (температура промывок может быть увеличена до 55°C при низкой степени жесткости). Средние условия жесткости, включают гибридизацию в $6 \times$ SSC при около 45°C , после чего следует одна или несколько промывок в $0,2 \times$ SSC, 0,1% SDS при 60°C , а условия высокой жесткости гибридизации включают гибридизацию в $6 \times$ SSC при температуре около 45°C , после чего следует одна или несколько промывок в $0,2 \times$ SSC, 0,1% SDS при 65°C .

Нуклеотидная или полинуклеотидная последовательность определяется в данном документе как нуклеотидный полимер, содержащий, по меньшей мере, 5 нуклеотидов или единиц нуклеиновой кислоты. Нуклеотид или нуклеиновая кислота, относятся к РНК и ДНК. Термины «нуклеиновая кислота» и «полинуклеотидная последовательность» используются в данном документе взаимозаменяемо.

«Пептид» относится к короткой цепи аминокислотных остатков, соединенных пептидной (амидной) связью. Самый короткий пептид, дипептид, состоит из 2-х аминокислот, соединенных одной пептидной связью.

Термин «полипептид» относится к молекуле, включающей аминокислотные остатки, соединенные пептидными связями и содержащей более пяти аминокислотных остатков. Термин «белок», используемый в данном документе, является синонимом термина «полипептид» и, возможно, также относятся к двум или более полипептидам. Таким образом, термины «протеин» и «полипептид» может быть использован

взаимозаменяемо. Полипептиды могут быть необязательно модифицированы (например, гликозилированы, фосфорилированы, ацилированы, фарнезилированы, пренилированы, сульфированы, и т.п.) для добавления функциональности. Полипептиды, проявляющие активность в присутствии специфического субстрата при определенных условиях могут быть отнесены к ферментам. Следует понимать, что, в результате вырожденности генетического кода может быть получено множество нуклеотидных последовательностей, кодирующих один полипептид.

«Выделенный нуклеотидный фрагмент» представляет собой фрагмент нуклеиновой кислоты, который не встречается в природе в виде фрагмента, и не может быть найден в естественном состоянии.

Термин «выделенный полипептид», используемый в данном документе, означает полипептид, который отделен от, по меньшей мере, одного компонента, например, другого полипептидного материала, с которым он связан в природе. Выделенный полипептид может быть свободным от каких-либо других примесей. Выделенный полипептид может быть, по меньшей мере, на 50% чистым, например, по меньшей мере, 60% чистым, по меньшей мере, 70% чистым, по меньшей мере, 75% чистым, по меньшей мере, 80% чистым, по меньшей мере, 85% чистым, по меньшей мере, 80% чистым, в не менее 90%, или, по меньшей мере, 95% чистым, 96%, 97%, 98%, 99%, 99,5%, 99,9%, что определяется с помощью SDS-PAGE или любого другого аналитического метода, пригодного для этой цели, и известного специалисту в данной области. Выделенный полипептид, может быть получен рекомбинантной клеткой-хозяином.

«Зрелый полипептид» определяется в данном документе как полипептид в его окончательной виде, и его получают после трансляции мРНК в полипептид и посттрансляционных модификаций указанного полипептида. Посттрансляционные модификации включают N-концевую обработку, C-концевое укорочение, гликозилирование, фосфорилирование и удаление расщеплением лидерных последовательностей, таких как сигнальные пептиды, пропептиды и/или препропептиды.

«Кодирующая последовательность зрелого полипептида» означает полинуклеотид, который кодирует зрелый полипептид.

Термин «нуклеотидная конструкция» в данном описании упоминается как молекула нуклеиновой кислоты, либо одно- или двухцепочечная, которая выделена из природного гена или которая была модифицирована для включения нуклеотидных сегментов, которые объединены и расположены рядом друг с другом таким образом, который не существует в природе. Термин нуклеотидная конструкция является синонимом термина «экспрессирующая кассета» или «экспрессирующий вектор», если

нуклеотидная конструкция содержит все регуляторные последовательности, необходимые для экспрессии кодирующей последовательности, где указанные регуляторные последовательности функционально связаны с указанной кодирующей последовательностью.

«Пролин-специфичная эндопротеаза» представляет собой протеазу, которая гидролизует белок или пептид, в положении, в котором белок или пептид содержит остаток пролина. Пролин-специфичная эндопротеаза может иметь пролин-специфичную эндопротеазную и/или олигопептидазную пролин-специфичную активности (ЕС3.4.21.26). Пролин-специфичная эндопротеаза предпочтительно представляет собой фермент, который гидролизует пептидную связь на карбоксильном конце остатков пролина, что дает пептид и/или полипептидный фрагмент с С-концевым пролином.

Термин «промотор» определяется в данном документе как ДНК-последовательность, которая связывается РНК-полимеразой, и направляет полимеразу к нижележащей нуклеотидной последовательности правильного сайта начала транскрипции для инициации транскрипции.

Термин «рекомбинантный», при использовании в отношении клетки, нуклеиновой кислоты, белка или вектора, указывает на то, что клетка, нуклеиновая кислота, белок или вектор, был модифицированы путем введения гетерологичной нуклеотидной последовательности или белка или изменения нативной нуклеиновой кислоты или белка, или что клетка происходит из клетки, модифицированной таким образом. Так, например, рекомбинантные клетки экспрессируют гены, которые не обнаруживаются в нативной (нерекомбинантной) форме клетки или экспрессируют нативные гены, которые в ином случае аномально экспрессируются, недостаточно экспрессируются или не экспрессируются вообще. Термин «рекомбинантный» является синонимом «генетически модифицированный» и «трансгенный».

«Идентичность последовательности», или гомология последовательности используются взаимозаменяемо в настоящем документе. Для целей настоящего изобретения, этот термин определен в данном документе для того, чтобы определить процент гомологии последовательности или идентичности последовательности из двух аминокислотных последовательностей или двух нуклеотидных последовательностей, которые выравнивают с целью оптимального сравнения. Для того чтобы оптимизировать выравнивание между двумя последовательностями могут быть введены разрывы в любую из этих двух сравниваемых последовательностей. Такое выравнивание может быть осуществлено по всей длине сравниваемых последовательностей. В качестве альтернативы, выравнивание может быть осуществлено на более коротком участке,

например, длиной около 20, около 50, около 100 или более нуклеотидов или аминокислот. Идентичность последовательности представляет собой процент идентичных совпадений между двумя последовательностями на указанном выровненном участке. Процент идентичности последовательности между двумя аминокислотными последовательностями или между двумя нуклеотидными последовательностями может быть определен с использованием алгоритма Нидлмана и Вунша для выравнивания двух последовательностей. (Needleman, S. B. and Wunsch, C. D. (1970) *J. Mol. Biol.* 48, 443-453). Как аминокислотные последовательности, так и нуклеотидные последовательности могут быть выровнены с помощью данного алгоритма. Алгоритм Нидлмана-Вунша был реализован в компьютерной программе NEEDLE. Для целей настоящего изобретения была использована программа NEEDLE из пакета EMBOSS: The European Molecular Biology Open Software Suite (2000) Rice, P. Longden, I. and Bleasby, A. *Trends in Genetics* 16, (6) pp276—277, <http://emboss.bioinformatics.nl/>). Для белковых последовательностей используется матрица замещения EBLOSUM62. Для нуклеотидных последовательностей используется EDNAFULL. Были использованы следующие опциональные параметры: штраф за открытие разрыва 10 и за расширение разрыва 0,5. Специалисту будет понятно, что все эти различные параметры будут давать несколько различные результаты, но что общий процент идентичности двух последовательностей существенно не изменится при использовании различных алгоритмов.

После выравнивания программой NEEDLE, как описано выше, процент идентичности последовательностей между последовательностью запроса и последовательностью по изобретению вычисляется следующим образом: количество соответствующих положений в выравнивании, демонстрирующих идентичную аминокислоту или идентичный нуклеотид в обеих последовательностях, разделенное на общую длину выравнивания после вычитания общего числа разрывов в выравнивании. Идентичность, как определено в данном документе, может быть получена в NEEDLE с помощью опции NOBRIEF и маркирована на выходе программы, как «самая длинная идентичность».

Нуклеотидные и белковые последовательности по настоящему изобретению, дополнительно могут быть использованы в качестве «последовательности запроса» для того, чтобы выполнить поиск в общедоступных базах данных, например, идентифицировать других членов семейства или родственные последовательности. Такие поиски могут быть выполнены с помощью программы NBLAST и XBLAST (версия 2.0) из Altschul, et al. (1990) *J. Mol. Biol.* 215:403—10. Нуклеотидные поиски BLAST могут быть выполнены с помощью программы NBLAST, балл = 100, длина «слова» = 12 для

получения нуклеотидных последовательностей, гомологичных нуклеотидным молекулам по изобретению. Поиск BLAST, белок может быть выполнен с помощью программы XBLAST, балл = 50, длина «слова» = 3, для получения аминокислотных последовательностей, гомологичных белковым молекулам по изобретению. Для получения выравненных с разрывами для целей сравнения, может быть использован Gapped BLAST, как описано в Altschul et al., (1997) *Nucleic Acids Res.* 25(17): 3389-3402. При использовании BLAST и Gapped BLAST, могут быть использованы параметры по умолчанию соответствующих программ (например, XBLAST и NBLAST). См. домашнюю страницу Национального центра биотехнологической информации на <http://www.ncbi.nlm.nih.gov/>.

Термин «по существу чистый» в отношении полипептидов относится к препарату полипептида, который содержит самое большее 50% по массе другого полипептидного вещества. Полипептиды, описанные в данном документе, предпочтительно по существу находятся в чистой форме. В частности, предпочтительно, чтобы полипептиды, описанные в данном документе, находились «по существу в чистой форме», т.е., чтобы препарат полипептида по существу не содержал другого полипептидного вещества. Необязательно, полипептид может также быть по существу быть свободным от неполипептидного вещества, такого как нуклеиновые кислоты, липиды, компоненты среды и тому подобное. В данном описании термин «по существу чистый полипептид» является синонимом терминов «выделенный полипептид» и «полипептид в выделенном виде». Термин «по существу чистый» в отношении полинуклеотида относится к препарату полинуклеотида, который содержит не более 50% по массе другого полинуклеотидного материала. Полинуклеотиды, описанные в данном документе, предпочтительно по существу находятся в чистой форме. В частности, предпочтительно, чтобы полинуклеотид, раскрытый в данном документе, был «по существу в чистой форме», т.е., чтобы препарат полинуклеотида по существу не содержал другого полинуклеотидного материала. Необязательно, полинуклеотид может также быть по существу свободным от не полинуклеотидных материалов, таких как полипептиды, липиды, компоненты среды и тому подобное. В данном описании термин «по существу чистый полинуклеотид» является синонимом терминов «выделенный полинуклеотид» и «полинуклеотид в выделенном виде».

«Замена», используемая в данном документе по отношению к полипептидам или нуклеиновым кислотам, означает замену одной или нескольких аминокислот в полипептидной последовательности или из одного или нескольких нуклеотидов в полинуклеотидной последовательности, соответственно, на другие аминокислоты или

нуклеотиды, соответственно. Например, замена указывает на то, что положение в полипептиде, раскрытое в данном описании, например, в варианном полипептиде, которое соответствует, по меньшей мере, одному положению, изложенному выше в SEQ ID NO: 1, содержит аминокислотный остаток, который отсутствует в этом же положении в исходном полипептиде (например, исходной последовательности SEQ ID NO: 1).

«Синтетическую молекулу», например, синтетическую нуклеиновую кислоту или синтетический полипептид получают *in vitro* химическим или ферментативным синтезом. Она включает, без ограничения указанным, варианты нуклеиновые кислоты, полученные с использованием кодонов, которые оптимальны для выбранных организмов-хозяев.

Синтетическая нуклеиновая кислота может быть оптимизирована по использованию кодонов, предпочтительно в соответствии со способами, описанными в WO2006/077258 и/или WO2008000632, которые включены в данное описание в виде ссылки. Документ WO2008/000632 посвящен кодовой оптимизации. Оптимизация кодовых пар представляет собой способ, в котором нуклеотидные последовательности, кодирующие полипептид, модифицируют по использованию кодонов, в частности, оптимизируют в отношении используемых кодовых пар для получения улучшенной экспрессии нуклеотидной последовательности, кодирующей полипептид, и/или улучшение выработки кодируемого полипептида. Кодовые пары определяются как совокупность двух последующих триплетов (кодонов) в кодирующей последовательности. Специалистам в данной области известно, что использование кодонов должно быть адаптировано в зависимости от вида хозяина, что возможно даст варианты со значительными отклонениями в гомологии от SEQ ID NO: 2, но при этом кодирующих полипептид согласно изобретению.

При использовании в данном описании термины «вариант», «производное», «мутант» или «гомолог» могут быть использованы взаимозаменяемо. Они могут относиться либо к полипептидам, либо к нуклеиновым кислотам. Варианты включают замены, вставки, делеции, усечения, трансверсии и/или инверсии, в одном или нескольких местах по отношению к эталонной последовательности. Варианты могут быть изготовлены, например, с помощью сайт-насыщающего мутагенеза, сканирующего мутагенеза, инсерционного мутагенеза, случайного мутагенеза, сайт-направленного мутагенеза и направленной эволюции, а также различных других подходов рекомбинации, известных специалисту в данной области. Вариантные гены нуклеиновых кислот могут быть синтезированы искусственно известными способами в данной области.

Чертежи

Фигура 1: Вектор pGBTOP-16, используемый для клонирования гена гамма-линоленовой кислоты. Вектор pGBTOP-16 получили из вектора pGBTOP-12, описанного в публикации WO 2011/009700. В дополнение к pGBTOP-12, он содержит ген *ccdB* из *E. coli* для положительного отбора на присутствие вставки между сайтами клонирования EcoRI и PstI. Сайт рестрикции PstI заменяет сайт рестрикции SnaBI, присутствующий в pGBTOP-12. Этот вектор перед трансформацией линейаризовали гидролизом NotI.

Фигура 2. Выравнивание эталонной пролин-специфичной эндопротеазы из *Aspergillus niger* с гомологичными пролин-специфичными эндопротеазами, полученными из *A. carbonarius*, *A. flavus*, *A. aculeatus* и *Rasamsonia emersonii*. Выравнивание осуществляли с помощью программы ClustalW, реализованной в программе BioEdit mbio Государственного университета Северной Каролины (NCSSU). (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>).

Последовательности

SEQ ID NO: 1: аминокислотная последовательность пролин-специфичной эндопротеазы *Aspergillus niger*, содержащая сигнальную последовательность пектинметилэстеразы.

SEQ ID NO: 2: нуклеотидная последовательность пролин-специфичной эндопротеазы *Aspergillus niger*, содержащая сигнальную последовательность пектинметилэстеразы.

SEQ ID NO: 3 Аминокислотная последовательность цитохрома С сердца лошади.

SEQ ID NO: 4: Фрагмент цитохрома С, гидролизованного с помощью PEP в соответствии с настоящим изобретением

SEQ ID NO: 5: Фрагмент цитохрома С, гидролизованного с помощью PEP в соответствии с настоящим изобретением

SEQ ID NO: 6: Фрагмент цитохрома С, гидролизованного с помощью PEP в соответствии с настоящим изобретением

SEQ ID NO: 7: Фрагмент цитохрома С, гидролизованного с помощью PEP в соответствии с настоящим изобретением

SEQ ID NO: 8: Фрагмент цитохрома С, гидролизованного с помощью PEP в соответствии с настоящим изобретением

SEQ ID NO: 9: Аминокислотная последовательность пролин-специфичной эндопротеазы (PEP) из *Aspergillus carbonarius* BC2G075 с сигнальной последовательностью пектинметилэстеразы из *A. niger* и пропоследовательностью PEP из *A. niger*

SEQ ID NO: 10: Аминокислотная последовательность пролин-специфичной эндопротеазы (PEP) из *Aspergillus Flavus* BC2G077 с сигнальной последовательностью пектинметилэстеразы из *A. niger* и пропоследовательностью PEP из *A. niger*

SEQ ID NO: 11: Аминокислотная последовательность пролин-специфичной эндопротеазы (PEP) из *Aspergillus Aculeatus* BC2G076 с сигнальной последовательностью пектинметилэстеразы из *A. niger* и пропоследовательностью PEP из *A. niger*

SEQ ID NO: 12: Аминокислотная последовательность пролин-специфичной эндопротеазы из *Rasamsonia emersonii*

SEQ ID NO: 13: Нуклеотидная последовательность пролин-специфичной эндопротеазы (PEP) из *Aspergillus carbonarius* BC2G075 с сигнальной последовательностью пектинметилэстеразы из *A. niger* и пропоследовательностью PEP из *A. niger*

SEQ ID NO: 14: Нуклеотидная последовательность пролин-специфичной эндопротеазы (PEP) из *Aspergillus flavus* BC2G077 с сигнальной последовательностью пектинметилэстеразы *A. niger* и пропоследовательностью PEP из *A. niger*

SEQ ID NO: 15: Нуклеотидная последовательность пролин-специфичной эндопротеазы (PEP) из *Aspergillus aculeatus* BC2G076 с сигнальной последовательностью пектинметилэстеразы *A. niger* и пропоследовательностью PEP *A. niger*

SEQ ID NO: 16: Нуклеотидная последовательность пролин-специфичной эндопротеазы из *Rasamsonia emersonii*

Подробное описание

В одном аспекте настоящее изобретение относится к полипептиду, имеющему активность пролин-специфичной эндопротеазы, где указанный полипептид содержит менее 70% остаточной активности при использовании ацетил-AlaAlaPro-паранитроанилина (Ac-AAP-PNA) в качестве субстрата после того, как полипептид был выдержан при температуре 65°C в течение 15 мин. Остаточную активность пролин-специфичной эндопротеазы измеряли с помощью ацетил-AlaAlaPro-паранитроанилина (Ac-AAP-PNA) при pH 4,5, например в натрий-ацетатном буфере при pH 4,5 при 20 градусах Цельсия. Удивительно, но полипептид, который содержит менее 55% остаточной активности полипептида после того, как полипептид был выдержан при температуре 65°C в течение 15 мин., предпочтительно может быть использован в таких приложениях, как пищевые и кормовые продукты, где желательна отсутствующая или незначительная остаточная активность. Предпочтительно, если полипептид предлагаемый изобретением содержит менее 70%, 60%, 55%, 50%, 45%, 40%, 30%, 20%, 15%, 10%, например, менее

5% остаточной активности после того, как полипептид был выдержан при температуре 65°C в течение 15 мин. Как определено в данном документе, менее чем 70%, 60%, 55%, или менее чем 50%, или менее чем 45%, 40%, 30%, 20%, 15%, 10% или 5% остаточной активности означает, что полипептид обладает менее чем 55% или менее 50%, или менее чем на 45%, 40%, 30%, 20%, 15%, 10% или 5%, соответственно, активности по сравнению с активностью полипептида перед выдержкой полипептида при 65°C в течение 15 мин. Предпочтительно, полипептид согласно данному изобретению не демонстрирует остаточную активность после того, как полипептид был выдержан при температуре 65°C в течение 15 мин.

