

(19)



Евразийское
патентное
ведомство

(21) 201791156 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2017.09.29

(51) Int. Cl. C12N 15/67 (2006.01)
C12N 15/81 (2006.01)

(22) Дата подачи заявки
2015.11.16

(54) ПОСЛЕДОВАТЕЛЬНОСТИ ОЛИГОНУКЛЕОТИДОВ ДЛЯ ПРИМЕНЕНИЯ В
КОНСТРУИРОВАНИИ ПУТЕЙ

(31) 14195026.1

(32) 2014.11.26

(33) EP

(86) PCT/EP2015/076715

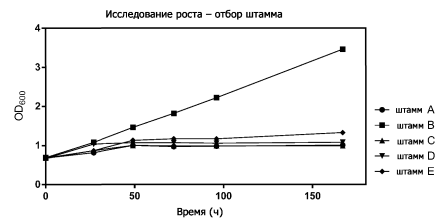
(87) WO 2016/083180 2016.06.02

(71) Заявитель:
КЛАРИАНТ ИНТЕРНЭШНЛ ЛТД.
(CN)

(72) Изобретатель:
Драгович Здравко, Райзингер
Кристоф, Диц Хайко (DE)

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

(57) Настоящее изобретение охватывает новые последовательности искусственных олигонуклеотидов, которые могут инициировать транскрипцию генов в различных условиях на высоком уровне. Кроме того, изобретение касается рекомбинантных фрагментов ДНК, включающих последовательности искусственных олигонуклеотидов, экспрессирующих плазмид, содержащих фрагменты рекомбинантной ДНК, и клеток-хозяев, трансформированных фрагментами рекомбинантной ДНК.



A1

201791156

201791156

A1

ПОСЛЕДОВАТЕЛЬНОСТИ ОЛИГОНУКЛЕОТИДОВ ДЛЯ ПРИМЕНЕНИЯ В КОНСТРУИРОВАНИИ ПУТЕЙ

Настоящее изобретение охватывает новые последовательности искусственных олигонуклеотидов, которые могут инициировать транскрипцию генов в различных условиях на высоком уровне. Кроме того, изобретение касается рекомбинантных фрагментов ДНК, включающих последовательности искусственных олигонуклеотидов, экспрессирующих плазмид, содержащих рекомбинантные фрагменты ДНК, и клеток-хозяев, трансформированных рекомбинантными фрагментами ДНК.

Дрожжи *Saccharomyces* и в более узком смысле вида *S. cerevisiae* применяются уже тысячи лет для производства хлеба и алкогольных напитков типа sake, вина или пива. За весь этот длительный период промышленного применения дрожжи приспособились к технологическим условиям и могут переносить механические воздействия в биореакторах, ингибирующие вещества и продукты ферментации. Кроме того, они устойчивы к колебаниям температуры и могут сбрасывать сахара при низких значениях pH, что сводит к минимуму риск загрязнения. Помимо этого, *S. cerevisiae* является ключевой лабораторной модельной системой и легко поддаются генетическим модификациям и вообще считаются безопасными – имеют статус GRAS. Имеется широкий набор генетических инструментов для *S. cerevisiae* и хорошо изучены многие внутриклеточные процессы, как-то метаболизм, секреция, транспорт, сигнализация и другие пути, что способствует успешному конструированию дрожжей для разнообразного применения.

Введение мультиферментных путей требует особенно точного контроля за уровнем экспрессии генов, в особенности ключевых ферментов, которые могут быть гетерологичными или собственными, чтобы довести до максимума утилизацию субстрата и/или образование продуктов. При этом транскрипционный контроль имеет место на уровне последовательности олигонуклеотида – промотора, находящегося в вышерасположенном участке гена. Таким образом, сила и регуляция промотора являются критическими точками для метаболической инженерии.

Поскольку эндогенные промоторы обычно не полностью осуществляют необходимый непрерывный транскрипционный контроль и поэтому не доводят до максимума уровни транскрипции, достижимые внутри клетки, то решающим шагом при инженерии дрожжей является правильный выбор промотора.

В данной области известны различные типы промоторов.

Индукцибельные или дерепрессируемые промоторы обеспечивают высокий уровень транскрипционного контроля, но они зависят от индуктора или заданных технологических

условий. Регулируемые промоторы лимитируются выработкой токсических белков или построением путей с токсическими промежуточными продуктами. Кроме того, известно, что большинство промоторов существующего уровня техники позволяют контролировать транскрипцию лишь одной конкретной группы генов.

Таким образом, промоторы, известные специалистам в данной области, всегда имеют определенные недостатки, которые серьезно ограничивают их использование лишь несколькими конкретными применениями.

Поэтому авторы настоящего изобретения поставили перед собой задачу разработать новые и улучшенные промоторы, обеспечивающие высокий уровень транскрипционного контроля для разнообразных генов, которые хорошо подходят для промышленного применения.

Авторы настоящего изобретения неожиданно обнаружили, что эта задача может быть решена при помощи:

последовательности олигонуклеотида, характеризующегося тем, что он повышает скорость транскрипции РНК, типа фрагментов матричной РНК, кодирующих белки из группы, состоящей из ферментов, структурных белков, коферментов, переносчиков, антител, гормонов и регуляторов, типа фрагментов регуляторной РНК, типа фрагментов ферментативно активной РНК или типа фрагментов транспортной РНК, причем данная последовательность олигонуклеотида по меньшей мере на 80% идентична последовательности SEQ ID NO: 1.

Номенклатура аминокислот, пептидов, нуклеотидов и нуклеиновых кислот в настоящей заявке соответствует рекомендациям IUPAC. Как правило, в этом документе аминокислоты именованы в соответствии с однобуквенным кодом.

Термин “олигонуклеотид” в настоящем изобретении следует понимать как одноцепочечную или двухцепочечную молекулу ДНК или РНК, содержащую от 2 до 1000 нуклеотидов, предпочтительно от 10 до 900 нуклеотидов, более предпочтительно от 50 до 850 нуклеотидов и наиболее предпочтительно от 100 до 820 нуклеотидов.

Термины “ДНК” и “РНК” хорошо известны специалистам в данной области. ДНК содержит дезоксирибозу, а РНК содержит рибозу (у дезоксирибозы нет гидроксильной группы, присоединенной к пентозному кольцу в положении 2'). Комплементарным основанием для аденина является не тимин, как в ДНК, а урацил, который является метилированной формой тимина.

Олигонуклеотиды по настоящему изобретению имеют нуклеотидную последовательность, которая по меньшей мере на 80%, предпочтительно по меньшей мере на 82%, более предпочтительно по меньшей мере на 85%, особенно предпочтительно по

меньшей мере на 90%, еще более предпочтительно по меньшей мере на 92%, еще предпочтительней по меньшей мере на 95%, еще более предпочтительно по меньшей мере на 98% и наиболее предпочтительно по меньшей мере на 99% идентична последовательности SEQ ID NO: 1.

В особенно предпочтительном воплощении нуклеотидную последовательность выбирают из группы, состоящей из SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6 и SEQ ID NO: 7.

Олигонуклеотиды по настоящему изобретению повышают скорость транскрипции некоторых фрагментов РНК. При этом термин “повышение скорости транскрипции” следует понимать как повышение по сравнению с олигонуклеотидом с активностью промотора из существующего уровня техники. “Повышение скорости транскрипции” обычно определяется следующим образом:

$$\frac{RTL_I}{RTL_S}$$

где: RTL_I – относительный уровень транскрипта в репортерной системе, контролируемой олигонуклеотидом по изобретению,

RTL_S – относительный уровень транскрипта в репортерной системе, контролируемой олигонуклеотидом из существующего уровня техники.

При этом относительный уровень транскрипта измеряется как концентрация РНК в клеточном экстракте репортерной системы относительно концентрации РНК гена домашнего хозяйства в том же клеточном экстракте.

При этом RTL_I и RTL_S определяются с использованием одного и того же типа клеток-хозяев, где клетки-хозяева трансформируют по меньшей мере одним рекомбинантным фрагментом ДНК, включающим соответствующий олигонуклеотид, и клетки-хозяева культивируют в идентичных условиях существующего уровня техники, причем клетки-хозяева собирают в экспоненциальной фазе роста.

В предпочтительном воплощении скорость транскрипции повышается по меньшей мере в 2 раза при выращивании дрожжевых клеток-хозяев, предпочтительно *S. cerevisiae*, трансформированных по меньшей мере одним рекомбинантным фрагментом ДНК, включающим олигонуклеотид по настоящему изобретению, как минимум на двух субстратах, выбранных из группы, состоящей из глюкозы, маннозы, фруктозы, галактозы, ксилозы, арабинозы, сахарозы, трегалозы, рафинозы, глицерина, этанола, ацетата и лактата. Повышение определяется следующим образом:

$$\frac{RTL_{Te}}{RTL_{Se}}$$

где: RTL_{Le} – относительный уровень транскрипта матричной РНК, кодирующей SEQ ID NO: 17, контролируемой олигонуклеотидом SEQ ID NO: 1,

RTL_{Se} – относительный уровень транскрипта матричной РНК, кодирующей SEQ ID NO: 17, контролируемой олигонуклеотидом SEQ ID NO: 9.

При этом относительный уровень транскрипта измеряется как концентрация матричной РНК, кодирующей SEQ ID NO: 17, в экстракте дрожжевых (*S. cerevisiae*) клеток относительно концентрации матричной РНК гена домашнего хозяйства, кодирующего актин в том же экстракте дрожжевых клеток.

При этом RTL_{Le} и RTL_{Se} определяются с использованием одного и того же типа дрожжевых клеток-хозяев (*S. cerevisiae*), где дрожжевые клетки-хозяева трансформируют по меньшей мере одним рекомбинантным фрагментом ДНК, включающим соответствующий олигонуклеотид, и дрожжевые и клетки-хозяева культивируют в идентичных условиях существующего уровня техники, причем дрожжевые клетки-хозяева собирают в экспоненциальной фазе роста.

В одном особенно предпочтительном воплощении настоящего изобретения скорость транскрипции гена в дрожжевых клетках-хозяевах, трансформированных по меньшей мере одним рекомбинантным фрагментом ДНК, включающим олигонуклеотид по настоящему изобретению, повышается по меньшей мере в 2 раза, более предпочтительно по меньшей мере в 4 раза, особенно предпочтительно по меньшей мере в 6 раз и наиболее предпочтительно по меньшей мере в 10 раз при выращивании дрожжевых клеток-хозяев как минимум на двух субстратах, выбранных из группы, состоящей из глюкозы, маннозы, фруктозы, галактозы, ксилозы, арабинозы, сахарозы, трегалозы, рафинозы, глицерина, этанола, ацетата и лактата.

В другом особенно предпочтительном воплощении настоящего изобретения скорость транскрипции гена в клетках-хозяевах повышается по меньшей мере в 2 раза, предпочтительно по меньшей мере в 4 раза, более предпочтительно по меньшей мере в 6 раз, особенно предпочтительно по меньшей мере в 8 раз и наиболее предпочтительно по меньшей мере в 10 раз при выращивании клеток-хозяев как минимум на двух субстратах, выбранных из группы, состоящей из глюкозы, маннозы, фруктозы, ксилозы, сахарозы, глицерина и этанола.

В еще одном особенно предпочтительном воплощении настоящего изобретения скорость транскрипции гена в клетках-хозяевах повышается по меньшей мере в 2 раза, предпочтительно по меньшей мере в 4 раза, более предпочтительно по меньшей мере в 6 раз, особенно предпочтительно по меньшей мере в 8 раз и наиболее предпочтительно по меньшей мере в 10 раз при выращивании клеток-хозяев как минимум на двух субстратах,

выбранных из группы, состоящей из глюкозы, маннозы, глицерина, этанола и ксилозы.

Другим преимуществом олигонуклеотидов по настоящему изобретению является то, что они повышают ферментативную активность ферментов, кодируемых РНК, контролируемой этими олигонуклеотидами. При этом термин “степень повышения активности фермента” следует понимать как повышение по сравнению с олигонуклеотидом с активностью промотора из существующего уровня техники. “Степень повышения активности фермента” обычно определяется следующим образом:

$$\frac{EA_I}{EA_S}$$

где: EA_I – активность фермента в репортерной системе, контролируемой олигонуклеотидом по изобретению,

EA_S – активность фермента в репортерной системе, контролируемой олигонуклеотидом из существующего уровня техники.

При этом активность фермента измеряется в виде количества субстрата, преобразованного за 1 минуту при заданном количестве клеточного экстракта, за вычетом фоновой активности репортерной системы.

При этом EA_I и EA_S определяются с использованием одного и того же типа клеток-хозяев, где клетки-хозяева трансформируют по меньшей мере одним рекомбинантным фрагментом ДНК, включающим соответствующий олигонуклеотид, и клетки-хозяева культивируют в идентичных условиях существующего уровня техники, причем клетки-хозяева собирают в экспоненциальной фазе роста.

В предпочтительном воплощении активность фермента повышается по меньшей мере в 2 раза при выращивании дрожжевых клеток-хозяев, предпочтительно *S. cerevisiae*, трансформированных по меньшей мере одним рекомбинантным фрагментом ДНК, включающим олигонуклеотид по настоящему изобретению, как минимум на двух субстратах, выбранных из группы, состоящей из глюкозы, маннозы, фруктозы, галактозы, ксилозы, арабинозы, сахарозы, трегалозы, рафинозы, глицерина, этанола, ацетата и лактата.

Повышение определяется следующим образом:

$$\frac{EA_{Ie}}{EA_{Se}}$$

где: EA_{Ie} – активность фермента у белка SEQ ID NO: 17, контролируемого олигонуклеотидом SEQ ID NO: 1, или его варианта с последовательностью, по меньшей мере на 80% идентичной SEQ ID NO: 1,

EA_{Se} – активность фермента у белка SEQ ID NO: 17, контролируемого

олигонуклеотидом SEQ ID NO: 9.

При этом активность фермента измеряется в виде количества ксилозы, преобразованного за 1 минуту при заданном количестве клеточного экстракта, за вычетом фоновой активности репортерной системы.

При этом EA_{1e} и EA_{Se} определяются с использованием одного и того же типа клеток-хозяев (*S. cerevisiae*), где клетки-хозяева трансформируют по меньшей мере одним рекомбинантным фрагментом ДНК, включающим соответствующий олигонуклеотид, и клетки-хозяева культивируют в идентичных условиях существующего уровня техники, причем клетки-хозяева собирают в экспоненциальной фазе роста.

В одном особенно предпочтительном воплощении настоящего изобретения активность фермента в дрожжевых клетках-хозяевах, трансформированных по меньшей мере одним рекомбинантным фрагментом ДНК, включающим олигонуклеотид по настоящему изобретению, повышается по меньшей мере в 2 раза, предпочтительно по меньшей мере в 4 раза, более предпочтительно по меньшей мере в 6 раз, особенно предпочтительно по меньшей мере в 8 раз и наиболее предпочтительно по меньшей мере в 10 раз при выращивании дрожжевых клеток-хозяев как минимум на двух субстратах, выбранных из группы, состоящей из глюкозы, маннозы, фруктозы, галактозы, ксилозы, арабинозы, сахарозы, трегалозы, рафинозы, глицерина, этанола, ацетата и лактата.

В другом особенно предпочтительном воплощении настоящего изобретения активность фермента в клетках-хозяевах повышается по меньшей мере в 2 раза, предпочтительно по меньшей мере в 4 раза, более предпочтительно по меньшей мере в 6 раз, особенно предпочтительно по меньшей мере в 8 раз и наиболее предпочтительно по меньшей мере в 10 раз при выращивании клеток-хозяев как минимум на двух субстратах, выбранных из группы, состоящей из глюкозы, маннозы, фруктозы, ксилозы, сахарозы, глицерина и этанола.

В еще одном особенно предпочтительном воплощении настоящего изобретения активность фермента в клетках-хозяевах повышается по меньшей мере в 2 раза, предпочтительно по меньшей мере в 4 раза, более предпочтительно по меньшей мере в 6 раз, особенно предпочтительно по меньшей мере в 8 раз и наиболее предпочтительно по меньшей мере в 10 раз при выращивании клеток-хозяев как минимум на двух субстратах, выбранных из группы, состоящей из глюкозы, маннозы, глицерина, этанола и ксилозы.

В настоящем изобретении термин “фрагмент регуляторной РНК” (фрагмент рРНК) следует понимать как цепочку РНК, которая обладает способностью подавлять экспрессию гена будучи комплементарной к какой-то части мРНК или ДНК гена. Примером “фрагментов рРНК” является микроРНК (миРНК), которая действует

посредством РНК-интерференции (RNAi), причем эффекторный комплекс миРНК с ферментами может расщеплять комплементарную мРНК, блокировать трансляцию мРНК или ускорять её деградацию. Сама мРНК тоже может содержать регуляторные элементы типа рибопереключателей в 5'-нетранслируемой области или 3'-нетранслируемой области; эти цис-регуляторные элементы регулируют активность этой мРНК. Нетранслируемые участки также могут содержать элементы, регулирующие другие гены.

В настоящем изобретении термин “фрагмент ферментативно активной РНК” следует понимать как РНК, которая входит в состав белкового комплекса, способного катализировать ферментативные реакции в клетке, типа рибосомной РНК или такой РНК, которая сама образует каталитически активный комплекс типа рибозима (фермента из рибонуклеиновой кислоты).

В настоящем изобретении термин “фрагмент транспортной РНК” (фрагмент тРНК) следует понимать как небольшую цепочку РНК примерно из 80 нуклеотидов, которая обладает способностью переносить определенную аминокислоту на растущую полипептидную цепь в рибосомальном сайте синтеза белка во время трансляции. Она содержит сайты для прикрепления аминокислот и антикодонный участок для распознавания кодона, который связывается с определенной последовательностью на цепочке матричной РНК посредством водородных связей.

В настоящем изобретении термин “фрагмент матричной РНК” (фрагмент мРНК) следует понимать как небольшую цепочку РНК, которая обладает способностью переносить информацию о последовательности белка на рибосомы. Каждые три нуклеотида (кодон) соответствуют одной аминокислоте. В эукариотических клетках, как только предшественница мРНК (пре-мРНК) транскрибируется из ДНК, она преобразуется в зрелую мРНК. При этом из нее удаляются интроны – некодирующие участки пре-мРНК. Затем мРНК экспортируется из ядра в цитоплазму, где она связывается с рибосомами и подвергается трансляции в соответствующую белковую форму с помощью тРНК. В прокариотических клетках, которые не имеют ядра и цитоплазматических компартментов, мРНК может связываться с рибосомами в то время, когда она транскрибируется с ДНК. По прошествии определенного времени матричная РНК подвергается деградации на ее составные нуклеотиды под действием рибонуклеаз.

В настоящем изобретении термин “структурные белки” относится к таким белкам, которые придают неподвижность и жесткость биологическим компонентам, находящимся в жидком состоянии. Предпочтительные структурные белки выбирают из группы, состоящей из таких волокнистых белков, как коллаген, эластин и кератин; и таких глобулярных белков, как актин и тубулин. Другие белки, которые выполняют

структурные функции и понимаются в качестве “структурных белков” в настоящем изобретении, – это моторные белки, как-то миозин, кинезин и динеин, которые способны генерировать механические силы.

Предпочтительные фрагменты РНК, кодирующие структурные белки, выбирают из группы, состоящей из актина, эластина, филамина, коллагена, миозина, ламина.

Предпочтительные фрагменты РНК, кодирующие коферменты, выбирают из числа фрагментов РНК, кодирующих полипептиды, которые подвергаются посттрансляционной модификации. Примерами являются триптофан-триптофилхинон (ТТQ) и 4-метилиден-имидазол-5-он (МИО).

Предпочтительные фрагменты РНК, кодирующие переносчики, выбирают из числа фрагментов РНК, кодирующих переносчики типа унипорта, симпорта и антипорта, протонные насосы, ионные каналы и аквапорины.

Предпочтительные фрагменты РНК, кодирующие антитела, выбирают из числа фрагментов РНК, кодирующих IgA, IgD, IgE, IgG, IgM, IgY и IgW.

Предпочтительные фрагменты РНК, кодирующие гормоны, выбирают из числа фрагментов РНК, кодирующих небольшие пептидные гормоны типа TRH и вазопрессина; инсулин; гормон роста; гликопротеидные гормоны, такие как лютеинизирующий гормон, фолликулостимулирующий гормон и тиреотропный гормон.

Предпочтительные фрагменты РНК, кодирующие регуляторы, выбирают из числа фрагментов РНК, кодирующих рецепторы, факторы транскрипции, метаболические рецепторы, зрительные рецепторы, электрорецепторы, механические рецепторы и преобразователи сигналов.

Предпочтительные фрагменты РНК, кодирующие ферменты, выбирают из числа фрагментов РНК, кодирующих ферменты, модифицирующие углеводы. В настоящем изобретении термин “модифицирующие углеводы ферменты” следует понимать как охватывающий любые ферменты, способные модифицировать углеводы любого типа, как-то (но без ограничения) ферменты, расщепляющие углеводы, окисляющие углеводы, восстанавливающие углеводы, декарбоксилирующие углеводы, деацетилирующие углеводы, ацетилирующие углеводы, метилирующие углеводы, деметилирующие углеводы, аминирующие углеводы, фосфорилирующие углеводы, дефосфорилирующие углеводы, изомеризующие углеводы, эпимеризующие углеводы и дезаминирующие углеводы.

В одном особенно предпочтительном воплощении настоящего изобретения модифицирующие углеводы ферменты выбирают из группы, состоящей из классов ЕС 5.1.3, ЕС 5.3.1, ЕС 2.7.1, ЕС 2.2.1, ЕС 2.2.1 и ЕС 1.1.1, предпочтительно выбирают из

группы, состоящей из ЕС 5.1.3.3, ЕС 5.3.1.5, ЕС 2.7.1.17, ЕС 2.2.1.2, ЕС 2.2.1.1 и ЕС 1.1.1.1. В другом особенно предпочтительном воплощении белок выбирают из группы, состоящей из SEQ ID NO: 11–53.

В следующем предпочтительном воплощении настоящего изобретения у олигонуклеотидов настоящего изобретения от 1 до 80 нуклеотидов являются “мутантными”. В настоящем изобретении термин “мутантный” следует понимать как содержащий “замену”, “делецию” или “вставку”. Термин “мутация” следует понимать как “замену”, “делецию” или “вставку”. Замены классифицируются как транзиции, при которых пурин заменяется на пурин (A <-> G) или пиримидин на пиримидин (C <-> T), либо трансверсии, при которых пурин заменяется на пиримидин и наоборот (C/T <-> A/G). При вставке в олигонуклеотид добавляется один или несколько дополнительных нуклеотидов (A, C, T или G). Удаление одного или нескольких нуклеотидов из ДНК называется делецией.

В следующем воплощении настоящего изобретения предусмотрены рекомбинантные фрагменты ДНК, включающие олигонуклеотиды по настоящему изобретению.

Особенно предпочтительные рекомбинантные фрагменты ДНК по настоящему изобретению включают олигонуклеотиды, входящие в группу, состоящую из SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6 и SEQ ID NO: 7, и фрагменты РНК, кодирующие белки, входящие в группу, состоящую из ферментов, структурных белков, коферментов, переносчиков, антител, гормонов и регуляторов. Также особенно предпочтительно, чтобы белок являлся ферментом, а фермент входил в группу, состоящую из классов ЕС 5.1.3, ЕС 5.3.1, ЕС 2.7.1, ЕС 2.2.1, ЕС 2.2.1 и ЕС 1.1.1, предпочтительно в группу, состоящую из ЕС 5.1.3.3, ЕС 5.3.1.5, ЕС 2.7.1.17, ЕС 2.2.1.2, ЕС 2.2.1.1 и ЕС 1.1.1.1. Другие особенно предпочтительные рекомбинантные фрагменты ДНК включают олигонуклеотиды, входящие в группу, состоящую из SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6 и SEQ ID NO: 7, и фрагменты РНК, входящие в группу, состоящую из SEQ ID NO: 11–53.

В следующем воплощении настоящего изобретения предусмотрены экспрессирующие плазмиды, содержащие по меньшей мере один рекомбинантный фрагмент ДНК по настоящему изобретению.

Настоящим изобретением также предусмотрены клетки-хозяева, трансформированные по меньшей мере одним рекомбинантным фрагментом ДНК, включающим олигонуклеотид по настоящему изобретению. Клетки-хозяева по настоящему изобретению предпочтительно применяются для конструирования путей или

для метаболического превращения содержащих ксилозу субстратов в предпочтительные метаболиты.

Рекомбинантные клетки-хозяева по настоящему изобретению предпочтительно выбирают из бактерий, дрожжевых или грибковых клеток. В особенно предпочтительном воплощении клетки-хозяева выбирают из группы, состоящей из *Escherichia*, *Klebsiella*, *Pseudomonas*, *Lactobacillus*, *Bacillus*, *Streptomyces*; *Saccharomyces*, *Kluveromyces*, *Schizosaccharomyces*, *Candida*, *Yarrowia*, *Komagataella*, *Pichia*, *Hansenula*, *Penicillium*, *Trichoderma*, *Hypocrea*, *Aspergillus*, *Cantharellus*, *Agricus*, *Boletus*, *Pleurotus*, *Trametes*, *Phanerochaete*, *Myceliophthora*, *Chaetomium*, *Humicola*, *Chrysosporium*, *Talaromyces* и *Neurospora*.

Особенно предпочтительно клетки-хозяева выбирают из группы, состоящей из *Lactococcus lactis*, *Lactobacillus brevis*, *Bacillus subtilis*, *Bacillus megaterium*, *Bacillus lentus*, *Bacillus amyloliquefaciens*, *Bacillus licheniformis*, *Pseudomonas fluorescens*, *Klebsiella planticola*, *Escherichia coli*, *Streptomyces lividans*, *Saccharomyces cerevisiae*, *Saccharomyces bayanus*, *Saccharomyces uravum*, *Saccharomyces pastorianus*, *Saccharomyces kudriavzevii*, *Saccharomyces mikatae*, *Saccharomyces carlsbergensis*, *Schizosaccharomyces pombe*, *Kluveromyces marxianus*, *Yarrowia lipolytica*, *Hansenula polymorpha*, *Pichia angusta*, *Komagataella pastoris*, *Pichia pastoris*, *Aspergillus niger*, *Aspergillus oryzae*, *Trichoderma reesei* и *Myceliophthora thermophila*.

Рекомбинантные клетки-хозяева по настоящему изобретению могут содержать одну или несколько плазмид по настоящему изобретению.

Примеры и фигуры

Далее настоящее изобретение будет описано на примерах и фигурах. Примеры и фигуры приводятся только в иллюстративных целях и не ограничивают объем настоящего изобретения и формулы изобретения никоим образом.

Пример 1. Исследование роста – отбор штамма для скрининга

Плазмидой, содержащей типичный олигонуклеотид для тестирования, трансформировали 5 различных штаммов клеток-хозяев.

Таблица 1. Перечень тестируемых клеток-хозяев

Порядковый номер	Название	Происхождение
Штамм А	Ethanol Red [®] (серия 897/1, дата продукции: 01.05.08)	Fermentis, Франция
Штамм В	Simi White [™] (партия: 02905340230601V)	Lallemand, Канада
Штамм С	Rhône 2226 [™] (партия: 025556551030Y9)	

Штамм D	CBS 7764	Centraalbureau voor Schimmelcultures (Институт королевской академии искусств и наук Нидерландов - KNAW), Нидерланды
Штамм E	CBS 6413	

Плазмиду конструировали путем рекомбинационного клонирования в *S. cerevisiae*. Дрожжевые клетки трансформировали продуктами ПЦР, перекрывающимися друг с другом на 45 п.н. Фрагментами служили дрожжевой маркер (pUG6 с 87 по 1559 п.н.), маркер *E. coli* и *ori* (pUG19 с 754 по 2534 п.н.), дрожжевой *ori* (хромосома IV *S. cerevisiae* S288C с 44978 по 449831 п.н. и хромосома II *S. cerevisiae* S288C с 63156 по 63454 п.н.) и функциональная часть (SEQ ID NO: 10, SEQ ID NO: 54, хромосома XI *S. cerevisiae* S288C с 326407 по 326108 п.н.). При этом данные части были фланкированы рестрикционными сайтами *SapI*, *SbfI*, *StuI* и *NotI*, соответственно.

Дрожжевые штаммы подвергали трансформации повторно выделенной плазмидой высокоэффективным методом LiAc согласно Gietz и Schiestl.

Затем клетки-хозяева культивировали в 50 мл содержащего ксилозу субстрата (10 г/л дрожжевого экстракта, 20 г/л пептона, 20 г/л ксилозы + 200 мг/л G418) в аэробных условиях в качалочной колбе на 300 мл при 30°C и 250 об/мин. Результаты представлены на фиг. 1.

Для дальнейшего тестирования был выбран штамм В вследствие отличных показателей роста.

Пример 2. Исследование роста – сравнение различных плазмид, содержащих различные олигонуклеотиды

Штамм В трансформировали 10 плазмидами. Плазмиды конструировали таким же образом, как описано в примере 1, и они содержали олигонуклеотиды SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9 и SEQ ID NO: 10, соответственно.

Штамм В трансформировали плазмидами, как описано в примере 1, и культивировали в таких же условиях. Результаты представлены на фиг. 2.

Клетки-хозяева, трансформированные плазмидами, содержащими олигонуклеотиды по настоящему изобретению (т.е. регулируемые SEQ ID NO: 1–7), проявляли значительно более высокие показатели роста, чем клетки-хозяева, трансформированные плазмидами, содержащими олигонуклеотиды из существующего уровня техники SEQ ID NO: 9, SEQ ID NO: 10 и SEQ ID NO: 8.

Пример 3. Уровень транскрипта – сравнение различных плазмид, содержащих различные олигонуклеотиды

Дрожжевые штаммы, содержащие различные плазмиды, описанные в примере 2,

культивировали в 100 мл субстрата, содержащего глюкозу, маннозу, этанол, глицерин или ксилозу (10 г/л дрожжевого экстракта, 20 г/л пептона, 20 г/л источника углерода + 200 мг/л G418) в качалочной колбе на 500 мл при 30°C и 250 об/мин. Собирали по 5 мл этих культур примерно при $OD_{600} = 2$, то есть центрифугировали, промывали водой и опять центрифугировали (два раза). После этого осадки культур хранили при -80°C.

Из клеток экстрагировали тотальную РНК с помощью набора RNeasy Mini Kit™, Qiagen Germany, в соответствии с инструкцией производителя. Затем 500 нг этой РНК подвергали транскрипции в кДНК с помощью набора ThermoScript™ RT-PCR Kit, Life Technologies USA, в соответствии с инструкцией производителя. Проводили амплификацию продуктов ПЦР длиной в 225 и 236 п.н., используя iQ™ SYBR® Green Supermix и iQ™ iCycler, BioRad Germany, и рассчитывали концентрации мРНК ACT1 и XylA, следуя информации производителя.

Относительные уровни транскрипта (концентрация РНК XylA, деленная на концентрацию РНК ACT1) представлены на фиг. 3.

