

(51) Classificação Internacional:

C12N 15/54 (2007.10) **C12N 9/10** (2007.10)
C12N 1/15 (2007.10) **C12N 1/19** (2007.10)
C12N 1/21 (2007.10) **C12N 5/10** (2007.10)
A01K 67/27 (2007.10) **A23D 9/00** (2007.10)
A23L 1/29 (2007.10) **A61K 31/185** (2007.10)
A61K 8/00 (2007.10)

(12) **FASCÍCULO DE PATENTE DE INVENÇÃO**

(22) Data de pedido: **1999.08.30**

(30) Prioridade(s): **1998.09.02 US 145828**

(43) Data de publicação do pedido: **2001.06.20**

(45) Data e BPI da concessão: **2008.04.16**
135/2008

(73) Titular(es):

ABBOTT LABORATORIES
CHAD 0377/AP6D-2, 100 ABBOTT PARK ROAD
ABBOTT PARK, IL 60064-6050 **US**

(72) Inventor(es):

PRADIP MUKERJI **US**
AMANDA EUN-YEONG LEONARD **US**
YUNG-SHENG HUANG **US**
JENNIFER THURMOND **US**
JENNIFER M. PARKER-BARNES **US**

(74) Mandatário:

ANTÓNIO JOÃO COIMBRA DA CUNHA FERREIRA
R DAS FLORES 74 4 AND 1249-235 LISBOA **PT**

(54) Epígrafe: **GENES DE ALONGASES E SUAS UTILIZAÇÕES**

(57) Resumo:

DESCRIÇÃO

"Genes de alongases e suas utilizações"

ANTECEDENTES DO INVENTO

Campo Técnico

O presente invento refere-se à identificação de vários genes envolvidos no alongamento de ácidos gordos poli-insaturados de cadeia longa (i.e., "alongases") e a utilizações destes. Em particular, a enzima alongase é utilizada na conversão de um ácido gordo noutro. Por exemplo, a alongase catalisa a conversão de ácido gama-linolénico (GLA) em ácido di-homo- γ -linolénico (DGLA, 20:3n-3) e a conversão de ácido estearidónico (STA, 18:4n-3) em ácido (n-3)-eicosatetraenónico (20:4n-3). A alongase catalisa também a conversão de ácido araquidónico (AA, 20:4n-6) em ácido adrénico (ADA, 22:4n-6), a conversão de ácido eicosapentaenónico (EPA, 20:5n-3) em ácido ω 3-docosapentaenónico (22:5n-3), e a conversão de ácido α -linolénico (ALA, 18:3n-3) em 20:3n-3. O DGLA pode, por exemplo, ser utilizado na produção de outros ácidos gordos poli-insaturados (AGPI), tais como o ácido araquidónico (AA) que pode ser adicionado a composições farmacêuticas, composições nutricionais, rações animais, bem como outros produtos tais cosméticos.

Informação Antecedente

As alongases que foram identificadas no passado diferem em termos dos substratos sobre os quais actuam. Para além disso, estão presentes tanto em animais como em plantas. As que se verificam em mamíferos têm a capacidade de actuar sobre ácidos gordos saturados, mono-insaturados e poli-insaturados. Em contraste, as que se verificam em plantas são específicas para ácidos gordos saturados ou mono-insaturados. Assim, para gerar ácidos gordos poli-insaturados em plantas, há a necessidade de uma alongase específica de AGPI.

Tanto em plantas como em animais, crê-se que o processo de alongamento seja o resultado de um mecanismo de quatro passos (Lassner *et al.*, *The Plant Cell* 8: 281-292, 1996). CoA

é o transportador acilo. O passo um envolve a condensação de malonil-CoA com um acil-CoA de cadeia longa para dar dióxido de carbono e um β -cetoacil-CoA no qual a porção acilo foi alongada em dois átomos de carbono. As reacções subsequentes incluem a redução em β -hidroxiacil-CoA, desidratação para um enoil-CoA, e uma segunda redução para produzir o acil-CoA alongado. A reacção de condensação inicial é não só o passo específico do substrato como também o passo limitante da velocidade.