В одном воплощении полипептид, описанный в данном документе, представляет собой полипептид, обладающий активностью пролин-специфичной эндопротеазы, где указанный полипептид имеет менее чем 90% остаточной активности при использовании ацетил-AlaAlaPro-паранитроанилина (Ac-AAP-PNA) в качестве субстрата после того, как полипептид был выдержан при температуре 60°C в течение 15 мин. Остаточная активность пролин-специфичной эндопротеазы измеряли при использовании ацетил-AlaAlaPro-паранитроанилина (Ac-AAP-PNA) при pH 4,5, например в натрий-ацетатном буфере при pH 4,5 при 20 градусах Цельсия. Удивительно, но полипептид, который содержит менее 90% остаточной активности после выдержки полипептида при температуре 60°C в течение 15 мин., предпочтительно может быть использован приложениях, например, пищевых или кормовых, где желательна отсутствующая или незначительная остаточная активность. Предпочтительно, если полипептид, предлагаемый в данном документе, имеет менее 85%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 15%, 10%, например, менее 5% остаточной активности после выдержки полипептида при температуре 60°C в течение 15 мин. Как определено в данном описании, менее, чем 90% 85%, 80%, 70%, или менее 60%, 50%, 45%, 40%, 30%, 20%, 15%, 10%, или 5%, остаточной активности означает, что полипептид обладает менее 70%, 60%, 50%, или менее чем 40%, 30%, 20%, 15%, 10% или 5% соответственно, активности по сравнению с активностью полипептида перед выдержкой полипептида при 60°C в течение 15 мин. В предпочтительном варианте полипептид согласно данному изобретению не проявляет остаточную активность после того, как полипептид был выдержан при температуре 60°C в течение 15 мин.

Изобретение также относится к полипептиду, обладающему активностью пролин-специфичной эндопротеазы, необязательно имеющему менее 70% остаточной активности при использовании ацетил-AlaAlaPro-паранитроанилина (Ac-AAP-PNA) в качестве субстрата после того, как полипептид был выдержан при температуре 65°C в течение 15

мин., или, необязательно, содержащий менее 90% остаточной активности при использовании ацетил-AlaAlaPro-паранитроанилина (Ac-AAP-PNA) в качестве субстрата после того, как полипептид был выдержан при температуре 60°C в течение 15 мин., где указанный полипептид выбранный из группы, включающей:

i. полипептид, который, при выравнивании относительно аминокислотной последовательности, соответствующей SEQ ID NO: 1, содержит аминокислотную замену в положении, соответствующем положению 469, и кроме того, необязательно, по меньшей мере, одну аминокислотную замену в положении 204, 304, 377, 466, и/или 477, где положение определяется относительно SEQ ID NO: 1;

ii. полипептид, который, при выравнивании относительно аминокислотной последовательности, соответствующей SEQ ID NO: 1, содержит аминокислоту, выбранную из группы, состоящей из Ala (A), Cys (C), Asp (D), Glu (E), Phe (F), Gly (G), His (H), Ile (I), Lys (K), Leu (L), Met (M), Asn (N), Gln (Q), Arg (R), Ser (S), Thr (T), Val (V), Trp (W) и Tyr (Y), в положении, соответствующем положению 469, и, необязательно, аминокислоты фенилаланин (F) в положении 204, Ser (S), в положении 304, аланин (A) в положении 377, Thr (T) в положении 466 и/или аланин (A) в положении 477, где положение определяется относительно SEQ ID NO: 1

iii. полипептид, содержащий аминокислотную последовательность, соответствующую SEQ ID NO: 1, где SEQ ID NO: 1 включает, по меньшей мере, одну аминокислотную замену, выбранную из группы, состоящей из P469A, P469C, P469D, P469E, P469F, P469G, P469H, P469I, P469K, P469L, P469M, P469N, P469Q, P469R, P469S, P469T, P469V, P469W и P469Y, и, возможно, аминокислотную замену I204F, P304S, P377A, P466T, и/или P477A, где аминокислотная замена определяется относительно SEQ ID NO: 1;

iv. полипептид по i)-iii), но утративший сигнальную последовательность и/или последовательность пробелка;

v. полипептид по i)-iv), где полипептид идентичен, по меньшей мере, на 60%, 70%, 75%, 80, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% аминокислотной последовательности по SEQ ID NO: 1;

vi. полипептид, кодируемый нуклеиновой кислотой, которая идентична, по меньшей мере, на 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, или 99% или 100% SEQ ID NO: 2, или зрелой кодирующей последовательности SEQ ID NO: 2, где последовательность SEQ ID NO: 2 содержит, по меньшей мере, одну мутацию, кодирующую, по меньшей мере, одну аминокислотную замену, выбранную из группы P469A, P469C, P469D, P469E, P469F, P469G, P469H, P469I, P469K, P469L, P469M, P469N, P469Q, P469R, P469S, P469T, P469V, P469W и P469Y, и необязательно аминокислотную

замену I204F, P304S, P377A, P466T, и/или P477A, где аминокислотные замены определяются относительно SEQ ID NO: 1.

При использовании в данном документе, когда полипептид выровнен с последовательностью пролин-специфичной эндопротеазы по SEQ ID NO: 1, полипептид по настоящему изобретению будет содержать, по меньшей мере, одну аминокислотную замену в положении, соответствующем 469 в SEQ ID NO: 1.

Эти положения в полипептиде по настоящему изобретению, который может быть рекомбинантным, синтетическим или вариантным полипептидом, соответствуют положениям, изложенным выше в SEQ ID NO: 1, и могут быть идентифицированы путем выравнивания последовательности полипептида согласно настоящему изобретению, с последовательностью по SEQ ID NO: 1 с использованием, например, программы выравнивания NEEDLE. Положения в полипептиде по настоящему изобретению, соответствующие положениям в SEQ ID NO: 1, как указано выше, могут быть, таким образом идентифицированы и упомянуты как те положения, которые определенные по последовательности SEQ ID NO: 1.

Настоящее изобретение также относится к полипептиду, который является выделенным, по существу чистым, чистым, рекомбинантным, синтетическим или вариантным полипептидом полипептида, описанного в данном документе.

Предпочтительно полипептид, предлагаемый изобретением, включает выбранную, по меньшей мере, одну аминокислоту из группы, состоящей из Ala (A), Cys (C), Asp (D), Glu (E), Phe (F), Gly (G), His (H), Ile (I), Lys (K), Leu (L), Met (M), Asn (N), Gln (Q), Arg (R), Ser (S), Thr (T), Val (V), Trp (W), и Tyr (Y) в положении, соответствующем положению 469, где положение определяется относительно SEQ ID NO: 1.

В одном воплощении полипептид, описанный в данном документе, дополнительно включает аминокислоту Phe (F) в положении 204, Ser (S) в положении 304, Ala (A) в положении 377, Thr (T) в положении 466 и/или Ala (A) в положении 477, где положение определяется ссылкой на последовательность SEQ ID NO: 1.

В одном воплощении в настоящем изобретении предлагается полипептид, обладающий активностью пролин-специфичной эндопротеазы, включающий аминокислотную последовательность по SEQ ID NO: 1, где последовательность SEQ ID NO: 1 содержит, по меньшей мере, одну аминокислотную замену, выбранную из группы, состоящей из P469A, P469C, P469D, P469E, P469F, P469G, P469H, P469I, P469K, P469L, P469M, P469N, P469Q, P469R, P469S, P469T, P469V, P469W и P469Y, и где аминокислотные замены определяются относительно SEQ ID NO: 1. В дополнительном воплощении полипептид, обладающий активностью пролин-специфичной эндопротеазы,

дополнительно включает аминокислотную замену I204F, P304S, P377A, P466T, и/или P477A. Предпочтительно, если настоящее изобретение относится к полипептиду, содержащему, по меньшей мере, одну аминокислотную замену(ы), соответствующую положению 469, и, возможно, положению 204, 304, 377, 466 и/или 477, как определено в данном документе выше, который идентичен, по меньшей мере, на 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% аминокислотной последовательности по SEQ ID NO: 1. Соответственно, настоящее изобретение относится к полипептиду, который идентичен, по меньшей мере, на 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% аминокислотной последовательности по SEQ ID NO: 1 или зрелому полипептиду по SEQ ID NO: 1, где SEQ ID NO: 1, содержит аминокислотную замену в положении 469, и, возможно, в положении 204, 304, 377, 466, и/или 477, как определено в данном документе выше.

В одном воплощении полипептид, обладающий активностью пролин-специфичной эндопротеазы, включает аминокислотные замены, выбранные из группы, состоящей из P469D и I204F; P469D и P377A; P469Q и P477A; P469Y и P304S и P377A; P469Q и I204F и P466T; и P469Q и P466T и P477A.

Полипептид, обладающий активностью пролин-специфичной эндопротеазы, которая, при выравнивании с аминокислотной последовательностью, в соответствии с SEQ ID NO: 1 содержит аминокислотную замену, соответствующем положению 469, как определено выше в настоящем изобретении, может включать дополнительные замены, делеции и/или вставки в одной или нескольких дополнительных аминокислотных положениях. Например, полипептид, описанный в данном документе, может представлять собой вариант полипептида или зрелого полипептида по SEQ ID NO: 1, включающий замену, делецию или вставку в положении 469, как определено в настоящем описании, и дополнительно имеет 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 или 12, или более дополнительных аминокислотных замен, делеций и/или вставок, в результате чего полипептид все еще обладает активностью или функцией полипептида согласно изобретению. Специалисту будет понятно, что эти незначительные изменения аминокислот в полипептиде по изобретению, могут присутствовать (например, природные мутации) или могут быть внесены (например, с использованием технологии р-ДНК) без потери функции или активности белка. В случае, если эти мутации присутствуют в связывающем домене, активном сайте или другом функциональном домене полипептида, свойство полипептида может измениться, но полипептид может сохранять свою активность. В том случае, если присутствует мутация, которая не находится близко к активному сайту, связывающему домену или другому функционально активному домену, можно ожидать меньший эффект.

Функциональные эквиваленты полипептида согласно изобретению, также могут быть идентифицированы, например, путем скрининга комбинаторных библиотек мутантов, например укороченных мутантов полипептида согласно изобретению на предмет биологической активности полипептида по изобретению. В одном воплощении смешанная библиотека образована комбинаторным мутагенезом на уровне нуклеиновой кислоты. Смешанная библиотека вариантов может быть получена, например, ферментативным лигированием смеси синтетических олигонуклеотидов в последовательности генов таким образом, чтобы вырожденный набор потенциальных последовательностей белка был представлен в виде отдельных полипептидов или, в ином случае, в виде набора более крупных слитых белков (например, для фагового дисплея). Существуют различные способы, которые могут быть использованы для получения библиотек потенциальных вариантов полипептидов по изобретению из вырожденной олигонуклеотидной последовательности. Способы синтеза вырожденных олигонуклеотидов известны в данной области (см., например, Narang (1983) *Tetrahedron* 39:3; Itakura et al. (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al. (1984) *Science* 198:1056; Ike et al. (1983) *Nucleic Acid Res.* 11:477). Термин «вырожденная нуклеотидная последовательность» или «вырожденная (олиго) нуклеотидная последовательность» обозначает последовательность нуклеиновых кислот, которая включает один или несколько вырожденных кодонов (по сравнению с молекулой эталонной молекулой нуклеиновой кислоты, которая кодирует полипептид). Вырожденные кодоны содержат различные триплеты нуклеиновых кислот, но кодируют один и тот же аминокислотный остаток (т.е., каждый из триплетов GAU и GAC кодирует Asp). Вырожденность кодонов относится к природе генетического кода, благодаря которой возможны вариации нуклеотидной последовательности, без оказания влияния на аминокислотную последовательность закодированного полипептида. Специалисту хорошо известно о «случайности использования синонимичных кодонов», проявляемой конкретной клеткой-хозяином при использовании кодонов нуклеиновых кислот для определения конкретной аминокислоты.

Кроме того, библиотеки фрагментов кодирующей последовательности полипептида по изобретению могут быть использованы для создания смешанной популяции полипептидов для скрининга и последующего отбора вариантов. Например, библиотека фрагментов кодирующей последовательности может быть получена путем обработки представляющего интерес двухцепочечного ПЦР-фрагмента кодирующей последовательности нуклеазой в условиях, в которых никирование происходит только один раз на молекулу, денатурации двухцепочечной ДНК, ренатурации ДНК с

образованием двухцепочечной ДНК, которая может включать смысловые/антисмысловые пары из различных никированных продуктов, удаление одноцепочечных частей из реформированных дуплексов путем обработки нуклеазой S1 и лигирования полученной библиотеки фрагментов в экспрессирующий вектор. С помощью этого способа можно получить экспрессионную библиотеку, которая кодирует N-концевые и внутренние фрагменты различных размеров из представляющего интерес белка.

В данной области известны несколько методов скрининга генных продуктов комбинаторных библиотек, полученных с помощью точечных мутаций укорочения, и для скрининга кДНК-библиотек на предмет генных продуктов, имеющих выбранное свойство. Наиболее широко используемые методы для скрининга больших библиотек генов, которые поддаются высокопроизводительному анализу, обычно включают клонирование библиотеки генов в реплицируемые экспрессирующие векторы, трансформацию подходящих клеток полученной библиотекой векторов и экспрессию комбинаторных генов в условиях, в которых детекция искомой активности облегчает выделение вектора, кодирующего ген, продукт которого был обнаружен. Рекурсивный множественный мутагенез (REM), метод, который повышает частоту функциональных мутантов в библиотеках, может быть использован в комбинации с анализами скринингом для идентификации вариантов белка по изобретению (Arkin and Yourvan (1992) Proc. Natl. Acad. Sci. USA 89:7811-7815; Delgrave et al. (1993) Protein Engineering 6(3): 327-331).

Полипептид, предлагаемый изобретением, может утратить сигнальную последовательность и/или последовательность пропротеина. Например, полипептид, предлагаемый в настоящем документе, может представлять собой полипептид, содержащий аминокислотную последовательность, в соответствии с SEQ ID NO: 1, включающий аминокислотную замену в положении, соответствующем положению 469 и лишенный первых 17 аминокислот, кодирующих сигнальную последовательность и/или лишенный следующих 19 аминокислот, кодирующих пропоследовательность. Соответственно полипептид, предлагаемый изобретением, может содержать зрелый полипептид по SEQ ID NO: 1, например, аминокислоты с 37 по 521 SEQ ID NO: 1 и содержать аминокислотную замену в положении, соответствующем положению 469, и, необязательно, в положениях 204, 304, 377, 466, и/или 477, как определено в настоящем описании, в котором аминокислота метионин в положении 1 в SEQ ID NO: 1, считается номером 1.

Полипептид, предлагаемый изобретением, может быть закодирован с помощью любой подходящей нуклеиновой кислоты, такой как нуклеиновая кислота, которая идентична, по меньшей мере, на 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%,

95%, 96%, 97%, 98%, 99% или 100% нуклеиновой кислоте по SEQ ID NO: 2, или зрелой кодирующей последовательности SEQ ID NO: 2, где последовательность SEQ ID NO: 2 содержит, по меньшей мере, одну мутацию, кодирующую, по меньшей мере, одну аминокислотную замену, выбранную из группы P469A, P469C, P469D, P469E, P469F, P469G, P469H, P469I, P469K, P469L, P469M, P469N, P469Q, P469R, P469S, P469T, P469V, P469W и P469Y, аминокислотную замену I204F, P304S, P377A, P466T, и/или P477A. где аминокислотные замены определяются относительно SEQ ID NO: 1.

В одном воплощении полипептид, описанный в данном документе, может быть, кодируемый нуклеиновой кислотой, которая идентична, по меньшей мере, на 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% нуклеиновой кислоте по SEQ ID NO: 2, или зрелой кодирующей последовательности SEQ ID NO: 2, где последовательность SEQ ID NO: 2 содержащая мутации, кодирующие полипептид, обладающий активностью пролин-специфичной эндопротеазы, включает аминокислотные замены, выбранные из группы, состоящей из P469D и I204F; P469D и P377A; P469Q и P477A; P469Y и P304S и P377A; P469Q и I204F и P466T; и P469Q и P466T и P477A.

Как правило, полинуклеотидная последовательность, как описано в данном документе, является оптимизированной по кодонам, или оптимизированной по кодонам последовательностью для оптимальной экспрессии полипептида, описанного в данном документе, в конкретной клетке-хозяине.

В одном воплощении настоящее изобретение представляет собой биологически активный фрагмент полипептида, описанный в данном документе.