Таблица 2. Степень изменения уровня транскрипта XylA (разы) под контролем различных олигонуклеотидов по сравнению с уровнем транскрипта XylA под контролем SEQ ID NO: 9

	SEQ ID NO: 1	SEQ ID NO: 2	SEQ ID NO: 3	SEQ ID NO: 4	SEQ ID NO: 5	SEQ ID NO: 6	SEQ ID NO: 7
Глюкоза	23,4 ± 5,7	23,9 ± 5,9	22,6 ± 5,8	24,4 ± 5,3	28,9 ± 5,1	26,4 ± 6,6	22,3 ± 4,5
Манноза	25,3 ± 6,5	27,6 ± 6,6	25,9 ± 7,0	23,3 ± 6,6	22,1 ± 4,0	22,4 ± 6,7	26,2 ± 7,3
Этанол	4,6 ± 0,8	5,6 ± 1,1	3,9 ± 0,7	3,9 ± 0,8	4,5 ± 0,6	4,7 ± 0,6	4,9 ± 1,1
Глицерин	8,3 ± 2,2	7,6 ± 3,4	7,6 ± 2,7	9,3 ± 2,7	6,6 ± 1,3	6,8 ± 2,3	7,0 ± 2,3
Ксилоза	6,8 ± 3,3	6,6 ± 5,3	5,8 ± 3,2	8,4 ± 4,1	7,1 ± 4,6	7,7 ± 4,5	7,9 ± 4,4

Репортерная система под контролем олигонуклеотидов по настоящему изобретению проявляла повышение уровня транскрипта в 4–29 раз. Уровни транскрипта варьируют при различных условиях роста, но существенных отличий в отношении контроля между олигонуклеотидами SEQ ID NO: 1 - SEQ ID NO: 7 не отмечалось.

Пример 4. Активность фермента – сравнение различных плазмид, содержащих различные олигонуклеотиды

Собирали по 50 мл культур, как указано в примере 3, примерно при $OD_{600} = 2$. После этого осадки культур хранили при -80°C. Кроме того, таким же образом обрабатывали культуру штамма В, несущего плазмиду, описанную в примере 2, но без функциональной части (пустая плаزمид).

Оттаявшие осадки суспендировали в 400 мкл буфера (100 мМ трис, рН 7,5, 10 мМ $MgCl_2$) и гомогенизировали. После лизиса клеток неочищенные экстракты разбавляли до общей концентрации белка в 1 мкг/мкл (измеряли методом Брэдфорда). Определение активности ксилозоизомеразы проводили в 100 мкл с 10% разбавленных неочищенных

экстрактов, 0,25 мМ NADH, 3 ед./мл сорбитолдегидрогеназы и 500 мМ ксилозы. Кинетику реакции отслеживали фотометрически при 340 нм.

Измеренные активности фермента (за вычетом фоновой активности – пустая плазида) представлены на фиг. 4.

Таблица 3. Степень изменения активности ксилозоизомеразы (разы) под контролем различных олигонуклеотидов по сравнению с активностью ксилозоизомеразы под контролем SEQ ID NO: 9

	SEQ ID NO: 1	SEQ ID NO: 2	SEQ ID NO: 3	SEQ ID NO: 4	SEQ ID NO: 5	SEQ ID NO: 6	SEQ ID NO: 7
Глюкоза	22,5 ± 3,6	23,7 ± 3,4	24,8 ± 3,5	24,9 ± 4,0	23,9 ± 3,9	22,8 ± 4,1	20,6 ± 3,2
Манноза	20,4 ± 4,4	20,4 ± 4,5	18,6 ± 5,4	18,9 ± 4,5	19,7 ± 4,7	19,7 ± 5,3	18,6 ± 4,2
Этанол	13,7 ± 2,2	13,0 ± 2,2	14,0 ± 2,3	15,2 ± 2,6	13,9 ± 2,3	14,6 ± 2,6	13,4 ± 2,7
Глицерин	14,8 ± 4,2	14,3 ± 3,8	16,9 ± 4,8	15,5 ± 4,3	15,4 ± 4,3	16,5 ± 4,6	15,3 ± 4,2
Ксилоза	14,3 ± 2,0	15,0 ± 2,5	16,7 ± 2,6	14,8 ± 2,4	14,8 ± 2,3	14,5 ± 2,2	14,5 ± 2,4

Репортерная система под контролем олигонуклеотидов по настоящему изобретению проявляла повышение активности фермента в 14–25 раз. Хотя активность фермента варьируется при различных условиях роста, но существенных отличий в отношении контроля между олигонуклеотидами SEQ ID NO: 1 - SEQ ID NO: 7 не отмечалось.

Краткое описание фигур

На фиг. 1 представлены показатели роста 5 различных штаммов дрожжевых клеток (от А до Е), которые были трансформированы плазмидой, содержащей ген, кодирующий SEQ ID NO: 17, регулируемый олигонуклеотидом из существующего уровня техники: SEQ ID NO: 10.

На фиг. 2 представлены показатели роста штамма В, трансформированного различными плазмидами, содержащими различные олигонуклеотиды (олигонуклеотиды из существующего уровня техники и олигонуклеотиды по настоящему изобретению).

На фиг. 3 представлено повышение в 4–29 раз уровня транскрипта в репортерной системе под контролем олигонуклеотидов по настоящему изобретению (SEQ ID NO: 1–7) по сравнению с олигонуклеотидами из существующего уровня техники SEQ ID NO 8, SEQ ID NO: 9 и SEQ ID NO: 10 на различных субстратах (глюкоза, манноза, этанол, глицерин и ксилоза).

На фиг. 4 представлено повышение в 14–25 раз активности фермента в репортерной системе под контролем олигонуклеотидов по настоящему изобретению (SEQ ID NO: 1–7) по сравнению с олигонуклеотидами из существующего уровня техники SEQ ID NO 8, SEQ ID NO: 9 и SEQ ID NO: 10 на различных субстратах (глюкоза, манноза, этанол, глицерин и ксилоза).

Перечень последовательностей с описанием

SEQ ID NO: 1: синтетический олигонуклеотид, состоящий из комбинации SEQ ID NO: 9 и SEQ ID NO: 8

ccattctctg ctgctttggt gcatcgtatc atgccccgtt tcagccagtc ccgcataccc 60
 cctttgcaac gttaacgfta ccgctagcgt ttaccatctc cacgctaaaa ggaaaaacgaa 120
 gattaagata aagttgggta aaatccgggg taagaggcaa gggggtagag aaaaaaaaaac 180
 cggagtcatt atatacgata ccgtccaggg taagacagtg atttctagct tccactttt 240
 tcaatttct ttttcgttc caaatggcgt ccaccctac atccggaatc tgacggcaca 300
 agagccgatt agtggaaagcc acggttacgt gattgcggtt ttttttct acgtataacg 360
 ctatgacggt agttgaatgt taaaaacgaa aacagagata ttgaattgac tcgtaggaac 420
 aatttcgggc cctgcgtgt tcttctgagg tcatctttt acatttgctt ctgctggata 480
 atttcagag gcaacaagga aaaattagat ggcaaaaagt cgtcttcaa ggaaaaatcc 540
 ccaccatctt tcgagatccc ctgtaactta ttggcaactg aaagaatgaa aaggaggaaa 600
 atacaaaata tactagaact gaaaaaaaaa agtataaat agagacgata tatgccaata 660
 ctccaatgt ttcgaatcta ttctcattt gcagctattg taaaataata aaacatcaag 720
 aacaacaag ctcaactgt ttttctaag aacaagaat aaacaaaaa acaaaaagtt 780
 ttttaattt taatcaaaaa 800

SEQ ID NO: 2: синтетический олигонуклеотид, соответствующий SEQ ID NO: 1 и содержащий следующие мутации: g315a, c322g, a623-, t679c

ccattctctg ctgctttggt gcatcgtatc atgccccgtt tcagccagtc ccgcataccc 60
 cctttgcaac gttaacgfta ccgctagcgt ttaccatctc cacgctaaaa ggaaaaacgaa 120
 gattaagata aagttgggta aaatccgggg taagaggcaa gggggtagag aaaaaaaaaac 180
 cggagtcatt atatacgata ccgtccaggg taagacagtg atttctagct tccactttt 240
 tcaatttct ttttcgttc caaatggcgt ccaccctac atccggaatc tgacggcaca 300
 agagccgatt agtgaagcc aggttacgt gattgcggtt ttttttct acgtataacg 360
 ctatgacggt agttgaatgt taaaaacgaa aacagagata ttgaattgac tcgtaggaac 420
 aatttcgggc cctgcgtgt tcttctgagg tcatctttt acatttgctt ctgctggata 480
 atttcagag gcaacaagga aaaattagat ggcaaaaagt cgtcttcaa ggaaaaatcc 540
 ccaccatctt tcgagatccc ctgtaactta ttggcaactg aaagaatgaa aaggaggaaa 600
 atacaaaata tactagaact gaaaaaaaaa agtataata gagacgatat atgccaatac 660
 ttcacaatgt tcgaatccat tcttcattt cagctattt aaaataataa aacatcaaga 720
 acaacaagc tcaactgtc ttttctaaga acaagaata aaacaaaaa caaaaagtt 780
 ttttaattt aatcaaaaa 799

SEQ ID NO: 3: синтетический олигонуклеотид, соответствующий SEQ ID NO: 1 и содержащий следующие мутации: t549c, a622-, a623-, c661a, t679c

ccattctctg ctgctttggt gcatcgtatc atcgcccgtt tcagccagtc ccgcataccc 60
cctttgcaac gttaacgta ccgctagcgt ttaccatctc cacgctaaaa ggaaaacgaa 120
gattaagata aagttgggta aatccgggg taagaggcaa ggggtagag aaaaaaaaaac 180
cggagtcatt atatacgata ccgtccaggg taagacagtg atttctagct tccactttt 240
tcaatttctt ttttcgttc caaatggcgt ccaccctgac atccggaatc tgacggcaca 300
agagccgatt agtggaaagcc acggttacgt gattgagggt ttttttct acgtataacg 360
ctatgacggg agttgaatgt taaaaacgaa aacagagata ttgaattgac tcgtaggaac 420
aatttcgggc cctgctgtgt tcttctgagg tcatctttt acatttgctt ctgctggata 480
attttcagag gcaacaagga aaaattagat ggcaaaaagt cgtctttcaa ggaaaaatcc 540
ccaccatcct tcgagatccc ctgtaactta ttggcaactg aaagaatgaa aaggaggaaa 600
atacaaaaata tactagaact gaaaaaaaa gtataaatag agacgatata tgccaataat 660
tcacaatgtt cgaatccatt cttcatttgc agctattgta aaataataaa acatcaagaa 720
сааасаагсг сааггггсггг тггггааа сааагаааа асааааааа аааааггггг 780
ттгааттга атсааааа 798

SEQ ID NO: 4: синтетический олигонуклеотид, соответствующий SEQ ID NO: 1 и содержащий следующие мутации: с495а, г674а, т679с, а687г

ccattctctg ctgctttggt gcatcgtatc atcgcccgtt tcagccagtc ccgcataccc 60
cctttgcaac gttaacgta ccgctagcgt ttaccatctc cacgctaaaa ggaaaacgaa 120
gattaagata aagttgggta aatccgggg taagaggcaa ggggtagag aaaaaaaaaac 180
cggagtcatt atatacgata ccgtccaggg taagacagtg atttctagct tccactttt 240
tcaatttctt ttttcgttc caaatggcgt ccaccctgac atccggaatc tgacggcaca 300
agagccgatt agtggaaagcc acggttacgt gattgagggt ttttttct acgtataacg 360
ctatgacggg agttgaatgt taaaaacgaa aacagagata ttgaattgac tcgtaggaac 420
aatttcgggc cctgctgtgt tcttctgagg tcatctttt acatttgctt ctgctggata 480
attttcagag gcaaaaagga aaaattagat ggcaaaaagt cgtctttcaa ggaaaaatcc 540
ccaccatcct tcgagatccc ctgtaactta ttggcaactg aaagaatgaa aaggaggaaa 600
atacaaaaata tactagaact gaaaaaaaa aagtataaat agagacgata tatgccaata 660
cttcacaatg ttcaatcca ttcttcgttt gcagctattg taaaataata aaacatcaag 720
ааааааааагсг гггггггггг гггггггггг гггггггггг гггггггггг гггггггггг 780
ттгггггггг гггггггггг гггггггггг гггггггггг гггггггггг гггггггггг 800

SEQ ID NO: 5: синтетический олигонуклеотид, соответствующий SEQ ID NO: 1 и содержащий следующие мутации: а171-, с322г, с430т, с431т, с495а

ccattctctg ctgctttggt gcatcgtatc atcgcccgtt tcagccagtc ccgcataccc 60
cctttgcaac gttaacgta ccgctagcgt ttaccatctc cacgctaaaa ggaaaacgaa 120
gattaagata aagttgggta aatccgggg taagaggcaa ggggtagag aaaaaaaccc 180

ggagtcatta tatacgatac cgtccagggt aagacagtga tttctagctt ccaacttttt 240
 caatttcttt ttttcgtcc aaatggcgtc cacccgtaca tccggaatct gacggcaca 300
 gagccgatta gtggaagcca gggttacgtg attgcggttt tttttccta cgtataacgc 360
 tatgacggta gttgaatgft aaaaacgaaa acagagatat tgaattgact cgtaggaaca 420
 atttcgggtt cctgcgtgtt cttctgaggt tcatctttta catttgcttc tgctggataa 480
 ttttcagagg caaaaaggaa aaattagatg gcaaaaagtc gtctttcaag gaaaaatccc 540
 caccatcttt cgagatcccc tgtaacttat tggcaactga aagaatgaaa aggaggaaaa 600
 taaaaaatat actagaactg aaaaaaaaa agtataaata gagacgatat atgccaatac 660
 ttcacaatgt tgaatctat tcttcatttg cagctattgt aaaataataa aacatcaaga 720
 acaacaagc tcaactgtc ttttcaaga acaagaata aacacaaaaa caaaaagttt 780
 ttttaatttt aatcaaaaa 799

SEQ ID NO: 6: синтетический олигонуклеотид, соответствующий SEQ ID NO: 1 и содержащий следующие мутации: c181a, c205t, c322g, t549c, a623-, t679c

ccattctctg ctgctttggt gcatcgtatc atcgcccgtt tcagccagtc ccgcataccc 60
 cctttgcaac gtaacgtta ccgctagcgt ttaccatctc cacgctaaaa ggaaaacgaa 120
 gattaagata aagttgggta aaatccgggg taagaggcaa gggggtagag aaaaaaaaaac 180
 aggagtcatt atatacgata ccgttcaggg taagacagtg atttctagct tccacttttt 240
 tcaatttctt ttttcgttc caaatggcgt ccaccctac atccggaatc tgacggcaca 300
 agagccgatt agtgaagcc agggttacgt gattgcggtt ttttttctt acgtataacg 360
 ctatgacggt agttgaatgt taaaaacgaa aacagagata ttgaattgac tcgtaggaac 420
 aatttcgggc ccctgcgtgt tcttctgagg tcatctttt acatttgctt ctgctggata 480
 attttcagag gcaacaagga aaaattagat ggcaaaaagt cgtctttcaa ggaaaaatcc 540
 ccaccatcct tcgagatccc ctgtaactta ttggcaactg aaagaatgaa aaggaggaaa 600
 atacaaaata tactagaact gaaaaaaaa agtataaata gagacgatat atgccaatac 660
 ttcacaatgt tgaatccat tcttcatttg cagctattgt aaaataataa aacatcaaga 720
 acaacaagc tcaactgtc ttttcaaga acaagaata aacacaaaaa caaaaagttt 780
 ttttaatttt aatcaaaaa 799

SEQ ID NO: 7: синтетический олигонуклеотид, соответствующий SEQ ID NO: 1 и содержащий следующие мутации: t186a, c322g, c430t, c431t, a623-, t679c

ccattctctg ctgctttggt gcatcgtatc atcgcccgtt tcagccagtc ccgcataccc 60
 cctttgcaac gtaacgtta ccgctagcgt ttaccatctc cacgctaaaa ggaaaacgaa 120
 gattaagata aagttgggta aaatccgggg taagaggcaa gggggtagag aaaaaaaaaac 180
 cggagacatt atatacgata ccgtccaggg taagacagtg atttctagct tccacttttt 240
 tcaatttctt ttttcgttc caaatggcgt ccaccctac atccggaatc tgacggcaca 300
 agagccgatt agtgaagcc agggttacgt gattgcggtt ttttttctt acgtataacg 360

ctatgacggt agttgaatgt taaaacgaa aacagagata ttgaattgac tcgtaggaac 420
aatttcgggt tcttgcgtgt tcttctgagg ttcatctttt acatttgctt ctgctggata 480
attttcagag gcaacaagga aaaattagat ggcaaaaagt cgtctttcaa ggaaaaatcc 540
ccaccatctt tcgagatccc ctgtaactta ttggcaactg aaagaatgaa aaggaggaaa 600
atacaaaata tactagaact gaaaaaaaaa agtataaata gagacgatat atgccaatac 660
ttcacaatgt tcgaatccat tcttcatttg cagctattgt aaaataataa aacatcaaga 720
acaacaagc tcaactgtc ttttctaaga acaagaata aacacaaaaa caaaaagttt 780
ttttaatttt aatcaaaaa 799

SEQ ID NO: 8: олигонуклеотид из *Saccharomyces cerevisiae* S288C, вышерасположенный участок PFK2

SEQ ID NO: 9: олигонуклеотид из *Saccharomyces cerevisiae* S288C, вышерасположенный участок HXT7

SEQ ID NO:10: олигонуклеотид из *Saccharomyces cerevisiae* S288C, вышерасположенный участок PGK1

SEQ ID NO: 11: белок с активностью альдозо-1-эпимеразы из *Saccharomyces cerevisiae* EC1118, 1O4_6579

SEQ ID NO: 12: белок с активностью альдозо-1-эпимеразы из *Saccharomyces cerevisiae* EC1119, 1F14_0056

SEQ ID NO: 13: белок с активностью альдозо-1-эпимеразы из *Saccharomyces cerevisiae* S288C, YHR210C

SEQ ID NO: 14: белок с активностью альдозо-1-эпимеразы из *Saccharomyces cerevisiae* × *Saccharomyces kudriavzevii* VIN7, типа YHR210C

SEQ ID NO: 15: белок с активностью альдозо-1-эпимеразы из *Saccharomyces cerevisiae* AWRI796, типа YHR210C

SEQ ID NO: 16: белок с активностью альдозо-1-эпимеразы из *Lactococcus lactis*, XylM

SEQ ID NO: 17: белок с активностью ксилозоизомеразы из *Eubacterium saburreum*, XylA

SEQ ID NO: 18: белок с активностью ксилозоизомеразы из *Piromyces* sp. E2, XylA

SEQ ID NO: 19: белок с активностью ксилозоизомеразы из *Orpinomyces* sp. ukk1, XylA

SEQ ID NO: 20: белок с активностью ксилозоизомеразы из *Clostridium phytofermentans*, XylA

SEQ ID NO: 21: белок с активностью ксилозоизомеразы из *Ruminococcus flavefaciens*, XylA

SEQ ID NO: 22: белок с активностью ксилоизомеразы из *Bacteroides uniformis*, XylA

SEQ ID NO: 23: белок с активностью ксилоизомеразы из *Clostridium cellulolyticum*, XylA

SEQ ID NO: 24: белок с активностью ксилоизомеразы из *Thermotoga maritima*, XylA

SEQ ID NO: 25: белок с активностью ксилоизомеразы из *Bacillus stearothermophilus*, XylA

SEQ ID NO: 26: белок с активностью ксилоизомеразы из *Bacteroides stercoris*, XylA

SEQ ID NO: 27: белок с активностью ксилоизомеразы из *Parabacteroides distasonis*, XylA

SEQ ID NO: 28: белок с активностью ксилоизомеразы из *Prevotella ruminicola*, XylA

SEQ ID NO: 29: белок с активностью ксилоизомеразы из *Agrobacterium tumefaciens*, XylA

SEQ ID NO: 30: белок с активностью ксилоизомеразы из *Clostridium cellulovorans*, XylA

SEQ ID NO: 31: белок с активностью ксилоизомеразы из *Burkholderia cenocepacia*, XylA

SEQ ID NO: 32: белок с активностью ксилоизомеразы из *Lactococcus lactis*, XylA

SEQ ID NO: 33: белок с активностью ксилоизомеразы из *Thermoanaerobacter thermohydrosulfuricus*, XylA

SEQ ID NO: 34: белок с активностью ксилоизомеразы из *Reticulitermes speratus*, XylA

SEQ ID NO: 35: белок с активностью ксилоизомеразы из некультивированных бактерий из рубца коровы, последовательность № 2 в WO 2014/164392

SEQ ID NO: 36: белок с активностью ксилоизомеразы из некультивированных бактерий из рубца коровы, последовательность № 1 в WO 2014/164392

SEQ ID NO: 37: белок с активностью ксилоизомеразы из *Lachnospiraceae bacterium ICM7*, XylA

SEQ ID NO: 38: белок с активностью ксилоизомеразы из *Lachnospiraceae bacterium oral taxon 107*, XylA

SEQ ID NO: 39: белок с активностью ксилоизомеразы из *Lachnospiraceae bacterium oral taxon 082*, XylA

SEQ ID NO: 40: белок с активностью ксилоизомеразы из некультивированных бактерий, XYM1, см. Parachin N.S. and Gorwa-Grauslund M.F. Isolation of xylose isomerases by sequence- and function-based screening from a soil metagenome library. *Biotechnol Biofuels* 4 (1), 9 (2011)

SEQ ID NO: 41: белок с активностью ксилоизомеразы из некультивированных бактерий, XYM2, см. Parachin N.S. and Gorwa-Grauslund M.F. Isolation of xylose isomerases by sequence- and function-based screening from a soil metagenome library. *Biotechnol Biofuels* 4 (1), 9 (2011)

SEQ ID NO: 42: белок с активностью ксилоизомеразы из *Thermus thermophilus*, XylA

SEQ ID NO: 43: белок с активностью ксилоизомеразы из *Escherichia coli*, XylA

SEQ ID NO: 44: белок с активностью ксилулокиназы из *Saccharomyces cerevisiae* S288C, XKS1

SEQ ID NO: 45: белок с активностью ксилулокиназы из *Scheffersomyces (Pichia) stipites* CBS 6054, XKS1

SEQ ID NO: 46: белок с активностью ксилулокиназы из *Trichoderma reesei* QM6a, TRIREDRAFT_123288

SEQ ID NO: 47: белок с активностью трансальдозазы из *Saccharomyces cerevisiae* S288C, TAL1

SEQ ID NO: 48: белок с активностью трансальдозазы из *Saccharomyces cerevisiae* S288C, NQM1

SEQ ID NO: 49: белок с активностью трансальдозазы из *Scheffersomyces (Pichia) stipites* CBS 6054, TAL1

SEQ ID NO: 50: белок с активностью транскетолазы из *Saccharomyces cerevisiae* S288C, TKL1

SEQ ID NO: 51: белок с активностью транскетолазы из *Saccharomyces cerevisiae* S288C, TKL2

SEQ ID NO: 52: белок с активностью транскетолазы из *Scheffersomyces (Pichia) stipites* CBS 6054, TKL1

SEQ ID NO: 53: белок с активностью алкогольдегидрогеназы из *Saccharomyces cerevisiae* S288C, ADH1

SEQ ID NO: 54: синтетическая последовательность ДНК, кодирующая SEQ ID NO:

17

atgaaggaat tcttcccagg tatttctcca gttaagttcg aaggtagaga ctctaagaac 60

ccattgtctt tcaagtacta cgacgctaag agagttatta tgggtaagac catggaagaa 120

cacttgctt tcgctatggc ttggtggcac aacttggtg cttgtggtg tgacatgtc 180
ggtcaaggta ccggtgacaa gtcttcggt gaatctctg gtacatgga acacgctaga 240
gctaagggtg acgctggtat tgaattcatg aagaagtgg gtattaagta ctactgttc 300
cacgacaccg acattgtcc agaagaccaa gaagacatta acgttaccia cgctagattg 360
gacgaaatta ccgactacat tttggaaaag accaaggaca ccgacattaa gtgtttgtgg 420
accacctgta acatgttctc taaccaaga tcatgaacg gtgctgggtc ttctaactct 480
gctgacgttt tctgttctgc tgctgctcaa gctaagaagg gtttgaaaa cgctgttaag 540
ttgggtgcta agggttctg tttctgggtt gtagagaag gttacgaaac cttgtgaac 600
accgacatga agttggaaga agaaaacatt gctacctgt tcacatgtg tagagactac 660
gtagatcta ttggttcat gggtgactc tacattgaac caaagccaaa ggaaccaatg 720
aagaccaat acgactcga cgctgctacc gctattggtt tctgagaaa gtacggttg 780
gacaaggact tcaagttaa cattgaagct aaccacgcta cctggctgg tcacacctc 840
caacacgaat tgagagttg tgctgtaac ggtatgatgg gttctgtga cgtaaccaa 900
ggtgacacct tgtgggttg ggacaccgac caattccaa ccaacgtta cgacaccacc 960
ttggctatgt acgaaatth gaaggctggt ggttgagag gtggttgaa cttcactct 1020
aagaacagaa gaccatctaa caccgctgac gacatgtct acggttcat tgctggtatg 1080
gacacctcg cttgggtt gattaaggct gctgaaatta tgaagacgg tagaattgac 1140
gacttcgta aggaaagata cgctcttac aactctgta ttgtaagaa gattagaaac 1200
agaaaggta cctgattga atgtctgaa tacgctgcta agttgaagaa gccagaattg 1260
ccagaatctg gtagacaaga atactggaa tctgtgtca acaacatth gttcgggtg 1320
tctggtta 1329

СПИСОК ПОСЛЕДОВАТЕЛЬНОСТЕЙ

<110> Clariant International Ltd

<120> Oligonucleotide sequence for use in pathway engineering

<130> 2014DE617-WO-PCT

<160> 54

<170> PatentIn version 3.5

<210> 1

<211> 800

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 1

```
ccattctctg ctgctttggt gcatcgtatc atcgcccgtt tcagccagtc ccgcataccc      60
cctttgcaac gttaacgtta ccgctagcgt ttaccatctc cacgctaaaa ggaaaacgaa      120
gattaagata aagttgggta aaatccgggg taagaggcaa gggggtagag aaaaaaaaaac      180
cggagtcatt atatacgata ccgtccaggg taagacagtg atttctagct tccacttttt      240
tcaatttctt tttttcgttc caaatggcgt ccacccgtac atccggaatc tgacggcaca      300
agagccgatt agtggaagcc acggttacgt gattgcggtt tttttttcct acgtataacg      360
ctatgacggt agttgaatgt taaaaacgaa aacagagata ttgaattgac tcgtaggaac      420
aatttcgggc ccctgcgtgt tcttctgagg ttcattcttt acatttgctt ctgctggata      480
attttcagag gcaacaagga aaaattagat ggcaaaaagt cgtctttcaa ggaaaaatcc      540
ccaccatctt tcgagatccc ctgtaactta ttggcaactg aaagaatgaa aaggaggaaa      600
atacaaaata tactagaact gaaaaaaaaa agtataaat agagacgata tatgccaata      660
cttcacaatg ttcgaatcta ttcttcattt gcagctattg taaaataata aaacatcaag      720
aacaacaag ctcaacttgt cttttctaag aacaagaat aaacacaaaa acaaaaagtt      780
tttttaattt taatcaaaaa                                     800
```

<210> 2

<211> 799

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 2

```
ccattctctg ctgctttggt gcatcgtatc atcgcccgtt tcagccagtc ccgcataccc      60
cctttgcaac gttaacgtta ccgctagcgt ttaccatctc cacgctaaaa ggaaaacgaa      120
gattaagata aagttgggta aaatccgggg taagaggcaa gggggtagag aaaaaaaaaac      180
cggagtcatt atatacgata ccgtccaggg taagacagtg atttctagct tccacttttt      240
tcaatttctt tttttcgttc caaatggcgt ccacccgtac atccggaatc tgacggcaca      300
agagccgatt agtgaaagcc agggttacgt gattgcggtt tttttttcct acgtataacg      360
```

ctatgacggg agttgaatgt taaaaacgaa aacagagata ttgaattgac tcgtaggaac 420
aatttcgggc ccctgcgtgt tcttctgagg ttcattctttt acatttgctt ctgctggata 480
attttcagag gcaacaagga aaaattagat ggcaaaaagt cgtctttcaa ggaaaaatcc 540
ccaccatctt tcgagatccc ctgtaactta ttggcaactg aaagaatgaa aaggaggaaa 600
atacaaaata tactagaact gaaaaaaaaa agtataaata gagacgatat atgccaatac 660
ttcacaatgt tcgaatccat tcttcatttg cagctattgt aaaataataa aacatcaaga 720
acaacaagc tcaacttgct ttttctaaga acaagaata aacacaaaaa caaaaagttt 780
ttttaatttt aatcaaaaa 799

<210> 3
<211> 798
<212> DNA
<213> *Saccharomyces cerevisiae*

<400> 3
ccattctctg ctgctttggt gcatcgtatc atcgcccgtt tcagccagtc ccgcataccc 60
cctttgcaac gttaacgtta ccgctagcgt ttaccatctc cacgctaaaa ggaaaaacgaa 120
gattaagata aagttgggta aatccgggg taagaggcaa gggggtagag aaaaaaaaaac 180
cggagtcatt atatacgata ccgtccaggg taagacagtg atttctagct tccacttttt 240
tcaatttctt tttttcgttc caaatggcgt ccaccctac atccggaatc tgacggcaca 300
agagccgatt agtgaagcc acggttacgt gattgcgggt tttttttcct acgtataacg 360
ctatgacggg agttgaatgt taaaaacgaa aacagagata ttgaattgac tcgtaggaac 420
aatttcgggc ccctgcgtgt tcttctgagg ttcattctttt acatttgctt ctgctggata 480
attttcagag gcaacaagga aaaattagat ggcaaaaagt cgtctttcaa ggaaaaatcc 540
ccaccatcct tcgagatccc ctgtaactta ttggcaactg aaagaatgaa aaggaggaaa 600
atacaaaata tactagaact gaaaaaaaaa gtataaatag agacgatata tgccaataat 660
tcacaatggt cgaatccatt cttcatttgc agctattgta aaataataaa acatcaagaa 720
caaacaagct caacttgctt tttctaagaa caagaataa acacaaaaac aaaaagtttt 780
tttaatttta atcaaaaa 798