Tal como observado anteriormente, as alongases, mais especificamente as que utilizam AGPI como substratos, são críticas na produção de ácidos gordos poli-insaturados de cadeia longa que têm muitas funções importantes. Por exemplo, os AGPI são componentes importantes da membrana plasmática de uma célula onde se encontram na forma de fosfolípidos. Servem também como precursores de prostaciclina de mamífero, eicosanóides, leucotrienos e prostaglandinas. Adicionalmente, os AGPI são necessários para o desenvolvimento correcto do cérebro infantil em desenvolvimento bem como para a formação e recuperação de tecidos. Tendo em vista a significância biológica dos AGPI, foram feitas tentativas para os produzir eficientemente bem como intermediários que conduzam à sua produção.

Várias enzimas estão envolvidas na biossíntese de AGPI incluindo alongases (elo) (ver Figura 1). Por exemplo, o ácido linoleico (LA, 18:2- Δ 9,12 ou 18:2n-6) é produzido a partir de ácido oleico (OA, 18:1- Δ 9 ou 18:1n-9) através de uma Δ 12-dessaturase. O GLA (18:3- Δ 6,9,12) é produzido a partir de ácido linoleico através de uma Δ 6-dessaturase. O AA (20:4- Δ 5,8,11,14) é produzido a partir de ácido di-homo- γ -linolénico (DGLA, 20:3- Δ 8,11,14) através de uma Δ 5-dessaturase. Tal como observado acima, o DGLA é produzido a partir de GLA através de uma alongase.

Deve notar-se que os animais não podem dessaturar para além da posição Δ 9 e portanto não podem converter ácido oleico em ácido linoleico. Igualmente, o ácido α -linolénico (ALA, 18:3- Δ 9,12,15 ou 18:3n-3) não pode ser sintetizado por mamíferos, uma vez que estes não têm actividade de Δ 15-dessaturase. No entanto, o ácido α -linolénico pode ser

convertido em ácido estearidónico (STA, 18:4- Δ 6,9,12,15) por uma Δ 6-dessaturase (ver Publicação PCT WO 96/13591; ver também Patente U.S. N.º 5552306), seguido de alongamento até ácido (n-3)-eicosatetraenóico (20:4- Δ 8,11,14,17 ou 20:4n-3) em mamíferos e algas. Este ácido gordo poli-insaturado (i.e., 20:4- Δ 8,11,14,17) pode então ser convertido em ácido eicosapentaenóico (EPA, 20:5- Δ 5,8,11,14,17) através de uma Δ 5-dessaturase. Outros eucariotas, incluindo fungos e plantas, têm enzimas que dessaturam nos carbonos 12 (ver publicação PCT WO 94/11516 e Patente U.S. N.º 5443974) e 15 (ver publicação PCT WO 93/11245). Os principais ácidos gordos poli-insaturados dos animais são portanto derivados da dieta e/ou de dessaturação e alongamento do ácido linoleico ou do ácido α -linolénico. Tendo em vista a incapacidade dos mamíferos para produzir estes ácidos gordos de cadeia longa essenciais, é de grande interesse isolar genes envolvidos na biossíntese de AGPI a partir de espécies que produzam naturalmente estes ácidos gordos e expressar esses genes num sistema microbiano, vegetal ou animal que possa ser alterado para proporcionar a produção de quantidades comerciais de um ou mais AGPI. Consequentemente, existe uma grande necessidade da enzima alongase, do gene codificando a enzima bem como de métodos recombinantes de produção desta enzima. Adicionalmente, existe a necessidade de óleos contendo níveis de AGPI para além dos naturalmente presentes bem como dos enriquecidos em novos AGPI. Tais óleos podem ser feitos através de isolamento e expressão do gene da alongase.

Um dos AGPI de cadeia longa mais importantes, observado acima, é o ácido araquidónico (AA). O AA é verificado em fungos filamentosos e pode também ser purificado a partir de tecidos de mamífero incluindo o fígado e as glândulas supra-renais. Tal como observado acima, a produção de AA a partir de DGLA é catalisada por uma Δ 5-dessaturase e a produção de DGLA a partir de ácido γ -linolénico (GLA) é catalisada por uma alongase. No entanto, até ao presente invento, não se tinha identificado nenhuma alongase que fosse activa em ácidos gordos substrato nas vias para a produção de AGPI de cadeia longa e, em particular, AA, ácido eicosapentaenóico (EPA), ácido adrenico, ácido docosa-hexaenóico (DHA, 22:6n-3), ácido ω 3-docosapentaenóico (22:5n-3) ou ácido ω 6-docosapentaenóico (22:5n-6).