Биологически активные фрагменты полипептида согласно изобретению, включают полипептиды, содержащие аминокислотные последовательности, в достаточной степени идентичные или полученные из аминокислотной последовательности белка пролин-специфичной эндопротеазы (например, зрелой аминокислотной последовательности SEQ ID NO: 1), которые включают меньше аминокислот, чем полноразмерный белок, но который проявляет, по меньшей мере, одну биологическую активность соответствующего полноразмерного белка. Как правило, биологически активные фрагменты содержат домен или мотив, по меньшей мере, с одной активностью белка пролин-специфичной эндопротеазы. Биологически активный фрагмент может, например, содержать каталитический домен. Биологически активный фрагмент белка по изобретению может представлять собой полипептид, который имеет, например, 10, 25, 50, 100 или более аминокислот в длину. Кроме того, другие биологически активные части, в которых удалены другие области белка, могут быть получены с помощью рекомбинантных

методов и оценены по одной или нескольких биологических активностей нативной формы полипептида согласно изобретению.

Изобретение также относится к фрагментам нуклеиновых кислот, которые кодируют указанные выше биологически активные фрагменты белка пролин-специфичной эндопротеазы.

Полипептид по настоящему изобретению может представлять собой слитый белок. Методы получения слитых полипептидов известны в данной области, и включают лигирование последовательностей, кодирующих полипептиды так, чтобы они находились в одной рамке. Экспрессия слитого полипептида находится под контролем одинакового промотора(ов) и терминатора. Гибридные полипептиды могут содержать комбинацию частичных или полных полипептидных последовательностей, полученных, по меньшей мере, из двух различных полипептидов, где один или несколько, может быть гетерологичными по отношению к клетке-хозяину. Такие слитые полипептиды, по меньшей мере, из двух различных полипептидов могут содержать связывающий домен из одного полипептида, функционально связанный с каталитическим доменом из второго полипептида. Примеры слитых полипептидов и химер слитых последовательностей, например, описаны в WO2010/121933, WO2013/007820 и WO2013/007821.

Полипептид по настоящему изобретению, может быть получен из любой подходящей эукариотической клетки. Эукариотическая клетка может быть клеткой млекопитающего, насекомого, растения, гриба или водоросли. Формулировка «полученный» или «производное» относительно к происхождению полипептида, описанного в настоящем документе, означает, что при проведении поиска BLAST с полипептидом в соответствии с настоящим изобретением, полипептид согласно данному изобретению, может быть получен из природного источника, такого как микробная клетка, с которой эндогенный полипептид демонстрирует самый высокий процент гомологии или идентичности с полипептидом, описанным в данном документе.

Полипептид по настоящему изобретению, может быть получен из мицелиальных грибных клеток или термофильного мицелиальных грибных клеток. Предпочтительные мицелиальные грибные клетки принадлежат к видам родов *Acremonium*, *Aspergillus*, *Chrysosporium*, *Myceliophthora*, *Penicillium*, *Talaromyces*, *Rasamsonia*, *Thielavia*, *Fusarium* или *Trichoderma*, *Amorphotheca*, *Pseudocercospora*, *Trametes*, *Rhizomucor*, *Calcarisporiella*, *Thermomyces*, *Thermoascus*, *Cornyascus*, *Myricoccum*, *Scytalidium*, *Chaetomium*, *Paecilomyces*, *Corynascus*, *Malbranchea*, *Stilbella*, *Thermomyces*, *Dactylomyces*, *Humicola*, *Chaetomium*, *Melanocarpus*, *Rhizomucor*, *Lentinula*, *Anaeromyces*, а наиболее предпочтительно принадлежат к видам *Aspergillus niger*, *Acremonium alabamense*,

Aspergillus awamori, *Aspergillus foetidus*, *Aspergillus sojae*, *Aspergillus fumigatus*, *Talaromyces emersonii*, *Rasamsonia emersonii*, *Aspergillus oryzae*, *Chrysosporium lucknowense*, *Fusarium oxysporum*, *Myceliophthora thermophila*, *Trichoderma reesei*, *Thielavia terrestris*, *Penicillium chrysogenum*, *Amorphotheca resinae*, *Aureobasidium pullulans*, *Pseudocercospora herpotrichoides*, *Trametes versicolor* 52J, *Rhizomucor pusillus*, *Calcarisporiella thermophila*, *Talaromyces thermophilus*, *Thermomyces lanuginosus*, *Thermoascus auratiacus*, *Corynascus thermophilus*, *Myricoccum thermophilum*, *Scytalidium thermophilum*, *Myceliophthora himmulea*, *Chaetomium thermophilum*, *Paecilomyces byssochlamydoides*, *Corynascus sepedonium*, *Malbranchea cinnamomea*, *Thielavia australiensis*, *Stilbella thermophila*, *Thermomyces stellatus*, *Talaromyces emersonii*, *Dactylomyces thermophilus*, *Humicola hyalothermophila*, *Acremonium thermophilum*, *Chaetomium olivicolor*, *Melanocarpus albomyces*, *Rhizomucor miehei*, *Lentinula edodes* или *Anaeromyces mucronatus*. Полипептид по настоящему изобретению может быть получен из *Aspergillus niger*, *Aspergillus Aculeatus*, *Aspergillus Flavus*, *Aspergillus carbonarius* или *Rasamsonia emersonii*.

Полипептид по настоящему изобретению, может быть природным или генетически модифицированным или рекомбинантным полипептидом.

Полипептид, описанный в данном документе, может быть очищен. Очистка белка известна специалисту в данной области. Хорошо известным способом очистки белков является высокоэффективная жидкостная хроматография.

Полипептиды по настоящему изобретению преимущественно обладают улучшенным свойством. Улучшенное свойство может быть повышенной специфической активностью и/или повышенной чувствительностью к температуре по сравнению с полипептидом, не включающим аминокислотную замену, определенную в данном описании, или любое другое улучшенное свойство, например желательное при обработке пищи или кормов. Предпочтительно полипептид, описанный в данном документе, имеет менее 70% остаточной активности на ацетил-AlaAlaPro-паранитроанилине (Ac-AAP-PNA) в качестве субстрата, когда полипептид был выдержан при температуре 65 °C в течение 15 мин.

Полипептиды согласно изобретению могут быть получены несколькими способами, известными специалисту в данной области, такими как:

1. ПЦР пониженной точности для введения случайных мутаций, с последующим скринингом полученных (вариантных) полипептидов и выделением (вариантного) полипептида(ов) с улучшенными свойствами
2. смешение семейств родственных вариантов генов, кодирующих полипептид

согласно изобретению, с последующим скринингом полученных вариантов и выделения вариантов с улучшенными свойствами

Варианты генов, кодирующих полипептид согласно настоящему изобретению, приводят к увеличению уровня мРНК и/или белка, что приводит к большей активности, которая может быть получена путем модификации полинуклеотидных последовательностей указанных генов. К таким модификациям относятся:

1. Улучшение использования кодонов таким образом, чтобы кодоны были (оптимально) адаптированы к родительскому микробному хозяину.
2. Улучшение использования кодоновых пар таким образом, что кодоны были (оптимально) адаптированы к родительскому микробному хозяину.
3. Добавление стабилизирующих последовательностей к геномной информации, кодирующей полипептид, в соответствии с изобретением, что дает молекулу мРНК с увеличенным периодом полужизни.

Способы выделения вариантов с улучшенными каталитическими свойствами или повышенными уровнями мРНК или белка описаны в WO03/010183 и WO03/01311. Способы оптимизации использования кодонов в родительских микробных штаммах описаны, например, в WO2008/000632. Способы добавления стабилизирующих элементов в гены, кодирующих полипептид согласно изобретению, описаны в WO2005/059149.

Соответственно, в одном аспекте настоящее изобретение относится к способу получения варианта полипептида, причем этот способ включает

i. выбор исходного полипептида, идентичного, по меньшей мере, на 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% аминокислотной последовательности по SEQ ID NO: 1, или зрелого полипептида последовательности по SEQ ID NO: 1; и

ii. замены, по меньшей мере, одной аминокислоты в положении, соответствующем положению 469, определенном относительно SEQ ID NO: 1, на аминокислоту, выбранную из группы, состоящей из Ala (A), Cys (C), Asp (D), Glu (E), Phe (F), Gly (G), His (H), Ile (I), Lys (K), Leu (L), Met (M), Asn (N), Gln (Q), Arg (R), Ser (S), Thr (T), Val (V), Trp (W) и Tyr (Y), и, необязательно, замены аминокислоты в положении 204 на Phe (F), в положении 304 на Ser (S), в положении 377 на Ala (A), в положении 466 на Thr (T), и/или в положении 477 на Ala (A); и

iii. получение вариантного полипептида,

в котором, возможно, вариантный полипептид имеет менее чем 70% остаточной активности при использовании Ac-AAP-PNA в качестве субстрата, после того, как полипептид был выдержан при температуре 65°C в течение 15 мин.

Получение вариантного полипептида, описанного в данном документе, может включать экспрессию гена, кодирующего вариантный полипептид в подходящей (рекомбинантной) клетке-хозяине, и культивирование клетки-хозяина для получения вариантного полипептида.

В другом аспекте настоящее изобретение относится к композиции, включающей полипептид, описанный в данном документе.

Композиция, описанная в данном документе, может включать носитель, наполнитель, вспомогательный фермент или другие соединения. Как правило, композиция или состав, включает соединение, с помощью которого может быть получена пролин-специфичная эндопротеаза.

Наполнитель, используемый в данном документе, является неактивным веществом, включенным в состав вместе с полипептидом, описанным в данном документе, например, представляет собой сахарозу или лактозу, глицерин, сорбит или хлорид натрия. Композиция, содержащая полипептид, описанный в данном документе, может быть жидкой композицией или твердой композицией. Жидкая композиция, как правило, включает воду. При изготовлении лекарственной формы в виде жидкой композиции, композиция, как правило, включает компоненты, которые понижают активность воды, такие как глицерин, сорбит или хлорид натрия (NaCl). Твердая композиция, содержащая полипептид, описанный в данном документе, может содержать гранулят, содержащий фермент или композиция содержит инкапсулированный полипептид в жидких матрицах, таких как липосомы или гели, такие как альгинат или каррагинаны. Есть много способов, известных в данной области, инкапсуляции или гранулирования полипептида или фермента (см., например, G.M.H. Meesters, «Encapsulation of Enzymes and Peptides», Chapter 9, in N.J. Zuidam and V.A. Nedović (eds.) «Encapsulation Technologies for Active Food Ingredients and food processing» 2010).

Композиция, описанная в данном документе, может также содержать носитель, содержащий полипептид, описанный в данном документе. Полипептид, раскрытый в настоящем описании, может быть связан или иммобилизован на носителе с помощью известных технологий в данной области.

Настоящее изобретение также относится к способу получения композиции, содержащей полипептид, раскрытый в настоящем описании, который может включать распылительную сушку ферментационной среды, содержащей полипептид или гранулирование, или инкапсулирование полипептида, описанного в данном документе, и получение композиции.

Настоящее изобретение также относится к упаковке, такой, как банка, бочонок или

бочка, содержащей полипептид или композицию, содержащую полипептид, описанный в данном документе.

В другом своем аспекте настоящее изобретение относится к нуклеиновой кислоте, кодирующей пролин-специфичную эндопротеазу, которая идентична, по меньшей мере, на 80%, 85%, 90%, 95%, 96%, 97%, 98% или 99% последовательности SEQ ID NO: 2, или зрелой кодирующей последовательности SEQ ID NO: 2, где SEQ ID NO: 2 содержит, по меньшей мере, одну мутацию, кодирующую, по меньшей мере, одну аминокислотную замену, выбранную из группы, состоящей из P469A, P469C, P469D, P469E, P469F, P469G, P469H, P469I, P469K, P469L, P469M, P469N, P469Q, P469R, P469S, P469T, P469V, P469W, P469Y, и необязательно, где SEQ ID NO: 2 включает, по меньшей мере, одну мутацию, кодирующую аминокислотную замену I204F, P304S, P377A, P466T, и/или P477A, где аминокислотные замены определяются относительно SEQ ID NO: 1.

В другом варианте выделенная молекула нуклеиновой кислоты по изобретению содержит молекулу нуклеиновой кислоты, которая является обратным комплементом нуклеотидной последовательности, показанной на SEQ ID NO: 2, или обратной комплементарной зрелой кодирующей последовательности SEQ ID NO: 2, содержащий, по меньшей мере, одну мутацию кодирование, по меньшей мере, одну аминокислотную замену, выбранную из группы, состоящей из P469A, P469C, P469D, P469E, P469F, P469G, P469H, P469I, P469K, P469L, P469M, P469N, P469Q, P469R, P469S, P469T, P469V, P469W, и P469Y, и, необязательно, в котором последовательность SEQ ID NO: 2 содержит, по меньшей мере, одну мутацию, кодирующую аминокислотную замену I204F, P304S, P377A, P466T, и/или P477A, где аминокислотные замены определяются относительно SEQ ID NO: 1, или варианта любой такой нуклеотидной последовательности.

Также раскрывается нуклеиновая кислота, которая гибридизуется в условиях средней жесткости, предпочтительно в условиях высокой жесткости с комплементарной цепью зрелого полипептида SEQ ID NO: 2, содержащая, по меньшей мере, одну мутацию, кодирующую, по меньшей мере, одну аминокислотную замену, выбранную из группы, P469A, P469C, P469D, P469E, P469F, P469G, P469H, P469I, P469K, P469L, P469M, P469N, P469Q, P469R, P469S, P469T, P469V, P469W, и P469Y, и, необязательно, где последовательность SEQ ID NO: 2 содержит, по меньшей мере, еще одну мутацию, кодирующую аминокислотную замену I204F, P304S, P377A, P466T, и/или P477A, где указанные замены определяясь с ссылкой на SEQ ID NO: 1.

Молекула нуклеиновой кислоты, которая комплементарна другой нуклеотидной последовательности, представляет собой последовательность, которая в достаточной степени комплементарна другой нуклеотидной последовательности таким образом, что

может гибридизоваться с другой нуклеотидной последовательностью с образованием устойчивого дуплекса. Термин «кДНК» (комплементарной ДНК) определяется в данном документе как молекула ДНК, которая может быть получена путем обратной транскрипции из молекулы мРНК. В прокариотах молекулы мРНК образуются при транскрипции геномной ДНК гена, присутствующего в клетке. В эукариотических клетках гены содержат как экзоны, т.е. кодирующие последовательности, так и интроны, т.е. промежуточные последовательности, расположенные между экзонами. Поэтому в эукариотических клетках исходная, первичная РНК, полученная транскрипцией геномной ДНК гена подвергается обработке в ряде стадий, прежде чем появляется в виде мРНК. Эти стадии включают удаление интронных последовательностей с помощью процесса, называемого сплайсинг. кДНК, полученной из мРНК содержат только кодирующие последовательности и могут быть непосредственно переведены в соответствующий полипептидный продукт.

В другом аспекте настоящее изобретение относится к вектору экспрессии, содержащему полинуклеотид, описанный в данном документе, функционально связанный, по меньшей мере, с одной контрольной последовательностью, которая направляет экспрессию полипептида в экспрессирующей клетке-хозяине.

Есть несколько способов вставки нуклеиновой кислоты в нуклеотидную конструкцию или экспрессирующий вектор, известных специалистам в данной области, см., например, Sambrook & Russell, *Molecular Cloning: A Laboratory Manual*, 3rd Ed., CSHL Press, Cold Spring Harbor, NY, 2001. Желательным может быть проведение манипуляции с нуклеиновой кислотой, кодирующей полипептид по настоящему изобретению с регуляторными последовательностями, такими как последовательности промотора и терминатора.

Промотор может быть любой приемлемой последовательностью промотора подходящей для эукариотической или прокариотической клетки-хозяина, которая проявляет транскрипционную активность, в том числе мутантные, усеченные и гибридные промоторы, и может быть получена из полинуклеотидов, кодирующих внеклеточные или внутриклеточные полипептиды в клетке, либо эндогенно (нативные) либо гетерологично (инородные). Промотор может быть конститутивным или индуцируемым промотором. В предпочтительном воплощении промотор представляет собой индуцируемый промотор, например крахмал-индуцируемый промотор. Промоторами, пригодными в мицелиальных грибах, являются промоторы, которые могут быть выбраны из группы, которая включает, без ограничения указанным, промоторы, полученные из полинуклеотидов, кодирующих ТАКА-амилазу *A. oryzae*, аспарагиновую протеазу *Rhizomucor miehei*, промотор *gpdA*

Aspergillus, нейтральную альфа-амилазу *A. niger*, кислотостабильную альфа-амилазу *A. niger*, глюкоамилазу (*glaA*) *A. niger* или *A. awamori*, эндоксиланазу (*xlnA*) или бета-ксилозидазу (*xlnD*) *A. niger* или *A. awamori*, целлобиогидролазу I (СВН) *T. reesei*, липазу *R. miehei*, щелочную протеазу *A. oryzae*, триозфосфатизомеразу *A. oryzae*, ацетомидазу *A. nidulans*, амилогликозидазу (WO 00/56900) *Fusarium venenatum*, *Fusarium venenatum* Dania (WO 00/56900), *Fusarium venenatum* Quinn (WO 00/56900), трипсинподобную протеазу *Fusarium oxysporum* (WO 96/00787), бета-глюкозидазу *Trichoderma reesei*, целлобиогидролазу I *Trichoderma reesei*, целлобиогидролазу II *Trichoderma reesei*, *Trichoderma reesei*, эндоглюконазу I *Trichoderma reesei*, эндоглюконазу II *Trichoderma reesei*, эндоглюконазу III *Trichoderma reesei*, эндоглюконазу IV *Trichoderma reesei*, эндоглюконазу V *Trichoderma reesei*, ксиланазу I *Trichoderma reesei*, ксиланазу II *Trichoderma reesei*, бета-ксилозидазу *Trichoderma reesei*, а также промотор NA2-tpi (гибрид промоторов из полинуклеотидов, кодирующих нейтральную альфа-амилазу *A. niger* и триозофосфат изомеразу *A. oryzae*), и их мутантные, укороченные и гибридные промоторы.