<210> 4
<211> 800
<212> DNA
<213> *Saccharomyces cerevisiae*

<400> 4
ccattctctg ctgctttggt gcatcgtatc atcgcccgtt tcagccagtc ccgcataccc 60
cctttgcaac gttaacgtta ccgctagcgt ttaccatctc cacgctaaaa ggaaaaacgaa 120
gattaagata aagttgggta aatccgggg taagaggcaa gggggtagag aaaaaaaaaac 180

cggagtcatt atatacgata cgcgccaggg taagacagtg atttctagct tccacttttt 240
 tcaatttctt tttttcgttc caaatggcgt ccaccggtac atccggaatc tgacggcaca 300
 agagccgatt agtggagcc acggttacgt gattgcggtt ttttttctc acgtataacg 360
 ctatgacggg agttgaatgt taaaaacgaa aacagagata ttgaattgac tcgtaggaac 420
 aatttcgggc ccctgcgtgt tcttctgagg ttcattctttt acatttgctt ctgctggata 480
 attttcagag gcaaaaagga aaaattagat ggcaaaaagt cgtctttcaa ggaaaaatcc 540
 ccaccatctt tcgagatccc ctgtaactta ttggcaactg aaagaatgaa aaggaggaaa 600
 atacaaaata tactagaact gaaaaaaaaa agtataaat agagacgata tatgccaata 660
 cttcacaatg ttcaaatcca ttcttcgttt gcagctattg taaaataata aaacatcaag 720
 aacaacaag ctcaacttgt cttttctaag aacaagaat aaacacaaaa acaaaaagtt 780
 tttttaattt taatcaaaaa 800

<210> 5
 <211> 799
 <212> DNA
 <213> *Saccharomyces cerevisiae*

<400> 5
 ccattctctg ctgctttggt gcatcgtatc atcgcccgtt tcagccagtc ccgcataccc 60
 cctttgcaac gttaacgta ccgctagcgt ttaccatctc cacgctaaaa ggaaaacgaa 120
 gattaagata aagttgggta aaatccgggg taagaggcaa gggggtagag aaaaaaaaaacc 180
 ggagtcatta tatacgatac cgtccagggg aagacagtga tttctagctt ccaactttttt 240
 caatttcttt ttttcgttcc aaatggcgtc caccggtaca tccggaatct gacggcacia 300
 gagccgatta gtggaagcca gggttacgtg attgcggttt ttttttctc cgtataacgc 360
 tatgacggta gttgaatggt aaaaacgaaa acagagatat tgaattgact cgtaggaaca 420
 atttcggggt cctgcgtggt cttctgaggt tcatctttta catttgcttc tgctggataa 480
 ttttcagagg caaaaaggaa aaattagatg gcaaaaagtc gtctttcaag gaaaaatccc 540
 caccatcttt cgagatcccc tgtaacttat tggcaactga aagaatgaaa aggaggaaaa 600
 tacaaaatat actagaactg aaaaaaaaaa agtataaata gagacgatat atgccaatac 660
 ttcacaatgt tcgaatctat tcttcatttg cagctattgt aaaataataa aacatcaaga 720
 acaacaagc tcaacttgct ttttctaaga acaagaata aacacaaaaa caaaaagttt 780
 ttttaatttt aatcaaaaa 799

<210> 6
 <211> 799
 <212> DNA
 <213> *Saccharomyces cerevisiae*

<400> 6
ccattctctg ctgctttggt gcatcgtatc atcgcccgtt tcagccagtc ccgcataccc 60
cctttgcaac gttaacgtta ccgctagcgt ttaccatctc cacgctaaaa ggaaaacgaa 120
gattaagata aagttgggta aaatccgggg taagaggcaa gggggtagag aaaaaaaaaac 180
aggagtcatt atatacgata ccgttcaggg taagacagtg atttctagct tccacttttt 240
tcaatttctt tttttcgttc caaatggcgt ccacccgtac atccggaatc tgacggcaca 300
agagccgatt agtggaagcc agggttacgt gattgcggtt tttttttcct acgtataacg 360
ctatgacggg agttgaatgt taaaaacgaa aacagagata ttgaattgac tcgtaggaac 420
aatttcgggc ccctgcgtgt tcttctgagg ttcacttttt acatttgctt ctgctggata 480
attttcagag gcaacaagga aaaattagat ggcaaaaagt cgtctttcaa ggaaaaatcc 540
ccaccatcct tcgagatccc ctgtaactta ttggcaactg aaagaatgaa aaggaggaaa 600
atacaaaata tactagaact gaaaaaaaaa agtataaata gagacgatat atgccaatac 660
ttcacaatgt tcgaatccat tcttcatttg cagctattgt aaaataataa aacatcaaga 720
acaacaagc tcaacttgtc ttttctaaga acaaagaata aacacaaaaa caaaaagttt 780
ttttaatttt aatcaaaaa 799

<210> 7
<211> 799
<212> DNA
<213> *Saccharomyces cerevisiae*

<400> 7
ccattctctg ctgctttggt gcatcgtatc atcgcccgtt tcagccagtc ccgcataccc 60
cctttgcaac gttaacgtta ccgctagcgt ttaccatctc cacgctaaaa ggaaaacgaa 120
gattaagata aagttgggta aaatccgggg taagaggcaa gggggtagag aaaaaaaaaac 180
cggagacatt atatacgata ccgtccaggg taagacagtg atttctagct tccacttttt 240
tcaatttctt tttttcgttc caaatggcgt ccacccgtac atccggaatc tgacggcaca 300
agagccgatt agtggaagcc agggttacgt gattgcggtt tttttttcct acgtataacg 360
ctatgacggg agttgaatgt taaaaacgaa aacagagata ttgaattgac tcgtaggaac 420
aatttcgggt tcctgcgtgt tcttctgagg ttcacttttt acatttgctt ctgctggata 480
attttcagag gcaacaagga aaaattagat ggcaaaaagt cgtctttcaa ggaaaaatcc 540
ccaccatctt tcgagatccc ctgtaactta ttggcaactg aaagaatgaa aaggaggaaa 600
atacaaaata tactagaact gaaaaaaaaa agtataaata gagacgatat atgccaatac 660
ttcacaatgt tcgaatccat tcttcatttg cagctattgt aaaataataa aacatcaaga 720
acaacaagc tcaacttgtc ttttctaaga acaaagaata aacacaaaaa caaaaagttt 780
ttttaatttt aatcaaaaa 799

<210> 8
<211> 409
<212> DNA
<213> *Saccharomyces cerevisiae*

<400> 8
ccattctctg ctgctttggt gcatcgtatc atcgcccgtt tcagccagtc ccgcataccc 60
cctttgcaac gttaacgtta ccgctagcgt ttaccatctc cacgctaaaa ggaaaacgaa 120
gattaagata aagttgggta aatccgggg taagaggcaa gggggtagag aaaaaaaaaac 180
cggagtcatt atatacgata ccgtccaggg taagacagtg atttctagct tccacttttt 240
tcaatttctt tttttcgttc caaatggcgt ccaccctac atccggaatc tgacggcaca 300
agagccgatt agtgaagcc acggttacgt gattgcgggt tttttttcct acgtataacg 360
ctatgacggg agttgaatgt taaaaacgaa aacagagata ttgaattga 409

<210> 9
<211> 391
<212> DNA
<213> *Saccharomyces cerevisiae*

<400> 9
ctcgtaggaa caatttcggg cccctgcgtg ttcttctgag gttcatcttt tacatttgc 60
tctgctggat aattttcaga ggcaacaagg aaaaattaga tggcaaaaag tcgtctttca 120
aggaaaaatc cccaccatct ttcgagatcc cctgtaactt attggcaact gaaagaatga 180
aaaggaggaa aatacaaaat atactagaac tgaaaaaaaa aaagtataaa tagagacgat 240
atatgccaat acttcacaat gttcgaatct attcttcatt tgcagctatt gtaaaataat 300
aaaacatcaa gaacaaaca gctcaacttg tcttttctaa gaacaaagaa taaacacaaa 360
aacaaaaagt ttttttaatt ttaatcaaaa a 391

<210> 10
<211> 500
<212> DNA
<213> *Saccharomyces cerevisiae*

<400> 10
tgtttgcaaa aagaacaaaa ctgaaaaaac ccagacacgc tcgacttcct gtcttcctat 60
tgattgcagc ttccaatttc gtcacacaac aaggctctag cgacggctca caggttttgt 120
aacaagcaat cgaaggttct ggaatggcgg gaaagggttt agtaccacat gctatgatgc 180
cactgtgat ctccagagca aagttcgttc gatcgtactg ttactctctc tctttcaaac 240
agaattgtcc gaatcgtgtg acaacaacag cctgtttctca cacactcttt tcttctaacc 300
aaggggggtg tttagtttag tagaacctcg tgaaacttac atttacaat atataaactt 360
gcataaattg gtcaatgcaa gaaatacata tttggctttt tctaattcgt agtttttcaa 420
gttcttagat gctttctttt tctctttttt acagatcatc aaggaagtaa ttatctactt 480

<210> 11
 <211> 342
 <212> PRT
 <213> Saccharomyces cerevisiae

<400> 11

Met Lys Asp Asn Gln Glu Asn His Asp Gln Lys Pro Phe Ile Ile Leu
 1 5 10 15

Gly Asp Lys Asn Lys Phe Glu Ala Thr Ile Ala Lys Val Gly Ala Thr
 20 25 30

Leu Val Asp Leu Lys Val Asn Gly Gln Ser Val Val Leu Gly Tyr Pro
 35 40 45

Asp Val Lys Gly Tyr Leu Asn Asp Pro Gly Asn Leu Val Gly Ala Asn
 50 55 60

Val Gly Arg Tyr Ala Asn Arg Ile Tyr Lys Gly Val Phe Asn Thr Pro
 65 70 75 80

Asp Gly Ala His Gln Leu Thr Val Asn Asn Cys Gly Asn Ala Asn His
 85 90 95

Ser Ser Ile Ser Cys Phe Asn Arg Lys Thr Phe Val Ala Ser Pro Val
 100 105 110

Glu Asn Ser Ser Arg Asp Val Tyr Ile Ala Lys Leu Thr Leu Leu Asp
 115 120 125

Asp His Thr Val Pro Asn Glu Phe Pro Gly Asp Leu Glu Val Thr Val
 130 135 140

Thr Tyr Thr Leu Asn Val Ala Glu Met Thr Leu Asp Leu Asp Tyr Arg
 145 150 155 160

Ala His Leu Val Lys Gly Asp Ala Thr Pro Ile Asn Met Thr Asn His
 165 170 175

Thr Tyr Phe Asn Leu Asn Lys Thr Arg Asn Glu Glu Ser Ile Ile Gly
 180 185 190

Thr Glu Ile Asn Ile Cys Ser Asp Lys Ser Leu Glu Val Thr Glu Gly
 195 200 205

Ala Leu Ile Pro Thr Gly Lys Ile Ile Lys Arg Asp Ile Ala Thr Phe

210

215

220

Gln Ser Ala His Pro Thr Thr Leu Gly Ser Lys Ala Pro Val Tyr Asp
225 230 235 240

Phe Cys Phe Ile Val Asp Ala Asn Lys Asp Leu Arg Ser Thr Asp Ser
245 250 255

Thr Ser Val Asn Lys Leu Val Pro Val Phe Lys Ala Tyr His Pro Glu
260 265 270

Ser Asn Ile Arg Leu Gln Val Ser Thr Thr Glu Pro Thr Val His Phe
275 280 285

Tyr Thr Gly Asp Ser Leu Gly Gly Lys Phe Val Pro Arg Ser Gly Phe
290 295 300

Ala Val Glu Gln Gly Arg Tyr Ile Asp Ala Ile Asn Arg Ser Glu Trp
305 310 315 320

Arg Asn Cys Val Leu Leu Lys Arg Gly Glu Val Tyr Thr Ser Lys Thr
325 330 335

Gln Tyr Cys Phe Lys Asn
340

<210> 12
<211> 359
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 12

Met Gly Ser Lys Asn Thr Leu Leu Val Phe His Leu Lys Gly Val Lys
1 5 10 15

Thr Asp Leu Ala Gly Ala Val Asn Asp Tyr Glu Asn Arg Phe Ile Thr
20 25 30

Ile Gly Ala Gly Ser Lys Phe Glu Ala Thr Ile Ala Asn Leu Gly Ala
35 40 45

Thr Leu Val Asp Leu Lys Val Asn Gly Gln Ser Val Val Leu Gly Tyr
50 55 60

Ser Arg Ala Glu Asp Tyr Asn Ser Asp Gly Gly Asn Tyr Ile Gly Ala
65 70 75 80

Thr Val Gly Arg Phe Ala Asn Arg Ile Lys Glu Gly Leu Phe Thr Leu
85 90 95

Gln Asp Gly Thr His Lys Leu Thr Val Asp Asn Cys Asp Asn Thr Asn
 100 105 110

His Ser Ser Ile Ser Ser Phe His Val Lys Lys Phe Leu Gly Pro Leu
 115 120 125

Ile Glu Asn Pro Ser Asp Glu Ile Tyr Thr Ala Glu Phe Leu Leu Leu
 130 135 140

Asp Asp His Ser Ile Pro Asn Glu Phe Pro Gly Asp Leu Glu Val Ile
 145 150 155 160

Val Lys Phe Thr Leu Asn Ile Ala Glu Met Ser Leu Lys Phe Ser Tyr
 165 170 175

Gln Ala Gln Leu Ile Asn Gly Glu Ala Thr Pro Ile Asn Met Thr Ser
 180 185 190

His Thr Tyr Phe Asn Leu Asn Lys Phe His Asn Glu Gln Ser Ile Ser
 195 200 205

Gly Thr Glu Val Arg Val Cys Ser Thr Lys Ser Leu Glu Val Ser Glu
 210 215 220

Gly Ala Leu Ile Pro Thr Gly Lys Val Ile Asp Arg Asp Val Ala Thr
 225 230 235 240

Phe Asp Ser Ala Gly Pro Thr Thr Leu Gly Gly Asp Gly Pro Thr Tyr
 245 250 255

Asp Tyr Cys Phe Ile Ser Asp Glu Asn Lys Gly Leu Lys Ser Pro Asp
 260 265 270

Ser Arg Ser Gln Asn Glu Leu Arg Pro Val Leu Lys Ala Tyr His Pro
 275 280 285

Glu Ser Lys Ile Thr Leu Glu Val Ser Thr Thr Glu Pro Ser Phe Val
 290 295 300

Leu Tyr Ser Gly Asp Asn Leu Phe Gly Lys Phe Ile Pro Arg Gln Gly
 305 310 315 320

Phe Cys Val Glu Gln Gly Arg Tyr Ile Asp Ala Ile Asn Arg Asp Asp
 325 330 335

Trp Lys Asp Cys Val Leu Leu Lys Arg Gly Asp Val Tyr Thr Ser Glu
 340 345 350

Thr Gln Tyr Arg Phe Glu Asn
355

<210> 13
<211> 341
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 13

Met Ser Asn Asn Lys Ala Gly Gly Glu Tyr Glu Val Ile Thr Ile Gly
1 5 10 15

Asp Ala Lys Lys Leu Gln Ala Thr Ile Ser Glu Leu Gly Ala Thr Leu
20 25 30

Leu Asp Leu Lys Val Asn Asn Glu Ser Ile Val Leu Gly Tyr Pro Asp
35 40 45

Ile His Gly Tyr Ile Ser Asp Gly Tyr Asn Tyr Ile Gly Ala Thr Val
50 55 60

Gly Arg Tyr Ala Asn Arg Ile Tyr Lys Gly Met Phe Ser Met Glu Asp
65 70 75 80

Gly Pro His Gln Leu Thr Val Asn Asn Cys Gly Asn Thr Asn His Ser
85 90 95

Ser Ile Ser Ser Phe His Leu Lys Lys Tyr Lys Ala Ser Lys Val Gln
100 105 110

Asn Pro Leu Asp Asp Leu Tyr Ile Val Glu Phe Thr Leu Leu Asp Asp
115 120 125

Arg Thr Leu Pro Asn Glu Phe Pro Gly Asp Leu Ala Val Asn Leu Lys
130 135 140

Tyr Thr Leu Asn Val Ala Asp Met Thr Leu Asp Leu Glu Tyr Glu Ala
145 150 155 160

Lys Leu Val Ser Gly Glu Ala Thr Pro Ile Asn Met Thr Asn His Thr
165 170 175

Tyr Phe Asn Leu Asn Lys Thr Met Asp Lys Lys Ser Ile Ser Gly Thr
180 185 190

Glu Val Arg Leu Cys Ser Asp Lys Ser Leu Glu Val Ser Glu Gly Ala
195 200 205

Leu Ile Pro Thr Gly Lys Ile Val Gln Arg Lys Ile Ala Thr Phe Asp
210 215 220

Ser Ser Lys Pro Thr Ile Leu Gln Asp Asp Gly Pro Ile Tyr Asp Tyr
225 230 235 240

Ala Phe Ile Val Asp Glu Asn Lys Asn Leu Lys Thr Thr Asp Ser Val
245 250 255

Ser Val Asn Lys Leu Val Pro Ala Phe Lys Ala Tyr His Pro Ala Ser
260 265 270

Arg Leu Ser Leu Glu Val Ser Thr Thr Glu Pro Thr Val Leu Phe Tyr
275 280 285

Thr Gly Asp Asn Leu Cys Asp Gly Phe Thr Pro Arg Ser Gly Phe Ala
290 295 300

Val Glu Gln Gly Arg Tyr Val Asp Ala Ile Asn Arg Asp Gly Trp Arg
305 310 315 320

Asp Cys Val Leu Leu Arg Arg Gly Glu Val Tyr Thr Ser Lys Thr Arg
325 330 335

Tyr Arg Phe Ala Val
340

<210> 14
<211> 360
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 14

Met Ser Asn Asp Gly Gln Gln Lys Lys Leu Ala Thr Ile Asp Asn Lys
1 5 10 15

Leu Glu Met Thr Asn Ser Asn Thr Glu Asn Lys Tyr Lys Val Ile Thr
20 25 30

Ile Gly Asp Glu Lys Arg Phe Gln Ala Thr Ile Ala Pro Leu Gly Ala
35 40 45

Thr Leu Val Asp Leu Lys Val Asn Asn Gln Ser Val Val Gln Gly Tyr
50 55 60

Ser Asn Val Gln Asp Tyr Leu Thr Asp Gly Asn Met Met Gly Ala Thr
65 70 75 80

Val Gly Arg Tyr Ala Asn Arg Ile Ala Lys Gly Val Phe Ser Leu Glu
 85 90 95

Asp Gly Pro His Lys Leu Thr Val Asn Asn Cys Gly Asn Thr Asn His
 100 105 110

Ser Ser Ile Ser Ser Leu Asn Leu Lys Gln Tyr Arg Ala Ser Pro Val
 115 120 125

Glu Asn Pro Ser Lys Asp Val Phe Val Val Glu Phe Lys Leu Leu Asp
 130 135 140

Asp His Thr Gln Pro Asn Pro Asn Glu Phe Pro Gly Asp Leu Glu Val
 145 150 155 160

Thr Val Lys Tyr Thr Leu Asn Val Ala Glu Met Thr Leu Gly Met Lys
 165 170 175

Tyr Gln Ala Gln Leu Val Arg Gly Asp Ala Thr Pro Ile Asn Met Thr
 180 185 190

Asn His Ser Tyr Phe Asn Leu Asn Lys Thr Lys Asn Glu Lys Ser Ile
 195 200 205

Ser Gly Thr Glu Val Lys Val Cys Ser Asn Lys Ser Leu Glu Val Thr
 210 215 220

Glu Gly Ala Leu Leu Pro Thr Glu Lys Ile Ile Glu Arg Lys Ile Ala
 225 230 235 240

Thr Phe Asp Ser Ala Lys Pro Thr Val Leu His Asp Asp Ala Pro Val
 245 250 255

Phe Asp Cys Thr Phe Ile Ile Asp Ala Asn Lys Asp Leu His Thr Thr
 260 265 270

Asp Ser Val Ser Val Asn Lys Leu Val Pro Val Phe Lys Ala Tyr His
 275 280 285

Pro Glu Ser Arg Ile Thr Phe Glu Val Ser Thr Thr Glu Pro Thr Val
 290 295 300

His Leu Tyr Thr Gly Asp Asn Leu Cys Gly Lys Phe Thr Pro Arg Ser
 305 310 315 320

Gly Phe Ala Val Gln Gln Gly Arg Tyr Val Asp Ala Ile Asn His Asp
 325 330 335

Lys Trp Arg Asp Cys Val Leu Leu Lys Arg Gly Glu Val Tyr Thr Ser
340 345 350

Glu Thr Gln Tyr Arg Phe Gly Ile
355 360

<210> 15
<211> 342
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 15

Met Ser Asn Ser Asn Gly Asp Asn Lys Tyr Gly Val Ile Thr Ile Gly
1 5 10 15

Asp Glu Lys Lys Phe Gln Ala Thr Ile Ala Pro Leu Gly Ala Thr Leu
20 25 30

Val Asp Leu Lys Val Asn Gly Gln Ser Val Val Gln Gly Tyr Ser Asn
35 40 45

Val Gln Asp Tyr Leu Thr Asp Gly Asn Met Met Gly Ala Thr Val Gly
50 55 60

Arg Tyr Ala Asn Arg Ile Ala Lys Gly Val Phe Ser Leu Asp Asp Gly
65 70 75 80

Pro His Lys Leu Thr Val Asn Asn Cys Gly Asn Thr Asn His Ser Ser
85 90 95

Ile Ser Ser Leu Asn Leu Lys Gln Tyr Lys Ala Ser Pro Val Glu Asn
100 105 110

Pro Ser Lys Gly Val Tyr Val Val Glu Phe Lys Leu Leu Asp Asp His
115 120 125

Thr Gln Pro Asn Pro Asn Glu Phe Pro Gly Asp Leu Glu Val Thr Val
130 135 140

Lys Tyr Thr Leu Asn Val Ala Glu Met Thr Leu Asp Met Glu Tyr Gln
145 150 155 160

Ala Gln Leu Val Arg Gly Asp Ala Thr Pro Ile Asn Met Thr Asn His
165 170 175

Ser Tyr Phe Asn Leu Asn Lys Val Lys Ser Glu Lys Ser Ile Arg Gly
180 185 190

Thr Glu Val Lys Val Cys Ser Asn Lys Ser Leu Glu Val Thr Glu Gly

195

200

205

Ala Leu Leu Pro Thr Gly Lys Ile Ile Glu Arg Asn Ile Ala Thr Phe
 210 215 220

Asp Ser Thr Lys Pro Thr Val Leu His Glu Asp Thr Pro Val Phe Asp
 225 230 235 240

Cys Thr Phe Ile Ile Asp Ala Asn Lys Asp Leu Lys Thr Thr Asp Ser
 245 250 255

Val Ser Val Asn Lys Leu Val Pro Val Phe Lys Ala Tyr His Pro Glu
 260 265 270

Ser His Ile Lys Phe Glu Val Ser Thr Thr Glu Pro Thr Val His Leu
 275 280 285

Tyr Thr Gly Asp Asn Leu Cys Gly Lys Phe Val Pro Arg Ser Gly Phe
 290 295 300

Ala Val Gln Gln Gly Arg Tyr Val Asp Ala Ile Asn Arg Asp Glu Trp
 305 310 315 320

Arg Gly Cys Val Leu Leu Lys Arg Gly Glu Val Tyr Thr Ser Lys Thr
 325 330 335

Gln Tyr Lys Phe Asp Ile
 340

<210> 16

<211> 334

<212> PRT

<213> Lactococcus lactis

<400> 16

Met Thr Phe Thr Ile Ser Lys Glu Ser Leu Pro Phe Arg Ala Asp Lys
 1 5 10 15

Ser Ile Ser Gln Ile Thr Leu Ser Asn Glu Arg Leu Thr Ile Val Val
 20 25 30

His Asp Tyr Gly Ala Arg Ala His Gln Leu Leu Thr Pro Asp Lys Asn
 35 40 45

Gly Thr Phe Glu Asn Ile Leu Leu Ser Lys Asn Asp Ser Glu Thr Tyr
 50 55 60

Ala Asn Asp Gly Gly Tyr Tyr Gly Val Ile Cys Gly Pro Val Ala Gly
 65 70 75 80

Arg Ile Ser Gly Ala Thr Tyr Asp Ser Val Ser Leu Glu Ala Asn Glu
 85 90 95

Gly Lys Asn Asn Leu His Ser Gly Ser His Gly Trp Glu Arg Gln Phe
 100 105 110

Trp Ser Tyr Glu Thr Phe Glu Thr Ala Ser Ser Leu Gly Ile Lys Leu
 115 120 125

Ser Leu Arg Asp Glu Glu Ser Gly Phe Pro Gly Gln Ile Gln Ala Glu
 130 135 140

Val Thr Tyr Lys Leu Thr Asp Asn Lys Leu Glu Val Thr Ile Ser Gly
 145 150 155 160

Leu Ser Val Thr Asp Thr Val Phe Asn Pro Ala Trp His Pro Tyr Phe
 165 170 175

Asn Leu Ser Ala Glu Leu Ser Thr Thr His Glu His Phe Ile Gln Ala
 180 185 190

Asn Val Asp Phe Leu Val Glu Thr Asn Gln Glu Asn Ile Pro Thr Gly
 195 200 205

Arg Leu Leu Thr Val Asp Asp Ser Ser Tyr Ser Ile Lys Glu Ser Val
 210 215 220

Ser Ile Lys Lys Leu Leu Lys Asp Asn Pro Glu Gly Leu Asp Asp Cys
 225 230 235 240

Phe Val Phe Asn Pro Lys Gly Asp Lys Ser Leu Met Leu Tyr Asp Pro
 245 250 255

Leu Ser Gly Arg Lys Leu Val Ala Gln Thr Asp Arg Gln Ala Val Val
 260 265 270

Ile Tyr Thr Ala Thr Asn Pro Glu Ile Glu Ser Met Ile Asn Gly Arg
 275 280 285

Pro Met Ser Lys Asn Arg Gly Ile Ala Ile Glu Phe Gln Glu Ile Pro
 290 295 300

Asp Leu Val His His Pro Glu Trp Gly Thr Ile Glu Leu Lys Ala Gly
 305 310 315 320

Gln Lys Lys Thr Phe Ile Thr Glu Tyr Leu Phe Thr Thr Asn
 325 330

<210> 17
<211> 442
<212> PRT
<213> Eubacterium saburreum

<400> 17

Met Lys Glu Phe Phe Pro Gly Ile Ser Pro Val Lys Phe Glu Gly Arg
1 5 10 15

Asp Ser Lys Asn Pro Leu Ser Phe Lys Tyr Tyr Asp Ala Lys Arg Val
20 25 30

Ile Met Gly Lys Thr Met Glu Glu His Leu Ser Phe Ala Met Ala Trp
35 40 45

Trp His Asn Leu Cys Ala Cys Gly Val Asp Met Phe Gly Gln Gly Thr
50 55 60

Val Asp Lys Ser Phe Gly Glu Ser Ser Gly Thr Met Glu His Ala Arg
65 70 75 80

Ala Lys Val Asp Ala Gly Ile Glu Phe Met Lys Lys Leu Gly Ile Lys
85 90 95

Tyr Tyr Cys Phe His Asp Thr Asp Ile Val Pro Glu Asp Gln Glu Asp
100 105 110

Ile Asn Val Thr Asn Ala Arg Leu Asp Glu Ile Thr Asp Tyr Ile Leu
115 120 125

Glu Lys Thr Lys Asp Thr Asp Ile Lys Cys Leu Trp Thr Thr Cys Asn
130 135 140

Met Phe Ser Asn Pro Arg Phe Met Asn Gly Ala Gly Ser Ser Asn Ser
145 150 155 160

Ala Asp Val Phe Cys Phe Ala Ala Ala Gln Ala Lys Lys Gly Leu Glu
165 170 175

Asn Ala Val Lys Leu Gly Ala Lys Gly Phe Val Phe Trp Gly Gly Arg
180 185 190

Glu Gly Tyr Glu Thr Leu Leu Asn Thr Asp Met Lys Leu Glu Glu Glu
195 200 205

Asn Ile Ala Thr Leu Phe Thr Met Cys Arg Asp Tyr Gly Arg Ser Ile
210 215 220

Gly Phe Met Gly Asp Phe Tyr Ile Glu Pro Lys Pro Lys Glu Pro Met
225 230 235 240

Lys His Gln Tyr Asp Phe Asp Ala Ala Thr Ala Ile Gly Phe Leu Arg
245 250 255

Lys Tyr Gly Leu Asp Lys Asp Phe Lys Leu Asn Ile Glu Ala Asn His
260 265 270

Ala Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Arg Val Cys Ala
275 280 285