Может быть использован любой терминатор, который является функциональным в клетке, описанной в данном документе, который известен специалисту в данной области. Примеры подходящих терминаторных последовательностей различных видов мицелиальных грибов включают терминаторные последовательности из гена мицелиального гриба, например, из генов *Aspergillus*, например из гена ТАКА-амилазы *A. oryzae*, из гена, кодирующего глюкоамилазу (*glaA*) *A. niger*, антранилатсинтазу *A. nidulans*, альфа-глюкозидазу *A. niger*, *trpC* и/или трипсиноподобную протеазу из *Fusarium oxysporum*.

В другом аспекте настоящее изобретение относится к клетке-хозяину, содержащей конструкцию нуклеиновой кислоты или вектор экспрессии, описанный в данном документе. Подходящая клетка-хозяин может быть клеткой млекопитающего, насекомого, растения, гриба водоросли или бактерии. Подходящая клетка-хозяин может быть грибной клеткой, например, из рода *Acremonium*, *Aspergillus*, *Chrysosporium*, *Fusarium*, *Myceliophthora*, *Penicillium*, *Rasamsonia*, *Talaromyces*, *Thielavia*, *Trichoderma*, *Saccaromyces*, *Kluyveromyces*, *Pichia*, например, *Aspergillus niger*, *Aspergillus Awamori*, *Aspergillus foetidus*, *A. oryzae*, *A. sojae*, *Talaromyces emersonii*, *Rasamsonia emersonii*, *Chrysosporium lucknowense*, *Fusarium oxysporum*, *Myceliophthora thermophila*, *Thielavia terrestris* или *Trichoderma reesei* или *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, *Pichia pastoris*.

Клетка-хозяин может быть рекомбинантной или трансгенной клеткой-хозяином.

Клетка-хозяин может быть генетически модифицирована с помощью нуклеотидной конструкции или экспрессирующего вектора, описанных в данном документе с помощью стандартных способов, известных в данной области, таких как электропорация, трансформация протопластов или конъюгация, например, как описано в Sambrook & Russell, *Molecular Cloning: A Laboratory Manual*, 3rd Ed., CSHL Press, Cold Spring Harbor, NY, 2001. Рекомбинантный хозяин может избыточно экспрессировать полипептид в соответствии с настоящим изобретением с помощью известных методов в данной области техники.

В одном из аспектов настоящее изобретение относится к способу получения полипептида, как описано в данном документе, включающем культивирование рекомбинантной клетки-хозяина в подходящей среде для ферментации в условиях, способствующих выработке полипептида и продуцирования полипептида. Специалистам в данной области известно, как осуществить способ получения полипептида, описанного в данном документе, в зависимости от используемой клетки-хозяина, pH, температуры и состава ферментационной среды. Клетки-хозяева могут культивироваться в колбах или в биореакторах, имеющих объем 0,5 или 1 литр или больше, от 10 до 100 или более кубических метров. Культивирование может быть осуществлено в аэробных или анаэробных условиях в зависимости от потребностей клетки-хозяина.

Предпочтительно, если полипептид, описанный в данном документе, извлекают или выделяют из ферментационной среды. Извлечение или выделение полипептида из ферментационной среды, например, может быть выполнено с помощью центрифугирования, фильтрации, и/или ультрафильтрации.

Полипептид, обладающий активностью пролин-специфичной эндопротеазы или композиция, содержащая полипептид, описанный в данном документе, могут быть использованы в самых различных применениях, например, в производстве продуктов питания или кормов, например, в производстве белкового гидролизата. Некоторые пищевые белки содержат высокоаллергенные субфракции, которые могут быть даже токсичными для конкретных лиц, такие как глютен, который содержит проламины с богатыми пролином пептидными последовательностями. Эти белки могут быть подвергнуты воздействию нового фермента для того, чтобы уменьшить их антигенность или токсичность.

Группа людей, для которых глютен является токсичным, представляет собой индивидуумов, страдающих от целиакии-спру. Целиакия-спру, также известная как целиакия, является аутоиммунным заболеванием тонкого кишечника, которое вызывается приемом в пищу глютенных белков из хлебных злаков, таких как альфа-глиадин из

пшеницы, гордеин из ячменя, секалин из ржи и авенин из овса.

Соответственно, полипептид, обладающий активностью пролин-специфичной эндопротеазы или композиция, содержащая полипептид, описанный в данном документе, могут быть использованы при приготовлении диетической добавки или в качестве лекарственного средства для лечения пациента, страдающего от целиакии-спру, или при лечении непереносящих глютен людей.

Полипептид, раскрытый в данном описании, также может быть использован в качестве технологической добавки для гидролиза глютена в пищевом продукте.

Соответственно, настоящее изобретение относится к способу приготовления пищевого или кормового продукта, который включает инкубирование промежуточной формы пищевого или кормового продукта с полипептидом или композицией, содержащей полипептид, описанный в данном документе, и приготовление пищевого или кормового продукта. Пищевой продукт в способе, раскрытом в описании, включает напиток, такой как пиво, вино или фруктовый сок, или хлебобулочное изделие, или молочный продукт, без ограничения перечисленным.

Пищевой продукт и/или промежуточная форма пищевого продукта может содержать глютен.

Было установлено, что полипептид, обладающий активностью пролин-специфичной эндопротеазы, описанный в данном документе, был способен к гидролизу токсичного эпитопа глютена на нетоксичные фрагменты.

Промежуточная форма пищевого продукта может быть любой подходящей формой пищевого продукта во время приготовления пищевого продукта. Например, промежуточная форма пива, может быть сушлом, а промежуточная форма хлеба может быть тестом или жидким тестом.

Способ приготовления пищевого продукта, в соответствии с настоящим изобретением может включать стадию пастеризации пищевого продукта. Пастеризация обычно включает нагревание пищевого продукта, или промежуточной формы пищевого продукта, например, путем доведения пищевого продукта или промежуточной формы пищевого продукта до температуры от 60 до 68°C в диапазоне от 10 до 20 мин., или от 12 до 18 мин., или до температуры между 70-74°C, например примерно 72°C в течение, по меньшей мере, 5, 10 или 15 секунд.

Пищевой продукт в способе, описанном в настоящем документе, также может быть белковым гидролизатом. Соответственно, настоящее изобретение относится к способу получения белкового гидролизата, включающий приведение в контакт с белковым субстратом полипептида или композиции, описанных в данном документе, и получения

белкового гидролизата. Белковый гидролизат может быть получен из любого подходящего белкового субстрата, например белкового субстрата, который богат остатками пролина, например, глютена зерновых или казеина коровьего молока.

В одном из аспектов настоящее изобретение относится к пищевому продукту, получаемому с помощью способа приготовления пищевого продукта, раскрытого в данном описании.

Следующие примеры иллюстрируют изобретение.

Примеры

Материалы и методы

Стандартные процедуры с ДНК проводили, как описано в Sambrook & Russell, 2001, *Molecular cloning: a laboratory manual*, 3rd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, если не указано иное. Последовательности ДНК были заказаны в DNA 2.0.

Пример 1. Клонирование, экспрессия и выделение (мутантной) пролин-специфичной эндопротеазы (PEP)

Пример 1.1. Клонирование и экспрессия

Последовательность белка пролин-специфичной эндопротеазы (PEP) из *A. niger* представлена в SEQ ID NO: 1, где первые 17 аминокислот представляют собой сигнальную последовательность пектинметилэстеразы *A. niger* (PMEA cc; SEQ ID NO: 2), а следующая часть состоит из 19 аминокислот пропоследовательности пролинэндопротеазы *A. niger* (SEQ ID NO: 4).

Разработали кодон-адаптированную последовательность ДНК для экспрессии этого белка в *Aspergillus niger*, содержащую дополнительные сайты рестрикции для субклонирования в экспрессирующем векторе для *Aspergillus*. Кодоновую адаптацию проводили, как описано в публикации WO 2008/000632. Кодон-оптимизированная последовательность ДНК для гена *A. niger*, кодирующая белок PEP с SEQ ID: NO: 1, представлена в SEQ ID NO: 2.

Подобным же образом мутантные пролин-специфические эндопротеазы из SEQ ID NO: 1, которые перечислены в таблице 1, таблице 2 и в таблице 3, были оптимизированы по кодонам для экспрессии в *Aspergillus niger*.

Аналогичным образом пролин-специфичной эндопротеазы из *A. Flavus*, *A. aculeatus* и *Rasamsonia emersonii*, показанные в SEQ ID NO: 10-12 и содержащие замену P469L в гомологическом положении относительно SEQ ID NO: 1, были оптимизированы по кодонам для экспрессии в *A. niger*, что дает нуклеотидные последовательности SEQ ID NO: 14-16, соответственно.

Последовательность инициации трансляции промотора глюкоамилазы *glaA* модифицировали в 5'-CACCGTCAAA ATG-3' и оптимальную последовательность терминации трансляции 5'-TAAA-3' использовали при получении экспрессирующей конструкции (как также подробно описано в WO2006/077258). Фрагмент ДНК, содержащий а.о. часть промотора глюкоамилазы и гена, кодирующего PEP, синтезировали полностью, очищали и гидролизовали с помощью *EcoRI* и *PacI*. Вектор pGBTOP-16 (Фиг. 1) линеаризовали гидролизом *EcoRI/PacI* и линеаризованный векторный фрагмент затем очищали с помощью гель-экстракции. ДНК-фрагмент, содержащий кодирующую область PEP клонировали в вектор pGBTOP-16, в результате чего получали pGBTOP-PEP. Затем *A. niger* GBA 306 ($\Delta glaA$, $\Delta pepA$, $\Delta hdfA$, адаптированный *Bam* HI ампликон, $\Delta amyBII$, $\Delta amyBI$, $\Delta amyA$ альфа-амилаза и отрицательная цепь глюкоамилазы) трансформировали вектором pGBTOP-PEP, линеаризованным гидролизом *Not I*, в контрансформационном протоколе с линеаризованным pGBAAS-4, с помощью штамма и способов, описанных в WO 2011/009700 и ссылок в этом документе, отбирали на ацетамид-содержащих средах и очищали колонии в соответствии со стандартными процедурами. Трансформацию и отбор осуществляли как описано в WO 98/46772 и WO 99/32617. Штаммы, содержащие ген PEP, отбирали с помощью праймеров ПЦР, специфичных для гена PEP для того, чтобы проверить наличие экспрессирующей кассеты pGBTOP-PEP. Трансформанты отбирали и далее пересевали репликой для получения инокулята одиночного штамма.

Пример 1.2. Получение (мутантного) PEP в штамме *A. Niger* PEP

Для каждой (мутантной) пролин-специфичной эндопротеазы PEP получали свежие споры *A. niger* PEP. 4 встряхиваемые колбы с 100 мл ферментационной среды 1 (10% масс./об. твердого вещества кукурузного экстракта, 1 масс./об. % глюкозы. H_2O , 0,1 % масс./об. $NaH_2PO_4 \cdot H_2O$, 0,05% масс./об. $MgSO_4 \cdot 7H_2O$, 0,025% масс./об. Basildon, pH 5,8) в 500 мл встряхиваемых колбах с дефлектором инокулировали 10^7 спор. Эти предварительные культуры инкубировали при 34°C и 170 оборотах в минуту в течение 16-24 часов. Из предварительных культур, 50 мл использовали для инокуляции 1 колбы с 1 литром ферментационной среды 2 (15% масс./об. мальтозы, 6% масс./об. Vasto-soytone, 1,5% масс./об. $(NH_4)_2SO_4$, 0,1% масс./об. $NaH_2PO_4 \cdot H_2O$, 0,1% масс./об. $MgSO_4 \cdot 7H_2O$, 0,1% масс./об. L-аргинина, 8 % масс./об. Tween-80, 2 % масс./об. Basildon, 2% масс./об. MES pH 5,1) в 5-литровой встряхиваемой колбе и встряхивали при 34 °C и 170 оборотах в минуту. Через 3, 4, 5, и 6 дней инкубации pH культуры понижали до pH 5,0 с помощью 2N HCl и образцы из каждой из этих временных точек анализировали на активность PEP. 50 мл образца отбирали и отделяли надосадочную жидкость от биомассы путем центрифугирования и последующей фильтрации. Образец с наибольшей активностью

использовали для описания полученного мутанта РЕР.

Пример 2. Измерения активности пролин-специфичной эндопротеазы (РЕР)

100 мкл культурального надосадочной жидкости, полученной в Примере 1, разводили в 0,1 М натрий-ацетатного буфера при рН 4.5 с 50 мМ NaCl, инкубировали 100 мкл 6 мМ Ac-AAP-pNA (ацетил-AlaAlaPro-паранитроанилин из Selleckchem или CPC Scientific; чистота > 95,0% на основании анализа ВЭЖХ) в 0,1 М NaAc буфере при рН 4.5 с 50 мМ NaCl, в 96 луночных плоскодонных МТР-планшетах (микротитрационных планшетах) Nunc. Через 60 минут при 20°C реакцию останавливали добавлением 40 мкл 1 М HCl. pNA, который были освобожден РЕР, измеряли спектрофотометром Tecan МТР при 405 нм (A405) (www.tecan.com). Пустые образцы получали смешиванием разбавленной культуральной надосадочной жидкости с раствором субстрата, который был смешан с раствором HCl заблаговременно. Активность выражается в рNASU.

1 рNASU - это количество фермента, которое высвобождает из Ac-AAP-pNA в течение 1 ч количество рNA, которое соответствует увеличению поглощения при 405 нм на 1 OD, в условиях, описанных выше. A405 не должно быть ниже пустого значения в начале реакции, или выше 2,5 в конце реакции, и также A405 не может превышать линейный диапазон используемого спектрофотометра.

Пример 3. Термическая стабильность пролин-специфичной эндопротеазы

Для оценки термостабильности родительского РЕР и мутантов, перечисленных в таблице 1, анализ активности проводили после инкубирования 100 мкл аликвоты десятикратного разведения надосадочной жидкости культуры, полученной в Примере 1, в буфере (0,1 М NaAc рН 4,5, с 50 мМ NaCl) при 55°C и 65°C в течение 15 мин. в ПЦР-планшете в машине для ПЦР. После 15 мин инкубации образцы быстро охлаждали до 25 °C в машине для ПЦР. Измеряли рNASU/мл каждого образца. Начальную активность, измеренную до инкубации при повышенной температуре (0 мин) использовали в качестве эталона (100%) для определения остаточной активности. Все активности измеряли четыре раза.

Таблица 1 показывает, что все пролин-специфические эндопротеазы, обладающие мутацией в положении 469, имеют значительную пониженную остаточную активность по сравнению с родительской пролин-специфичной эндопротеазой после выдерживания ферментов при 65°C в течение 15 мин.

Таблица 1. Остаточная активность мутантов пролин-специфичной эндопротеазы по сравнению с исходной пролин-специфичной эндопротеазой после выдержки при температуре 55°C и 65°C в течение 15 мин.

| PEP Клон | Замены по отношению к родительской SEQ ID NO: 1 | Остаточная активность (pNASU) при указанной T после 15' | | |
|-------------|--|--|------|------|
| | | 55°C | 60°C | 65°C |
| PEP | Родительская | 100% | 93% | 80% |
| P469A | P469A | 97% | 67% | 51% |
| P469C | P469C | 93% | 61% | 18% |
| P469D | P469D | 77% | 35% | <2% |
| P469E | P469E | 99% | 72% | 30% |
| P469F | P469F | 72% | 37% | <2% |
| P469G | P469G | 89% | 80% | 21% |
| P469H | P469H | 103% | 27% | 0% |
| P469I | P469I | 103% | 66% | 15% |
| P469K | P469K | 127% | 70% | 15% |
| P469L | P469L | 88% | 58% | 7% |
| P469M | P469M | 92% | 85% | 30% |
| P469N | P469N | 90% | 41% | <2% |
| P469Q | P469Q | 84% | 78% | 33% |
| P469R | P469R | 100% | 59% | 13% |
| P469S | P469S | 104% | 63% | 25% |
| P469T | P469T | 84% | 65% | 11% |
| P469V | P469V | 92% | 80% | 21% |
| P469W | P469W | 72% | 47% | <2% |
| P469Y | P469Y | 58% | 39% | <2% |

Пример 4. Термостойкость пролин-специфичной эндопротеазы, содержащей мутацию в положении P469 и дополнительную мутацию

Мутации в положении P469 в значительной степени способствуют снижению термостабильности пролин-специфичную эндопротеазы из *Aspergillus niger*. Другие аминокислотные замены, т.е. I204F, P377A, P477A, P304S и P377A, P466T и I204F, и P466T и P477A, вводили в последовательность пролин-специфичной эндопротеазы для исследования возможного дополнительного снижения термостабильности пролин-специфичной эндопротеазы из *A. niger*.

Клонирование, экспрессию и извлечение мутантов проводили, как описано в Примере 1. Определение остаточной активности проводили, как описано в Примере 3. В таблице 2 показана термостабильность родительской пролин-специфичной эндопротеазы и мутантной пролин-специфичной эндопротеазы, содержащей одну замену в положении 204, 377 и 477.

Таблица 2. Остаточная активность родительской пролин-специфичной эндопротеазы (PEP) из *Aspergillus niger* и мутантной PEP, содержащей одну замену

| | Замены по отношению к родительской SEQ ID NO 1 | Остаточная активность при указанной T после 15 ' | | |
|-------|--|--|------|------|
| | | 55°C | 60°C | 65°C |
| PEP | Родительская | 100% | 93% | 80% |
| PEP4 | I204F | 83% | 58% | 8% |
| PEP8 | P377A | 89% | 56% | 11% |
| PEP10 | P477A | 86% | 37% | 1% |

Таблица 3. Остаточная активность мутантной пролин-специфичную эндопротеазы из *Aspergillus niger*, которая несет замену в положении P469 и дополнительные замены

| | Замены по отношению к родительской SEQ ID NO: 1 | Остаточная активность при указанной T после 15 ' | | | | | |
|----------|---|--|--------|--------|--------|--------|--------|
| | | 60,4°C | 57,3°C | 52,5°C | 47,0°C | 42,4°C | 39,4°C |
| PEP-5_18 | P469D + I204F | 0% | 0% | 4% | 64% | 99% | 100% |
| PEP-5_57 | P477A + P469Q | 0% | 0% | 10% | 70% | 93% | 99% |
| PEP-5_22 | P469D + P377A | 0% | 5% | 39% | 79% | 99% | 100% |
| PEP-5_15 | P377A + P304S + P469Y | 0% | 5% | 37% | 82% | 100% | 99% |
| PEP-5_73 | P466T + P469Q + I204F | 1% | 1% | 12% | 66% | 91% | 98% |
| PEP-5_74 | P466T + P469Q + P477A | 0% | 0% | 3% | 46% | 80% | 96% |

Результаты в таблице 3 показывают, что термостабильность пролин-специфичной эндопротеазы, включающий замену в положении 469, может быть дополнительно снижена путем замены в одном или нескольких дополнительных положениях аминокислот, без потери активности фермента при более низких температурах.

Пример 5. Термическая стабильность мутантов гомологичной пролин-специфичной эндопротеазы

Для того, чтобы установить, является ли мутации в положении P469 в пролин-специфичной эндопротеазы *Aspergillus niger* в более общем смысле применимой для снижения термостабильности в гомологичном положении в другой пролин-специфичной эндопротеазе, мутацию в положении, гомологичном 469 в SEQ ID NO: 1, вводили в пролин-специфичные эндопротеазы, полученные из *Aspergillus flavus*, *Aspergillus aculeatus* и *Rasamsonia emersonii*.

Для того, чтобы установить положение в пролин-специфичной эндопротеазе, которое соответствует мутированному положению в эталонной пролин-специфичной

эндопротеазе из *Aspergillus niger*, последовательности выравняли. Выравнивание показано на фиг. 2, и проценты идентичности приведены в таблице 7. Идентичность была определена с помощью NEEDLE с использованием настроек NOBRIEF. Самая длинная идентичность (Longest_identity), т.е. включая препоследовательность, была принята в качестве меры идентичности между двумя последовательностями.

Таблица 7. Идентичность последовательности гомологичных пролин-специфичных эндопротеаз, полученный из *A. carbonarius*, *A. Flavus*, *A. aculeatus* и *Rasamsonia emersonii* по отношению к пролин-специфичной эндопротеазе из *A. niger*.

| Происхождение пролин-специфичной эндопротеазы | Аминокислотная последовательность (в том числе пре-про последовательности) | Longest_identity определена с помощью NEEDLE (NOBRIEF) |
|---|--|--|
| <i>Aspergillus carbonarius</i> | SEQ ID NO: 9 | 91,4% |
| <i>Aspergillus Flavus</i> | SEQ ID NO: 10 | 81,0% |
| <i>Aspergillus aculeatus</i> | SEQ ID NO: 11 | 81,0% |
| <i>Rasamsonia emersonii</i> | SEQ ID NO: 12 | 62,1% |

Замену вводили P469L в *Aspergillus Flavus*, *Aspergillus aquileatus* и *Rasamsonia emersonii*. Для оценки термостабильности родительской пролин-специфичной эндопротеазы и P469L мутантов анализу активности предшествовало инкубирование 100 мкл аликвот десятикратного разведения культуральной надосадочной жидкости в буфере (0,1 М NaAc pH 4,5, 50 mM NaCl) при 60°C в течение 15 мин. в ПЦР-планшете в машине для ПЦР. После 15 мин. инкубации образцы быстро охлаждали до 25°C в машине для ПЦР. Измеряли рNASU/мл для каждого образца. Начальную активность, измеренную до инкубации при повышенной температуре (0 мин) использовали в качестве эталона (100%) для определения остаточной активности.

Таблица 8. Остаточная активность грибных пролин-специфичных эндопротеаз, имеющих аминокислотную замену, гомологичную P469L в PEP из *A. niger*, после инкубации в течение 15' при 60°C

| Родительская | Родительская активность | P469L |
|---------------------|-------------------------|-------|
| <i>A. niger</i> | 100% | 62% |
| <i>A. flavus</i> | 100% | 3% |
| <i>A. aculeatus</i> | 100% | 18% |
| <i>R. emersonii</i> | 100% | 85% |

Результаты, представленные в таблице 8, показывают, что замещение аминокислоты в положении 469, т.е. P469L, снижает термостабильность пролин-специфичной эндопротеазы не только из *Aspergillus niger*, но также и в случае

гомологичных положений в пролин-специфических эндопротеазах из *Aspergillus Flavus*, *Aspergillus Aculeatus* и *Rasamsonia emersonii*.

Пример 6. Субстратная специфичность пролин-специфичных эндопротеаз

Для того, чтобы подтвердить, что мутанты являются пролин-специфичными эндопротеазами, оценивали субстратную специфичность различных мутантов PEP с помощью цитохрома *c* из сердца лошади. Разведения культуральной надосадочной жидкости, полученной в Примере 1, инкубировали с цитохромом *c* из сердца лошади (Sigma), который имеет аминокислотную последовательность SEQ ID NO: 3. Субстрат получали растворением 1 мг/мл цитохрома *C* в 100 mM натрий-ацетатного буфера, pH 4,5 и нагревании при 95°C в течение 15 мин. Культуральную надосадочную жидкость разбавляли в 100 mM NaAc буфере, pH 4,5, и инкубировали с раствором субстрата цитохрома *c* при 50°C в течение 3 часов. Реакцию прекращали путем разведения в воде и добавлением 0,4 M NaOH для повышения pH до 10. Инкубированные реакционные смеси анализировали на Accela UHPLC (Thermo Electron, Бреда, Нидерланды) в сочетании с масс-спектрометром LTQ-Orbitrap Fourier Transform Mass Spectrometer (Thermo Electron, Бремен, Германия). Хроматографическое разделение проводили на колонке C-18 Eclipse XDB Zorbax 2,1 × 50 мм, с размером частиц в 1,8 мкм и размером пор 80Å (Agilent, Санта-Клара, Калифорния, США), с использованием градиентного элюирования водой класса LC-MS (A), содержащей 0,1% муравьиной кислоты и (B) ацетонитрила класса LC-MS, содержащего 0,1% раствор муравьиной кислоты (Biosolve BV, Нидерланды) в качестве подвижной фазы. 25 мин. градиент начинали от 0%, выдерживали в течение 1 минуты, а затем линейно повышали до 40% (B) в течение 14 мин, после чего промывали 80% (B) в течение 4 мин. и повторно уравнивали с 0% (B) в течение 5 мин. Скорость потока поддерживали на уровне 0,4 мл/мин, с использованием инжецируемого объема 5 мкл, и температуры колонки 50°C. Получение данных масс-спектрометрии осуществляли с Top 3 зависимым от данных получением с использованием опций «Хроматографии» и «Динамическое исключения», которые были включены только и состояния заряда 2 и 3 были включены. Разрешение для сканирования FT MS составляло 15000 и сканировали в диапазоне 210-2000 *m/z*, при этом эксперименты MS/MS проводили в ионной ловушке. Ширина изоляции была установлена на уровне 3,0 *m/z*, а нормализованная энергия столкновений было установлена на уровне 35. Идентификацию пептида проводили с использованием точной массы и MS/MS De Novo секвенированных данных.

Когда основные пептиды SEQ ID NO: 4-SEQ ID NO: 8, в частности SEQ ID NO: 4-SEQ ID NO: 6 были получены, это показало, что мутантные пролин-специфические эндопротеазы имели активность пролин-специфичной эндопротеазы, т.е. имеют

предпочтение разрезать в положении, в котором в пептиде находится остаток пролина. Это было подтверждено для P469D, P469F, P469G, P469H, P469N, P469Q, P469S, P469T, и P469W (результаты не показаны).

СПИСОК ПОСЛЕДОВАТЕЛЬНОСТЕЙ

<110> DSM IP Assets B.V.

<120> Proline-specific endoprotease and use thereof

<130> 30378-WO-PCT

<160> 16

<170> BiSSAP 1.2

<210> 1

<211> 521

<212> PRT

<213> Artificial Sequence

<220>

<223> A. niger proline-specific endoprotease, with
Pectinemethylesterase signal sequence

<400> 1

```

Met Val Lys Ser Ile Leu Ala Ser Val Phe Phe Ala Ala Thr Ala Leu
1          5          10      15
Ala Ala Arg Pro Arg Leu Val Pro Lys Pro Val Ser Arg Pro Ala Ser
20          25          30
Ser Lys Ser Ala Ala Thr Thr Gly Glu Ala Tyr Phe Glu Gln Leu Leu
35          40          45
Asp His His Asn Pro Glu Lys Gly Thr Phe Ser Gln Arg Tyr Trp Trp
50          55          60
Ser Thr Glu Tyr Trp Gly Gly Pro Gly Ser Pro Val Val Leu Phe Thr
65          70          75          80
Pro Gly Glu Val Ser Ala Asp Gly Tyr Glu Gly Tyr Leu Thr Asn Glu
85          90          95
Thr Leu Thr Gly Val Tyr Ala Gln Glu Ile Gln Gly Ala Val Ile Leu
100         105         110
Ile Glu His Arg Tyr Trp Gly Asp Ser Ser Pro Tyr Glu Val Leu Asn
115         120         125
Ala Glu Thr Leu Gln Tyr Leu Thr Leu Asp Gln Ala Ile Leu Asp Met
130         135         140
Thr Tyr Phe Ala Glu Thr Val Lys Leu Gln Phe Asp Asn Ser Thr Arg
145         150         155         160
Ser Asn Ala Gln Asn Ala Pro Trp Val Met Val Gly Gly Ser Tyr Ser
165         170         175
Gly Ala Leu Thr Ala Trp Thr Glu Ser Val Ala Pro Gly Thr Phe Trp
180         185         190
Ala Tyr His Ala Thr Ser Ala Pro Val Glu Ala Ile Tyr Asp Tyr Trp
195         200         205
Gln Tyr Phe Tyr Pro Ile Gln Gln Gly Met Ala Gln Asn Cys Ser Lys
210         215         220
Asp Val Ser Leu Val Ala Glu Tyr Val Asp Lys Ile Gly Lys Asn Gly
225         230         235         240
Thr Ala Lys Glu Gln Gln Ala Leu Lys Glu Leu Phe Gly Leu Gly Ala
245         250         255
Val Glu His Phe Asp Asp Phe Ala Ala Val Leu Pro Asn Gly Pro Tyr
260         265         270
Leu Trp Gln Asp Asn Asp Phe Ala Thr Gly Tyr Ser Ser Phe Phe Gln
275         280         285
Phe Cys Asp Ala Val Glu Gly Val Glu Ala Gly Ala Ala Val Thr Pro
290         295         300
Gly Pro Glu Gly Val Gly Leu Glu Lys Ala Leu Ala Asn Tyr Ala Asn
305         310         315         320
Trp Phe Asn Ser Thr Ile Leu Pro Asp Tyr Cys Ala Ser Tyr Gly Tyr
325         330         335

```

Trp Thr Asp Glu Trp Ser Val Ala Cys Phe Asp Ser Tyr Asn Ala Ser
 340 345 350
 Ser Pro Ile Tyr Thr Asp Thr Ser Val Gly Asn Ala Val Asp Arg Gln
 355 360 365
 Trp Glu Trp Phe Leu Cys Asn Glu Pro Phe Phe Tyr Trp Gln Asp Gly
 370 375 380
 Ala Pro Glu Gly Thr Ser Thr Ile Val Pro Arg Leu Val Ser Ala Ser
 385 390 395 400
 Tyr Trp Gln Arg Gln Cys Pro Leu Tyr Phe Pro Glu Thr Asn Gly Tyr
 405 410 415
 Thr Tyr Gly Ser Ala Lys Gly Lys Asn Ala Ala Thr Val Asn Ser Trp
 420 425 430
 Thr Gly Gly Trp Asp Met Thr Arg Asn Thr Thr Arg Leu Ile Trp Thr
 435 440 445
 Asn Gly Gln Tyr Asp Pro Trp Arg Asp Ser Gly Val Ser Ser Thr Phe
 450 455 460
 Arg Pro Gly Gly Pro Leu Ala Ser Thr Ala Asn Glu Pro Val Gln Ile
 465 470 475 480
 Ile Pro Gly Gly Phe His Cys Ser Asp Leu Tyr Met Ala Asp Tyr Tyr
 485 490 495
 Ala Asn Glu Gly Val Lys Lys Val Val Asp Asn Glu Val Lys Gln Ile
 500 505 510
 Lys Glu Trp Val Glu Glu Tyr Tyr Ala
 515 520

<210> 2

<211> 1566

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<222> 1..1566

<223> /organism="Artificial Sequence"

 /note="A. niger proline-specific endoprotease, with
 Pectinemethylesterase signal sequence"

 /mol_type="unassigned DNA"

<400> 2

atggtc aagt cc atcctggc ctccgtcttc ttcgctgcca ctgctcttgc tgcaaggcct 60
 cgtctcgttc ccaagcccgt ttctcgtccc gccagctcca agtccgctgc tactactggt 120
 gaggcctact ttgaacagct gttggaccac cacaaccctg agaagggtac tttctcgcaa 180
 agatactggt ggagcaccga gtactgggggt ggtcccggat cccccgttgt cctgttctact 240
 cccggtgagg tcagcgtga tggctacgag ggttatctga ccaacgagac tctcaccggt 300
 gtctacgcc aggagattca ggggtgctgtc atcctgatcg aacaccgata ctgggggtgac 360
 tcgtctccct acgagggtgct gaacgccgag actctccagt acttgaccct cgaccaggct 420
 atccttgata tgacctactt cgccgaaacc gtcaagctcc agtttgacaa ctccaccgcg 480
 tccaacgctc agaacgctcc ttgggttatg gtcggcgcca gctacagcgg tgctctgact 540
 gcttgaccg agtccgttgc tcccggcacc ttctgggctt accacgccac ctctgctcct 600
 gttgaggcca tctacgacta ctggcaatac ttctacccca ttcagcaggg tatggctcag 660
 aactgctcca aagatgtctc tctttagtga gaatacgtcg acaagatcgg caagaacggc 720
 actgccaagg agcaacaggc tctgaaggag cttttcggcc taggagcagt ggagcacttc 780

gacgacttcg ccgctgttct gcccaacggg ccttacctct ggcaagacaa cgactttgcc 840
accggttact cttctttctt ccagttctgt gatgccgtcg aggggtgtcga ggctgggtgct 900
gccgtcacc cgggtcctga aggtgttggt ctggaaaagg cccttgctaa ctacgcgaac 960
tggttcaact ctaccatcct ccccgattac tgcgccagct acggctactg gactgacgag 1020
tggtcctcgc cctgcttcga ctctacaac gcctcctctc ctatatacac cgacaccagc 1080
gttggtaacg ccgtcgaccg tcagtgggag tggttcctct gcaatgagcc cttcttttac 1140
tggcaggacg gtgccccga ggggtacttca acgatagtag cccgcttagt gtccgcctcc 1200
tactggcagc gtcaatgtcc gttgtacttc cccgagacta acggttacac ctacggctcc 1260
gccaagggaa agaacgccgc caccgtcaac agctggaccg gtggctggga catgaccctg 1320
aacaccacc gtctgatctg gacgaacggc caatacgacc cctggcgtga ctccgggtgc 1380
tcttccacct tccgccccgg tgggtcccctc gcttcgaccg ccaacgagcc cgtccagata 1440
ataccgggtg gtttccattg ctccgacctc tacatggcag actactacgc caacgagggc 1500
gtcaagaagg ttgtcgacaa cgaagtcaaa caaatcaagg agtggggtga ggaatactac 1560
gcgtaa 1566

<210> 3
<211> 104
<212> PRT
<213> Artificial Sequence

<220>
<223> Amino acid sequence horse heart cytochrome C

<400> 3
Gly Asp Val Glu Lys Gly Lys Lys Ile Phe Val Gln Lys Cys Ala Gln
1 5 10 15
Cys His Thr Val Glu Lys Gly Gly Lys His Lys Thr Gly Pro Asn Leu
20 25 30
His Gly Leu Phe Gly Arg Lys Thr Gly Gln Ala Pro Gly Phe Thr Tyr
35 40 45
Thr Asp Ala Asn Lys Asn Lys Gly Ile Thr Trp Lys Glu Glu Thr Leu
50 55 60
Met Glu Tyr Leu Glu Asn Pro Lys Lys Tyr Ile Pro Gly Thr Lys Met
65 70 75 80
Ile Phe Ala Gly Ile Lys Lys Lys Thr Glu Arg Glu Asp Leu Ile Ala
85 90 95
Tyr Leu Lys Lys Ala Thr Asn Glu
100

<210> 4
<211> 30
<212> PRT
<213> Artificial Sequence

<220>
<223> Fragment of cytochrome C digested with PEP

<400> 4
Gly Asp Val Glu Lys Gly Lys Lys Ile Phe Val Gln Lys Cys Ala Gln

<223> BC2G075 Aspergillus carbonarius proline-specific endoprotease (PEP) with A. niger pectinmethylesterase signal sequence and A.niger PEP prosequence