Val Asn Gly Met Met Gly Ser Val Asp Ala Asn Gln Gly Asp Thr Leu
290 295 300

Leu Gly Trp Asp Thr Asp Gln Phe Pro Thr Asn Val Tyr Asp Thr Thr
305 310 315 320

Leu Ala Met Tyr Glu Ile Leu Lys Ala Gly Gly Leu Arg Gly Gly Leu
325 330 335

Asn Phe Asp Ser Lys Asn Arg Arg Pro Ser Asn Thr Ala Asp Asp Met
340 345 350

Phe Tyr Gly Phe Ile Ala Gly Met Asp Thr Phe Ala Leu Gly Leu Ile
355 360 365

Lys Ala Ala Glu Ile Ile Glu Asp Gly Arg Ile Asp Asp Phe Val Lys
370 375 380

Glu Arg Tyr Ala Ser Tyr Asn Ser Gly Ile Gly Lys Lys Ile Arg Asn
385 390 395 400

Arg Lys Val Thr Leu Ile Glu Cys Ala Glu Tyr Ala Ala Lys Leu Lys
405 410 415

Lys Pro Glu Leu Pro Glu Ser Gly Arg Gln Glu Tyr Leu Glu Ser Val
420 425 430

Val Asn Asn Ile Leu Phe Gly Gly Ser Gly
435 440

<210> 18

<211> 437

<212> PRT

<213> Piromyces sp. E2

<400> 18

Met Ala Lys Glu Tyr Phe Pro Gln Ile Gln Lys Ile Lys Phe Glu Gly
 1 5 10 15

Lys Asp Ser Lys Asn Pro Leu Ala Phe His Tyr Tyr Asp Ala Glu Lys
 20 25 30

Glu Val Met Gly Lys Lys Met Lys Asp Trp Leu Arg Phe Ala Met Ala
 35 40 45

Trp Trp His Thr Leu Cys Ala Glu Gly Ala Asp Gln Phe Gly Gly Gly
 50 55 60

Thr Lys Ser Phe Pro Trp Asn Glu Gly Thr Asp Ala Ile Glu Ile Ala
 65 70 75 80

Lys Gln Lys Val Asp Ala Gly Phe Glu Ile Met Gln Lys Leu Gly Ile
 85 90 95

Pro Tyr Tyr Cys Phe His Asp Val Asp Leu Val Ser Glu Gly Asn Ser
 100 105 110

Ile Glu Glu Tyr Glu Ser Asn Leu Lys Ala Val Val Ala Tyr Leu Lys
 115 120 125

Glu Lys Gln Lys Glu Thr Gly Ile Lys Leu Leu Trp Ser Thr Ala Asn
 130 135 140

Val Phe Gly His Lys Arg Tyr Met Asn Gly Ala Ser Thr Asn Pro Asp
 145 150 155 160

Phe Asp Val Val Ala Arg Ala Ile Val Gln Ile Lys Asn Ala Ile Asp
 165 170 175

Ala Gly Ile Glu Leu Gly Ala Glu Asn Tyr Val Phe Trp Gly Gly Arg
 180 185 190

Glu Gly Tyr Met Ser Leu Leu Asn Thr Asp Gln Lys Arg Glu Lys Glu
 195 200 205

His Met Ala Thr Met Leu Thr Met Ala Arg Asp Tyr Ala Arg Ser Lys
 210 215 220

Gly Phe Lys Gly Thr Phe Leu Ile Glu Pro Lys Pro Met Glu Pro Thr
 225 230 235 240

Lys His Gln Tyr Asp Val Asp Thr Glu Thr Ala Ile Gly Phe Leu Lys
 245 250 255

Ala His Asn Leu Asp Lys Asp Phe Lys Val Asn Ile Glu Val Asn His
260 265 270

Ala Thr Leu Ala Gly His Thr Phe Glu His Glu Leu Ala Cys Ala Val
275 280 285

Asp Ala Gly Met Leu Gly Ser Ile Asp Ala Asn Arg Gly Asp Tyr Gln
290 295 300

Asn Gly Trp Asp Thr Asp Gln Phe Pro Ile Asp Gln Tyr Glu Leu Val
305 310 315 320

Gln Ala Trp Met Glu Ile Ile Arg Gly Gly Gly Phe Val Thr Gly Gly
325 330 335

Thr Asn Phe Asp Ala Lys Thr Arg Arg Asn Ser Thr Asp Leu Glu Asp
340 345 350

Ile Ile Ile Ala His Val Ser Gly Met Asp Ala Met Ala Arg Ala Leu
355 360 365

Glu Asn Ala Ala Lys Leu Leu Gln Glu Ser Pro Tyr Thr Lys Met Lys
370 375 380

Lys Glu Arg Tyr Ala Ser Phe Asp Ser Gly Ile Gly Lys Asp Phe Glu
385 390 395 400

Asp Gly Lys Leu Thr Leu Glu Gln Val Tyr Glu Tyr Gly Lys Lys Asn
405 410 415

Gly Glu Pro Lys Gln Thr Ser Gly Lys Gln Glu Leu Tyr Glu Ala Ile
420 425 430

Val Ala Met Tyr Gln
435

<210> 19

<211> 437

<212> PRT

<213> Orpinomyces sp. PC-2

<400> 19

Met Thr Lys Glu Tyr Phe Pro Thr Ile Gly Lys Ile Arg Phe Glu Gly
1 5 10 15

Lys Asp Ser Lys Asn Pro Met Ala Phe His Tyr Tyr Asp Ala Glu Lys
20 25 30

Glu Val Met Gly Lys Lys Met Lys Asp Trp Leu Arg Phe Ala Met Ala

35

40

45

Trp Trp His Thr Leu Cys Ala Asp Gly Ala Asp Gln Phe Gly Val Gly
50 55 60

Thr Lys Ser Phe Pro Trp Asn Glu Gly Thr Asp Pro Ile Ala Ile Ala
65 70 75 80

Lys Gln Lys Val Asp Ala Gly Phe Glu Ile Met Thr Lys Leu Gly Ile
85 90 95

Glu His Tyr Cys Phe His Asp Val Asp Leu Val Ser Glu Gly Asn Ser
100 105 110

Ile Glu Glu Tyr Glu Ser Asn Leu Lys Gln Val Val Ala Tyr Leu Lys
115 120 125

Gln Lys Gln Gln Glu Thr Gly Ile Lys Leu Leu Trp Ser Thr Ala Asn
130 135 140

Val Phe Gly Asn Pro Arg Tyr Met Asn Gly Ala Ser Thr Asn Pro Asp
145 150 155 160

Phe Asp Val Val Ala Arg Ala Ile Val Gln Ile Lys Asn Ala Met Asp
165 170 175

Ala Gly Ile Glu Leu Gly Ala Glu Asn Tyr Val Phe Trp Gly Gly Arg
180 185 190

Glu Gly Tyr Met Ser Leu Leu Asn Thr Asp Gln Lys Arg Glu Lys Glu
195 200 205

His Met Ala Thr Met Leu Thr Met Ala Arg Asp Tyr Ala Arg Ser Lys
210 215 220

Gly Phe Lys Gly Thr Phe Leu Ile Glu Pro Lys Pro Met Glu Pro Thr
225 230 235 240

Lys His Gln Tyr Asp Val Asp Thr Glu Thr Val Ile Gly Phe Leu Arg
245 250 255

Ala His Asn Leu Asp Lys Asp Phe Lys Val Asn Ile Glu Val Asn His
260 265 270

Ala Thr Leu Ala Gly His Thr Phe Glu His Glu Leu Ala Cys Ala Val
275 280 285

Asp Ala Gly Met Leu Gly Ser Ile Asp Ala Asn Arg Gly Asp Tyr Gln

290

295

300

Asn Gly Trp Asp Thr Asp Gln Phe Pro Ile Asp Gln Tyr Glu Leu Val
305 310 315 320

Gln Ala Trp Met Glu Ile Ile Arg Gly Gly Gly Phe Val Thr Gly Gly
325 330 335

Thr Asn Phe Asp Ala Lys Thr Arg Arg Asn Ser Thr Asp Leu Glu Asp
340 345 350

Ile Ile Ile Ala His Ile Ser Gly Met Asp Ala Met Ala Arg Ala Leu
355 360 365

Glu Asn Ala Ala Lys Leu Leu Gln Glu Ser Pro Tyr Cys Asn Met Lys
370 375 380

Lys Glu Arg Tyr Ala Ser Phe Asp Ser Gly Ile Gly Lys Asp Phe Glu
385 390 395 400

Asp Gly Lys Leu Thr Leu Glu Gln Val Tyr Glu Tyr Gly Lys Lys Asn
405 410 415

Gly Glu Pro Lys Val Thr Ser Gly Lys Gln Glu Leu Tyr Glu Ala Ile
420 425 430

Val Ala Met Tyr Gln
435

- <210> 20
- <211> 438
- <212> PRT
- <213> Clostrium phytofermentans

<400> 20

Met Lys Asn Tyr Phe Pro Asn Val Pro Glu Val Lys Tyr Glu Gly Pro
1 5 10 15

Asn Ser Thr Asn Pro Phe Ala Phe Lys Tyr Tyr Asp Ala Glu Arg Ile
20 25 30

Val Ala Gly Lys Thr Met Lys Glu His Cys Arg Phe Ala Leu Ser Trp
35 40 45

Trp His Thr Leu Cys Ala Gly Gly Ala Asp Pro Phe Gly Val Thr Thr
50 55 60

Met Asp Arg Ser Tyr Gly Asn Ile Thr Asp Pro Met Glu Phe Ala Lys
65 70 75 80

Ala Lys Val Asp Ala Gly Phe Glu Leu Met Thr Lys Leu Gly Ile Glu
85 90 95

Tyr Phe Cys Phe His Asp Ala Asp Ile Ala Pro Glu Gly Glu Asn Phe
100 105 110

Glu Glu Ser Lys Lys Asn Leu Phe Val Ile Val Asp Tyr Ile Lys Glu
115 120 125

Lys Met Asp Gln Thr Gly Ile Lys Leu Leu Trp Gly Thr Ala Asn Asn
130 135 140

Phe Gly His Pro Arg Phe Met His Gly Ala Ser Thr Ser Cys Asn Ala
145 150 155 160

Asp Val Phe Ala Tyr Ala Ala Ala Lys Ile Lys Asn Ala Leu Asp Ala
165 170 175

Thr Ile Lys Leu Gly Gly Lys Gly Tyr Val Phe Trp Gly Gly Arg Glu
180 185 190

Gly Tyr Glu Thr Leu Leu Asn Thr Asp Leu Gly Leu Glu Leu Asp Asn
195 200 205

Met Ala Arg Leu Met Lys Met Ala Val Glu Tyr Gly Arg Ala Asn Gly
210 215 220

Phe Asp Gly Asp Phe Tyr Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys
225 230 235 240

His Gln Tyr Asp Phe Asp Thr Ala Thr Val Leu Gly Phe Leu Arg Lys
245 250 255

Tyr Gly Leu Glu Lys Asp Phe Lys Met Asn Ile Glu Ala Asn His Ala
260 265 270

Thr Leu Ala Gly His Thr Phe Glu His Glu Leu Ala Leu Ala Arg Val
275 280 285

Asn Gly Val Phe Gly Ser Val Asp Ala Asn Gln Gly Asp Pro Asn Leu
290 295 300

Gly Trp Asp Thr Asp Gln Phe Pro Thr Asp Val His Ser Ala Thr Leu
305 310 315 320

Ala Met Leu Glu Val Leu Lys Ala Gly Gly Phe Thr Asn Gly Gly Leu
325 330 335

Asn Phe Asp Ala Lys Val Arg Arg Gly Ser Phe Glu Phe Asp Asp Ile
340 345 350

Ala Tyr Gly Tyr Ile Ala Gly Met Asp Thr Phe Ala Leu Gly Leu Ile
355 360 365

Lys Ala Ala Glu Ile Ile Glu Asp Gly Arg Ile Ala Lys Phe Val Glu
370 375 380

Asp Arg Tyr Ala Ser Tyr Lys Thr Gly Ile Gly Lys Ala Ile Val Asp
385 390 395 400

Gly Thr Thr Ser Leu Glu Glu Leu Glu Gln Tyr Val Leu Thr His Asn
405 410 415

Glu Pro Val Met Gln Ser Gly Arg Gln Glu Val Leu Glu Ser Ile Val
420 425 430

Asn Asn Ile Leu Phe Arg
435

<210> 21
<211> 438
<212> PRT
<213> Ruminococcus flavefaciens

<400> 21

Met Glu Phe Phe Ser Asn Ile Gly Lys Ile Gln Tyr Gln Gly Pro Lys
1 5 10 15

Ser Thr Asp Pro Leu Ser Phe Lys Tyr Tyr Asn Pro Glu Glu Val Ile
20 25 30

Asn Gly Lys Thr Met Arg Glu His Leu Lys Phe Ala Leu Ser Trp Trp
35 40 45

His Thr Met Gly Gly Asp Gly Thr Asp Met Phe Gly Cys Gly Thr Thr
50 55 60

Asp Lys Thr Trp Gly Gln Ser Asp Pro Ala Ala Arg Ala Lys Ala Lys
65 70 75 80

Val Asp Ala Ala Phe Glu Ile Met Asp Lys Leu Ser Ile Asp Tyr Tyr
85 90 95

Cys Phe His Asp Arg Asp Leu Ser Pro Glu Tyr Gly Ser Leu Lys Ala
100 105 110

Thr Asn Asp Gln Leu Asp Ile Val Thr Asp Tyr Ile Lys Glu Lys Gln
 115 120 125

Gly Asp Lys Phe Lys Cys Leu Trp Gly Thr Ala Lys Cys Phe Asp His
 130 135 140

Pro Arg Phe Met His Gly Ala Gly Thr Ser Pro Ser Ala Asp Val Phe
 145 150 155 160

Ala Phe Ser Ala Ala Gln Ile Lys Lys Ala Leu Glu Ser Thr Val Lys
 165 170 175

Leu Gly Gly Asn Gly Tyr Val Phe Trp Gly Gly Arg Glu Gly Tyr Glu
 180 185 190

Thr Leu Leu Asn Thr Asn Met Gly Leu Glu Leu Asp Asn Met Ala Arg
 195 200 205

Leu Met Lys Met Ala Val Glu Tyr Gly Arg Ser Ile Gly Phe Lys Gly
 210 215 220

Asp Phe Tyr Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys His Gln Tyr
 225 230 235 240

Asp Phe Asp Thr Ala Thr Val Leu Gly Phe Leu Arg Lys Tyr Gly Leu
 245 250 255

Asp Lys Asp Phe Lys Met Asn Ile Glu Ala Asn His Ala Thr Leu Ala
 260 265 270

Gln His Thr Phe Gln His Glu Leu Arg Val Ala Arg Asp Asn Gly Val
 275 280 285

Phe Gly Ser Ile Asp Ala Asn Gln Gly Asp Val Leu Leu Gly Trp Asp
 290 295 300

Thr Asp Gln Phe Pro Thr Asn Ile Tyr Asp Thr Thr Met Cys Met Tyr
 305 310 315 320

Glu Val Ile Lys Ala Gly Gly Phe Thr Asn Gly Gly Leu Asn Phe Asp
 325 330 335

Ala Lys Ala Arg Arg Gly Ser Phe Thr Pro Glu Asp Ile Phe Tyr Ser
 340 345 350

Tyr Ile Ala Gly Met Asp Ala Phe Ala Leu Gly Phe Arg Ala Ala Leu
 355 360 365

Lys Leu Ile Glu Asp Gly Arg Ile Asp Lys Phe Val Ala Asp Arg Tyr
370 375 380

Ala Ser Trp Asn Thr Gly Ile Gly Ala Asp Ile Ile Ala Gly Lys Ala
385 390 395 400

Asp Phe Ala Ser Leu Glu Lys Tyr Ala Leu Glu Lys Gly Glu Val Thr
405 410 415

Ala Ser Leu Ser Ser Gly Arg Gln Glu Met Leu Glu Ser Ile Val Asn
420 425 430

Asn Val Leu Phe Ser Leu
435

<210> 22
<211> 438
<212> PRT
<213> Bacteroides uniformis

<400> 22

Met Ala Thr Lys Glu Tyr Phe Pro Gly Ile Gly Lys Ile Lys Phe Glu
1 5 10 15

Gly Lys Glu Ser Lys Asn Pro Met Ala Phe Arg Tyr Tyr Asp Ala Asp
20 25 30

Lys Val Ile Met Gly Lys Lys Met Ser Glu Trp Leu Lys Phe Ala Met
35 40 45

Ala Trp Trp His Thr Leu Cys Ala Glu Gly Gly Asp Gln Phe Gly Gly
50 55 60

Gly Thr Lys Lys Phe Pro Trp Asn Gly Glu Ala Asp Lys Val Gln Ala
65 70 75 80

Ala Lys Asn Lys Met Asp Ala Gly Phe Glu Phe Met Gln Lys Met Gly
85 90 95

Ile Glu Tyr Tyr Cys Phe His Asp Val Asp Leu Cys Glu Glu Ala Glu
100 105 110

Thr Ile Glu Glu Tyr Glu Ala Asn Leu Lys Glu Ile Val Ala Tyr Ala
115 120 125

Lys Gln Lys Gln Ala Glu Thr Gly Ile Lys Leu Leu Trp Gly Thr Ala
130 135 140

Asn Val Phe Gly His Ala Arg Tyr Met Asn Gly Ala Ala Thr Asn Pro
 145 150 155 160

Asp Phe Asp Val Val Ala Arg Ala Ala Ile Gln Ile Lys Asn Ala Ile
 165 170 175

Asp Ala Thr Ile Glu Leu Gly Gly Ser Asn Tyr Val Phe Trp Gly Gly
 180 185 190

Arg Glu Gly Tyr Met Ser Leu Leu Asn Thr Asp Gln Lys Arg Glu Lys
 195 200 205

Glu His Leu Ala Gln Met Leu Thr Ile Ala Arg Asp Tyr Ala Arg Ala
 210 215 220

Arg Gly Phe Lys Gly Thr Phe Leu Ile Glu Pro Lys Pro Met Glu Pro
 225 230 235 240

Thr Lys His Gln Tyr Asp Val Asp Thr Glu Thr Val Ile Gly Phe Leu
 245 250 255

Lys Ala His Asn Leu Asp Lys Asp Phe Lys Val Asn Ile Glu Val Asn
 260 265 270

His Ala Thr Leu Ala Gly His Thr Phe Glu His Glu Leu Ala Val Ala
 275 280 285

Val Asp Asn Gly Met Leu Gly Ser Ile Asp Ala Asn Arg Gly Asp Tyr
 290 295 300

Gln Asn Gly Trp Asp Thr Asp Gln Phe Pro Ile Asp Asn Phe Glu Leu
 305 310 315 320

Thr Gln Ala Met Met Gln Ile Ile Arg Asn Gly Gly Phe Gly Asn Gly
 325 330 335

Gly Thr Asn Phe Asp Ala Lys Thr Arg Arg Asn Ser Thr Asp Leu Glu
 340 345 350

Asp Ile Phe Ile Ala His Ile Ala Gly Met Asp Val Met Ala Arg Ala
 355 360 365

Leu Glu Ser Ala Ala Lys Leu Leu Glu Glu Ser Pro Tyr Lys Lys Met
 370 375 380

Leu Ala Asp Arg Tyr Ala Ser Phe Asp Ser Gly Lys Gly Lys Glu Phe
 385 390 395 400

Glu Asp Gly Lys Leu Thr Leu Glu Asp Leu Val Ala Tyr Ala Lys Ala
405 410 415

Asn Gly Glu Pro Lys Gln Thr Ser Gly Lys Gln Glu Leu Tyr Glu Ala
420 425 430

Ile Val Asn Met Tyr Cys
435

<210> 23
<211> 439
<212> PRT
<213> Clostridium cellulolyticum

<400> 23

Met Ser Glu Val Phe Ser Gly Ile Ser Asn Ile Lys Phe Glu Gly Ser
1 5 10 15

Gly Ser Asp Asn Pro Leu Ala Phe Lys Tyr Tyr Asp Pro Lys Ala Val
20 25 30

Ile Gly Gly Lys Thr Met Glu Glu His Leu Arg Phe Ala Val Ala Tyr
35 40 45

Trp His Thr Phe Ala Ala Pro Gly Ala Asp Met Phe Gly Ala Gly Ser
50 55 60

Tyr Val Arg Pro Trp Asn Thr Met Ser Asp Pro Leu Glu Ile Ala Lys
65 70 75 80

Tyr Lys Val Glu Ala Asn Phe Glu Phe Ile Glu Lys Leu Gly Ala Pro
85 90 95

Phe Phe Ala Phe His Asp Arg Asp Ile Ala Pro Glu Gly Asp Thr Leu
100 105 110

Ala Glu Thr Asn Lys Asn Leu Asp Thr Ile Val Ser Val Ile Lys Asp
115 120 125

Arg Met Lys Ser Ser Pro Val Lys Leu Leu Trp Gly Thr Thr Asn Ala
130 135 140

Phe Gly Asn Pro Arg Phe Met His Gly Ala Ser Thr Ser Pro Asn Ala
145 150 155 160

Asp Ile Phe Ala Tyr Ala Ala Ala Gln Val Lys Lys Ala Met Glu Ile
165 170 175

Thr Lys Glu Leu Gly Gly Glu Asn Tyr Val Phe Trp Gly Gly Arg Glu

180

185

190

Gly Tyr Glu Thr Leu Leu Asn Thr Asp Met Lys Leu Glu Leu Asp Asn
 195 200 205

Leu Ala Arg Phe Leu Lys Met Ala Val Asp Tyr Ala Lys Glu Ile Gly
 210 215 220

Phe Asp Gly Gln Phe Leu Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys
 225 230 235 240

His Gln Tyr Asp Phe Asp Thr Ala Thr Val Ile Gly Phe Leu Lys Thr
 245 250 255

Tyr Gly Leu Asp Pro Tyr Phe Lys Met Asn Ile Glu Ala Asn His Ala
 260 265 270

Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Ala Met Cys Arg Ile
 275 280 285

Asn Asp Met Leu Gly Ser Ile Asp Ala Asn Gln Gly Asp Val Met Leu
 290 295 300

Gly Trp Asp Thr Asp Gln Phe Pro Thr Asn Leu Tyr Asp Ala Thr Leu
 305 310 315 320

Ala Met Val Glu Val Leu Lys Ala Gly Gly Leu Lys Lys Gly Gly Leu
 325 330 335

Asn Phe Asp Ser Lys Val Arg Arg Gly Ser Phe Glu Pro Ser Asp Leu
 340 345 350

Phe Tyr Gly His Ile Ala Gly Met Asp Thr Phe Ala Lys Gly Leu Ile
 355 360 365

Ile Ala Asn Lys Ile Val Glu Asp Gly Lys Phe Asp Ala Phe Val Ala
 370 375 380

Asp Arg Tyr Ser Ser Tyr Thr Asn Gly Ile Gly Lys Asp Ile Val Glu
 385 390 395 400

Gly Lys Val Gly Phe Lys Glu Leu Glu Gln Tyr Ala Leu Thr Ala Lys
 405 410 415

Ile Gln Asn Lys Ser Gly Arg Gln Glu Met Leu Glu Ala Leu Leu Asn
 420 425 430

Gln Tyr Ile Leu Glu Thr Lys

<210> 24
 <211> 444
 <212> PRT
 <213> *Thermotoga maritima*

<400> 24

Met Ala Glu Phe Phe Pro Glu Ile Pro Lys Ile Gln Phe Glu Gly Lys
 1 5 10 15

Glu Ser Thr Asn Pro Leu Ala Phe Arg Phe Tyr Asp Pro Asn Glu Val
 20 25 30

Ile Asp Gly Lys Pro Leu Lys Asp His Leu Lys Phe Ser Val Ala Phe
 35 40 45

Trp His Thr Phe Val Asn Glu Gly Arg Asp Pro Phe Gly Asp Pro Thr
 50 55 60

Ala Glu Arg Pro Trp Asn Arg Phe Ser Asp Pro Met Asp Lys Ala Phe
 65 70 75 80

Ala Arg Val Asp Ala Leu Phe Glu Phe Cys Glu Lys Leu Asn Ile Glu
 85 90 95

Tyr Phe Cys Phe His Asp Arg Asp Ile Ala Pro Glu Gly Lys Thr Leu
 100 105 110

Arg Glu Thr Asn Lys Ile Leu Asp Lys Val Val Glu Arg Ile Lys Glu
 115 120 125

Arg Met Lys Asp Ser Asn Val Lys Leu Leu Trp Gly Thr Ala Asn Leu
 130 135 140

Phe Ser His Pro Arg Tyr Met His Gly Ala Ala Thr Thr Cys Ser Ala
 145 150 155 160

Asp Val Phe Ala Tyr Ala Ala Ala Gln Val Lys Lys Ala Leu Glu Ile
 165 170 175

Thr Lys Glu Leu Gly Gly Glu Gly Tyr Val Phe Trp Gly Gly Arg Glu
 180 185 190

Gly Tyr Glu Thr Leu Leu Asn Thr Asp Leu Gly Leu Glu Leu Glu Asn
 195 200 205

Leu Ala Arg Phe Leu Arg Met Ala Val Glu Tyr Ala Lys Lys Ile Gly
 210 215 220

Phe Thr Gly Gln Phe Leu Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys
225 230 235 240

His Gln Tyr Asp Phe Asp Val Ala Thr Ala Tyr Ala Phe Leu Lys Asn
245 250 255

His Gly Leu Asp Glu Tyr Phe Lys Phe Asn Ile Glu Ala Asn His Ala
260 265 270

Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Arg Met Ala Arg Ile
275 280 285

Leu Gly Lys Leu Gly Ser Ile Asp Ala Asn Gln Gly Asp Leu Leu Leu
290 295 300

Gly Trp Asp Thr Asp Gln Phe Pro Thr Asn Ile Tyr Asp Thr Thr Leu
305 310 315 320

Ala Met Tyr Glu Val Ile Lys Ala Gly Gly Phe Thr Lys Gly Gly Leu
325 330 335

Asn Phe Asp Ala Lys Val Arg Arg Ala Ser Tyr Lys Val Glu Asp Leu
340 345 350

Phe Ile Gly His Ile Ala Gly Met Asp Thr Phe Ala Leu Gly Phe Lys
355 360 365

Ile Ala Tyr Lys Leu Ala Lys Asp Gly Val Phe Asp Lys Phe Ile Glu
370 375 380

Glu Lys Tyr Arg Ser Phe Lys Glu Gly Ile Gly Lys Glu Ile Val Glu
385 390 395 400

Gly Lys Thr Asp Phe Glu Lys Leu Glu Glu Tyr Ile Ile Asp Lys Glu
405 410 415

Asp Ile Glu Leu Pro Ser Gly Lys Gln Glu Tyr Leu Glu Ser Leu Leu
420 425 430

Asn Ser Tyr Ile Val Lys Thr Ile Ala Glu Leu Arg
435 440

<210> 25
<211> 441
<212> PRT
<213> Bacillus stearothermophilus

<400> 25

Met Pro Tyr Phe Asp Asn Ile Ser Thr Ile Ala Tyr Glu Gly Pro Ala
 1 5 10 15

Ser Lys Asn Pro Leu Ala Phe Lys Phe Tyr Asn Pro Glu Glu Lys Val
 20 25 30

Gly Asp Lys Thr Met Glu Glu His Leu Arg Phe Ser Val Ala Tyr Trp
 35 40 45

His Thr Phe Thr Gly Asp Gly Ser Asp Pro Phe Gly Ala Gly Asn Met
 50 55 60

Ile Arg Pro Trp Asn Lys Tyr Ser Gly Met Asp Leu Ala Lys Ala Arg
 65 70 75 80

Val Glu Ala Ala Phe Glu Phe Phe Glu Lys Leu Asn Ile Pro Phe Phe
 85 90 95

Cys Phe His Asp Val Asp Ile Ala Pro Glu Gly Glu Thr Leu Lys Glu
 100 105 110

Thr Tyr Lys Asn Leu Asp Ile Ile Val Asp Met Ile Glu Glu Tyr Met
 115 120 125

Lys Thr Ser Lys Thr Lys Leu Leu Trp Asn Thr Ala Asn Leu Phe Thr
 130 135 140

His Pro Arg Phe Val His Gly Ala Ala Thr Ser Cys Asn Ala Asp Val
 145 150 155 160

Phe Ala Tyr Ala Ala Ala Lys Val Lys Lys Gly Leu Glu Ile Ala Lys
 165 170 175

Arg Leu Gly Ala Glu Asn Tyr Val Phe Trp Gly Gly Arg Glu Gly Tyr
 180 185 190

Glu Thr Leu Leu Asn Thr Asp Met Lys Leu Glu Leu Asp Asn Leu Ala
 195 200 205

Arg Phe Leu His Met Ala Val Asp Tyr Ala Lys Glu Ile Gly Phe Asp
 210 215 220

Gly Gln Phe Leu Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys His Gln
 225 230 235 240