<400> 9

Met Val Lys Ser Ile Leu Ala Ser Val Phe Phe Ala Ala Thr Ala Leu
1 5 10 15
Ala Ala Arg Pro Arg Leu Val Pro Lys Pro Val Ser Arg Pro Ala Ser
20 25 30
Ser Thr Ser Ala Ala Thr Thr Gly Glu Ala Tyr Phe Glu Gln Leu Val
35 40 45
Asp His His Asn Pro Glu Lys Gly Thr Phe Ser Gln Arg Tyr Trp Trp
50 55 60
Ser Thr Glu Tyr Trp Gly Gly Pro Gly Ser Pro Val Val Leu Phe Thr
65 70 75 80
Pro Gly Glu Val Ser Ala Asp Gly Tyr Glu Gly Tyr Leu Thr Asn Asp
85 90 95
Thr Leu Thr Gly Val Tyr Ala Gln Glu Ile Gln Gly Ala Val Val Leu
100 105 110
Ile Glu His Arg Tyr Trp Gly Asp Ser Ser Pro Tyr Glu Val Leu Asn
115 120 125
Ala Glu Thr Leu Gln Tyr Leu Thr Leu Asp Gln Ala Val Leu Asp Met
130 135 140
Thr Tyr Phe Ala Glu Thr Val Lys Phe Gln Phe Asp Asn Ser Thr Arg
145 150 155 160
Ser Asn Ala Gln Asn Ala Pro Trp Val Met Val Gly Gly Ser Tyr Ser
165 170 175
Gly Ala Leu Thr Ala Trp Val Glu Ser Val Ala Pro Gly Thr Phe Trp
180 185 190
Ala Tyr His Ala Thr Ser Ala Pro Val Glu Ala Ile Tyr Asp Phe Trp
195 200 205
Gln Tyr Phe Tyr Pro Ile Ser Gln Gly Met Ala Gln Asn Cys Ser Lys
210 215 220
Asp Val Ser Arg Val Ala Glu His Val Asp Lys Val Gly Lys Ser Gly
225 230 235 240
Thr Ala Glu Glu Gln Gln Lys Leu Lys Glu Leu Phe Gly Leu Gly Ala
245 250 255
Leu Glu His Tyr Asp Asp Phe Ala Ala Val Leu Pro Asn Gly Pro Tyr
260 265 270
Leu Trp Gln Asp Asn Asp Phe Ala Thr Gly Tyr Ser Glu Phe Phe Gln
275 280 285
Phe Cys Asp Ala Val Glu Gly Val Glu Ala Gly Ala Ala Val Thr Pro
290 295 300
Gly Pro Glu Gly Val Gly Leu Glu Lys Ala Leu Ala Asn Tyr Ala Tyr
305 310 315 320
Trp Phe Asn Ser Thr Leu Leu Pro Asn Tyr Cys Ala Ser Tyr Gly Tyr
325 330 335
Trp Ser Asp Glu Trp Ser Val Ala Cys Phe Asp Ser Tyr Asn Ala Ser
340 345 350
Ser Pro Leu Phe Thr Asp Thr Ser Val Asp Asn Ala Val Asp Arg Gln
355 360 365
Trp Glu Trp Phe Leu Cys Asn Glu Pro Phe Phe Trp Trp Gln Asp Gly
370 375 380
Ala Pro Glu Asp Val Thr Thr Ile Val Pro Arg Leu Val Asn Ala Glu
385 390 395 400
Tyr Trp Gln Arg Gln Cys Ser Leu Tyr Phe Pro Glu Thr Asn Gly Tyr
405 410 415
Thr Phe Gly Ser Ala Lys Asn Lys Thr Ala Ala Thr Val Asn Asp Trp
420 425 430
Thr Gly Gly Trp Phe Glu Thr Arg Asn Thr Thr Arg Leu Ile Trp Thr
435 440 445
Asn Gly Gln Tyr Asp Pro Trp Arg Asp Ser Gly Val Ser Ser Thr Phe
450 455 460
Arg Pro Gly Gly Gln Leu Val Ser Thr Ala Asn Glu Pro Val Gln Ile

| | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 465 | | | | | 470 | | | | | 475 | | | | 480 |
| Ile | Pro | Gly | Gly | Phe | His | Cys | Ser | Asp | Leu | Tyr | Met | Ala | Asp | Tyr |
| | | | | 485 | | | | | 490 | | | | | 495 |
| Ala | Asn | Ala | Gly | Val | Arg | Lys | Val | Val | Asp | Asn | Glu | Val | Ala | Gln |
| | | | 500 | | | | | 505 | | | | | 510 | Ile |
| Lys | Lys | Trp | Val | Ala | Glu | Tyr | Tyr | Ala | | | | | | |
| | | 515 | | | | | 520 | | | | | | | |

<210> 10
 <211> 521
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> BC2G077 Aspergillus_flavus proline-specific endoprotease (PEP)
 with A. niger pectinemethylesterase signal sequence and A.niger
 PEP prosequence

<400> 10

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Val | Lys | Ser | Ile | Leu | Ala | Ser | Val | Phe | Phe | Ala | Ala | Thr | Ala | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Ala | Ala | Arg | Pro | Arg | Leu | Val | Pro | Lys | Pro | Val | Ser | Arg | Pro | Ala | Ser |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Ser | Lys | Ser | Ala | Ala | Thr | Thr | Gly | Glu | Ala | Tyr | Phe | Glu | Gln | Leu | Leu |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Asp | His | His | Asp | Ser | Ser | Lys | Gly | Thr | Phe | Ser | Gln | Arg | Tyr | Trp | Trp |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Ser | Thr | Glu | Tyr | Trp | Gly | Gly | Pro | Gly | Ser | Pro | Val | Val | Leu | Phe | Thr |
| 65 | | | | | 70 | | | | 75 | | | | | 80 | |
| Pro | Gly | Glu | Ala | Ser | Ala | Asp | Gly | Tyr | Glu | Gly | Tyr | Leu | Thr | Asn | Asn |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Thr | Leu | Thr | Gly | Leu | Tyr | Ala | Gln | Glu | Ile | Gln | Gly | Ala | Val | Ile | Leu |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Ile | Glu | His | Arg | Tyr | Trp | Gly | Asp | Ser | Ser | Pro | Tyr | Glu | Glu | Leu | Thr |
| | | 115 | | | | | 120 | | | | | 125 | | | |
| Ala | Glu | Thr | Leu | Gln | Tyr | Leu | Thr | Leu | Glu | Gln | Ser | Ile | Leu | Asp | Leu |
| | | 130 | | | | 135 | | | | | 140 | | | | |
| Thr | His | Phe | Ala | Glu | Thr | Val | Gln | Leu | Glu | Phe | Asp | Thr | Ser | Asn | Ser |
| 145 | | | | | 150 | | | | | 155 | | | | 160 | |
| Ser | Asn | Ala | Pro | Lys | Ala | Pro | Trp | Val | Leu | Val | Gly | Gly | Ser | Tyr | Ser |
| | | | | 165 | | | | | 170 | | | | | 175 | |
| Gly | Ala | Leu | Ala | Ala | Trp | Thr | Ala | Ala | Val | Ala | Pro | Gly | Thr | Phe | Trp |
| | | | 180 | | | | 185 | | | | | | 190 | | |
| Ala | Tyr | His | Ala | Thr | Ser | Ala | Pro | Val | Gln | Ala | Ile | Asp | Asp | Phe | Trp |
| | | 195 | | | | | 200 | | | | | 205 | | | |
| Gln | Tyr | Phe | Asp | Pro | Ile | Arg | His | Gly | Met | Ala | Pro | Asn | Cys | Ser | Arg |
| | 210 | | | | | 215 | | | | | 220 | | | | |
| Asp | Val | Ser | Leu | Val | Ala | Asn | His | Ile | Asp | Thr | Val | Gly | Lys | Asn | Gly |
| 225 | | | | | 230 | | | | | 235 | | | | 240 | |
| Ser | Ala | Ala | Asp | Gln | Leu | Ala | Leu | Lys | Glu | Leu | Phe | Gly | Leu | Glu | Ala |
| | | | 245 | | | | | | 250 | | | | | 255 | |
| Leu | Glu | His | Tyr | Asp | Asp | Phe | Ala | Ala | Ala | Leu | Pro | Thr | Gly | Pro | Tyr |
| | | | 260 | | | | 265 | | | | | | 270 | | |
| Leu | Trp | Gln | Ser | Asn | Thr | Phe | Val | Thr | Gly | Tyr | Ser | Asn | Phe | Phe | Ala |
| | | 275 | | | | | 280 | | | | | 285 | | | |
| Phe | Cys | Asp | Ala | Val | Glu | Asn | Val | Glu | Ala | Gly | Ala | Ala | Val | Val | Pro |
| 290 | | | | | | 295 | | | | | 300 | | | | |
| Gly | Pro | Glu | Gly | Val | Gly | Leu | Gln | Lys | Ala | Leu | Thr | Gly | Tyr | Ala | Asn |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 |
| Trp | Phe | Asn | Ser | Thr | Ile | Ile | Pro | Gly | Tyr | Cys | Ala | Ser | Tyr | Gly | Tyr |
| | | | | 325 | | | | | 330 | | | | | 335 | |
| Trp | Thr | Asp | Asn | Arg | Thr | Val | Ala | Cys | Phe | Asp | Thr | His | Asn | Pro | Ser |
| | | | 340 | | | | | 345 | | | | | 350 | | |
| Ser | Ala | Ile | Phe | Thr | Asp | Thr | Ser | Val | Asp | Asn | Ala | Val | Asp | Arg | Gln |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | 355 | | | | | | 360 | | | | | 365 | | | |
| Trp | Gln | Trp | Phe | Leu | Cys | Asn | Glu | Pro | Phe | Phe | Trp | Trp | Gln | Asp | Gly |
| | 370 | | | | | 375 | | | | | 380 | | | | |
| Ala | Pro | Glu | Gly | Val | Pro | Thr | Ile | Val | Pro | Arg | Thr | Ile | Asn | Ala | Glu |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 |
| Tyr | Trp | Gln | Arg | Gln | Cys | Ser | Leu | Tyr | Phe | Pro | Glu | Val | Asn | Gly | Tyr |
| | | | | 405 | | | | | 410 | | | | | 415 | |
| Thr | Tyr | Gly | Ser | Ala | Lys | Gly | Lys | Thr | Ala | Ala | Thr | Val | Asn | Thr | Trp |
| | | | 420 | | | | | 425 | | | | | 430 | | |
| Thr | Gly | Gly | Trp | Ser | Asp | Ser | Lys | Asn | Thr | Ser | Arg | Leu | Leu | Trp | Val |
| | | 435 | | | | | 440 | | | | | 445 | | | |
| Asn | Gly | Gln | Tyr | Asp | Pro | Trp | Arg | Asp | Ser | Gly | Val | Ser | Ser | Thr | His |
| | 450 | | | | | 455 | | | | | 460 | | | | |
| Arg | Pro | Gly | Gly | Pro | Leu | Thr | Ser | Thr | Ala | Asp | Glu | Pro | Val | Gln | Val |
| 465 | | | | | 470 | | | | | 475 | | | | | 480 |
| Ile | Pro | Gly | Gly | Phe | His | Cys | Ser | Asp | Leu | Tyr | Leu | Lys | Asp | Tyr | Phe |
| | | | | 485 | | | | 490 | | | | | | 495 | |
| Ala | Asn | Ala | Gly | Val | Lys | Gln | Val | Val | Asp | Asn | Ala | Val | Ala | Gln | Ile |
| | | | 500 | | | | | 505 | | | | | 510 | | |
| Lys | Ser | Trp | Val | Ala | Glu | Tyr | Tyr | Lys | | | | | | | |
| | | 515 | | | | | 520 | | | | | | | | |

<210> 11

<211> 521

<212> PRT

<213> Artificial Sequence

<220>

<223> BC2G076 *Aspergillus_aculeatus* proline-specific endoprotease (PEP) with *A. niger* pectinemethylsterase signal sequence and *A.niger* PEP prosequence

<400> 11

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Val | Lys | Ser | Ile | Leu | Ala | Ser | Val | Phe | Phe | Ala | Ala | Thr | Ala | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Ala | Ala | Arg | Pro | Arg | Leu | Val | Pro | Lys | Pro | Val | Ser | Arg | Pro | Ala | Ser |
| | | | 20 | | | | | 25 | | | | | 30 | | |
| Ser | Lys | Ser | Ala | Ala | Thr | Thr | Gly | Glu | Ala | Tyr | Phe | Glu | Gln | Leu | Ile |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Asp | His | Ser | Asp | Pro | Ser | Lys | Gly | Thr | Phe | Ser | Gln | Arg | Tyr | Trp | Tyr |
| | 50 | | | | | 55 | | | | | 60 | | | | |
| Ser | Ala | Gln | Tyr | Trp | Gly | Gly | Pro | Gly | Ser | Pro | Val | Val | Leu | Phe | Thr |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 |
| Pro | Gly | Glu | Val | Ser | Ala | Asp | Gly | Tyr | Gln | Gly | Tyr | Leu | Thr | Asn | Ala |
| | | | | 85 | | | | | 90 | | | | | 95 | |
| Thr | Leu | Thr | Gly | Val | Tyr | Ala | Gln | Gln | Leu | Gln | Gly | Ala | Val | Val | Leu |
| | | | 100 | | | | | 105 | | | | | 110 | | |
| Val | Glu | His | Arg | Tyr | Trp | Gly | Gly | Ser | Ser | Pro | Tyr | Thr | Asn | Leu | Thr |
| | | 115 | | | | 120 | | | | | | | 125 | | |
| Ala | Glu | Thr | Leu | Gln | Tyr | Leu | Thr | Leu | Glu | Gln | Ser | Val | Leu | Asp | Leu |
| | 130 | | | | | 135 | | | | | 140 | | | | |
| Thr | Tyr | Phe | Ala | Glu | Asn | Val | Lys | Leu | Gly | Phe | Asp | Asn | Ser | Thr | Ser |
| 145 | | | | | 150 | | | | | 155 | | | | | 160 |
| Ser | Asn | Ala | Pro | His | Val | Pro | Trp | Val | Leu | Val | Gly | Gly | Ser | Tyr | Ser |
| | | | | 165 | | | | | | 170 | | | | 175 | |
| Gly | Ala | Leu | Thr | Ala | Trp | Thr | Glu | His | Leu | Ala | Pro | Gly | Thr | Phe | Trp |
| | | | 180 | | | | | 185 | | | | | 190 | | |
| Ala | Tyr | His | Ala | Thr | Ser | Ala | Pro | Val | Glu | Ser | Ile | Tyr | Asp | Phe | Trp |
| | | 195 | | | | | 200 | | | | | | 205 | | |
| Gln | Tyr | Phe | Arg | Pro | Ile | Gln | Asp | Gly | Met | Ala | Lys | Asn | Cys | Ser | Lys |
| | 210 | | | | | 215 | | | | | 220 | | | | |
| Asp | Val | Ser | Leu | Val | Ala | Glu | His | Val | Asp | Lys | Ile | Gly | Lys | Thr | Gly |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 |
| Thr | Lys | Ala | Gln | Gln | Thr | Glu | Leu | Lys | Lys | Leu | Phe | Gly | Leu | Glu | Ala |

245 250 255
 Leu Glu His Phe Asp Asp Phe Ala Ala Val Leu Pro Ile Gly Pro Tyr
 260 265 270
 Leu Trp Gln Asp Asn Thr Phe Ala Thr Gly Tyr Ser Asp Phe Phe Ala
 275 280 285
 Phe Cys Asp Ala Val Glu Asn Val Glu Ala Gly Ala Val Thr Pro
 290 295 300
 Gly Ala Glu Gly Val Gly Leu Glu Lys Ala Leu Thr Gly Tyr Ala Asn
 305 310 315 320
 Trp Phe Lys Asn Glu Ile Phe Pro Gly Tyr Cys Ala Ser Tyr Gly Tyr
 325 330 335
 Trp Ser Asp Glu Tyr Ser Val Ala Cys Tyr Asp Thr Tyr Asn Thr Thr
 340 345 350
 Ser Pro Leu Phe Thr Asp Thr Ser Val Asp Asn Ala Val Asp Arg Gln
 355 360 365
 Trp Gln Trp Phe Leu Cys Asn Glu Pro Phe Phe Trp Trp Gln Asp Gly
 370 375 380
 Ala Pro Ser Ser Glu Thr Thr Ile Val Pro Arg Leu Val Ser Ala Asp
 385 390 395 400
 Tyr Trp Gln Arg Gln Cys Ala Leu Tyr Phe Pro Glu Val Asn Gly Tyr
 405 410 415
 Thr Tyr Gly Ser Ala Lys Gly Lys Ser Ala Asn Thr Phe Asn Ala Trp
 420 425 430
 Thr Asp Gly Trp Phe Met Asn Gly Asn Ser Thr Arg Leu Ile Trp Thr
 435 440 445
 Asn Gly Gln Tyr Asp Pro Trp Arg Asp Ala Thr Val Ser Ser Thr Phe
 450 455 460
 Arg Pro Gly Gly Pro Leu Ala Ser Thr Pro Ser Glu Pro Val Gln Ile
 465 470 475 480
 Ile Pro Gly Gly Phe His Cys Ser Asp Leu Tyr Ile Ser Asp Ser Val
 485 490 495
 Val Asn Ala Gly Val Lys Lys Val Val Asp Asn Glu Val Ala Gln Ile
 500 505 510
 Lys Ala Trp Val Ala Glu Phe Tyr Ala
 515 520

<210> 12

<211> 526

<212> PRT

<213> *Rasamsonia emersonii*

<400> 12

Met Pro Ser Leu Ser Ser Leu Val Ala Leu Thr Ala Ser Leu Val Ser
 1 5 10 15
 Leu Ala Ala Ala Ala Ala Pro Arg Leu Pro Leu Pro Pro Arg Pro Pro
 20 25 30
 Leu Pro Pro Arg Asp Pro Leu His Gly Pro Thr Asn Ala Ser Ala Thr
 35 40 45
 Phe Gln Gln Leu Ile Asp His Asn His Pro Glu Leu Gly Thr Phe Ser
 50 55 60
 Gln Arg Tyr Trp Trp Asn Asp Glu Phe Trp Lys Gly Pro Gly Ser Pro
 65 70 75 80
 Val Val Leu Phe Thr Pro Gly Glu Glu Asp Ala Ser Gly Tyr Val Gly
 85 90 95
 Tyr Leu Lys Asn Thr Thr Ile Thr Gly Leu Ile Ala Gln Thr Ile Gly
 100 105 110
 Gly Ala Val Ile Val Leu Glu His Arg Tyr Trp Gly Gln Ser Ser Pro
 115 120 125
 Tyr Asp Ser Leu Thr Thr Lys Asn Leu Gln Tyr Leu Thr Leu Lys Gln
 130 135 140
 Ser Ile Ala Asp Leu Thr Tyr Phe Ala Lys Thr Val Lys Leu Pro Phe
 145 150 155 160
 Asp Arg Asn Gly Ser Ser Asn Ala Asp Lys Ala Pro Trp Val Leu Ser

| | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|
| | | | | 165 | | | | | 170 | | | | | 175 | | | |
| Gly | Gly | Ser | Tyr | Ser | Gly | Ala | Leu | Ser | Ala | Trp | Thr | Ala | Ser | Thr | Ser | | |
| | | | 180 | | | | | 185 | | | | | 190 | | | | |
| Pro | Gly | Thr | Phe | Trp | Ala | Tyr | His | Ala | Ser | Ser | Ala | Pro | Val | Glu | Ala | | |
| | | 195 | | | | | 200 | | | | | | 205 | | | | |
| Ile | Tyr | Asp | Tyr | Trp | Gln | Tyr | Phe | Ala | Pro | Val | Gln | Asp | Gly | Leu | Pro | | |
| | 210 | | | | | 215 | | | | | 220 | | | | | | |
| Ala | Asn | Cys | Ser | Lys | Asp | Leu | Ser | Arg | Val | Val | Asp | Tyr | Ile | Asp | Ser | | |
| 225 | | | | | 230 | | | | | 235 | | | | | 240 | | |
| Val | Leu | Gln | Ser | Gly | Asn | Ala | Thr | Ala | Lys | Gln | Gln | Leu | Lys | Asp | Leu | | |
| | | | | 245 | | | | | 250 | | | | | 255 | | | |
| Phe | Gly | Leu | Gly | Ala | Leu | Glu | His | Asp | Asp | Asp | Phe | Ala | Ser | Ala | Leu | | |
| | | | 260 | | | | | 265 | | | | | | 270 | | | |
| Glu | Asn | Gly | Pro | Trp | Leu | Trp | Gln | Ser | Asn | Ser | Phe | Tyr | Asp | Pro | Tyr | | |
| | | 275 | | | | | 280 | | | | | 285 | | | | | |
| Pro | Pro | Val | Tyr | Glu | Phe | Cys | Asp | Tyr | Val | Glu | Asn | Ala | Tyr | Ala | Ser | | |
| | 290 | | | | | 295 | | | | | 300 | | | | | | |
| Pro | Pro | Val | Ala | Ala | Gly | Pro | Asp | Gly | Val | Gly | Leu | Glu | Lys | Ala | Leu | | |
| 305 | | | | | 310 | | | | | 315 | | | | | 320 | | |
| Ser | Gly | Tyr | Ala | Thr | Trp | Trp | Asn | Lys | Val | Phe | Phe | Pro | Gly | Tyr | Cys | | |
| | | | | 325 | | | | | 330 | | | | | | 335 | | |
| Ala | Thr | Tyr | Gly | Tyr | Trp | Ser | Ser | Asn | Asp | Ser | Ile | Ala | Cys | Phe | Asp | | |
| | | | 340 | | | | | 345 | | | | | | 350 | | | |
| Thr | Tyr | Asn | Gln | Ser | Ser | Pro | Met | Phe | Thr | Asp | Leu | Ser | Val | Ser | Asn | | |
| | | 355 | | | | | 360 | | | | | | 365 | | | | |
| Thr | Ile | Asn | Arg | Gln | Trp | Asn | Trp | Phe | Leu | Cys | Asn | Glu | Pro | Phe | Phe | | |
| | 370 | | | | | 375 | | | | | | 380 | | | | | |
| Tyr | Trp | Gln | Asp | Gly | Ala | Pro | Lys | Asn | Val | Pro | Ser | Ile | Val | Ser | Arg | | |
| 385 | | | | | 390 | | | | | 395 | | | | | 400 | | |
| Leu | Val | Thr | Ala | Glu | Tyr | Trp | Gln | Arg | Gln | Cys | Pro | Leu | Phe | Phe | Pro | | |
| | | | | 405 | | | | | 410 | | | | | | 415 | | |
| Glu | Glu | Asp | Gly | Tyr | Thr | Tyr | Gly | Ser | Ala | Lys | Gly | Lys | Thr | Ala | Ala | | |
| | | | 420 | | | | | 425 | | | | | | 430 | | | |
| Asp | Val | Asn | Ala | Trp | Thr | Lys | Gly | Trp | Phe | Leu | Thr | Asn | Thr | Thr | Arg | | |
| | | 435 | | | | | 440 | | | | | | 445 | | | | |
| Leu | Ile | Trp | Thr | Asn | Gly | Glu | Leu | Asp | Pro | Trp | Arg | Ser | Ala | Gly | Val | | |
| | 450 | | | | | 455 | | | | | 460 | | | | | | |
| Ser | Ser | Lys | Phe | Arg | Pro | Gly | Gly | Pro | Leu | Gln | Ser | Thr | Pro | Gln | Ala | | |
| 465 | | | | | 470 | | | | | | 475 | | | | 480 | | |
| Pro | Leu | Gln | Leu | Ile | Pro | Glu | Gly | Val | His | Cys | Tyr | Asp | Leu | Ile | Leu | | |
| | | | | 485 | | | | | | 490 | | | | | 495 | | |
| Lys | Asn | Ala | Glu | Ala | Asn | Ala | Gly | Val | Gln | Arg | Val | Val | Thr | Asn | Glu | | |
| | | | 500 | | | | | 505 | | | | | | 510 | | | |
| Val | Ala | Gln | Ile | Lys | Ala | Trp | Val | Asn | Glu | Tyr | Tyr | Arg | Lys | | | | |
| | | 515 | | | | | 520 | | | | | | 525 | | | | |

<210> 13

<211> 1565

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<222> 1..1565

<223> /organism="Artificial Sequence"

/note="BC2G075_Aspergillus_carbonarius proline-specific
endoprotease (PEP) with A. niger pectinemethylesterase signal
sequence and A.niger PEP prosequence "
/mol_type="unassigned DNA"

<400> 13

atgggtcaagt ccactctggc ctccgtcttc ttcgctgcca ctgctcttgc tgctcgctct 60

cgcttagtgc ccaagcccgt gtctcgcccc gctccagca cttctgccgc caccaccggt 120

gaggcctact tcgagcagct ggttgaccac cacaaccccc agaagggcac cttctcccag 180
cgctactggg ggagcactga gtactggggg ggtccccggc cccccgtcgt cctcttcacc 240
cccggggaag tctctgccga tggctacgag ggctacctga ccaacgacac cctgaccggg 300
gtctacgctc aggagatcca ggggtgctgt gtgttgattg agcaccgtta ctggggcgac 360
agcagcccct acgaggctct caacgccgag actctccagt acctgaccct cgaccaggct 420
gtccttgaca tgacctactt cgctgagact gtcaagttcc agttcgacaa ctcgaccggc 480
agcaacgccc agaacgctcc ttgggtgatg gtcgggtgaa gctactctgg tgctctcacc 540
gcctggggtg agtccgctgc tcctggaacc ttctgggcct accacgccac ctcggtctct 600
gttgaggcca tctacgactt ctggcagtac ttctaccca tctcccaggg tatggcccag 660
aactgctcca aggatgtctc ccgtggtgct gagcacgttg acaaggtcgg caagtctggc 720
actgctgagg agcagcagaa gctcaaggaa ctcttcggtc ttgggtgctct tgagcactac 780
gatgacttcg ctgctgtcct ccccaacggc ccctacctt ggcaggacaa cgacttcgcc 840
actggatact ccgagttctt ccagttctgc gatgccgtcg aggggtgtgga agccgggtgct 900
gctgtgacc cgggccccga ggggtgttgg cttagaagg ctctggccaa ctacgcctac 960
tggttcaact cgactctct tcccaactac tgcgcttct acggctactg gtcggatgag 1020
tggccggtt cctgcttcga ctctacaac gcctcctct ctctgttcac cgacacctcc 1080
gttgacaacg ccggtgaccg tcagtgaggaa tggttcctct gcaacgaacc tttcttctgg 1140
tggcaggatg gtgctcccga ggatgtcacc accattgtgc ctctctggt caacgcggaa 1200
tactggcagc gtcagtgtc tctgtacttc cccgaaacca acggctacac cttcgggttc 1260
gccaagaaca agactgtgc caccgtcaac gactggactg gtggctgggt cgaaacccgc 1320
aacaccacc gtctcatctg gaccaacggc cagtacgacc cctggcgcgga cagcgggtgc 1380
tcttcacct tccgtcctgg tggccagctc gtcagcactg ccaacgagcc tgtccagatc 1440
atccccggg gtttccactg ctcgatctg tacatggccg actactacgc caacgcgggt 1500
gtccgcaagg tcgtcgacaa cgaggttgct cagatcaaga agtgggttgc tgagtactac 1560
gccta 1565

<210> 14

<211> 1565

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<222> 1..1565

<223> /organism="Artificial Sequence"

/note="A.flavus proline-specific endoprotease (PEP) with A. niger pectinemethylesterase signal sequence and A.niger PEP prosequence"

"

/mol_type="unassigned DNA"

<400> 14
atgggtcaagt ccacacctggc ctccgtcttc ttccgcccga ctgctctggc tgctcgcccc 60
cgcttggttc ccaagcccgt ctctcgctcc gccagcagca agtcggctgc caccaccggt 120
gaagcctact tcgagcagct ccttgaccac cacgactctt ccaagggcac cttctcccag 180
cgttactggg ggagcactga gtactggggg ggtcctggaa gccctgttgt cctcttcaact 240
cccgggtgagg cctccgcccga tggctacgag ggatacctga ccaacaacac cctgaccggt 300
ctgtacgctc aggagatcca ggggtgccgtc atcttgattg agcaccgtta ctggggcgat 360
tcttctccct acgaggagct gaccgctgaa accctccagt acctcacctt ggagcagctc 420
atcttgatc tgaccactt cgctgagact gtccagcttg agttcgacac cagcaactcc 480
tccaacgccc ccaagggccc ctgggttctc gtcggtgga gctactctgg tgctcttggc 540
gcctggactg ctgctgttgc tcctggaacc ttctgggctt accacgccac ctgggctcct 600
gtgcaggcca ttgatgactt ctggcagtac ttcgaccca tccgtcacgg aatggctccc 660
aactgctctc gtgatgtctc cctcgtcgcc aaccacatcg acaccgtcgg caagaacggc 720
tctgctgccc accagcttgc tctgaaggag ctgttcggtc ttgaggctct cgaacactac 780
gatgacttgc ctgctgcctt tcccactggg ccctacctct ggcagtccea caccttctgc 840
accggtact ccaacttctt cgctttctgc gatgccgttg agaacgtcga ggctgggtgc 900
gctgtgggtc ctggccccga ggggtgttgg ctgcagaagg ccttgactgg ctacgccaac 960
tggttcaact cgaccatcat ccctggctac tgcgcttctt acggctactg gactgacaac 1020
cgcaccgttg cttgcttctga caccacaac cccagctctg ccatcttca cgcacacctc 1080
gtcgataacg ccgtcgaccg ccagtggcag tggttcctct gcaacgagcc cttcttctgg 1140
tggcaggatg gtgctcctga ggggtgttctt accattgtgc ctgcacat caacgccgaa 1200
tactggcagc gccagtgtc tctctacttc cccgaagtca acggctacac ctacggctcc 1260
gccaagggca agactgtctc cactgtcaac acctggactg gtggctggtc cgacagcaag 1320
aacaccagcc gtctcctctg ggtgaacggc cagtacgacc cctggcgtga ctccgggtgc 1380
tcctccacc accgccctgg tggctcctct accagcactg ccgatgagcc cgtgcaggtc 1440
atccccggtg gtttccactg ctccgacctc tacctcaagg actacttctc caacgccggt 1500
gtcaagcagg tcgtcgaaa cgccgttctc cagatcaagt cctgggttgc tgaatactac 1560
aaata 1565

<210> 15

<211> 1565

<212> DNA

<213> Artificial Sequence

<220>

<221> source
<222> 1..1565
<223> /organism="Artificial Sequence"
/note="BC2G076_ Aspergillus_aculeatus proline-specific
endoprotease (PEP) with A. niger pectinemethylesterase signal
sequence and A.niger PEP prosequence "
/mol_type="unassigned DNA"

<400> 15
atgggtcaagt ccattcttgc ctccgtcttc ttcgctgcca ctgctcttgc tgctcgtccc 60
cgtctcgtcc ccaagcccgt ctcccgcccc gccagcagca agtctgctgc caccaccggt 120
gaagcctact tcgagcagct gatcgaccac tccgaccctt ccaagggtagc tttctcccag 180
cgctactggt actctgctca gtactgggggt ggtcctggca gccctgtcgt cctcttccact 240
cctgggtgaag tgtctgccga tggctaccag ggctacctca ccaacgccac cctgaccggt 300
gtctacgctc agcagctcca gggtgccggt gtccctcgtcg agcaccgcta ctggggcgggc 360
agctctccct acaccaactt gactgctgag actctccagt acttgacttt ggagcagagc 420
gtgcttgacc tgacctactt cgctgagaac gtcaagctcg gtttcgacaa ctcgacctcc 480
tccaacgctc ctcacgtccc ctgggtgctg gtcggtggaa gctactctgg tgctctgacc 540
gcctggaccg aacacctggc tcctggcacc ttctgggctt accacgccac ctcggctccc 600
gtggagagca tctacgactt ctggcagtagc ttccgtccca tccaggatgg catggccaag 660
aactgctcca aggatgtctc gctagttgct gaacacgctg acaagatcgg caagaccggc 720
accaaggccc agcagaccga gctgaagaag ctcttcgggtc tggaagccct tgaacacttc 780
gatgacttcg ctgctgtcct tcccattggt ccctacctct ggcaggacaa caccttcgcc 840
actggatact ccgacttctt cgccttctgt gatgccgctg agaacgctga ggctgggtgct 900
gctgtcacc cgggtgctga ggggtgttggc cttgagaagg ccctcaccgg atacgccaac 960
tggttcaaga acgagatctt ccccggatac tgcgcttctt acggctactg gtcagatgag 1020
tactctgttg cctgctacga cacctacaac accacttctc ccctcttcac cgacacctcc 1080
gtcgacaacg ccggtgaccg ccagtggcag tggttcctgt gcaacgagcc cttcttctgg 1140
tggcaggatg gtgctccctc ctccgagact accattgtgc ctcgtctcgt ctctgcccac 1200
tactggcagc gccagtgcgc tctctacttc cccgaggctc acggatacac ctacggctct 1260
gccaagggca agtccgcaa cactttcaac gcctggactg atggctgggt catgaatggc 1320
aacagcacc gtctgatctg gaccaacggc cagtacgacc cttggcgtga tgccaccgctc 1380
tcctccacct tccgccccgg tggctcctctg gccagcactc cttccgagcc tgtgcagatc 1440
atccccggtg gattccactg ctcggatctc tacatctccg actccgctcgt caacgccggt 1500
gtcaagaagg ttgttgacaa cgaggttgct cagatcaagg cctggggtgc tgagttctat 1560
gcata 1565

<210> 16
<211> 1580
<212> DNA
<213> *Rasamsonia emersonii*

<220>
<221> source
<222> 1..1580
<223> /organism="Rasamsonia emersonii"
/mol_type="unassigned DNA"

<400> 16
atgccctccc tctcctcctt cgttgccctg actgcttctc ttgtctctct ggctgctgcc 60
gctgctcctc gtctccctct tctcctcctc cctcccttgc ctccccgtga ccccttgcac 120
ggacctacca acgcctccgc cactttccag cagctcatcg accacaacca ccccgagctt 180
ggcaccttct cccagcgccta ctgggtggaat gatgagttct ggaaggggtcc cggctctccc 240
gttgtccttt tcacccccgg tgaagaagat gccagcgggtt acgtgggcta cctgaagaac 300
accaccatca ccggtctgat cgctcagacc atcgggtggtg ccgtcatcgt cctcgaacac 360
cgctactggg gccagtcctc cccctacgac tctctgacca ccaagaacct gcagtacctg 420
accctcaagc agtccattgc cgacctcacc tacttgcgca agaccgtcaa gctccccctc 480
gaccgcaacg gcagctccaa cgccgacaag gctccctggg ttctcagcgg tggaagctac 540
tctgggtgctc tctccgctg gactgccagc acctcccccg gtactttctg ggcctaccac 600
gccagctctg ctctgttga ggccatctac gattactggc agtacttctc tcccgtgcag 660
gatggattgc ctgccaaactg ctcgaaggac ctctcccgtg tcgtcgacta catcgactcc 720
gttctccagt ccggcaacgc cactgccaaag caacagctca aggacctttt cggctctgggt 780
gctctggagc acgacgatga cttcgcctcc gctcttgaga acggcccttg gctctggcag 840
tcgaactcgt tctacgacct ctaccctcct gtctacgagt tctgcgacta cgttgagaac 900
gcctacgcca gccctcccgt tgctgctggt cccgatggtg ttggctctgga gaaggctctg 960
tctggctacg ccacctgggtg gaacaaggtc ttcttccccg gctactgcgc tacctacggc 1020
tactggctcct ccaacgactc cattgcctgc ttcgacacct acaaccagtc gtcgccccatg 1080
ttcaccgacc tttccgtctc caaactatc aaccgccagt ggaactgggt cctctgcaac 1140
gagcccttct tctactggca ggatggtgct cccaagaacg tccccagcat tgtctctcgt 1200
ctggtcactg ctgagtactg gcagcgcagc tgccccctgt tcttccctga agaggatggc 1260
tacacctacg gaagcgccaa gggcaagact gctgccgatg tcaacgcctg gaccaagggc 1320
tggttcttga ctaacaccac ccgtctgatc tggaccaacg gcgagcttga cccctggcgc 1380
tctgctggtg tcagcagcaa gttccgtccc ggtgggtcccc tccagtccac tccccaggct 1440
cctctgcagc tcattccccg ggggtgtccac tgctacgatc tgatcctcaa gaacgcccag 1500
gccaacgccc gtgtgcagcg tgttgtcacc aacgaggttg ctcagatcaa ggcctggggtg 1560

aacgaatact accgtaagta

1580

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Полипептид, обладающий активностью пролин-специфичной эндопротеазы, где указанный полипептид имеет менее чем 70% остаточной активности на ацетил-AlaAlaPro-паранитроанилине (Ac-AAP-PNA) в качестве субстрата после выдержки полипептида при температуре 65°C в течение 15 мин.