Tyr Asp Phe Asp Val Ala Thr Ala Leu Ala Phe Leu Gln Thr Tyr Gly
 245 250 255

Leu Lys Asp Tyr Phe Lys Phe Asn Ile Glu Ala Asn His Ala Thr Leu
260 265 270

Ala Gly His Thr Phe Glu His Glu Leu Arg Val Ala Arg Ile His Gly
275 280 285

Met Leu Gly Ser Val Asp Ala Asn Gln Gly Asp Met Leu Leu Gly Trp
290 295 300

Asp Thr Asp Glu Phe Pro Thr Asp Leu Tyr Ser Thr Thr Leu Ala Met
305 310 315 320

Tyr Glu Ile Leu Lys Asn Gly Gly Leu Gly Arg Gly Gly Leu Asn Phe
325 330 335

Asp Ala Lys Val Arg Arg Gly Ser Phe Glu Pro Glu Asp Leu Phe Tyr
340 345 350

Ala His Ile Ala Gly Met Asp Ser Phe Ala Val Gly Leu Lys Val Ala
355 360 365

His Arg Leu Ile Glu Asp Arg Val Phe Asp Glu Phe Ile Glu Glu Arg
370 375 380

Tyr Lys Ser Tyr Thr Glu Gly Ile Gly Arg Glu Ile Val Glu Gly Thr
385 390 395 400

Ala Asp Phe His Lys Leu Glu Ala His Ala Leu Gln Leu Gly Glu Ile
405 410 415

Gln Asn Gln Ser Gly Arg Gln Glu Arg Leu Lys Thr Leu Leu Asn Gln
420 425 430

Tyr Leu Leu Glu Val Cys Ala Ala Arg
435 440

<210> 26
<211> 438
<212> PRT
<213> Bacteroides stercoris

<400> 26

Met Ala Thr Lys Glu Tyr Phe Pro Gly Ile Gly Lys Ile Lys Phe Glu
1 5 10 15

Gly Lys Glu Ser Lys Asn Pro Met Ala Phe Arg Tyr Tyr Asp Ala Glu
20 25 30

Lys Val Ile Met Gly Lys Lys Met Lys Asp Trp Leu Lys Phe Ser Met
 35 40 45

Ala Trp Trp His Thr Leu Cys Ala Glu Gly Gly Asp Gln Phe Gly Gly
 50 55 60

Gly Thr Lys His Phe Pro Trp Asn Gly Asp Ala Asp Lys Leu Gln Ala
 65 70 75 80

Ala Lys Asn Lys Met Asp Ala Gly Phe Glu Phe Met Gln Lys Met Gly
 85 90 95

Ile Glu Tyr Tyr Cys Phe His Asp Val Asp Leu Cys Asp Glu Ala Asp
 100 105 110

Thr Ile Glu Glu Tyr Glu Ala Asn Leu Lys Ala Ile Val Ala Tyr Ala
 115 120 125

Lys Gln Lys Gln Glu Glu Thr Gly Ile Lys Leu Leu Trp Gly Thr Ala
 130 135 140

Asn Val Phe Gly His Ala Arg Tyr Met Asn Gly Ala Ala Thr Asn Pro
 145 150 155 160

Asp Phe Asp Val Val Ala Arg Ala Ala Val Gln Ile Lys Asn Ala Ile
 165 170 175

Asp Ala Thr Ile Glu Leu Gly Gly Ser Asn Tyr Val Phe Trp Gly Gly
 180 185 190

Arg Glu Gly Tyr Met Ser Leu Leu Asn Thr Asp Gln Lys Arg Glu Lys
 195 200 205

Glu His Leu Ala Gln Met Leu Thr Ile Ala Arg Asp Tyr Ala Arg Ala
 210 215 220

Arg Gly Phe Lys Gly Thr Phe Leu Ile Glu Pro Lys Pro Met Glu Pro
 225 230 235 240

Thr Lys His Gln Tyr Asp Ala Asp Thr Glu Thr Val Val Gly Phe Leu
 245 250 255

Lys Ala His Gly Leu Asp Lys Asp Phe Lys Val Asn Ile Glu Val Asn
 260 265 270

His Ala Thr Leu Ala Gly His Thr Phe Glu His Glu Leu Ala Val Ala
 275 280 285

Val Asp Asn Gly Met Leu Gly Ser Ile Asp Ala Asn Arg Gly Asp Tyr
 290 295 300

Gln Asn Gly Trp Asp Thr Asp Gln Phe Pro Ile Asp Asn Tyr Glu Leu
 305 310 315 320

Thr Gln Ala Met Met Gln Ile Ile Arg Asn Gly Gly Phe Gly Asp Gly
 325 330 335

Gly Thr Asn Phe Asp Ala Lys Thr Arg Arg Asn Ser Thr Asp Leu Glu
 340 345 350

Asp Ile Phe Ile Ala His Ile Ala Gly Met Asp Val Met Ala Arg Ala
 355 360 365

Leu Glu Ser Ala Ala Lys Leu Leu Glu Glu Ser Pro Tyr Lys Lys Met
 370 375 380

Leu Ala Asp Arg Tyr Ala Ser Phe Asp Ser Gly Lys Gly Lys Glu Phe
 385 390 395 400

Glu Glu Gly Lys Leu Thr Leu Glu Asp Val Val Ala Tyr Ala Lys Ala
 405 410 415

Asn Gly Glu Pro Lys Gln Thr Ser Gly Lys Gln Glu Leu Tyr Glu Ala
 420 425 430

Ile Val Asn Met Tyr Cys
 435

<210> 27
 <211> 442
 <212> PRT
 <213> Parabacteroides distasonis

<400> 27

Met Ser Tyr Phe Lys Gly Glu Lys Glu Phe Phe Pro Gly Ile Gly Gln
 1 5 10 15

Ile Gln Phe Glu Gly Arg Glu Ser Lys Asn Pro Leu Ala Phe His Tyr
 20 25 30

Tyr Asp Ala Asp Lys Val Val Met Gly Lys Thr Leu Lys Asp His Leu
 35 40 45

Arg Phe Ala Met Ala Tyr Trp His Thr Leu Cys Ala Glu Gly Gly Asp
 50 55 60

Gln Phe Gly Gly Gly Thr Lys Thr Phe Pro Trp Asn Asp Ser Thr Asp

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

325

330

335

Leu Thr Thr Gly Gly Thr Asn Phe Asp Ala Lys Thr Arg Arg Asn Ser
 340 345 350

Thr Asp Leu Glu Asp Ile Phe Ile Ala His Ile Gly Gly Met Asp Ala
 355 360 365

Phe Ala Arg Ala Leu Met Ile Ala Ala Asp Ile Leu Glu Asn Ser Asp
 370 375 380

Tyr Arg Lys Met Arg Ala Glu Arg Tyr Ala Ser Phe Asp Ala Gly Glu
 385 390 395 400

Gly Lys Ala Phe Glu Asp Gly Lys Leu Thr Leu Glu Asp Leu Arg Thr
 405 410 415

Ile Ala Leu Arg Asp Gly Glu Pro Lys Gln Ile Ser Gly Lys Gln Glu
 420 425 430

Leu Tyr Glu Met Ile Val Asn Leu His Ile
 435 440

<210> 28

<211> 439

<212> PRT

<213> *Prevotella ruminicola*

<400> 28

Met Ala Lys Glu Tyr Phe Pro Phe Thr Gly Lys Ile Pro Phe Glu Gly
 1 5 10 15

Lys Asp Ser Lys Asn Val Met Ala Phe His Tyr Tyr Glu Pro Glu Lys
 20 25 30

Val Val Met Gly Lys Lys Met Lys Asp Trp Leu Lys Phe Ala Met Ala
 35 40 45

Trp Trp His Thr Leu Gly Gly Ala Ser Ala Asp Gln Phe Gly Gly Gln
 50 55 60

Thr Arg Ser Tyr Glu Trp Asp Lys Ala Ala Asp Ala Val Gln Arg Ala
 65 70 75 80

Lys Asp Lys Met Asp Ala Gly Phe Glu Ile Met Asp Lys Leu Gly Ile
 85 90 95

Glu Tyr Phe Cys Phe His Asp Val Asp Leu Val Glu Glu Gly Glu Thr
 100 105 110

Ile Ala Glu Tyr Glu Arg Arg Met Lys Glu Ile Thr Asp Tyr Ala Leu
115 120 125

Val Lys Met Lys Glu Tyr Pro Asn Ile Lys Leu Leu Trp Gly Thr Ala
130 135 140

Asn Val Phe Gly Asn Lys Arg Tyr Ala Asn Gly Ala Ser Thr Asn Pro
145 150 155 160

Asp Phe Asp Val Val Ala Arg Ala Ile Val Gln Ile Lys Asn Ala Ile
165 170 175

Asp Ala Thr Ile Lys Leu Gly Gly Thr Asn Tyr Val Phe Trp Gly Gly
180 185 190

Arg Glu Gly Tyr Met Ser Leu Leu Asn Thr Asp Gln Lys Arg Glu Lys
195 200 205

Glu His Met Ala Thr Met Leu Thr Met Ala Arg Asp Tyr Ala Arg Ala
210 215 220

Lys Gly Phe Lys Gly Thr Phe Leu Ile Glu Pro Lys Pro Met Glu Pro
225 230 235 240

Ser Lys His Gln Tyr Asp Val Asp Thr Glu Thr Val Cys Gly Phe Leu
245 250 255

Arg Ala His Gly Leu Asp Lys Asp Phe Lys Val Asn Ile Glu Val Asn
260 265 270

His Ala Thr Leu Ala Gly His Thr Phe Glu His Glu Leu Ala Cys Ala
275 280 285

Val Asp Asn Gly Met Leu Gly Ser Ile Asp Ala Asn Arg Gly Asp Ala
290 295 300

Gln Asn Gly Trp Asp Thr Asp Gln Phe Pro Ile Asp Asn Phe Glu Leu
305 310 315 320

Thr Gln Ala Met Leu Glu Ile Ile Arg Asn Gly Gly Leu Gly Asn Gly
325 330 335

Gly Thr Asn Phe Asp Ala Lys Ile Arg Arg Asn Ser Thr Asp Leu Glu
340 345 350

Asp Leu Phe Ile Ala His Ile Ser Gly Met Asp Ala Met Ala Arg Ala
355 360 365

Leu Met Asn Ala Ala Ala Ile Leu Glu Glu Ser Glu Leu Pro Lys Met
370 375 380

Lys Lys Glu Arg Tyr Ala Ser Phe Asp Asn Gly Ile Gly Lys Asp Phe
385 390 395 400

Glu Asp Gly Lys Leu Thr Leu Glu Gln Ala Tyr Glu Tyr Gly Lys Lys
405 410 415

Val Glu Glu Pro Lys Gln Thr Ser Gly Lys Gln Glu Lys Tyr Glu Thr
420 425 430

Thr Val Ala Leu Tyr Cys Lys
435

<210> 29
<211> 436
<212> PRT
<213> Agrobacterium tumefaciens

<400> 29

Met Ser Thr Gly Phe Phe Gly Asp Ile Thr Lys Ile Lys Tyr Glu Gly
1 5 10 15

Pro Asp Ser Thr Asn Pro Leu Ala Phe Arg His Tyr Asn Pro Asp Glu
20 25 30

Ile Val Ala Gly Lys Arg Met Glu Asp His Leu Arg Phe Ala Val Ala
35 40 45

Tyr Trp His Thr Phe Thr Trp Pro Gly Gly Asp Pro Phe Gly Gly Gln
50 55 60

Thr Phe Gln Arg Pro Trp Phe Glu Asp Thr Met Gln Ala Ala Lys Leu
65 70 75 80

Lys Ala Asp Val Ala Phe Glu Phe Phe Ser Leu Leu Gly Ser Pro Phe
85 90 95

Tyr Cys Phe His Asp Ala Asp Val Arg Pro Glu Gly Lys Asn Phe Ala
100 105 110

Glu Asn Thr Lys Asn Leu Asn Glu Ile Val Asp Tyr Phe Ala Glu Lys
115 120 125

Gln Ala Gln Thr Gly Val Lys Leu Leu Trp Gly Thr Ala Asn Leu Phe
130 135 140

Ser Asn Arg Arg Phe Met Ser Gly Ala Ala Thr Asn Pro Asp Pro Asp
145 150 155 160

Val Phe Ala Phe Ser Ala Ala Thr Val Lys Thr Cys Leu Asp Ala Thr
165 170 175

Lys Lys Leu Gly Gly Ala Asn Tyr Val Leu Trp Gly Gly Arg Glu Gly
180 185 190

Tyr Glu Thr Leu Leu Asn Thr Asp Leu Ser Arg Glu Leu Asp Gln Met
195 200 205

Gly Arg Phe Leu Asn Leu Val Val Glu Tyr Lys Tyr Lys Ile Gly Phe
210 215 220

Glu Gly Thr Ile Leu Ile Glu Pro Lys Pro Gln Glu Pro Thr Lys His
225 230 235 240

Gln Tyr Asp Tyr Asp Val Ala Thr Val Tyr Ala Phe Leu Gln Lys Asn
245 250 255

Gly Leu Glu Lys Glu Val Lys Val Asn Ile Glu Gln Gly His Ala Ile
260 265 270

Leu Ala Gly His Ser Phe Glu His Glu Leu Ala Leu Ala Asn Ala Phe
275 280 285

Gly Ile Phe Gly Ser Ile Asp Met Asn Arg Asn Asp Tyr Gln Ser Gly
290 295 300

Trp Asp Thr Asp Gln Phe Pro Asn Asn Val Pro Glu Met Ala Leu Ala
305 310 315 320

Tyr Tyr His Val Leu Ala Gly Gly Gly Phe Lys Asn Gly Gly Thr Asn
325 330 335

Phe Asp Ala Lys Leu Arg Arg Gln Ser Leu Asp Pro Gln Asp Leu Leu
340 345 350

Ile Gly His Ile Gly Gly Met Asp Cys Cys Ala Arg Gly Leu Lys Ala
355 360 365

Ala Ala Lys Met Ile Glu Asp Gly Ala Leu Ser Lys Pro Leu Ser Glu
370 375 380

Arg Tyr Ala Lys Trp Asp Ser Ala Glu Ala Gln Lys Met Leu Arg Gly
385 390 395 400

Glu Leu Lys Leu Asp Asp Ile Ala Ala Leu Val Glu Arg Glu Asp Ile
405 410 415

Asn Pro Glu Pro Lys Ser Gly Arg Gln Glu Tyr Leu Glu Asn Val Val
420 425 430

Asn Arg Tyr Val
435

<210> 30

<211> 439

<212> PRT

<213> Clostridium cellulovorans

<400> 30

Met Arg Glu Tyr Phe Ala Asn Val Pro Lys Ile Lys Tyr Glu Gly Lys
1 5 10 15

Asp Ser Lys Asn Pro Leu Ala Phe Lys Tyr Tyr Asn Pro Asp Glu Val
20 25 30

Val Gly Gly Lys Thr Met Lys Glu His Leu Arg Phe Thr Leu Ser Tyr
35 40 45

Trp His Thr Leu Thr Gly Ala Gly Ser Asp Pro Phe Gly Val Gly Thr
50 55 60

Met Leu Arg Pro Trp Asp Cys Ala Glu Asp Glu Met Glu Leu Ala Lys
65 70 75 80

Met Arg Met Glu Ala Asn Phe Glu Leu Met Asp Lys Leu Gly Ile Glu
85 90 95

Tyr Phe Ala Phe His Asp Arg Asp Ile Ala Pro Glu Gly Lys Thr Leu
100 105 110

Ala Asp Thr Asn Glu Lys Leu Asp Glu Ile Val Ala Tyr Cys Lys Glu
115 120 125

Leu Met Gln Lys His Gly Lys Lys Leu Leu Trp Gly Thr Ala Asn Met
130 135 140

Phe Gly Asn Pro Arg Phe Val His Gly Ala Ala Thr Thr Cys Asn Ala
145 150 155 160

Asp Val Phe Ala Tyr Ala Ala Ala Gln Thr Lys Lys Ala Met Asp Val
165 170 175

Thr Lys Glu Leu Gly Gly Glu Asn Tyr Val Phe Trp Gly Gly Arg Glu
 180 185 190

Gly Tyr Glu Thr Leu Leu Asn Thr Asp Leu Gly Leu Glu Gln Asp Asn
 195 200 205

Leu Ala Arg Phe Phe Gln Met Ala Val Asp Tyr Ala Lys Lys Ile Gly
 210 215 220

Phe Thr Gly Gln Phe Leu Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys
 225 230 235 240

His Gln Tyr Asp Phe Asp Val Ala Thr Val Leu Gly Phe Leu Arg Lys
 245 250 255

Tyr Asn Leu Glu Lys Tyr Phe Lys Met Asn Ile Glu Ala Asn His Ala
 260 265 270

Thr Leu Ala Gln His Thr Phe Gln His Glu Val Ala Val Ala Arg Val
 275 280 285

Asn Gly Val Leu Gly Ser Leu Asp Val Asn Gln Gly Asp Pro Asn Leu
 290 295 300

Gly Trp Asp Thr Asp Gln Phe Pro Thr Asn Ile Tyr Asp Ala Thr Met
 305 310 315 320

Val Met Tyr Glu Val Leu Lys Asn Gly Gly Ile Ala Pro Gly Gly Leu
 325 330 335

Asn Phe Asp Ala Lys Thr Arg Arg Ala Ser Phe Glu Pro Glu Asp Leu
 340 345 350

Phe Leu Ser Tyr Ile Ala Gly Met Asp Thr Met Ala Lys Gly Leu Arg
 355 360 365

Val Ala Tyr Ser Leu Leu Asp Asp Ala Val Leu Glu Asn Asn Thr Ser
 370 375 380

Glu Arg Tyr Lys Thr Phe Ser Glu Gly Ile Gly Lys Asp Ile Val Glu
 385 390 395 400

Gly Lys Val Asp Phe Glu Ser Leu Glu Lys Tyr Ala Leu Glu Asn Ser
 405 410 415

Val Ile Ser Asn Lys Ser Gly Arg Gln Glu Tyr Leu Glu Ser Val Val
 420 425 430

Asn Gln Tyr Ile Phe Asn Asp
435

<210> 31

<211> 440

<212> PRT

<213> Burkholderia cenocepacia

<400> 31

Met Ser Tyr Phe Glu His Ile Pro Ala Ile Arg Tyr Glu Gly Pro Gln
1 5 10 15

Ser Asp Asn Pro Leu Ala Tyr His His Tyr Asp Pro Asp Lys Arg Val
20 25 30

Leu Gly Lys Thr Leu Ala Glu His Leu Arg Ile Ala Val Cys Tyr Trp
35 40 45

His Thr Phe Val Trp Pro Gly His Asp Ile Phe Gly Gln Gly Ala Phe
50 55 60

Gln Arg Pro Trp Gln Gln Pro Gly Asp Ala Leu Glu Arg Ala Arg Gln
65 70 75 80

Lys Ala Asp Ala Ala Phe Glu Phe Phe Thr Lys Leu Gly Thr Pro Phe
85 90 95

Tyr Thr Phe His Asp Thr Asp Val Ala Pro Glu Gly Asp Ser Leu Arg
100 105 110

Asp Tyr Ala Ala Asn Phe Ala Arg Met Val Asp Tyr Leu Gly Glu Arg
115 120 125

Gln Gln Ala Ser Gly Val Arg Leu Leu Trp Gly Thr Ala Asn Leu Phe
130 135 140

Ser His Pro Arg Phe Ala Ala Gly Ala Ala Thr Asn Pro Asn Pro Asp
145 150 155 160

Val Phe Ala Trp Ala Ala Thr Gln Val Cys His Ala Leu Asp Ala Thr
165 170 175

His Arg Leu Gly Gly Glu Asn Tyr Val Leu Trp Gly Gly Arg Glu Gly
180 185 190

Tyr Glu Thr Leu Leu Asn Thr Asp Leu Lys Arg Glu Arg Asp Gln Phe
195 200 205

Ala Arg Phe Leu Ser Met Val Val Glu His Lys His Arg Ile Gly Phe

210

215

220

Lys Gly Ala Leu Leu Ile Glu Pro Lys Pro Gln Glu Pro Thr Lys His
225 230 235 240

Gln Tyr Asp Tyr Asp Val Ala Thr Val His Gly Phe Leu Val Gln Tyr
245 250 255

Gly Leu Gln Asn Glu Ile Arg Val Asn Ile Glu Ala Asn His Ala Thr
260 265 270

Leu Ala Gly His Ser Phe His His Glu Ile Ala Asn Ala Phe Ala Leu
275 280 285

Gly Val Phe Gly Ser Val Asp Ala Asn Arg Gly Asp Pro Gln Asn Gly
290 295 300

Trp Asp Thr Asp Gln Phe Pro Asn Ser Val Glu Glu Leu Thr Leu Ala
305 310 315 320

Phe Tyr Glu Ile Leu Arg His Gly Gly Phe Thr Thr Gly Gly Met Asn
325 330 335

Phe Asp Ala Lys Val Arg Arg Gln Ser Ile Asp Pro Glu Asp Leu Phe
340 345 350

Tyr Gly His Val Gly Ala Ile Asp Val Leu Ala Leu Ala Leu Glu Arg
355 360 365

Ala Ala Val Leu Val Glu Asn Asp Arg Leu Asp Ala Leu Arg Arg Gln
370 375 380

Arg Tyr Ala Gln Trp Asp Asp Ala Phe Gly Arg Lys Ile Leu Ser Gly
385 390 395 400

Gly Tyr Thr Leu Glu Ser Leu Ala Ala Asp Ala Leu Ala Arg Gly Val
405 410 415

Asn Pro Arg His Ala Ser Gly Ala Gln Glu Arg Leu Glu Asn Ile Val
420 425 430

Asn Gln Ala Ile Tyr Gly Leu Arg
435 440

- <210> 32
- <211> 439
- <212> PRT
- <213> Lactococcus lactis

<400> 32

Met Ala Tyr Phe Asn Asp Ile Ala Pro Ile Lys Tyr Glu Gly Thr Lys
1 5 10 15

Thr Lys Asn Met Phe Ala Phe Arg His Tyr Asn Pro Glu Gly Val Val
20 25 30

Ala Gly Lys Thr Met Glu Glu Gln Leu His Phe Ala Leu Ala Phe Trp
35 40 45

His Thr Ile Thr Met Asp Gly Ser Asp Pro Phe Gly Gly Ala Thr Met
50 55 60

Glu Arg Pro Trp Asp Leu Glu Gly Gly Ser Glu Leu Asp Arg Ala His
65 70 75 80

Arg Arg Val Asp Ala Phe Phe Glu Ile Ala Glu Lys Leu Gly Val Lys
85 90 95

Tyr Tyr Cys Phe His Asp Ile Asp Ile Ala Pro Thr Gly Asn Ser Leu
100 105 110

Lys Glu Phe Tyr Ala Asn Leu Asp Glu Ile Thr Asp His Leu Leu Glu
115 120 125

Lys Gln Lys Ala Thr Gly Ile Lys Leu Leu Trp Asn Thr Ala Asn Met
130 135 140

Phe Ser Asn Pro Arg Tyr Met Asn Gly Val Ser Thr Ser Asn Arg Ala
145 150 155 160

Glu Val Phe Ala Tyr Gly Ala Ala Gln Val Lys Lys Gly Leu Glu Leu
165 170 175

Ser Lys Lys Leu Gly Gly Glu Asn Tyr Val Phe Trp Gly Gly Arg Glu
180 185 190

Gly Tyr Glu Ser Leu Leu Asn Thr Asp Met Gly Leu Glu Met Asp His
195 200 205

Met Ala Lys Phe Phe His Leu Ala Ile Asp Tyr Ala Lys Ser Ile Asn
210 215 220

His Leu Pro Ile Phe Leu Ile Glu Pro Lys Pro Lys Glu Pro Met Thr
225 230 235 240

His Gln Tyr Asp Phe Asp Ser Ala Thr Ala Leu Ala Phe Leu Gln Lys
245 250 255

Tyr Asp Leu Asp Lys Tyr Phe Lys Leu Asn Leu Glu Thr Asn His Ala
260 265 270

Trp Leu Ala Gly His Thr Phe Glu His Glu Leu Asn Thr Ala Arg Thr
275 280 285

Phe Asn Ala Leu Gly Ser Ile Asp Ala Asn Gln Gly Asn Tyr Leu Leu
290 295 300

Gly Trp Asp Thr Asp Glu Phe Pro Thr Leu Val Ile Asp Ile Thr Leu
305 310 315 320

Ala Met His Gln Ile Leu Leu Asn Gly Gly Leu Gly Lys Gly Gly Ile
325 330 335

Asn Phe Asp Ala Lys Val Arg Arg Thr Ser Phe Lys Ala Glu Asp Leu
340 345 350

Ile Leu Ala His Ile Ala Gly Met Asp Thr Tyr Ala Arg Ala Leu Lys
355 360 365

Gly Ala Ala Ala Ile Ile Glu Asp Lys Phe Leu Ser Asp Ile Val Asp
370 375 380

Glu Arg Tyr Ser Ser Tyr Arg Asn Thr Glu Val Gly Gln Ser Ile Glu
385 390 395 400

Asn Gly Thr Ala Thr Phe Glu Ser Leu Ala Ala Phe Ala Leu Glu Tyr
405 410 415

Gly Asp Asp Ile Glu Leu Asp Ser Asn His Leu Glu Tyr Ile Lys Ser
420 425 430

Val Leu Asn Asp Tyr Leu Val
435

<210> 33
<211> 438
<212> PRT
<213> Thermoanaerobacter thermohydrosulfuricus

<400> 33

Met Glu Tyr Phe Lys Asn Val Pro Gln Ile Lys Tyr Glu Gly Pro Lys
1 5 10 15

Ser Asn Asn Pro Tyr Ala Phe Lys Phe Tyr Asn Pro Asp Glu Ile Ile
20 25 30

Asp Gly Lys Pro Leu Lys Glu His Leu Arg Phe Ser Val Ala Tyr Trp
 35 40 45

His Thr Phe Thr Ala Asn Gly Thr Asp Pro Phe Gly Ala Pro Thr Met
 50 55 60

Gln Arg Pro Trp Asp His Phe Thr Asp Pro Met Asp Ile Ala Lys Ala
 65 70 75 80

Arg Val Glu Ala Ala Phe Glu Leu Phe Glu Lys Leu Asp Val Pro Phe
 85 90 95

Phe Cys Phe His Asp Arg Asp Ile Ala Pro Glu Gly Glu Thr Leu Arg
 100 105 110

Glu Thr Asn Lys Asn Leu Asp Thr Ile Val Ala Met Ile Lys Asp Tyr
 115 120 125

Leu Lys Thr Ser Lys Thr Lys Val Leu Trp Gly Thr Ala Asn Leu Phe
 130 135 140

Ser Asn Pro Arg Phe Val His Gly Ala Ala Thr Ser Cys Asn Ala Asp
 145 150 155 160

Val Phe Ala Tyr Ala Ala Ala Gln Val Lys Lys Ala Leu Glu Ile Thr
 165 170 175

Lys Glu Leu Gly Gly Gln Asn Tyr Val Phe Trp Gly Gly Arg Glu Gly
 180 185 190

Tyr Glu Thr Leu Leu Asn Thr Asp Met Glu Leu Glu Leu Asp Asn Leu
 195 200 205

Ala Arg Phe Leu His Met Ala Val Glu Tyr Ala Gln Glu Ile Gly Phe
 210 215 220

Glu Gly Gln Phe Leu Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys His
 225 230 235 240

Gln Tyr Asp Phe Asp Ala Ala Ser Val His Ala Phe Leu Lys Lys Tyr
 245 250 255

Asp Leu Asp Lys Tyr Phe Lys Leu Asn Ile Glu Ala Asn His Ala Thr
 260 265 270

Leu Ala Gly His Asp Phe Gln His Glu Leu Arg Tyr Ala Arg Ile Asn
 275 280 285

Asn Met Leu Gly Ser Ile Asp Ala Asn Met Gly Asp Met Leu Leu Gly
290 295 300

Trp Asp Thr Asp Gln Tyr Pro Thr Asp Ile Arg Met Thr Thr Leu Ala
305 310 315 320

Met Tyr Glu Val Ile Lys Met Gly Gly Phe Asn Lys Gly Gly Leu Asn
325 330 335

Phe Asp Ala Lys Val Arg Arg Ala Ser Phe Glu Pro Glu Asp Leu Phe
340 345 350

Leu Gly His Ile Ala Gly Met Asp Ala Phe Ala Lys Gly Phe Lys Val
355 360 365

Ala Tyr Lys Leu Val Lys Asp Gly Val Phe Asp Arg Phe Ile Glu Glu
370 375 380

Arg Tyr Lys Ser Tyr Arg Glu Gly Ile Gly Ala Glu Ile Val Ser Gly
385 390 395 400

Lys Ala Asn Phe Lys Thr Leu Glu Glu Tyr Ala Leu Asn Asn Pro Lys
405 410 415

Ile Glu Asn Lys Ser Gly Lys Gln Glu Leu Leu Glu Ser Ile Leu Asn
420 425 430

Gln Tyr Leu Phe Ser Glu
435

<210> 34

<211> 439

<212> PRT

<213> Reticulitermes speratus

<400> 34

Met Ser Gln Ile Phe Lys Asp Ile Pro Val Ile Lys Tyr Glu Gly Pro
1 5 10 15

Ala Ser Lys Asn Pro Leu Ser Phe Lys Tyr Tyr Asp Ala Asn Lys Val
20 25 30

Ile Asp Gly Lys Pro Met Lys Glu His Leu Arg Tyr Ala Met Ala Trp
35 40 45

Trp His Asn Leu Cys Ala Thr Gly Gln Asp Met Phe Gly Pro Gly Thr
50 55 60

Ala Asp Lys Ser Phe Gly Ser Lys Thr Val Gly Thr Met Glu His Ala
 65 70 75 80

His Ala Lys Val Asp Ala Gly Phe Glu Phe Met Ser Lys Leu Gly Val
 85 90 95

Glu Tyr Phe Cys Phe His Asp Ala Asp Leu Val Pro Glu Ala Asp Thr
 100 105 110

Leu Ser Glu Thr Asn Lys Arg Leu Asp Glu Ile Ala Glu His Ile Val
 115 120 125

Ala Lys Gln Lys Ala Thr Gly Ile Lys Cys Leu Trp Gly Thr Ala Asn
 130 135 140

Leu Phe Ser Asn Pro Arg Phe Leu Asn Gly Ser Gly Ser Ser Asn Ser
 145 150 155 160

Ala Asp Val Tyr Ala Tyr Ala Ala Ala Gln Ile Lys Lys Ala Leu Asp
 165 170 175

Leu Thr Val Lys Phe Gly Gly Val Gly Tyr Val Phe Trp Gly Gly Arg
 180 185 190

Glu Gly Tyr Glu Thr Leu Leu Asn Thr Asp Val Lys Phe Glu Gln Glu
 195 200 205

Asn Ile Ala Asn Leu Met His Leu Ala Val Thr Tyr Gly Arg Ser Ile
 210 215 220

Gly Phe Lys Gly Asp Phe Tyr Ile Glu Pro Lys Pro Lys Glu Pro Thr
 225 230 235 240

Lys His Gln Tyr Asp Phe Asp Ala Ala Thr Thr Ile Gly Phe Ile Arg
 245 250 255

Gln Tyr Gly Leu Glu Lys Asp Phe Lys Leu Asn Ile Glu Ala Asn His
 260 265 270

Ala Thr Leu Ala Gly His Thr Phe Gln His Asp Leu Arg Ile Ser Ala
 275 280 285

Ile Asn Gly Met Leu Gly Ser Val Asp Ala Asn Thr Gly Asp Pro Leu
 290 295 300

Leu Gly Trp Asp Thr Asp Glu Phe Pro Tyr Ser Val Tyr Asp Thr Thr
 305 310 315 320

Leu Ala Met Tyr Glu Ile Ile Lys Ala Gly Gly Leu Thr Gly Gly Leu
 325 330 335

Asn Phe Asp Ser Lys Val Arg Arg Pro Ser Tyr Thr His Glu Asp Leu
 340 345 350

Phe Tyr Gly Phe Ile Leu Gly Met Asp Ser Phe Ala Leu Gly Leu Ile
 355 360 365

Lys Ala Lys Ala Leu Ile Ala Asp Gly Arg Leu Asp Ser Phe Val Lys
 370 375 380

Asp Arg Tyr Ala Ser Tyr Gly Ser Gly Ile Gly Ala Lys Ile Arg Asp
 385 390 395 400

His Ser Ala Thr Leu Glu Glu Leu Ala Ala Tyr Ala Leu Ala Lys Asp
 405 410 415

Thr Val Ala Leu Pro Gly Ser Gly Arg Gln Glu Tyr Leu Glu Ser Ile
 420 425 430

Ile Asn Gln Ile Leu Phe Gln
 435

<210> 35
 <211> 438
 <212> PRT
 <213> uncultured bacteria from cow rumen

<400> 35

Met Ser Glu Ile Phe Ala Asn Ile Pro Val Ile Pro Tyr Glu Gly Pro
 1 5 10 15

Gln Ser Lys Asn Pro Leu Ala Phe Lys Phe Tyr Asp Ala Asp Lys Val
 20 25 30

Ile Leu Gly Lys Lys Met Ser Glu His Leu Pro Phe Ala Met Ala Trp
 35 40 45

Trp His Asn Leu Cys Ala Gly Gly Thr Asp Met Phe Gly Arg Asp Thr
 50 55 60

Ala Asp Lys Ser Phe Gly Ala Glu Lys Gly Thr Met Ala His Ala Arg
 65 70 75 80

Ala Lys Val Asp Ala Gly Phe Glu Phe Met Lys Lys Val Gly Val Lys
 85 90 95

Tyr Phe Cys Phe His Asp Val Asp Leu Val Pro Glu Ala Asp Asp Ile

100

105

110

Lys Glu Thr Asn Arg Arg Leu Asp Glu Ile Ser Asp Tyr Ile Leu Glu
 115 120 125

Lys Met Lys Gly Thr Asp Ile Lys Cys Leu Trp Gly Thr Ala Asn Met
 130 135 140

Phe Gly Asn Pro Arg Tyr Met Asn Gly Ala Gly Ser Thr Asn Ser Ala
 145 150 155 160

Asp Val Phe Cys Phe Ala Ala Ala Gln Ile Lys Lys Ala Leu Asp Leu
 165 170 175

Thr Val Lys Leu Gly Gly Arg Gly Tyr Val Phe Trp Gly Gly Arg Glu
 180 185 190

Gly Tyr Glu Thr Leu Leu Asn Thr Asp Met Lys Phe Glu Gln Glu Asn
 195 200 205

Ile Ala Arg Leu Met His Leu Ala Val Asp Tyr Gly Arg Ser Ile Gly
 210 215 220

Phe Thr Gly Asp Phe Tyr Ile Glu Pro Lys Pro Lys Glu Pro Met Lys
 225 230 235 240

His Gln Tyr Asp Phe Asp Ala Ala Thr Ala Ile Gly Phe Leu Arg Gln
 245 250 255

Tyr Gly Leu Asp Lys Asp Phe Lys Met Asn Ile Glu Ala Asn His Ala
 260 265 270

Thr Leu Ala Gly His Thr Phe Gln His Asp Leu Arg Ile Ser Ala Ile
 275 280 285

Asn Gly Met Leu Gly Ser Ile Asp Ala Asn Gln Gly Asp Leu Leu Leu
 290 295 300

Gly Trp Asp Thr Asp Glu Phe Pro Phe Asn Val Tyr Glu Ala Thr Leu
 305 310 315 320

Cys Met Tyr Glu Val Leu Lys Ala Gly Gly Leu Thr Gly Gly Phe Asn
 325 330 335

Phe Asp Ser Lys Thr Arg Arg Pro Ser Tyr Thr Leu Glu Asp Met Phe
 340 345 350

His Ala Tyr Ile Leu Gly Met Asp Thr Phe Ala Leu Gly Leu Ile Lys

355

360

365

Ala Ala Ala Leu Ile Glu Asp Gly Arg Leu Asp Gln Phe Val Ala Asp
370 375 380

Arg Tyr Ala Ser Tyr Lys Thr Gly Ile Gly Ala Lys Ile Arg Ser Gly
385 390 395 400

Glu Thr Thr Leu Ala Glu Leu Ala Ala Tyr Ala Asp Lys Leu Gly Ala
405 410 415

Pro Ala Leu Pro Ser Ser Gly Arg Gln Glu Tyr Leu Glu Ser Ile Val
420 425 430

Asn Ser Ile Leu Phe Gly
435

<210> 36

<211> 438

<212> PRT

<213> uncultured bacteria from cow rumen

<400> 36

Met Ala Glu Ile Phe Lys Gly Ile Pro Glu Ile Arg Tyr Glu Gly Pro
1 5 10 15

Asn Ser Thr Asn Pro Leu Ser Phe Lys Tyr Tyr Asp Pro Asp Lys Val
20 25 30

Ile Leu Gly Lys Pro Met Lys Glu His Leu Pro Phe Ala Met Ala Trp
35 40 45

Trp His Asn Leu Gly Ala Ala Gly Thr Asp Met Phe Gly Arg Asp Thr
50 55 60

Ala Asp Lys Ser Phe Gly Ala Glu Lys Gly Thr Met Glu His Ala Lys
65 70 75 80

Ala Lys Val Asp Ala Gly Phe Glu Phe Met Lys Lys Leu Gly Ile Arg
85 90 95

Tyr Phe Cys Phe His Asp Val Asp Leu Val Pro Glu Ala Asp Asp Ile
100 105 110

Lys Val Thr Asn Ala Arg Leu Asp Glu Ile Ser Asp Tyr Ile Leu Glu
115 120 125

Lys Met Lys Gly Thr Asp Ile Lys Cys Leu Trp Gly Thr Ala Asn Met
130 135 140

Phe Ser Asn Pro Arg Phe Met Asn Gly Ala Gly Ser Thr Asn Ser Ala
145 150 155 160

Asp Val Phe Cys Phe Ala Ala Ala Gln Val Lys Lys Ala Leu Asp Ile
165 170 175

Thr Val Lys Leu Gly Gly Lys Gly Tyr Val Phe Trp Gly Gly Arg Glu
180 185 190

Gly Tyr Glu Thr Leu Leu Asn Thr Asp Val Lys Phe Glu Gln Glu Asn
195 200 205

Ile Ala Lys Leu Met His Leu Ala Val Asp Tyr Gly Arg Ser Ile Gly
210 215 220

Phe Lys Gly Asp Phe Phe Ile Glu Pro Lys Pro Lys Glu Pro Met Lys
225 230 235 240

His Gln Tyr Asp Phe Asp Ala Ala Thr Ala Ile Gly Phe Val Arg Gln
245 250 255

Tyr Gly Leu Asp Lys Asp Phe Lys Met Asn Ile Glu Ala Asn His Ala
260 265 270

Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Arg Ile Ser Ala Ile
275 280 285

Asn Gly Met Leu Gly Ser Ile Asp Ala Asn Gln Gly Asp Met Leu Leu
290 295 300

Gly Trp Asp Thr Asp Glu Phe Pro Phe Asn Val Tyr Asp Thr Thr Leu
305 310 315 320

Cys Met Tyr Glu Val Leu Lys Asn Gly Gly Ile Pro Gly Gly Phe Asn
325 330 335

Phe Asp Ala Lys Asn Arg Arg Pro Ser Tyr Thr Ala Glu Asp Met Phe
340 345 350

Tyr Gly Phe Ile Leu Gly Met Asp Ser Phe Ala Leu Gly Leu Ile Lys
355 360 365

Ala Ala Lys Leu Ile Glu Asp Gly Arg Ile Asp Lys Phe Val Glu Glu
370 375 380

Arg Tyr Ala Ser Tyr Lys Asp Gly Ile Gly Lys Lys Ile Arg Asp Gly
385 390 395 400

Glu Thr Thr Leu Ala Glu Leu Ala Ala Tyr Ala Asp Gln Leu Gly Ala
405 410 415

Pro Lys Leu Pro Gly Ser Gly Arg Gln Glu Asp Leu Glu Ser Val Phe
420 425 430

Asn Gln Val Leu Phe Gly
435

<210> 37

<211> 442

<212> PRT

<213> Lachnospiraceae bacterium

<400> 37

Met Lys Glu Phe Phe Pro Ser Ile Ser Pro Ile Lys Phe Glu Gly Ser
1 5 10 15

Glu Ser Lys Asn Pro Leu Ser Phe Lys Tyr Tyr Asp Ala Lys Arg Val
20 25 30

Ile Met Gly Lys Thr Met Glu Glu His Leu Ser Phe Ala Met Ala Trp
35 40 45

Trp His Asn Leu Cys Ala Ser Gly Val Asp Met Phe Gly Gln Gly Thr
50 55 60

Ala Asp Lys Gly Phe Gly Glu Asn Leu Gly Thr Met Glu His Ala Lys
65 70 75 80

Ala Lys Val Asp Ala Gly Ile Glu Phe Met Gln Lys Leu Gly Ile Lys
85 90 95

Tyr Tyr Cys Phe His Asp Thr Asp Ile Val Pro Glu Asp Gln Glu Asp
100 105 110

Ile Asn Val Thr Asn Ala Arg Leu Asp Glu Ile Thr Asp Tyr Ile Leu
115 120 125

Glu Lys Thr Lys Gly Thr Asp Ile Lys Cys Leu Trp Ala Thr Cys Asn
130 135 140

Met Phe Ser Asn Pro Arg Phe Met Asn Gly Ala Gly Ser Ser Asn Ser
145 150 155 160

Ala Asp Val Phe Cys Phe Ala Ala Ala Gln Ala Lys Lys Gly Leu Glu
165 170 175

Asn Ala Val Lys Leu Gly Ala Lys Gly Phe Val Phe Trp Gly Gly Arg
 180 185 190

Glu Gly Tyr Glu Thr Leu Leu Asn Thr Asp Met Lys Leu Glu Glu Glu
 195 200 205

Asn Ile Ala Thr Leu Phe Thr Met Cys Arg Asp Tyr Gly Arg Ser Ile
 210 215 220

Gly Phe Lys Gly Asp Phe Tyr Ile Glu Pro Lys Pro Lys Glu Pro Met
 225 230 235 240

Lys His Gln Tyr Asp Phe Asp Ala Ala Thr Ala Ile Gly Phe Leu Arg
 245 250 255

Lys Tyr Gly Leu Asp Lys Asp Phe Lys Met Asn Ile Glu Ala Asn His
 260 265 270

Ala Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Arg Val Ser Ala
 275 280 285

Ile Asn Gly Met Leu Gly Ser Val Asp Ala Asn Gln Gly Asp Thr Leu
 290 295 300

Leu Gly Trp Asp Thr Asp Gln Phe Pro Thr Asn Val Tyr Asp Thr Thr
 305 310 315 320

Leu Ala Met Tyr Glu Ile Leu Lys Ala Gly Gly Leu Ser Gly Gly Leu
 325 330 335

Asn Phe Asp Ser Lys Asn Arg Arg Pro Ser Asn Thr Ala Glu Asp Met
 340 345 350

Phe Tyr Gly Phe Ile Ala Gly Met Asp Thr Phe Ala Leu Gly Leu Ile
 355 360 365

Lys Ala Ala Gln Ile Ile Glu Asp Gly Arg Ile Asp Glu Phe Val Lys
 370 375 380

Glu Arg Tyr Ser Ser Tyr Asn Ser Gly Ile Gly Glu Lys Ile Arg Asn
 385 390 395 400

Arg Ser Val Thr Leu Val Glu Cys Ala Glu Tyr Ala Leu Lys Met Lys
 405 410 415

Lys Pro Glu Leu Pro Glu Ser Gly Arg Gln Glu Tyr Leu Glu Thr Val
 420 425 430

Val Asn Asn Ile Phe Phe Asn Ser Lys Leu
435 440

<210> 38
<211> 442
<212> PRT
<213> Lachnospiraceae bacterium

<400> 38

Met Lys Glu Phe Phe Pro Gly Ile Ser Pro Val Lys Phe Glu Gly Ser
1 5 10 15

Glu Ser Lys Asn Pro Leu Ser Phe Lys Tyr Tyr Asp Ala Lys Arg Val
20 25 30

Ile Met Gly Lys Thr Met Glu Glu His Leu Ser Phe Ala Met Ala Trp
35 40 45

Trp His Asn Leu Cys Ala Ser Gly Val Asp Met Phe Gly Gln Gly Thr
50 55 60

Ala Asp Lys Gly Phe Gly Glu Ser Ser Gly Thr Met Gly His Ala Lys
65 70 75 80

Ala Lys Val Asp Ala Gly Ile Glu Phe Met Lys Lys Leu Gly Ile Lys
85 90 95

Tyr Tyr Cys Phe His Asp Thr Asp Ile Val Pro Glu Asp Gln Glu Asp
100 105 110

Ile Asn Val Thr Asn Ala Arg Leu Asp Glu Ile Thr Asp Tyr Ile Leu
115 120 125

Glu Lys Thr Lys Gly Ser Asp Ile Lys Cys Leu Trp Thr Thr Cys Asn
130 135 140

Met Phe Gly Asn Pro Arg Phe Met Asn Gly Ala Gly Ser Ser Asn Ser
145 150 155 160

Ala Asp Val Phe Cys Phe Ala Ala Ala Gln Ala Lys Lys Gly Leu Glu
165 170 175

Asn Ala Val Lys Leu Gly Ala Lys Gly Phe Val Phe Trp Gly Gly Arg
180 185 190

Glu Gly Tyr Glu Thr Leu Leu Asn Thr Asp Met Lys Leu Glu Glu Glu
195 200 205

Asn Ile Ala Thr Leu Phe Thr Met Cys Arg Asp Tyr Gly Arg Ser Ile
 210 215 220

Gly Phe Lys Gly Asp Phe Tyr Ile Glu Pro Lys Pro Lys Glu Pro Met
 225 230 235 240

Lys His Gln Tyr Asp Phe Asp Ala Ala Thr Ala Ile Gly Phe Leu Arg
 245 250 255

Lys Tyr Gly Leu Asp Lys Asp Phe Lys Leu Asn Ile Glu Ala Asn His
 260 265 270

Ala Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Arg Val Ser Ala
 275 280 285