2. Полипептид, обладающий активностью пролин-специфичной эндопротеазы, необязательно по п. 1, отличающийся тем, что выбран из группы, состоящей из:

i. полипептида, который, когда выровнен относительно аминокислотной последовательности по SEQ ID NO: 1, содержит аминокислотную замену в положении, соответствующем положению 469, и кроме того, необязательно, по меньшей мере, одну аминокислотную замену в положении 204, 304, 377, 466, и/или 477, где положение определяется относительно SEQ ID NO: 1;

ii. полипептида, который, когда выровнен относительно аминокислотной последовательности по SEQ ID NO: 1, содержит аминокислоту, выбранную из группы, состоящей из Ala (A), Cys (C), Asp (D), Glu (E), Phe (F), Gly (G), His (H), Ile (I), Lys (K), Leu (L), Met (M), Asn (N), Gln (Q), Arg (R), Ser (S), Thr (T), Val (V), Trp (W), и Tyr (Y), в положении, соответствующем положению 469, и, необязательно, аминокислоту Phe (F) в положении 204, Ser (S), в положении 304, Ala (A) в положении 377, Thr (T) в положении 466 и/или Ala (A) в положении 477, где положение определяется относительно SEQ ID NO: 1;

iii. полипептида, содержащего аминокислотную последовательность, в соответствии с SEQ ID NO: 1, где SEQ ID NO: 1 содержит, по меньшей мере, одну аминокислотную замену, выбранную из группы, состоящей из P469A, P469C, P469D, P469E, P469F, P469G, P469H, P469I, P469K, P469L, P469M, P469N, P469Q, P469R, P469S, P469T, P469V, P469W и P469Y, и, возможно, аминокислотную замену I204F, P304S, P377A, P466T, и/или P477A, где аминокислотная замена определяется относительно SEQ ID NO: 1;

iv. полипептида по i)-iii), но утратившего сигнальную последовательность и/или последовательности пропротеина;

v. полипептида по i)-iv), где полипептид идентичен, по меньшей мере, на 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% или 99% аминокислотной последовательности по SEQ ID NO: 1; и,

vi. полипептида, кодируемого нуклеиновой кислотой, которая идентична, по меньшей мере, на 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, или 99% SEQ ID NO: 2, или зрелой кодирующей последовательности SEQ ID NO: 2, где

последовательность SEQ ID NO: 2 содержит, по меньшей мере, одну мутацию, кодирующую, по меньшей мере, одну аминокислотную замену, выбранную из группы P469A, P469C, P469D, P469E, P469F, P469G, P469H, P469I, P469K, P469L, P469M, P469N, P469Q, P469R, P469S, P469T, P469V, P469W и P469Y, и необязательно, где SEQ ID NO: 2 содержит, по меньшей мере, одну дополнительную мутацию, кодирующую аминокислотные замены I204F, P304S, P377A, P466T, и/или P477A, где аминокислотные замены определяются относительно SEQ ID NO: 1.

3. Полипептид, который представляет собой выделенный, по существу, чистый, чистый, рекомбинантный, синтетический или вариантный полипептид полипептида по п. 1 или 2.

4. Способ получения вариантного полипептида, который включает

i. отбор исходного полипептида, идентичного, по меньшей мере, на 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% аминокислотной последовательности по SEQ ID NO: 1, или зрелому полипептиду по SEQ ID NO: 1; и

ii. замены, по меньшей мере, одной аминокислоты в положении, соответствующем положению 469, на аминокислоту, выбранную из группы, состоящей из Ala (A), Cys (C), Asp (D), Glu (E), Phe (F), Gly (G), His (H), Ile (I), Lys (K), Leu (L), Met (M), Asn (N), Gln (Q), Arg (R), Ser (S), Thr (T), Val (V), Trp (W), и Tyr (Y), и, необязательно, замены аминокислоты в положении 204 на Phe (F), в положении 304 на Ser (S), в положении 377 на Ala (A), в положении 466 в Thr (T), и/или в положении 477 в Ala (A); и

iii. получения вариантного полипептида,

где, необязательно, вариантный полипептид содержит менее 70% остаточной активности на ацетил-AlaAlaPro-паранитроанилине (Ac-AAP-PNA) в качестве субстрата после того, как полипептид был выдержан при температуре 65°C в течение 15 мин.

5. Композиция, содержащая полипептид по любому из п.п. 1-3.

6. Композиция по п. 5, содержащая носитель, наполнитель или вспомогательный фермент.

7. Нуклеиновая кислота, кодирующая пролин-специфичную эндопротеазу, которая идентична, по меньшей мере, на 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% или 99% SEQ ID NO: 2, или зрелой кодирующей последовательности по SEQ ID NO: 2, где последовательность SEQ ID NO: 2 содержит, по меньшей мере, одну мутацию, кодирующую, по меньшей мере, одну аминокислотную замену, выбранную из группы, состоящей из P469A, P469C, P469D, P469E, P469F, P469G, P469H, P469I, P469K, P469L, P469M, P469N, P469Q, P469R, P469S, P469T, P469V, P469W, P469Y и, необязательно, в

котором последовательность SEQ ID NO: 2 содержит, по меньшей мере, одну мутацию, кодирующую аминокислотную замену I204F, P304S, P377A, P466T, и/или P477A, где аминокислотные замены определяются относительно SEQ ID NO: 1.

8. Нуклеиновая кислота, которая представляет собой выделенную, по существу чистую, чистую, рекомбинантную, синтетическую или вариантную нуклеиновую кислоту нуклеиновой кислоты по п. 7.

9. Экспрессирующий вектор, содержащий нуклеиновую кислоту по п. 7 или 8, функционально связанный, по меньшей мере, с одной регуляторной последовательностью, которая управляет экспрессией полипептида в клетке-хозяине.

10. Рекомбинантная клетка-хозяин, содержащая нуклеиновую кислоту по п. 7 или 8, или экспрессирующий вектор по п. 9.

11. Способ получения полипептида по п. 1 или 3, включающий культивирование клетки-хозяина по п. 10 в подходящей среде для ферментации, в условиях, обеспечивающих экспрессию полипептида, и получение полипептида и необязательно выделение полипептида.

12. Способ приготовления пищевого или кормового продукта, включающий взаимодействие промежуточной формы пищевого или кормового продукта с полипептидом по п. 1-3, или композиции по п. 5 или 6, и приготовление пищевого или кормового продукта.

13. Способ по п. 12, в котором пищевой продукт представляет собой напиток, предпочтительно пиво.

14. Способ по п. 12 или 13, отличающийся тем, что пищевой продукт содержит глютен.

15. Пищевой или кормовой продукт, получаемый способом по любому из п.п. 12-14.

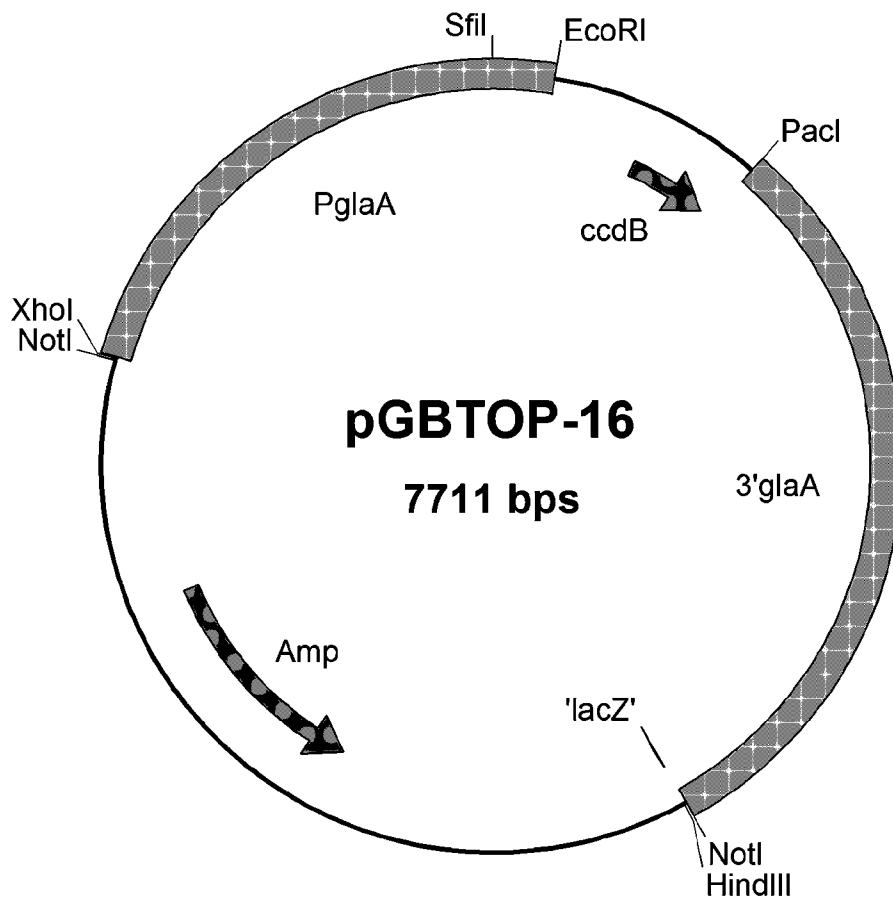


Fig. 1

```

0          10          20          30          40          50          60
.....|.....|.....|.....|.....|.....|.....|
Aspergillus_niger(FEP)  -----MVKSILASVFFAATALAARFRLVPKPVSRRPASSKSAATTGEAYFEQLLDHNNPEKGTFSQRYWWS
Aspergillus_carbonarius -----MVKSILASVFFAATALAARFRLVPKPVSRRPASSKSAATTGEAYFEQLVDHNNPEKGTFSQRYWWS
Aspergillus_flavus      -----MVKSILASVFFAATALAARFRLVPKPVSRRPASSKSAATTGEAYFEQLLDHDSKSGTFSQRYWWS
Aspergillus_aculeatus   -----MVKSILASVFFAATALAARFRLVPKPVSRRPASSKSAATTGEAYFEQLIDHSDPSKGTFSQRYWWS
Rasamsoni_emersonii     MPSLSSLVALTASLVSLAAAAAPRLPLPRFPPLFRDPLHGPTNASATFQQLIDHNNPELGTFSQRYWWS

```

```

70          80          90          100         110         120         130
.....|.....|.....|.....|.....|.....|.....|
Aspergillus_niger(FEP)  TEYWGGPGSPVVLFTPGEVSADGYEGYLNETLTGVYAQEIQQGAVVILIEHRYWGDSSPYEVLNAETLQYL
Aspergillus_carbonarius TEYWGGPGSPVVLFTPGEVSADGYEGYLNTDNLTLGVYAQEIQQGAVVILIEHRYWGDSSPYEVLNAETLQYL
Aspergillus_flavus      TEYWGGPGSPVVLFTPGEASADGYEGYLNTNLTGLYAQEIQQGAVVILIEHRYWGDSSPYEELTAEATLQYL
Aspergillus_aculeatus   AQYWGGPGSPVVLFTPGEVSADGYQGYLTNATLTGVYAQQLQGAVVILVEHRYWGDSSPYINLTAEATLQYL
Rasamsoni_emersonii     DEFWKGPSPVVLFTPGEEDASGYVGYLKNNTITGLIAQTIQQGAVVILVEHRYWGDSSPYDSLTTKLNQYL

```

```

140         150         160         170         180         190         200
.....|.....|.....|.....|.....|.....|.....|
Aspergillus_niger(FEP)  TLDQAILDMTYFAETVKLQFDNSTRSNAQNAPFVWVGGSYSGALTAWTESVAPGTFWAYHATSAPVEAIY
Aspergillus_carbonarius TLDQAVLDMTYFAETVKLQFDNSTRSNAQNAPFVWVGGSYSGALTAWVESVAPGTFWAYHATSAPVEAIY
Aspergillus_flavus      TLEQSIIDLTHFAETVQLQFDNSTRSNAQNAPFVWVGGSYSGALAAWTAAVPCTFWAYHATSAPVQAIY
Aspergillus_aculeatus   TLEQSVLDLTYFAENVKLFQFDNSTRSNAQNAPFVWVGGSYSGALTAWTEHLAPGTFWAYHATSAPVESIY
Rasamsoni_emersonii     TLKQSIADLTYFAETVKLQFDNSTRSNAQNAPFVWVGGSYSGALSAWTAASVPGTFWAYHATSAPVEAIY

```

```

210         220         230         240         250         260         270
.....|.....|.....|.....|.....|.....|.....|
Aspergillus_niger(FEP)  DYWQYFFPIQGGMAQNCSDKVSLVAEYVDKIGKNGTAKEQQALKELFGLGAVEHFDFAAVLPLNGPYLWQ
Aspergillus_carbonarius DFWQYFFPIQGGMAQNCSDKVSRVAEHVDKVGKSGTAEEQQKLELFGLEALEHYDDFAAVLPLNGPYLWQ
Aspergillus_flavus      DFWQYFDPRIRHGMAPNCSRVDVSLVANHIDTVGKNGSAADQLALKELFGLGAELEHYDDFAAALPTGPFYLWQ
Aspergillus_aculeatus   DFWQYFRPIQDGMANCSKDVSLVAEYVDKIGKGTGKAQQTELKLELFGLEALEHFDFAAVLPIGPFYLWQ
Rasamsoni_emersonii     DYWQYFAPVQDGLPANCSKDLRSRVVDYIDSVLQSGNATAKQQLKDLFGLGAELEHDDDFASALENGFWLWQ

```

```

280         290         300         310         320         330         340
.....|.....|.....|.....|.....|.....|.....|
Aspergillus_niger(FEP)  DNDFATGYSSFFQPCDAVEGVVEAGAAVTPGPEGVGLKALANYANWFNSTILPDCASYGYWTDDEWSVAC
Aspergillus_carbonarius DNDFATGYSEFFQPCDAVEGVVEAGAAVTPGPEGVGLKALANYAYWFNSTILLPNYCASGYWSDDEWSVAC
Aspergillus_flavus      SNTFVTGYSNFFAPCDAVENVEAGAAVTPGPEGVGLQKALTGYANWFNSTIIPGYCASGYWTDNRVTAC
Aspergillus_aculeatus   DNTFATGYSDFFAPCDAVENVEAGAAVTPGAEVGLKALTGYANWFKNIEIPGYCASGYWSDDEYSVAC
Rasamsoni_emersonii     SNSFYDYPYPPVVEPCDYVENAYASFPVAAGPDGVGLEKALSGYATWNNKVFPPGYCATYGYWSSNDSIAC

```

```

350         360         370         380         390         400         410
.....|.....|.....|.....|.....|.....|.....|
Aspergillus_niger(FEP)  FDSYNASSPIYTDTSVGNVAVDRQWQWFLCNEPPFFWQDGAPEGSTIVPRLVSASYWQRQCPLYPFETNG
Aspergillus_carbonarius FDSYNASSPLFTDTSVDNAVDRQWQWFLCNEPPFFWQDGAPEVDITIVPRLVNAEYWQRQCPLYPFETNG
Aspergillus_flavus      FDTNHPSSAIFTDTSVDNAVDRQWQWFLCNEPPFFWQDGAPEGVPTIVPRTINAEYWQRQCPLYPFETNG
Aspergillus_aculeatus   YDTYNTTSPFLTDTSDNAVDRQWQWFLCNEPPFFWQDGAPESETTIVPRLVSADYWQRQCPLYPFETNG
Rasamsoni_emersonii     FDTYNTTSPFLTDLVSNNTINRQWQWFLCNEPPFFWQDGAPEKVPVIVSRLVTAEYWQRQCPLYPFETNG

```

```

420         430         440         450         460         470         480
.....|.....|.....|.....|.....|.....|.....|
Aspergillus_niger(FEP)  YTYGSAGKNAATVNSWTGGWDMTRNTTRLIWINGQYDPWRDSGVSSSTRPFGGPLASTANEPVQIIPGGF
Aspergillus_carbonarius YTFGSANKNTAATVNDWTGGWFETRNTTRLIWINGQYDPWRDSGVSSSTRPFGGPLASTANEPVQIIPGGF
Aspergillus_flavus      YTYGSAGKTAATVNTWTGGWSDSKNTSRLLVVNGQYDPWRDSGVSSSTRPFGGPLASTANEPVQIIPGGF
Aspergillus_aculeatus   YTYGSAGKSAANTFNAWTDGWFMMNGNSTRLIWINGQYDPWRDATVSSSTRPFGGPLASTANEPVQIIPGGF
Rasamsoni_emersonii     YTYGSAGKTAADVNAWTKGWFLLT-NTTRLIWINGELDPWRSAGVSSKFRPFGGPLASTANEPVQIIPGGF

```

```

490         500         510         520
.....|.....|.....|.....|
Aspergillus_niger(FEP)  HCSLDLYMADYYANEGVKKVVDNEVKQIKWVVEEYVA-
Aspergillus_carbonarius HCSLDLYMADYYANAGVVRKVVVDNEVAQIKWVAEYVA-
Aspergillus_flavus      HCSLDLYLKDYPANAGVKKVVDNAVAQIKSWVAEYVK-
Aspergillus_aculeatus   HCSLDLYISDSVNVAGVKKVVDNEVAQIKAWVAEYVA-
Rasamsoni_emersonii     HCYDLILKNAEANAGVQRVVVTNEVAQIKAWVNEYYRK

```

Fig. 2