Ile Asn Gly Met Leu Gly Ser Val Asp Ala Asn Gln Gly Asp Thr Leu
 290 295 300

Leu Gly Trp Asp Thr Asp Gln Phe Pro Thr Asn Ile Tyr Asp Thr Thr
 305 310 315 320

Phe Ala Met Tyr Glu Ile Leu Lys Ala Gly Gly Leu Ser Gly Gly Leu
 325 330 335

Asn Phe Asp Ser Lys Asn Arg Arg Pro Ser Asn Thr Ala Glu Asp Met
 340 345 350

Phe Tyr Gly Phe Ile Ala Gly Met Asp Thr Phe Ala Leu Gly Leu Ile
 355 360 365

Lys Ala Ala Gln Ile Ile Glu Asp Gly Arg Ile Asp Glu Phe Ile Lys
 370 375 380

Glu Arg Tyr Ser Ser Tyr Ser Thr Gly Ile Gly Glu Lys Ile Arg Asn
 385 390 395 400

Lys Ser Val Thr Leu Glu Glu Cys Ala Glu Tyr Ala Ala Lys Leu Lys
 405 410 415

Lys Pro Glu Leu Pro Glu Ser Gly Arg Gln Glu Tyr Leu Glu Thr Val
 420 425 430

Val Asn Asn Ile Leu Phe Asn Ser Lys Leu
 435 440

<210> 39
 <211> 439
 <212> PRT
 <213> Lachnospiraceae bacterium

<400> 39

Met Lys Glu Phe Phe Pro Gly Ile Ser Pro Val Lys Phe Glu Gly Lys
1 5 10 15

Asp Ser Lys Asn Pro Leu Ser Phe Lys Tyr Tyr Asp Ala Lys Arg Val
20 25 30

Ile Met Gly Lys Thr Met Glu Glu His Leu Ser Phe Ala Met Ala Trp
35 40 45

Trp His Asn Leu Cys Ala Cys Gly Val Asp Met Phe Gly Gln Gly Thr
50 55 60

Ile Asp Lys Ser Phe Gly Ala Leu Pro Gly Thr Met Glu His Ala Lys
65 70 75 80

Ala Lys Val Asp Ala Gly Ile Glu Phe Met Gln Lys Leu Gly Ile Lys
85 90 95

Tyr Tyr Cys Phe His Asp Thr Asp Ile Val Pro Glu Asp Gln Glu Asp
100 105 110

Ile Asn Val Thr Asn Ala Arg Leu Asp Glu Ile Thr Asp Tyr Ile Leu
115 120 125

Glu Lys Thr Lys Gly Thr Asp Ile Lys Cys Leu Trp Thr Thr Cys Asn
130 135 140

Met Phe Ser Asn Pro Arg Phe Met Asn Gly Ala Gly Ser Ser Asn Ser
145 150 155 160

Ala Asp Val Phe Cys Phe Ala Ala Ala Gln Ala Lys Lys Gly Leu Glu
165 170 175

Asn Ala Val Lys Leu Gly Ala Lys Gly Phe Val Phe Trp Gly Gly Arg
180 185 190

Glu Gly Tyr Glu Thr Leu Leu Asn Thr Asp Met Lys Leu Glu Glu Glu
195 200 205

Asn Ile Ala Thr Leu Phe Thr Met Cys Arg Asp Tyr Gly Arg Ser Ile
210 215 220

Gly Phe Met Gly Asp Phe Tyr Ile Glu Pro Lys Pro Lys Glu Pro Met
225 230 235 240

Lys His Gln Tyr Asp Phe Asp Ala Ala Thr Ala Ile Gly Phe Leu Arg

245

250

255

Lys Tyr Gly Leu Glu Lys Asp Phe Lys Met Asn Ile Glu Ala Asn His
260 265 270

Ala Thr Leu Ala Gly His Thr Phe Gln His Glu Leu Arg Val Cys Ala
275 280 285

Val Asn Gly Met Ile Gly Ser Val Asp Ala Asn Gln Gly Asp Thr Leu
290 295 300

Leu Gly Trp Asp Thr Asp Gln Phe Pro Thr Asn Val Tyr Asp Thr Thr
305 310 315 320

Leu Ala Met Tyr Glu Ile Leu Lys Ala Gly Gly Leu Arg Gly Gly Leu
325 330 335

Asn Phe Asp Ser Lys Asn Arg Arg Pro Ser Asn Thr Ala Asp Asp Met
340 345 350

Phe Tyr Gly Phe Ile Ala Gly Met Asp Ala Phe Ala Leu Gly Leu Ile
355 360 365

Lys Ala Ala Glu Ile Ile Glu Asp Gly Arg Ile Asp Glu Phe Val Lys
370 375 380

Glu Arg Tyr Ser Ser Tyr Asn Ser Gly Ile Gly Glu Lys Ile Arg Asn
385 390 395 400

Arg Ala Val Thr Leu Val Glu Cys Ala Glu Tyr Ala Ala Lys Leu Lys
405 410 415

Lys Pro Glu Leu Pro Asp Ser Gly Lys Gln Glu Tyr Leu Glu Ser Val
420 425 430

Val Asn Asn Ile Leu Phe Gly
435

<210> 40
<211> 443
<212> PRT
<213> uncultured bacterium

<400> 40

Met Ser Val Val Leu Gly Asp Lys Glu Tyr Phe Pro Gly Val Gly Lys
1 5 10 15

Ile Ala Tyr Glu Gly Pro Glu Ser Asp Asn Pro Leu Ser Phe Lys Trp
20 25 30

Tyr Asp Glu Asn Arg Val Val Ala Gly Lys Thr Leu Lys Asp His Phe
 35 40 45

Lys Phe Ala Val Cys Tyr Trp His Thr Phe Cys Gly Ala Gly His Asp
 50 55 60

Ser Phe Gly Pro Gly Pro Phe Val Phe Pro Trp Gly Ala Gly Ser Asp
 65 70 75 80

Ala Leu Ser Arg Ala Lys Met Lys Ala Asp Ala Ala Phe Glu Phe Ile
 85 90 95

Thr Lys Leu Gly Val Pro Tyr Tyr Cys Phe His Asp Ile Asp Leu Ile
 100 105 110

Glu Glu Gly Ser Ser Arg Ala Glu Thr Ala Lys Arg Val Met Asp Ile
 115 120 125

Val Glu Tyr Ala Lys Gln Lys Gln Ala Ala Ser Gly Val Lys Leu Leu
 130 135 140

Trp Gly Thr Ala Asn Leu Phe Ser Asn Pro Arg Tyr Ile Asn Gly Ala
 145 150 155 160

Ser Thr Asn Pro Asp Phe Ala Val Val Ala His Ala Gly Ala Gln Leu
 165 170 175

Lys Asp Ala Leu Asp Ala Thr Ile Ala Leu Gly Gly Glu Asn Tyr Val
 180 185 190

Phe Trp Gly Gly Arg Glu Gly Tyr Met Ser Pro Leu Asn Thr Asp Met
 195 200 205

Lys Arg Glu Val Glu Arg Phe Ala Arg Phe Leu Thr Met Ala Arg Asp
 210 215 220

Tyr Ala Arg Gly Gln Gly Phe Lys Gly Val Phe Phe Ile Glu Pro Lys
 225 230 235 240

Pro Met Glu Pro Ser Lys His Gln Tyr Asp Phe Asp Cys Ala Thr Val
 245 250 255

Ile Gly Phe Leu Arg Gln Tyr Gly Leu Asp Lys Asp Phe Lys Leu Asn
 260 265 270

Ile Glu Thr Asn His Ala Thr Leu Ala Gly His Thr Met Glu His Glu
 275 280 285

Met Gln Val Ala Ala Asp Ala Gly Met Leu Gly Ser Ile Asp Ala Asn
290 295 300

Arg Gly Asp Tyr Gln Asn Ala Trp Asp Thr Asp Gln Phe Pro Asn Asn
305 310 315 320

Ile Asn Glu Thr Thr Glu Met Met Leu Val Leu Leu Arg Ser Gly Gly
325 330 335

Phe Gln Gly Gly Gly Val Asn Phe Asp Ala Lys Arg Arg Arg Asn Ser
340 345 350

Thr Asp Pro Asn Asp Thr Phe His Gly His Ile Gly Gly Met Asp Thr
355 360 365

Phe Ala Arg Ser Leu Ile Ile Ala Gly Asp Ile Leu Glu Lys Ser Pro
370 375 380

Ile Glu Lys Met Arg Lys Glu Arg Tyr Ala Ser Phe Asp Ala Gly Lys
385 390 395 400

Gly Ala Glu Phe Glu Ala Gly Lys Leu Ser Leu Val Gln Leu Ala Glu
405 410 415

Leu Gly Asn Lys Gly Gly Glu Pro Thr Gln Lys Ser Gly Arg Gln Glu
420 425 430

Leu Tyr Glu Asn Ile Met Asn Arg Trp Ile Arg
435 440

<210> 41
<211> 442
<212> PRT
<213> uncultured bacterium

<400> 41

Met Lys Leu Thr Val Gly Asp Lys Glu Tyr Phe Lys Gly Ile Lys Pro
1 5 10 15

Ile Lys Phe Glu Gly Lys Asp Ser Asp Asn Pro Leu Ala Phe Lys Tyr
20 25 30

Tyr Asn Pro Ser Gln Lys Val Gly Lys Lys Thr Met Glu Glu His Phe
35 40 45

Arg Phe Ala Ile Ala Tyr Trp His Thr Phe Cys Gly Thr Gly Gly Asp
50 55 60

Pro Phe Gly Pro Gly Thr Lys Thr Phe Pro Trp Leu Gln Asn Ser Asp
65 70 75 80

Ala Val Gln Arg Ala Tyr Asp Lys Met Asp Ala Ala Phe Glu Phe Ile
85 90 95

Thr Lys Ile Gly Ala Pro Phe Tyr Cys Phe His Asp Tyr Asp Leu Val
100 105 110

Asp Glu Gly Pro Thr Leu Lys Glu Ser Glu Ser Arg Leu Gln Lys Val
115 120 125

Val Asp Tyr Ala Lys Lys Lys Gln Lys Ala Ser Gly Val Lys Leu Leu
130 135 140

Trp Gly Thr Ala Asn Leu Phe Ser His Pro Arg Tyr Met Asn Gly Ala
145 150 155 160

Ala Thr Asn Pro Asp Phe Asp Val Val Cys Tyr Ala Ala Ser Gln Val
165 170 175

Lys Asn Ala Leu Asp Ala Thr Ile Ala Leu Gly Gly Ala Asn Tyr Val
180 185 190

Phe Trp Gly Gly Arg Glu Gly Tyr Met Ser Leu Leu Asn Thr Asn Met
195 200 205

Lys Arg Glu Gln Glu His Met Ala Lys Phe Leu His Met Ala Lys Asp
210 215 220

Tyr Ala Arg Ala Asn Gly Phe Lys Gly Thr Phe Phe Ile Glu Pro Lys
225 230 235 240

Pro Met Glu Pro Ser Lys His Gln Tyr Asp Phe Asp Ser Ala Thr Val
245 250 255

Ile Gly Phe Leu Arg Gln Phe Asp Leu Leu Gly Asp Phe Lys Leu Asn
260 265 270

Ile Glu Val Asn His Ala Thr Leu Ala His His Thr Phe Gln His Glu
275 280 285

Leu Gln Val Ala Ala Asp Ala Gly Ala Leu Gly Ser Ile Asp Ala Asn
290 295 300

Arg Gly Asp Tyr Gln Asn Gly Trp Asp Thr Asp Gln Phe Pro Asn Asn
305 310 315 320

Leu Tyr Glu Leu Ala Glu Ser Met Leu Val Ile Leu Glu Ala Gly Gly
325 330 335

Phe Lys Ser Gly Gly Val Asn Phe Asp Ala Lys Thr Arg Arg Asn Ser
340 345 350

Thr Asp Leu Val Asp Ile Phe His Ala His Ile Gly Gly Met Asp Thr
355 360 365

Phe Ala Arg Ser Leu Leu Ile Ala Gln Ala Val Leu Asp Asn Gly Glu
370 375 380

Tyr Thr Lys Ile Arg Lys Asp Arg Tyr Ser Ser Phe Asp Ser Gly Lys
385 390 395 400

Gly Lys Gln Phe Asp Gln Gly Lys Leu Ser Leu Glu Asp Leu Arg Asn
405 410 415

Leu Ala His Lys Ala Gly Glu Pro Lys Gln Leu Ser Gly Lys Gln Glu
420 425 430

Tyr Ile Glu Asn Leu Ile Ser Arg Phe Ile
435 440

<210> 42
<211> 387
<212> PRT
<213> *Thermus thermophilus*

<400> 42

Met Tyr Glu Pro Lys Pro Glu His Arg Phe Thr Phe Gly Leu Trp Thr
1 5 10 15

Val Gly Asn Val Gly Arg Asp Pro Phe Gly Asp Ala Val Arg Glu Arg
20 25 30

Leu Asp Pro Val Tyr Val Val His Lys Leu Ala Glu Leu Gly Ala Tyr
35 40 45

Gly Val Asn Leu His Asp Glu Asp Leu Ile Pro Arg Gly Thr Pro Pro
50 55 60

Gln Glu Arg Asp Gln Ile Val Arg Arg Phe Lys Lys Ala Leu Asp Glu
65 70 75 80

Thr Gly Leu Lys Val Pro Met Val Thr Ala Asn Leu Phe Ser Asp Pro
85 90 95

Ala Phe Lys Asp Gly Ala Phe Thr Ser Pro Asp Pro Trp Val Arg Ala
 100 105 110

Tyr Ala Leu Arg Lys Ser Leu Glu Thr Met Asp Leu Gly Ala Glu Leu
 115 120 125

Gly Ala Glu Ile Tyr Val Val Trp Pro Gly Arg Glu Gly Ala Glu Val
 130 135 140

Glu Ala Thr Gly Lys Ala Arg Lys Val Trp Asp Trp Val Arg Glu Ala
 145 150 155 160

Leu Asn Phe Met Ala Ala Tyr Ala Glu Asp Gln Gly Tyr Gly Tyr Arg
 165 170 175

Phe Ala Leu Glu Pro Lys Pro Asn Glu Pro Arg Gly Asp Ile Tyr Phe
 180 185 190

Ala Thr Val Gly Ser Met Leu Ala Phe Ile His Thr Leu Asp Arg Pro
 195 200 205

Glu Arg Phe Gly Leu Asn Pro Glu Phe Ala His Glu Thr Met Ala Gly
 210 215 220

Leu Asn Phe Val His Ala Val Ala Gln Ala Leu Asp Ala Gly Lys Leu
 225 230 235 240

Phe His Ile Asp Leu Asn Asp Gln Arg Met Ser Arg Phe Asp Gln Asp
 245 250 255

Leu Arg Phe Gly Ser Glu Asn Leu Lys Ala Ala Phe Phe Leu Val Asp
 260 265 270

Leu Leu Glu Ser Ser Gly Tyr Gln Gly Pro Arg His Phe Asp Ala His
 275 280 285

Ala Leu Arg Thr Glu Asp Glu Glu Gly Val Trp Ala Phe Ala Arg Gly
 290 295 300

Cys Met Arg Thr Tyr Leu Ile Leu Lys Glu Arg Ala Glu Ala Phe Arg
 305 310 315 320

Glu Asp Pro Glu Val Lys Glu Leu Leu Ala Ala Tyr Tyr Gln Glu Asp
 325 330 335

Pro Ala Ala Leu Ala Leu Leu Gly Pro Tyr Ser Arg Glu Lys Ala Glu
 340 345 350

Ala Leu Lys Arg Ala Glu Leu Pro Leu Glu Ala Lys Arg Arg Arg Gly
355 360 365

Tyr Ala Leu Glu Arg Leu Asp Gln Leu Ala Val Glu Tyr Leu Leu Gly
370 375 380

Val Arg Gly
385

<210> 43
<211> 440
<212> PRT
<213> Escherichia coli

<400> 43

Met Gln Ala Tyr Phe Asp Gln Leu Asp Arg Val Arg Tyr Glu Gly Ser
1 5 10 15

Lys Ser Ser Asn Pro Leu Ala Phe Arg His Tyr Asn Pro Asp Glu Leu
20 25 30

Val Leu Gly Lys Arg Met Glu Glu His Leu Arg Phe Ala Ala Cys Tyr
35 40 45

Trp His Thr Phe Cys Trp Asn Gly Ala Asp Met Phe Gly Val Gly Ala
50 55 60

Phe Asn Arg Pro Trp Gln Gln Pro Gly Glu Ala Leu Ala Leu Ala Lys
65 70 75 80

Arg Lys Ala Asp Val Ala Phe Glu Phe Phe His Lys Leu His Val Pro
85 90 95

Phe Tyr Cys Phe His Asp Val Asp Val Ser Pro Glu Gly Ala Ser Leu
100 105 110

Lys Glu Tyr Ile Asn Asn Phe Ala Gln Met Val Asp Val Leu Ala Gly
115 120 125

Lys Gln Glu Glu Ser Gly Val Lys Leu Leu Trp Gly Thr Ala Asn Cys
130 135 140

Phe Thr Asn Pro Arg Tyr Gly Ala Gly Ala Ala Thr Asn Pro Asp Pro
145 150 155 160

Glu Val Phe Ser Trp Ala Ala Thr Gln Val Val Thr Ala Met Glu Ala
165 170 175

Thr His Lys Leu Gly Gly Glu Asn Tyr Val Leu Trp Gly Gly Arg Glu

180

185

190

Gly Tyr Glu Thr Leu Leu Asn Thr Asp Leu Arg Gln Glu Arg Glu Gln
195 200 205

Leu Gly Arg Phe Met Gln Met Val Val Glu His Lys His Lys Ile Gly
210 215 220

Phe Gln Gly Thr Leu Leu Ile Glu Pro Lys Pro Gln Glu Pro Thr Lys
225 230 235 240

His Gln Tyr Asp Tyr Asp Ala Ala Thr Val Tyr Gly Phe Leu Lys Gln
245 250 255

Phe Gly Leu Glu Lys Glu Ile Lys Leu Asn Ile Glu Ala Asn His Ala
260 265 270

Thr Leu Ala Gly His Ser Phe His His Glu Ile Ala Thr Ala Ile Ala
275 280 285

Leu Gly Leu Phe Gly Ser Val Asp Ala Asn Arg Gly Asp Ala Gln Leu
290 295 300

Gly Trp Asp Thr Asp Gln Phe Pro Asn Ser Val Glu Glu Asn Ala Leu
305 310 315 320

Val Met Tyr Glu Ile Leu Lys Ala Gly Gly Phe Thr Thr Gly Gly Leu
325 330 335

Asn Phe Asp Ala Lys Val Arg Arg Gln Ser Thr Asp Lys Tyr Asp Leu
340 345 350

Phe Tyr Gly His Ile Gly Ala Met Asp Thr Met Ala Leu Ala Leu Lys
355 360 365

Ile Ala Ala Arg Met Ile Glu Asp Gly Glu Leu Asp Lys Arg Ile Ala
370 375 380

Gln Arg Tyr Ser Gly Trp Asn Ser Glu Leu Gly Gln Gln Ile Leu Lys
385 390 395 400

Gly Gln Met Ser Leu Ala Asp Leu Ala Lys Tyr Ala Gln Glu His His
405 410 415

Leu Ser Pro Val His Gln Ser Gly Arg Gln Glu Gln Leu Glu Asn Leu
420 425 430

Val Asn His Tyr Leu Phe Asp Lys

435

440

<210> 44
 <211> 600
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 44

Met Leu Cys Ser Val Ile Gln Arg Gln Thr Arg Glu Val Ser Asn Thr
 1 5 10 15

Met Ser Leu Asp Ser Tyr Tyr Leu Gly Phe Asp Leu Ser Thr Gln Gln
 20 25 30

Leu Lys Cys Leu Ala Ile Asn Gln Asp Leu Lys Ile Val His Ser Glu
 35 40 45

Thr Val Glu Phe Glu Lys Asp Leu Pro His Tyr His Thr Lys Lys Gly
 50 55 60

Val Tyr Ile His Gly Asp Thr Ile Glu Cys Pro Val Ala Met Trp Leu
 65 70 75 80

Glu Ala Leu Asp Leu Val Leu Ser Lys Tyr Arg Glu Ala Lys Phe Pro
 85 90 95

Leu Asn Lys Val Met Ala Val Ser Gly Ser Cys Gln Gln His Gly Ser
 100 105 110

Val Tyr Trp Ser Ser Gln Ala Glu Ser Leu Leu Glu Gln Leu Asn Lys
 115 120 125

Lys Pro Glu Lys Asp Leu Leu His Tyr Val Ser Ser Val Ala Phe Ala
 130 135 140

Arg Gln Thr Ala Pro Asn Trp Gln Asp His Ser Thr Ala Lys Gln Cys
 145 150 155 160

Gln Glu Phe Glu Glu Cys Ile Gly Gly Pro Glu Lys Met Ala Gln Leu
 165 170 175

Thr Gly Ser Arg Ala His Phe Arg Phe Thr Gly Pro Gln Ile Leu Lys
 180 185 190

Ile Ala Gln Leu Glu Pro Glu Ala Tyr Glu Lys Thr Lys Thr Ile Ser
 195 200 205

Leu Val Ser Asn Phe Leu Thr Ser Ile Leu Val Gly His Leu Val Glu
 210 215 220

Leu Glu Glu Ala Asp Ala Cys Gly Met Asn Leu Tyr Asp Ile Arg Glu
225 230 235 240

Arg Lys Phe Ser Asp Glu Leu Leu His Leu Ile Asp Ser Ser Ser Lys
245 250 255

Asp Lys Thr Ile Arg Gln Lys Leu Met Arg Ala Pro Met Lys Asn Leu
260 265 270

Ile Ala Gly Thr Ile Cys Lys Tyr Phe Ile Glu Lys Tyr Gly Phe Asn
275 280 285

Thr Asn Cys Lys Val Ser Pro Met Thr Gly Asp Asn Leu Ala Thr Ile
290 295 300

Cys Ser Leu Pro Leu Arg Lys Asn Asp Val Leu Val Ser Leu Gly Thr
305 310 315 320

Ser Thr Thr Val Leu Leu Val Thr Asp Lys Tyr His Pro Ser Pro Asn
325 330 335

Tyr His Leu Phe Ile His Pro Thr Leu Pro Asn His Tyr Met Gly Met
340 345 350

Ile Cys Tyr Cys Asn Gly Ser Leu Ala Arg Glu Arg Ile Arg Asp Glu
355 360 365

Leu Asn Lys Glu Arg Glu Asn Asn Tyr Glu Lys Thr Asn Asp Trp Thr
370 375 380

Leu Phe Asn Gln Ala Val Leu Asp Asp Ser Glu Ser Ser Glu Asn Glu
385 390 395 400

Leu Gly Val Tyr Phe Pro Leu Gly Glu Ile Val Pro Ser Val Lys Ala
405 410 415

Ile Asn Lys Arg Val Ile Phe Asn Pro Lys Thr Gly Met Ile Glu Arg
420 425 430

Glu Val Ala Lys Phe Lys Asp Lys Arg His Asp Ala Lys Asn Ile Val
435 440 445

Glu Ser Gln Ala Leu Ser Cys Arg Val Arg Ile Ser Pro Leu Leu Ser
450 455 460

Asp Ser Asn Ala Ser Ser Gln Gln Arg Leu Asn Glu Asp Thr Ile Val
465 470 475 480

Lys Phe Asp Tyr Asp Glu Ser Pro Leu Arg Asp Tyr Leu Asn Lys Arg
485 490 495

Pro Glu Arg Thr Phe Phe Val Gly Gly Ala Ser Lys Asn Asp Ala Ile
500 505 510

Val Lys Lys Phe Ala Gln Val Ile Gly Ala Thr Lys Gly Asn Phe Arg
515 520 525

Leu Glu Thr Pro Asn Ser Cys Ala Leu Gly Gly Cys Tyr Lys Ala Met
530 535 540

Trp Ser Leu Leu Tyr Asp Ser Asn Lys Ile Ala Val Pro Phe Asp Lys
545 550 555 560

Phe Leu Asn Asp Asn Phe Pro Trp His Val Met Glu Ser Ile Ser Asp
565 570 575

Val Asp Asn Glu Asn Trp Asp Arg Tyr Asn Ser Lys Ile Val Pro Leu
580 585 590

Ser Glu Leu Glu Lys Thr Leu Ile
595 600

<210> 45
<211> 623
<212> PRT
<213> *Pichia stipitis*

<400> 45

Met Thr Thr Thr Pro Phe Asp Ala Pro Asp Lys Leu Phe Leu Gly Phe
1 5 10 15

Asp Leu Ser Thr Gln Gln Leu Lys Ile Ile Val Thr Asp Glu Asn Leu
20 25 30

Ala Ala Leu Lys Thr Tyr Asn Val Glu Phe Asp Ser Ile Asn Ser Ser
35 40 45

Val Gln Lys Gly Val Ile Ala Ile Asn Asp Glu Ile Ser Lys Gly Ala
50 55 60

Ile Ile Ser Pro Val Tyr Met Trp Leu Asp Ala Leu Asp His Val Phe
65 70 75 80

Glu Asp Met Lys Lys Asp Gly Phe Pro Phe Asn Lys Val Val Gly Ile
85 90 95

Ser Gly Ser Cys Gln Gln His Gly Ser Val Tyr Trp Ser Arg Thr Ala
 100 105 110

Glu Lys Val Leu Ser Glu Leu Asp Ala Glu Ser Ser Leu Ser Ser Gln
 115 120 125

Met Arg Ser Ala Phe Thr Phe Lys His Ala Pro Asn Trp Gln Asp His
 130 135 140

Ser Thr Gly Lys Glu Leu Glu Glu Phe Glu Arg Val Ile Gly Ala Asp
 145 150 155 160

Ala Leu Ala Asp Ile Ser Gly Ser Arg Ala His Tyr Arg Phe Thr Gly
 165 170 175

Leu Gln Ile Arg Lys Leu Ser Thr Arg Phe Lys Pro Glu Lys Tyr Asn
 180 185 190

Arg Thr Ala Arg Ile Ser Leu Val Ser Ser Phe Val Ala Ser Val Leu
 195 200 205

Leu Gly Arg Ile Thr Ser Ile Glu Glu Ala Asp Ala Cys Gly Met Asn
 210 215 220

Leu Tyr Asp Ile Glu Lys Arg Glu Phe Asn Glu Glu Leu Leu Ala Ile
 225 230 235 240

Ala Ala Gly Val His Pro Glu Leu Asp Gly Val Glu Gln Asp Gly Glu
 245 250 255

Ile Tyr Arg Ala Gly Ile Asn Glu Leu Lys Arg Lys Leu Gly Pro Val
 260 265 270

Lys Pro Ile Thr Tyr Glu Ser Glu Gly Asp Ile Ala Ser Tyr Phe Val
 275 280 285

Thr Arg Tyr Gly Phe Asn Pro Asp Cys Lys Ile Tyr Ser Phe Thr Gly
 290 295 300

Asp Asn Leu Ala Thr Ile Ile Ser Leu Pro Leu Ala Pro Asn Asp Ala
 305 310 315 320

Leu Ile Ser Leu Gly Thr Ser Thr Thr Val Leu Ile Ile Thr Lys Asn
 325 330 335

Tyr Ala Pro Ser Ser Gln Tyr His Leu Phe Lys His Pro Thr Met Pro
 340 345 350

Asp His Tyr Met Gly Met Ile Cys Tyr Cys Asn Gly Ser Leu Ala Arg
 355 360 365

Glu Lys Val Arg Asp Glu Val Asn Glu Lys Phe Asn Val Glu Asp Lys
 370 375 380

Lys Ser Trp Asp Lys Phe Asn Glu Ile Leu Asp Lys Ser Thr Asp Phe
 385 390 395 400

Asn Asn Lys Leu Gly Ile Tyr Phe Pro Leu Gly Glu Ile Val Pro Asn
 405 410 415

Ala Ala Ala Gln Ile Lys Arg Ser Val Leu Asn Ser Lys Asn Glu Ile
 420 425 430

Val Asp Val Glu Leu Gly Asp Lys Asn Trp Gln Pro Glu Asp Asp Val
 435 440 445

Ser Ser Ile Val Glu Ser Gln Thr Leu Ser Cys Arg Leu Arg Thr Gly
 450 455 460

Pro Met Leu Ser Lys Ser Gly Asp Ser Ser Ala Ser Ser Ser Ala Ser
 465 470 475 480

Pro Gln Pro Glu Gly Asp Gly Thr Asp Leu His Lys Val Tyr Gln Asp
 485 490 495

Leu Val Lys Lys Phe Gly Asp Leu Tyr Thr Asp Gly Lys Lys Gln Thr
 500 505 510

Phe Glu Ser Leu Thr Ala Arg Pro Asn Arg Cys Tyr Tyr Val Gly Gly
 515 520 525

Ala Ser Asn Asn Gly Ser Ile Ile Arg Lys Met Gly Ser Ile Leu Ala
 530 535 540

Pro Val Asn Gly Asn Tyr Lys Val Asp Ile Pro Asn Ala Cys Ala Leu
 545 550 555 560

Gly Gly Ala Tyr Lys Ala Ser Trp Ser Tyr Glu Cys Glu Ala Lys Lys
 565 570 575

Glu Trp Ile Gly Tyr Asp Gln Tyr Ile Asn Arg Leu Phe Glu Val Ser
 580 585 590

Asp Glu Met Asn Ser Phe Glu Val Lys Asp Lys Trp Leu Glu Tyr Ala
 595 600 605

Asn Gly Val Gly Met Leu Ala Lys Met Glu Ser Glu Leu Lys His
610 615 620

<210> 46
<211> 575
<212> PRT
<213> Hypocrea jecorina

<400> 46

Met Ser Glu Glu Lys Gly Pro Leu Tyr Leu Gly Phe Asp Leu Ser Thr
1 5 10 15

Gln Gln Leu Lys Ala Ile Val Val Asn Ser Asn Leu Lys Ser Ile Ala
20 25 30

Glu Ala Lys Val Asp Phe Asp Gln Asp Phe Gly Pro Gln Tyr Gly Ile
35 40 45

Gln Lys Gly Val His Val Arg Glu Ser Thr Gly Glu Val Phe Ala Pro
50 55 60

Val Ala Leu Trp Leu Glu Ser Leu Asp Leu Val Leu Ser Arg Leu Ser
65 70 75 80

Lys Ala Met His Pro Leu Pro Met Ser Arg Ile Arg Gly Val Ser Gly
85 90 95

Ser Gly Gln Gln His Gly Ala Val Phe Trp Asn Ala Ser Ala Glu Glu
100 105 110

Leu Leu Gly Gly Leu Asp Ala Ala Lys Gly Ser Leu Val Glu Gln Leu
115 120 125

Arg Gly Ala Leu Ala His Glu Phe Ala Pro Asn Trp Gln Asp His Ser
130 135 140

Thr Gln Glu Glu Leu Val Ala Phe Asp Ala Glu Leu Gly Asp Arg Glu
145 150 155 160

Lys Leu Ala Glu Val Thr Gly Ser Gly Ala His His Arg Phe Thr Gly
165 170 175

Leu Gln Ile Met Arg Ile Arg Arg Val Leu Pro Gln Val Tyr Ala Asn
180 185 190

Ala Lys Arg Ile Ser Leu Val Ser Ser Trp Leu Ala Ser Val Leu Met
195 200 205

Gly Ser Ile Ala Pro Leu Asp Val Ser Asp Val Cys Gly Met Asn Leu
 210 215 220

Trp Asp Ile Pro Asn Gln Ala Trp Ser Glu Lys Leu Leu Ala Leu Ser
 225 230 235 240

Gly Gly Gly Gly Leu Asp Gly Ala Ala Asn Leu Arg Arg Lys Leu Gly
 245 250 255

Glu Pro Arg Met Asp Gly Gly Gly Ser Met Gly Ser Ile Ser Arg Tyr
 260 265 270

Tyr Val Ser Lys Tyr Gly Phe Ser Pro Glu Cys Gln Ile Thr Pro Phe
 275 280 285

Thr Gly Asp Asn Pro Ala Thr Ile Leu Ala Leu Pro Leu Arg Pro Leu
 290 295 300

Asp Ala Ile Val Ser Leu Gly Thr Ser Thr Thr Phe Leu Met Asn Thr
 305 310 315 320

Pro Ala Tyr Lys Pro Asp Gly Ser Tyr His Phe Phe Asn His Pro Thr
 325 330 335

Thr Pro Gly Asn Tyr Met Phe Met Leu Cys Tyr Lys Asn Gly Gly Leu
 340 345 350

Ala Arg Glu Lys Val Arg Asp Thr Leu Pro Lys Pro Glu Gly Gly Ala
 355 360 365

Thr Gly Trp Glu Thr Phe Asn Glu Ala Ile Met Ala Thr Lys Pro Leu
 370 375 380

Gly Ile Glu Ser Asp Gly Asp Arg Ala Lys Leu Gly Leu Tyr Phe Tyr
 385 390 395 400

Leu Arg Glu Thr Val Pro Asn Ile Arg Ala Gly Thr Trp Arg Phe Thr
 405 410 415

Cys Arg Gln Asp Gly Ser Asp Leu Gln Glu Ala Arg Glu Ala Trp Pro
 420 425 430

Lys Glu Ala Asp Ala Arg Ala Ile Val Glu Ser Gln Ala Leu Ser Met
 435 440 445

Arg Leu Arg Ser Gln Lys Leu Val His Ser Pro Arg Asp Gly Leu Pro
 450 455 460

Ala Gln Pro Arg Arg Ile Tyr Val Val Gly Gly Gly Ser Leu Asn Pro
 465 470 475 480

Ala Ile Thr Arg Val Leu Gly Glu Val Leu Gly Gly Ala Asp Gly Val
 485 490 495

Tyr Lys Leu Asp Val Gly Gly Asn Ala Cys Ala Leu Gly Gly Ala Tyr
 500 505 510

Lys Ala Leu Trp Ala Leu Glu Arg Lys Asp Gly Glu Thr Phe Asp Asp
 515 520 525

Leu Ile Gly Gly Arg Trp Thr Glu Glu Gly Ser Ile Asp Lys Val Asp
 530 535 540

Val Gly Tyr Arg Glu Gly Thr Tyr Glu Arg Tyr Gly Lys Val Leu Gly
 545 550 555 560

Ala Phe Glu Glu Met Glu Arg Arg Leu Leu Ala Glu Glu Glu His
 565 570 575

<210> 47
 <211> 335
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 47

Met Ser Glu Pro Ala Gln Lys Lys Gln Lys Val Ala Asn Asn Ser Leu
 1 5 10 15

Glu Gln Leu Lys Ala Ser Gly Thr Val Val Val Ala Asp Thr Gly Asp
 20 25 30

Phe Gly Ser Ile Ala Lys Phe Gln Pro Gln Asp Ser Thr Thr Asn Pro
 35 40 45

Ser Leu Ile Leu Ala Ala Ala Lys Gln Pro Thr Tyr Ala Lys Leu Ile
 50 55 60

Asp Val Ala Val Glu Tyr Gly Lys Lys His Gly Lys Thr Thr Glu Glu
 65 70 75 80

Gln Val Glu Asn Ala Val Asp Arg Leu Leu Val Glu Phe Gly Lys Glu
 85 90 95

Ile Leu Lys Ile Val Pro Gly Arg Val Ser Thr Glu Val Asp Ala Arg
 100 105 110

Leu Ser Phe Asp Thr Gln Ala Thr Ile Glu Lys Ala Arg His Ile Ile

115

120

125

Lys Leu Phe Glu Gln Glu Gly Val Ser Lys Glu Arg Val Leu Ile Lys
 130 135 140

Ile Ala Ser Thr Trp Glu Gly Ile Gln Ala Ala Lys Glu Leu Glu Glu
 145 150 155 160

Lys Asp Gly Ile His Cys Asn Leu Thr Leu Leu Phe Ser Phe Val Gln
 165 170 175

Ala Val Ala Cys Ala Glu Ala Gln Val Thr Leu Ile Ser Pro Phe Val
 180 185 190

Gly Arg Ile Leu Asp Trp Tyr Lys Ser Ser Thr Gly Lys Asp Tyr Lys
 195 200 205

Gly Glu Ala Asp Pro Gly Val Ile Ser Val Lys Lys Ile Tyr Asn Tyr
 210 215 220

Tyr Lys Lys Tyr Gly Tyr Lys Thr Ile Val Met Gly Ala Ser Phe Arg
 225 230 235 240

Ser Thr Asp Glu Ile Lys Asn Leu Ala Gly Val Asp Tyr Leu Thr Ile
 245 250 255

Ser Pro Ala Leu Leu Asp Lys Leu Met Asn Ser Thr Glu Pro Phe Pro
 260 265 270

Arg Val Leu Asp Pro Val Ser Ala Lys Lys Glu Ala Gly Asp Lys Ile
 275 280 285

Ser Tyr Ile Ser Asp Glu Ser Lys Phe Arg Phe Asp Leu Asn Glu Asp
 290 295 300

Ala Met Ala Thr Glu Lys Leu Ser Glu Gly Ile Arg Lys Phe Ser Ala
 305 310 315 320

Asp Ile Val Thr Leu Phe Asp Leu Ile Glu Lys Lys Val Thr Ala
 325 330 335

<210> 48

<211> 333

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 48

Met Ser Glu Pro Ser Glu Lys Lys Gln Lys Val Ala Thr Ser Ser Leu
 1 5 10 15

Glu Gln Leu Lys Lys Ala Gly Thr His Val Val Ala Asp Ser Gly Asp
 20 25 30

Phe Glu Ala Ile Ser Lys Tyr Glu Pro Gln Asp Ser Thr Thr Asn Pro
 35 40 45

Ser Leu Ile Leu Ala Ala Ser Lys Leu Glu Lys Tyr Ala Arg Phe Ile
 50 55 60

Asp Ala Ala Val Glu Tyr Gly Arg Lys His Gly Lys Thr Asp His Glu
 65 70 75 80

Lys Ile Glu Asn Ala Met Asp Lys Ile Leu Val Glu Phe Gly Thr Gln
 85 90 95

Ile Leu Lys Val Val Pro Gly Arg Val Ser Thr Glu Val Asp Ala Arg
 100 105 110

Leu Ser Phe Asp Lys Lys Ala Thr Val Lys Lys Ala Leu His Ile Ile
 115 120 125

Lys Leu Tyr Lys Asp Ala Gly Val Pro Lys Glu Arg Val Leu Ile Lys
 130 135 140

Ile Ala Ser Thr Trp Glu Gly Ile Gln Ala Ala Arg Glu Leu Glu Val
 145 150 155 160

Lys His Gly Ile His Cys Asn Met Thr Leu Leu Phe Ser Phe Thr Gln
 165 170 175

Ala Val Ala Cys Ala Glu Ala Asn Val Thr Leu Ile Ser Pro Phe Val
 180 185 190

Gly Arg Ile Met Asp Phe Tyr Lys Ala Leu Ser Gly Lys Asp Tyr Thr
 195 200 205

Ala Glu Thr Asp Pro Gly Val Leu Ser Val Lys Lys Ile Tyr Ser Tyr
 210 215 220

Tyr Lys Arg His Gly Tyr Ala Thr Glu Val Met Ala Ala Ser Phe Arg
 225 230 235 240

Asn Leu Asp Glu Leu Lys Ala Leu Ala Gly Ile Asp Asn Met Thr Leu
 245 250 255

Pro Leu Asn Leu Leu Glu Gln Leu Tyr Glu Ser Thr Asp Pro Ile Glu
 260 265 270

Asn Lys Leu Asn Ser Glu Ser Ala Lys Glu Glu Gly Val Glu Lys Val
275 280 285

Ser Phe Ile Asn Asp Glu Pro His Phe Arg Tyr Val Leu Asn Glu Asp
290 295 300

Gln Met Ala Thr Glu Lys Leu Ser Asp Gly Ile Arg Lys Phe Ser Ala
305 310 315 320

Asp Ile Glu Ala Leu Tyr Lys Leu Val Glu Glu Lys Met
325 330

<210> 49

<211> 323

<212> PRT

<213> Scheffersomyces stipitis

<400> 49

Met Ser Ser Asn Ser Leu Glu Gln Leu Lys Ala Thr Gly Thr Val Ile
1 5 10 15

Val Thr Asp Thr Gly Glu Phe Asp Ser Ile Ala Lys Tyr Thr Pro Gln
20 25 30

Asp Ala Thr Thr Asn Pro Ser Leu Ile Leu Ala Ala Ala Lys Lys Pro
35 40 45

Glu Tyr Ala Lys Val Ile Asp Val Ala Ile Glu Tyr Ala Lys Asp Lys
50 55 60

Gly Ser Ser Lys Lys Glu Lys Ala Glu Ile Ala Leu Asp Arg Leu Leu
65 70 75 80

Ile Glu Phe Gly Lys Asn Ile Leu Ala Ile Val Pro Gly Arg Val Ser
85 90 95

Thr Glu Val Asp Ala Arg Leu Ser Phe Asp Lys Glu Ala Thr Ile Lys
100 105 110

Lys Ala Leu Glu Leu Ile Ala Leu Tyr Glu Ser Gln Gly Ile Ser Lys
115 120 125

Asp Arg Ile Leu Ile Lys Ile Ala Ser Thr Trp Glu Gly Ile Gln Ala
130 135 140

Ala Arg Glu Leu Glu Ala Lys His Gly Ile His Cys Asn Leu Thr Leu
145 150 155 160

Leu Phe Ser Phe Val Gln Ala Val Ala Cys Ala Glu Ala Lys Val Thr
165 170 175

Leu Ile Ser Pro Phe Val Gly Arg Ile Leu Asp Trp Tyr Lys Ala Ser
180 185 190

Thr Gly Lys Thr Tyr Glu Gly Asp Glu Asp Pro Gly Val Ile Ser Val
195 200 205

Arg Ala Ile Tyr Asn Tyr Tyr Lys Lys Tyr Gly Tyr Lys Thr Ile Val
210 215 220

Met Gly Ala Ser Phe Arg Asn Thr Gly Glu Ile Lys Ala Leu Ala Gly
225 230 235 240

Cys Asp Tyr Leu Thr Val Ala Pro Lys Leu Leu Glu Glu Leu Leu Asn
245 250 255

Ser Thr Glu Pro Val Pro Gln Val Leu Asp Ala Ala Ser Ala Ser Ala
260 265 270

Thr Asp Val Glu Lys Val Ser Tyr Val Asp Asp Glu Ala Thr Phe Arg
275 280 285

Tyr Leu Phe Asn Glu Asp Ala Met Ala Thr Glu Lys Leu Ala Gln Gly
290 295 300

Ile Arg Ala Phe Gly Lys Asp Ala Val Thr Leu Leu Glu Gln Leu Glu
305 310 315 320

Ala Arg Phe

<210> 50
<211> 680
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 50

Met Thr Gln Phe Thr Asp Ile Asp Lys Leu Ala Val Ser Thr Ile Arg
1 5 10 15

Ile Leu Ala Val Asp Thr Val Ser Lys Ala Asn Ser Gly His Pro Gly
20 25 30

Ala Pro Leu Gly Met Ala Pro Ala Ala His Val Leu Trp Ser Gln Met
35 40 45

Arg Met Asn Pro Thr Asn Pro Asp Trp Ile Asn Arg Asp Arg Phe Val
 50 55 60

Leu Ser Asn Gly His Ala Val Ala Leu Leu Tyr Ser Met Leu His Leu
 65 70 75 80

Thr Gly Tyr Asp Leu Ser Ile Glu Asp Leu Lys Gln Phe Arg Gln Leu
 85 90 95

Gly Ser Arg Thr Pro Gly His Pro Glu Phe Glu Leu Pro Gly Val Glu
 100 105 110

Val Thr Thr Gly Pro Leu Gly Gln Gly Ile Ser Asn Ala Val Gly Met
 115 120 125

Ala Met Ala Gln Ala Asn Leu Ala Ala Thr Tyr Asn Lys Pro Gly Phe
 130 135 140

Thr Leu Ser Asp Asn Tyr Thr Tyr Val Phe Leu Gly Asp Gly Cys Leu
 145 150 155 160

Gln Glu Gly Ile Ser Ser Glu Ala Ser Ser Leu Ala Gly His Leu Lys
 165 170 175

Leu Gly Asn Leu Ile Ala Ile Tyr Asp Asp Asn Lys Ile Thr Ile Asp
 180 185 190

Gly Ala Thr Ser Ile Ser Phe Asp Glu Asp Val Ala Lys Arg Tyr Glu
 195 200 205

Ala Tyr Gly Trp Glu Val Leu Tyr Val Glu Asn Gly Asn Glu Asp Leu
 210 215 220

Ala Gly Ile Ala Lys Ala Ile Ala Gln Ala Lys Leu Ser Lys Asp Lys
 225 230 235 240

Pro Thr Leu Ile Lys Met Thr Thr Thr Ile Gly Tyr Gly Ser Leu His
 245 250 255

Ala Gly Ser His Ser Val His Gly Ala Pro Leu Lys Ala Asp Asp Val
 260 265 270

Lys Gln Leu Lys Ser Lys Phe Gly Phe Asn Pro Asp Lys Ser Phe Val
 275 280 285

Val Pro Gln Glu Val Tyr Asp His Tyr Gln Lys Thr Ile Leu Lys Pro
 290 295 300

Gly Val Glu Ala Asn Asn Lys Trp Asn Lys Leu Phe Ser Glu Tyr Gln
 305 310 315 320

Lys Lys Phe Pro Glu Leu Gly Ala Glu Leu Ala Arg Arg Leu Ser Gly
 325 330 335

Gln Leu Pro Ala Asn Trp Glu Ser Lys Leu Pro Thr Tyr Thr Ala Lys
 340 345 350

Asp Ser Ala Val Ala Thr Arg Lys Leu Ser Glu Thr Val Leu Glu Asp
 355 360 365

Val Tyr Asn Gln Leu Pro Glu Leu Ile Gly Gly Ser Ala Asp Leu Thr
 370 375 380

Pro Ser Asn Leu Thr Arg Trp Lys Glu Ala Leu Asp Phe Gln Pro Pro
 385 390 395 400

Ser Ser Gly Ser Gly Asn Tyr Ser Gly Arg Tyr Ile Arg Tyr Gly Ile
 405 410 415

Arg Glu His Ala Met Gly Ala Ile Met Asn Gly Ile Ser Ala Phe Gly
 420 425 430

Ala Asn Tyr Lys Pro Tyr Gly Gly Thr Phe Leu Asn Phe Val Ser Tyr
 435 440 445

Ala Ala Gly Ala Val Arg Leu Ser Ala Leu Ser Gly His Pro Val Ile
 450 455 460

Trp Val Ala Thr His Asp Ser Ile Gly Val Gly Glu Asp Gly Pro Thr
 465 470 475 480

His Gln Pro Ile Glu Thr Leu Ala His Phe Arg Ser Leu Pro Asn Ile
 485 490 495

Gln Val Trp Arg Pro Ala Asp Gly Asn Glu Val Ser Ala Ala Tyr Lys
 500 505 510

Asn Ser Leu Glu Ser Lys His Thr Pro Ser Ile Ile Ala Leu Ser Arg
 515 520 525

Gln Asn Leu Pro Gln Leu Glu Gly Ser Ser Ile Glu Ser Ala Ser Lys
 530 535 540

Gly Gly Tyr Val Leu Gln Asp Val Ala Asn Pro Asp Ile Ile Leu Val
 545 550 555 560

Ala Thr Gly Ser Glu Val Ser Leu Ser Val Glu Ala Ala Lys Thr Leu
565 570 575

Ala Ala Lys Asn Ile Lys Ala Arg Val Val Ser Leu Pro Asp Phe Phe
580 585 590

Thr Phe Asp Lys Gln Pro Leu Glu Tyr Arg Leu Ser Val Leu Pro Asp
595 600 605

Asn Val Pro Ile Met Ser Val Glu Val Leu Ala Thr Thr Cys Trp Gly
610 615 620

Lys Tyr Ala His Gln Ser Phe Gly Ile Asp Arg Phe Gly Ala Ser Gly
625 630 635 640

Lys Ala Pro Glu Val Phe Lys Phe Phe Gly Phe Thr Pro Glu Gly Val
645 650 655

Ala Glu Arg Ala Gln Lys Thr Ile Ala Phe Tyr Lys Gly Asp Lys Leu
660 665 670

Ile Ser Pro Leu Lys Lys Ala Phe
675 680

<210> 51
<211> 681
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 51

Met Ala Gln Phe Ser Asp Ile Asp Lys Leu Ala Val Ser Thr Leu Arg
1 5 10 15

Leu Leu Ser Val Asp Gln Val Glu Ser Ala Gln Ser Gly His Pro Gly
20 25 30

Ala Pro Leu Gly Leu Ala Pro Val Ala His Val Ile Phe Lys Gln Leu
35 40 45

Arg Cys Asn Pro Asn Asn Glu His Trp Ile Asn Arg Asp Arg Phe Val
50 55 60

Leu Ser Asn Gly His Ser Cys Ala Leu Leu Tyr Ser Met Leu His Leu
65 70 75 80

Leu Gly Tyr Asp Tyr Ser Ile Glu Asp Leu Arg Gln Phe Arg Gln Val
85 90 95

Asn Ser Arg Thr Pro Gly His Pro Glu Phe His Ser Ala Gly Val Glu

			100					105						110		
Ile	Thr	Ser	Gly	Pro	Leu	Gly	Gln	Gly	Ile	Ser	Asn	Ala	Val	Gly	Met	
		115					120					125				
Ala	Ile	Ala	Gln	Ala	Asn	Phe	Ala	Ala	Thr	Tyr	Asn	Glu	Asp	Gly	Phe	
	130					135					140					
Pro	Ile	Ser	Asp	Ser	Tyr	Thr	Phe	Ala	Ile	Val	Gly	Asp	Gly	Cys	Leu	
145					150					155					160	
Gln	Glu	Gly	Val	Ser	Ser	Glu	Thr	Ser	Ser	Leu	Ala	Gly	His	Leu	Gln	
				165					170					175		
Leu	Gly	Asn	Leu	Ile	Thr	Phe	Tyr	Asp	Ser	Asn	Ser	Ile	Ser	Ile	Asp	
			180					185					190			
Gly	Lys	Thr	Ser	Tyr	Ser	Phe	Asp	Glu	Asp	Val	Leu	Lys	Arg	Tyr	Glu	
		195					200					205				
Ala	Tyr	Gly	Trp	Glu	Val	Met	Glu	Val	Asp	Lys	Gly	Asp	Asp	Asp	Met	
	210					215					220					
Glu	Ser	Ile	Ser	Ser	Ala	Leu	Glu	Lys	Ala	Lys	Leu	Ser	Lys	Asp	Lys	
225					230					235					240	
Pro	Thr	Ile	Ile	Lys	Val	Thr	Thr	Thr	Ile	Gly	Phe	Gly	Ser	Leu	Gln	
				245					250					255		
Gln	Gly	Thr	Ala	Gly	Val	His	Gly	Ser	Ala	Leu	Lys	Ala	Asp	Asp	Val	
			260					265					270			
Lys	Gln	Leu	Lys	Lys	Arg	Trp	Gly	Phe	Asp	Pro	Asn	Lys	Ser	Phe	Val	
		275					280					285				
Val	Pro	Gln	Glu	Val	Tyr	Asp	Tyr	Tyr	Lys	Lys	Thr	Val	Val	Glu	Pro	
	290					295					300					
Gly	Gln	Lys	Leu	Asn	Glu	Glu	Trp	Asp	Arg	Met	Phe	Glu	Glu	Tyr	Lys	
305				310						315					320	
Thr	Lys	Phe	Pro	Glu	Lys	Gly	Lys	Glu	Leu	Gln	Arg	Arg	Leu	Asn	Gly	
				325					330					335		
Glu	Leu	Pro	Glu	Gly	Trp	Glu	Lys	His	Leu	Pro	Lys	Phe	Thr	Pro	Asp	
			340					345					350			
Asp	Asp	Ala	Leu	Ala	Thr	Arg	Lys	Thr	Ser	Gln	Gln	Val	Leu	Thr	Asn	

355

360

365

Met Val Gln Val Leu Pro Glu Leu Ile Gly Gly Ser Ala Asp Leu Thr
 370 375 380

Pro Ser Asn Leu Thr Arg Trp Glu Gly Ala Val Asp Phe Gln Pro Pro
 385 390 395 400

Ile Thr Gln Leu Gly Asn Tyr Ala Gly Arg Tyr Ile Arg Tyr Gly Val
 405 410 415

Arg Glu His Gly Met Gly Ala Ile Met Asn Gly Ile Ser Ala Phe Gly
 420 425 430

Ala Asn Tyr Lys Pro Tyr Gly Gly Thr Phe Leu Asn Phe Val Ser Tyr
 435 440 445

Ala Ala Gly Ala Val Arg Leu Ala Ala Leu Ser Gly Asn Pro Val Ile
 450 455 460

Trp Val Ala Thr His Asp Ser Ile Gly Leu Gly Glu Asp Gly Pro Thr
 465 470 475 480

His Gln Pro Ile Glu Thr Leu Ala His Leu Arg Ala Ile Pro Asn Met
 485 490 495

His Val Trp Arg Pro Ala Asp Gly Asn Glu Thr Ser Ala Ala Tyr Tyr
 500 505 510

Ser Ala Ile Lys Ser Gly Arg Thr Pro Ser Val Val Ala Leu Ser Arg
 515 520 525

Gln Asn Leu Pro Gln Leu Glu His Ser Ser Phe Glu Lys Ala Leu Lys
 530 535 540

Gly Gly Tyr Val Ile His Asp Val Glu Asn Pro Asp Ile Ile Leu Val
 545 550 555 560

Ser Thr Gly Ser Glu Val Ser Ile Ser Ile Asp Ala Ala Lys Lys Leu
 565 570 575

Tyr Asp Thr Lys Lys Ile Lys Ala Arg Val Val Ser Leu Pro Asp Phe
 580 585 590

Tyr Thr Phe Asp Arg Gln Ser Glu Glu Tyr Arg Phe Ser Val Leu Pro
 595 600 605

Asp Gly Val Pro Ile Met Ser Phe Glu Val Leu Ala Thr Ser Ser Trp

610

615

620

Gly Lys Tyr Ala His Gln Ser Phe Gly Leu Asp Glu Phe Gly Arg Ser
625 630 635 640

Gly Lys Gly Pro Glu Ile Tyr Lys Leu Phe Asp Phe Thr Ala Asp Gly
645 650 655

Val Ala Ser Arg Ala Glu Lys Thr Ile Asn Tyr Tyr Lys Gly Lys Gln
660 665 670

Leu Leu Ser Pro Met Gly Arg Ala Phe
675 680

<210> 52

<211> 677

<212> PRT

<213> *Pichia stipitis*

<400> 52

Met Ser Ser Val Asp Gln Lys Ala Ile Ser Thr Ile Arg Leu Leu Ala
1 5 10 15

Val Asp Ala Val Ala Ala Ala Asn Ser Gly His Pro Gly Ala Pro Leu
20 25 30

Gly Leu Ala Pro Ala Ala His Ala Val Phe Lys Lys Met Arg Phe Asn
35 40 45

Pro Lys Asp Thr Lys Trp Ile Asn Arg Asp Arg Phe Val Leu Ser Asn
50 55 60

Gly His Ala Cys Ala Leu Leu Tyr Ser Met Leu Val Leu Tyr Gly Tyr
65 70 75 80

Asp Leu Thr Val Glu Asp Leu Lys Lys Phe Arg Gln Leu Gly Ser Lys
85 90 95

Thr Pro Gly His Pro Glu Asn Thr Asp Val Pro Gly Ala Glu Val Thr
100 105 110

Thr Gly Pro Leu Gly Gln Gly Ile Cys Asn Gly Val Gly Ile Ala Leu
115 120 125

Ala Gln Ala Gln Phe Ala Ala Thr Tyr Asn Lys Pro Asp Phe Pro Ile
130 135 140

Ser Asp Ser Tyr Thr Tyr Val Phe Leu Gly Asp Gly Cys Leu Met Glu
145 150 155 160

Gly Val Ser Ser Glu Ala Ser Ser Leu Ala Gly His Leu Gln Leu Gly
 165 170 175

Asn Leu Ile Ala Phe Trp Asp Asp Asn Lys Ile Ser Ile Asp Gly Ser
 180 185 190

Thr Glu Val Ala Phe Thr Glu Asp Val Ile Ala Arg Tyr Lys Ser Tyr
 195 200 205

Gly Trp His Ile Val Glu Val Ser Asp Ala Asp Thr Asp Ile Thr Ala
 210 215 220

Ile Ala Ala Ala Ile Asp Glu Ala Lys Lys Val Thr Asn Lys Pro Thr
 225 230 235 240

Leu Val Arg Leu Thr Thr Thr Ile Gly Phe Gly Ser Leu Ala Gln Gly
 245 250 255

Thr His Gly Val His Gly Ala Pro Leu Lys Ala Asp Asp Ile Lys Gln
 260 265 270

Leu Lys Thr Lys Trp Gly Phe Asn Pro Glu Glu Ser Phe Ala Val Pro
 275 280 285

Ala Glu Val Thr Ala Ser Tyr Asn Glu His Val Ala Glu Asn Gln Lys
 290 295 300

Ile Gln Gln Gln Trp Asn Glu Leu Phe Ala Ala Tyr Lys Gln Lys Tyr
 305 310 315 320

Pro Glu Leu Gly Ala Glu Leu Gln Arg Arg Leu Asp Gly Lys Leu Pro
 325 330 335

Glu Asn Trp Asp Lys Ala Leu Pro Val Tyr Thr Pro Ala Asp Ala Ala
 340 345 350

Val Ala Thr Arg Lys Leu Ser Glu Ile Val Leu Ser Lys Ile Ile Pro
 355 360 365

Glu Val Pro Glu Ile Ile Gly Gly Ser Ala Asp Leu Thr Pro Ser Asn
 370 375 380

Leu Thr Lys Ala Lys Gly Thr Val Asp Phe Gln Pro Ala Ala Thr Gly
 385 390 395 400

Leu Gly Asp Tyr Ser Gly Arg Tyr Ile Arg Tyr Gly Val Arg Glu His
 405 410 415

Ala Met Gly Ala Ile Met Asn Gly Ile Ala Ala Phe Gly Ala Asn Tyr
 420 425 430

Lys Asn Tyr Gly Gly Thr Phe Leu Asn Phe Val Ser Tyr Ala Ala Gly
 435 440 445

Ala Val Arg Leu Ser Ala Leu Ser Glu Phe Pro Ile Thr Trp Val Ala
 450 455 460

Thr His Asp Ser Ile Gly Leu Gly Glu Asp Gly Pro Thr His Gln Pro
 465 470 475 480

Ile Glu Thr Leu Ala His Phe Arg Ala Thr Pro Asn Ile Ser Val Trp
 485 490 495

Arg Pro Ala Asp Gly Asn Glu Thr Ser Ala Ala Tyr Lys Ser Ala Ile
 500 505 510

Glu Ser Thr His Thr Pro His Ile Leu Ala Leu Thr Arg Gln Asn Leu
 515 520 525

Pro Gln Leu Glu Gly Ser Ser Ile Glu Lys Ala Ser Lys Gly Gly Tyr
 530 535 540

Thr Leu Val Gln Gln Asp Lys Ala Asp Ile Ile Ile Val Ala Thr Gly
 545 550 555 560

Ser Glu Val Ser Leu Ala Val Asp Ala Leu Lys Val Leu Glu Gly Gln
 565 570 575

Gly Ile Lys Ala Gly Val Val Ser Leu Pro Asp Gln Leu Thr Phe Asp
 580 585 590

Lys Gln Ser Glu Glu Tyr Lys Leu Ser Val Leu Pro Asp Gly Val Pro
 595 600 605

Ile Leu Ser Val Glu Val Met Ser Thr Phe Gly Trp Ser Lys Tyr Ser
 610 615 620

His Gln Gln Phe Gly Leu Asn Arg Phe Gly Ala Ser Gly Lys Ala Pro
 625 630 635 640

Glu Ile Phe Lys Leu Phe Glu Phe Thr Pro Glu Gly Val Ala Glu Arg
 645 650 655

Ala Ala Lys Thr Val Ala Phe Tyr Lys Gly Lys Asp Val Val Ser Pro
 660 665 670

Leu Arg Ser Ala Phe
675

<210> 53
<211> 348
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 53

Met Ser Ile Pro Glu Thr Gln Lys Gly Val Ile Phe Tyr Glu Ser His
1 5 10 15

Gly Lys Leu Glu Tyr Lys Asp Ile Pro Val Pro Lys Pro Lys Ala Asn
20 25 30

Glu Leu Leu Ile Asn Val Lys Tyr Ser Gly Val Cys His Thr Asp Leu
35 40 45

His Ala Trp His Gly Asp Trp Pro Leu Pro Val Lys Leu Pro Leu Val
50 55 60

Gly Gly His Glu Gly Ala Gly Val Val Val Gly Met Gly Glu Asn Val
65 70 75 80

Lys Gly Trp Lys Ile Gly Asp Tyr Ala Gly Ile Lys Trp Leu Asn Gly
85 90 95

Ser Cys Met Ala Cys Glu Tyr Cys Glu Leu Gly Asn Glu Ser Asn Cys
100 105 110

Pro His Ala Asp Leu Ser Gly Tyr Thr His Asp Gly Ser Phe Gln Gln
115 120 125

Tyr Ala Thr Ala Asp Ala Val Gln Ala Ala His Ile Pro Gln Gly Thr
130 135 140

Asp Leu Ala Gln Val Ala Pro Ile Leu Cys Ala Gly Ile Thr Val Tyr
145 150 155 160

Lys Ala Leu Lys Ser Ala Asn Leu Met Ala Gly His Trp Val Ala Ile
165 170 175

Ser Gly Ala Ala Gly Gly Leu Gly Ser Leu Ala Val Gln Tyr Ala Lys
180 185 190

Ala Met Gly Tyr Arg Val Leu Gly Ile Asp Gly Gly Glu Gly Lys Glu
195 200 205

Glu Leu Phe Arg Ser Ile Gly Gly Glu Val Phe Ile Asp Phe Thr Lys
210 215 220

Glu Lys Asp Ile Val Gly Ala Val Leu Lys Ala Thr Asp Gly Gly Ala
225 230 235 240

His Gly Val Ile Asn Val Ser Val Ser Glu Ala Ala Ile Glu Ala Ser
245 250 255

Thr Arg Tyr Val Arg Ala Asn Gly Thr Thr Val Leu Val Gly Met Pro
260 265 270

Ala Gly Ala Lys Cys Cys Ser Asp Val Phe Asn Gln Val Val Lys Ser
275 280 285

Ile Ser Ile Val Gly Ser Tyr Val Gly Asn Arg Ala Asp Thr Arg Glu
290 295 300

Ala Leu Asp Phe Phe Ala Arg Gly Leu Val Lys Ser Pro Ile Lys Val
305 310 315 320

Val Gly Leu Ser Thr Leu Pro Glu Ile Tyr Glu Lys Met Glu Lys Gly
325 330 335

Gln Ile Val Gly Arg Tyr Val Val Asp Thr Ser Lys
340 345

- <210> 54
- <211> 1329
- <212> DNA
- <213> Eubacterium saburreum

<400> 54
atgaaggaat tcttcccagg tattttctcca gttaagttcg aaggtagaga ctctaagaac 60
ccattgtctt tcaagtacta cgacgctaag agagttatta tgggtaagac catggaagaa 120
cacttgtctt tcgctatggc ttggtggcac aacttgtgtg cttgtggtgt tgacatgttc 180
ggtaaggta ccggtgacaa gtctttcggt gaatcttctg gtacatgga acacgctaga 240
gctaaggttg acgctggtat tgaattcatg aagaagttgg gtattaagta ctactgtttc 300
cacgacaccg acattgttcc agaagaccaa gaagacatta acgttaccaa cgctagattg 360
gacgaaatta ccgactacat tttggaaaag accaaggaca ccgacattaa gtgtttgtgg 420
accacctgta acatgttctc taaccaaga ttcataaacg gtgctgggtc ttctaactct 480
gctgacgttt tctgtttcgc tgctgctcaa gctaagaagg gtttggaaaa cgctgttaag 540
ttgggtgcta agggtttctg tttctggggg ggtagagaag gttacgaaac cttgttgaac 600
accgacatga agttggaaga agaaaacatt gctaccttgt tcacatgtg tagagactac 660

ggtagatcta ttggtttcat gggtgacttc tacattgaac caaagccaaa ggaaccaatg	720
aagcaccaat acgacttcga cgctgctacc gctattgggt tcttgagaaa gtacggtttg	780
gacaaggact tcaagttgaa cattgaagct aaccacgcta ccttggtctgg tcacaccttc	840
caacacgaat tgagagtttg tgctgttaac ggtatgatgg gttctggtga cgctaaccaa	900
ggtgacacct tgttgggttg ggacaccgac caattcccaa ccaacgttta cgacaccacc	960
ttggctatgt acgaaatfff gaaggctggg ggtttgagag gtggtttgaa cttcgactct	1020
aagaacagaa gaccatctaa caccgctgac gacatgttct acggtttcat tgctgggatg	1080
gacaccttcg ctttgggttt gattaaggct gctgaaatta ttgaagacgg tagaattgac	1140
gacttcgtta aggaaagata cgcttcttac aactctggta ttggtaagaa gattagaaac	1200
agaaaggtta ccttgattga atgtgctgaa tacgctgcta agttgaagaa gccagaattg	1260
ccagaatctg gtagacaaga atacttggaa tctgttgtca acaacatfff gttcgggtgg	1320
tctgggtaa	1329

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Последовательность олигонуклеотида, характеризующегося тем, что он повышает скорость транскрипции РНК типа фрагмента матричной РНК, кодирующего белок, выбранный из группы, состоящей из ферментов, структурных белков, коферментов, переносчиков, антител, гормонов и регуляторов, типа фрагмента регуляторной РНК, типа фрагмента ферментативно активной РНК или типа фрагмента транспортной РНК, причем последовательность данного олигонуклеотида по меньшей мере на 80% идентична последовательности SEQ ID NO: 1.

2. Последовательность олигонуклеотида по п. 1, где последовательность данного олигонуклеотида по меньшей мере на 85% идентична последовательности SEQ ID NO: 1.

3. Последовательность олигонуклеотида по п. 1 или 2, при этом скорость транскрипции фрагмента РНК в дрожжевых клетках-хозяевах повышается по меньшей мере в 2 раза при выращивании дрожжевых клеток-хозяев, трансформированных по меньшей мере одним рекомбинантным фрагментом ДНК, включающим этот олигонуклеотид, как минимум на двух субстратах, выбранных из группы, состоящей из глюкозы, маннозы, фруктозы, галактозы, ксилозы, арабинозы, сахарозы, трегалозы, рафинозы, глицерина, этанола, ацетата и лактата.

4. Последовательность олигонуклеотида по п. 1 или 2, при этом активность фермента, кодируемого фрагментом РНК, контролируемым этим олигонуклеотидом в дрожжевых клетках-хозяевах, повышается по меньшей мере в 2 раза при выращивании дрожжевых клеток-хозяев, трансформированных по меньшей мере одним рекомбинантным фрагментом ДНК, включающим этот олигонуклеотид, как минимум на двух субстратах, выбранных из группы, состоящей из глюкозы, маннозы, фруктозы, галактозы, ксилозы, арабинозы, сахарозы, трегалозы, рафинозы, глицерина, этанола, ацетата и лактата.

5. Последовательность олигонуклеотида по любому из предыдущих пунктов, при этом фермент является модифицирующим углеводы ферментом.

6. Последовательность олигонуклеотида по любому из предыдущих пунктов, при этом фермент является модифицирующим углеводы ферментом, выбранным из группы, состоящей из ЕС 5.1.3, ЕС 5.3.1, ЕС 2.7.1, ЕС 2.2.1, ЕС 2.2.1 и ЕС 1.1.1.

7. Последовательность олигонуклеотида по любому из предыдущих пунктов, при этом белок выбран из группы, состоящей из SEQ ID NO: 11–53.

8. Последовательность олигонуклеотида по любому из предыдущих пунктов, при этом от 1 до 80 нуклеотидов подверглись замене, делеции или вставке.

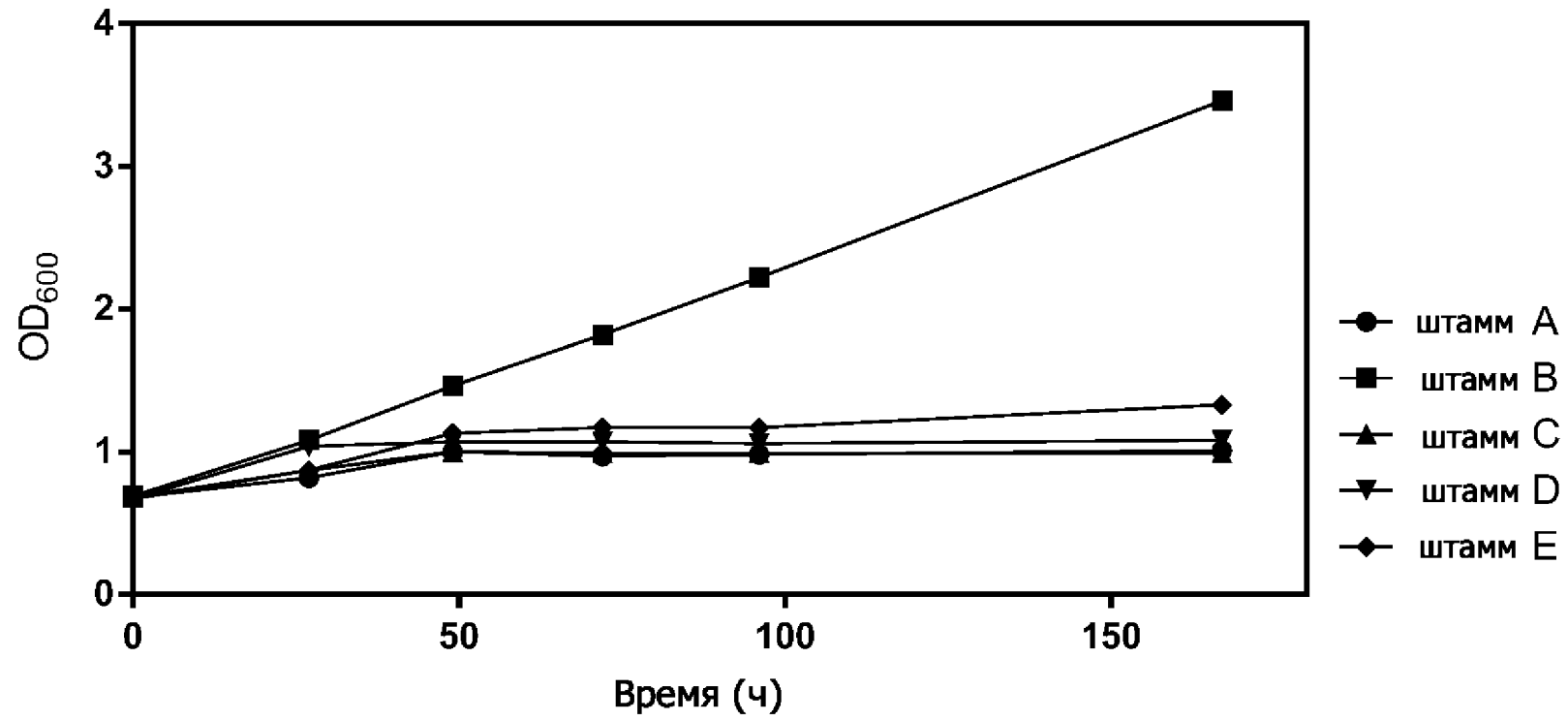
9. Рекомбинантный фрагмент ДНК, включающий последовательность

олигонуклеотида по любому из п.п. 1–8.

10. Экспрессирующая плазмида, содержащая по меньшей мере один рекомбинантный фрагмент ДНК по п. 9.

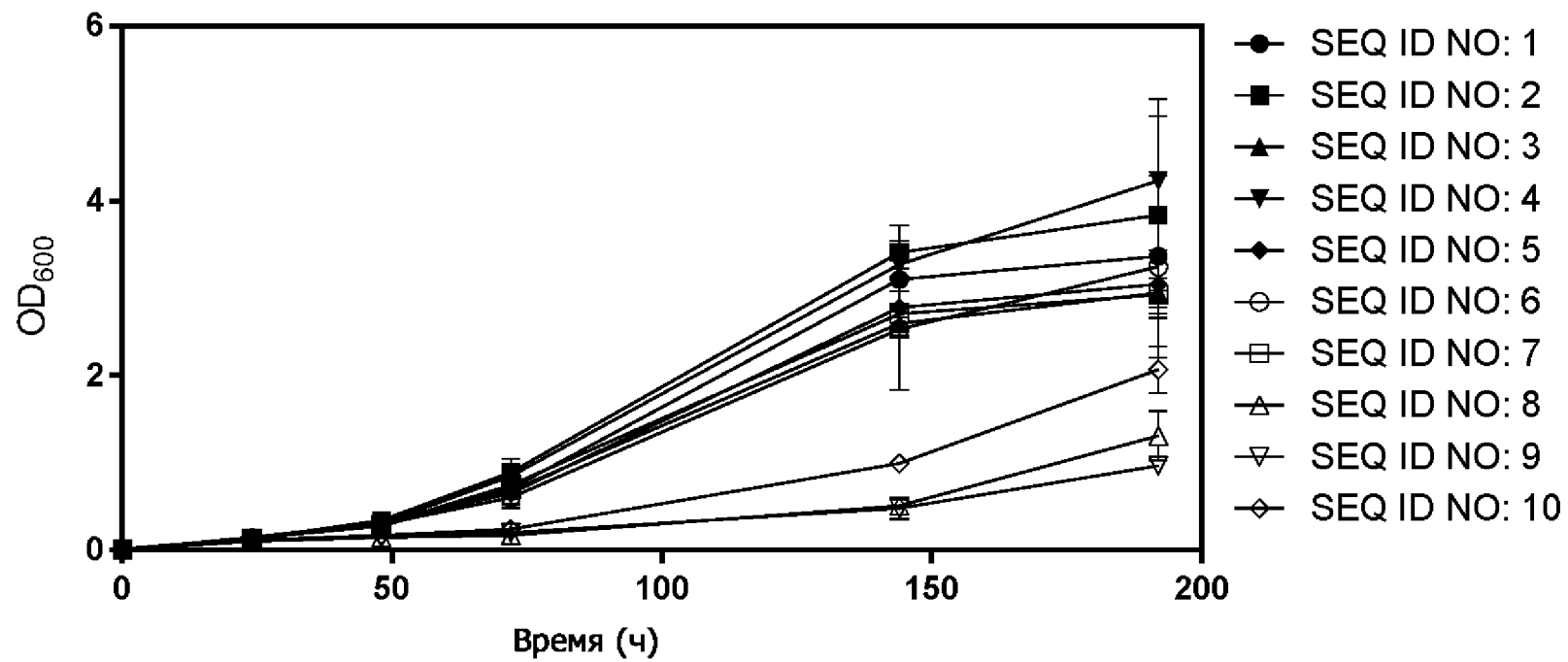
11. Клетка-хозяин, трансформированная по меньшей мере одним рекомбинантным фрагментом ДНК по п. 9 либо трансформированная по меньшей мере одной экспрессирующей плазмидой по п. 10.

Исследование роста – отбор штамма



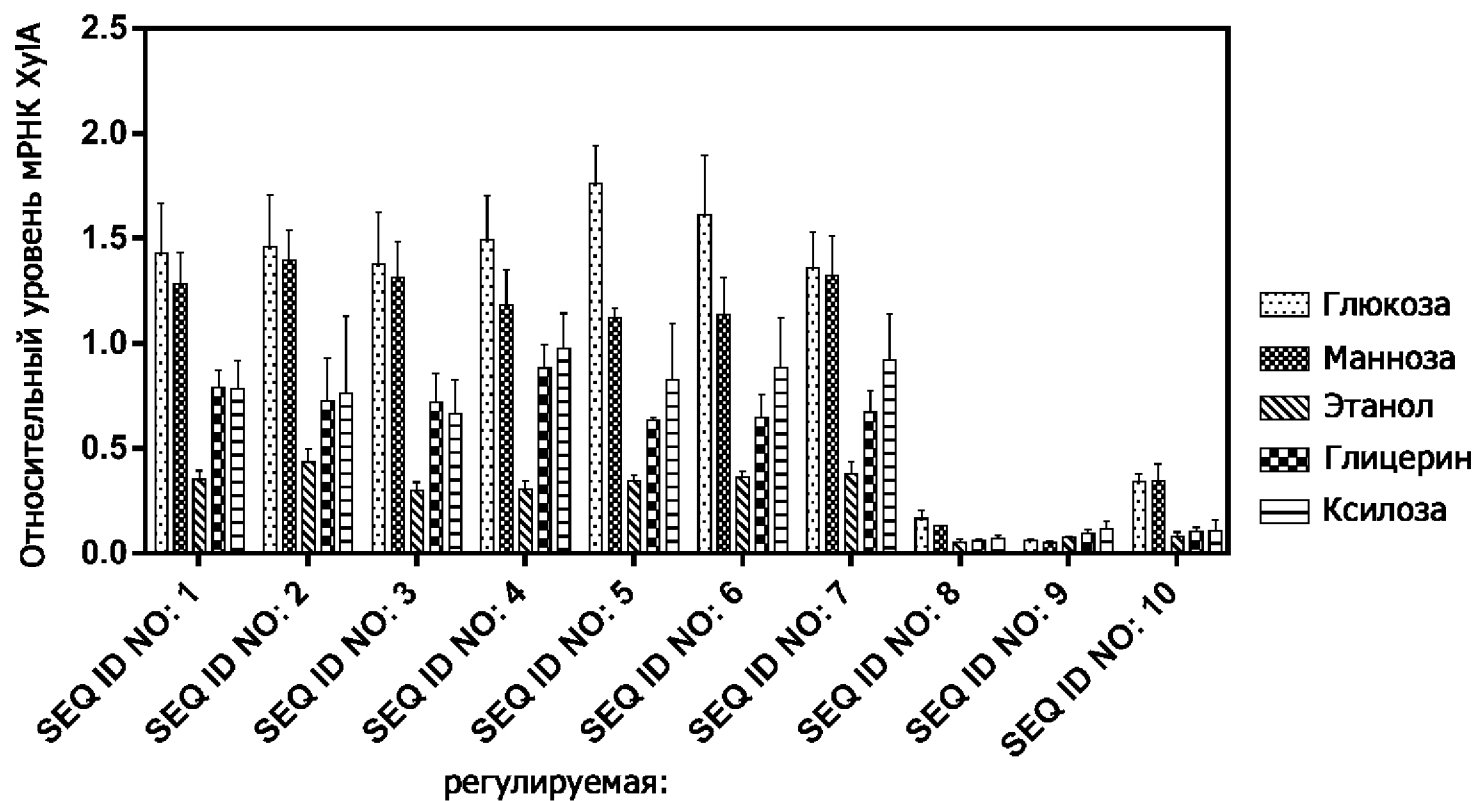
ФИГ. 1

Исследование роста – репортерная система



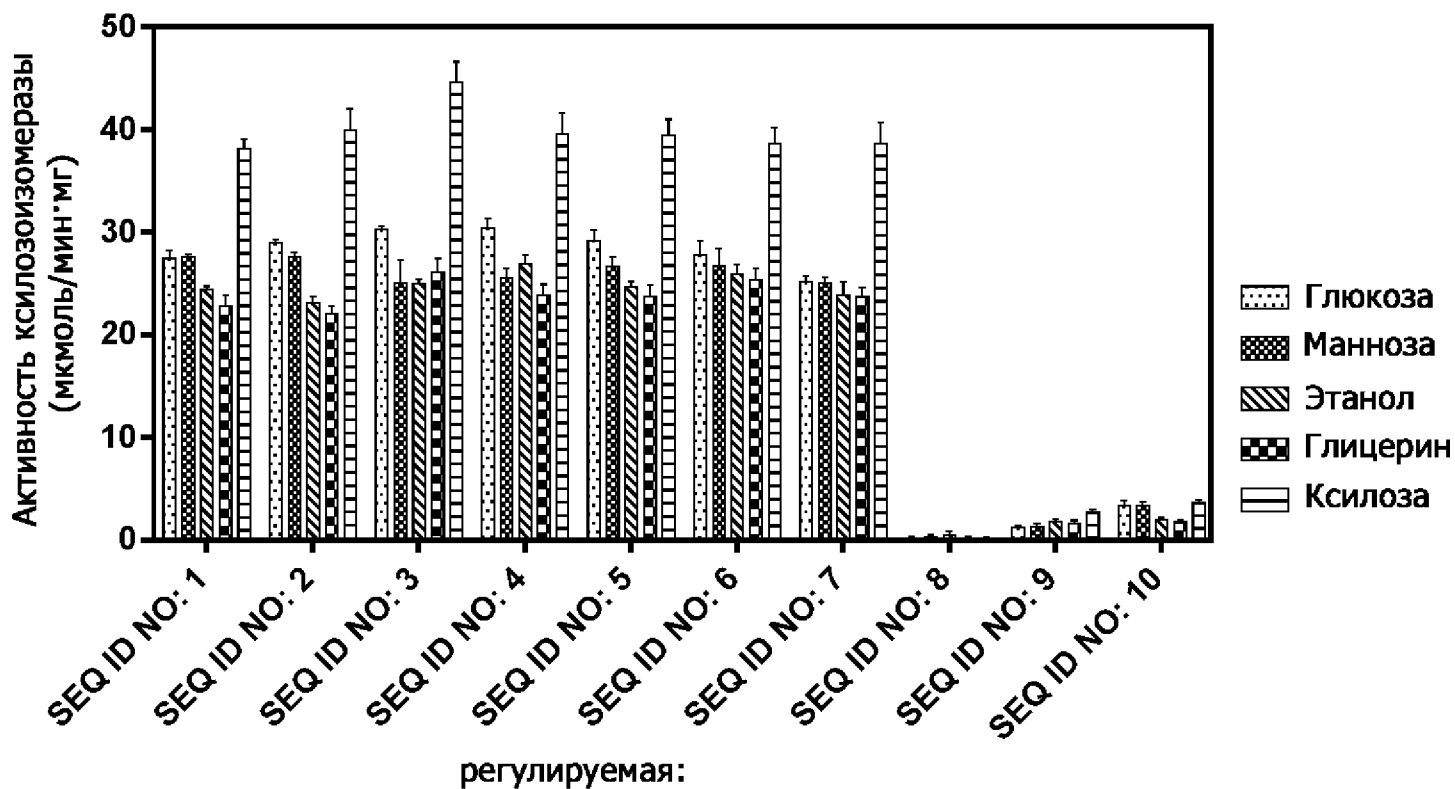
ФИГ. 2

Уровень транскрипции– репортерная система



ФИГ. 3

Активность фермента – репортерная система



ФИГ. 4