Dois genes pareciam ter interesse na presente pesquisa do gene da alongase. Em particular, a β -cetoacil-coenzima A-sintetase (KCS) de jojoba ou KCS de jojoba (Acesso GenBank # U37088), catalisa a reacção inicial da via de alongamento de acil-CoA gordo (i.e., a condensação de malonil-CoA com acil-CoA de cadeia longa (Lassner *et al.*, *The Plant Cell* 8: 281-292, 1996). O substrato de preferência de KCS de jojoba é 18:0, 20:0, 20:1, 18:1, 22:1, 22:0 e 16:0. A alongase de *Saccharomyces cerevisiae* (ELO2) catalisa também a conversão de ácidos gordos de cadeia longa saturados e mono-insaturados, produzindo níveis elevados de 22:0, 24:0 e também 18:0, 18:1, 20:0, 20:1, 22:0, 22:1 e 24:1 (Oh *et al.*, *The Journal of Biological Chemistry* 272(28): 17376-17384, 1997); ver também Patente U.S. N.º 5484724 para uma sequência nucleotídica que inclui a sequência de ELO2; ver o pedido PCT WO 88/07577 para uma discussão da sequência de um factor de inibição da glicosilação que é descrito no Exemplo V). A pesquisa de uma alongase específica de AGPI de cadeia longa em *Mortierella alpina* começou com base numa revisão das homologias partilhadas entre estes dois genes e através de pesquisa por expressão da actividade de alongase de AGPI.

SUMÁRIO DO INVENTO

O presente invento refere-se a uma sequência nucleotídica isolada correspondente ou complementar a pelo menos 50% da sequência nucleotídica mostrada na Figura 6.

A sequência codifica uma alongase funcionalmente activa que utiliza um ácido gordo poli-insaturado ou um ácido gordo mono-insaturado como substrato. Em particular, a sequência pode ser derivada de um fungo do género *Mortierella* e pode ser especificamente isolada de *Mortierella alpina*.

O presente invento inclui também uma proteína purificada codificada pela sequência nucleotídica de cima.

Adicionalmente, o presente invento engloba um método de produção de uma enzima alongase compreendendo os passos de: a) isolamento da sequência nucleotídica representada pela Figura 6; b) construção de um vector compreendendo: i) a

sequência nucleotídica isolada operativamente ligada a ii) um promotor; e c) introdução do vector numa célula hospedeira durante um tempo e sob condições suficientes para expressão da enzima alongase. A célula hospedeira pode ser uma célula eucariótica ou uma célula procariótica.

A célula procariótica pode ser, por exemplo, uma célula de *E. coli*, uma célula de cianobactéria ou uma célula de *B. subtilis*. A célula eucariótica pode ser, por exemplo, uma célula de mamífero, uma célula de insecto, uma célula vegetal ou uma célula fúngica. A célula fúngica pode ser, por exemplo, *Saccharomyces* sps., *Candida* sps., *Lipomyces* sps., *Yarrowia* sps., *Kluyveromyces* sps., *Hansenula* sps., *Aspergillus* sps., *Penicillium* sps., *Neurospora* sps., *Trichoderma* sps. ou *Pichia* sps. Em particular, a célula fúngica pode ser uma célula de levedura tal como *Saccharomyces* sps., em particular, *Saccharomyces cerevisiae*, *Candida* sps., *Hansenula* sps. ou *Pichia* sps.

O presente invento inclui também um vector compreendendo: a) uma sequência nucleotídica tal como representada pela Figura 6 operativamente ligada a b) um promotor, bem como uma célula hospedeira compreendendo este vector. O hospedeiro pode ser uma célula procariótica ou uma célula eucariótica. Exemplos adequados de células procarióticas incluem células de *E. coli*, *Cyanobacteria* e *B. subtilis*. Exemplos adequados de células eucarióticas incluem uma célula de mamífero, uma célula de insecto, uma célula vegetal e uma célula fúngica. A célula fúngica pode ser, por exemplo, *Saccharomyces* sps., *Candida* sps., *Lipomyces* sps., *Yarrowia* sps., *Kluyveromyces* sps., *Hansenula* sps., *Aspergillus* sps., *Penicillium* sps., *Neurospora* sps., *Trichoderma* sps. e *Pichia* sps. Em particular, a célula fúngica pode ser, por exemplo, uma célula de levedura tal como, por exemplo, *Saccharomyces* sps., em particular, *Saccharomyces cerevisiae*, *Candida* sps., *Hansenula* sps. e *Pichia* sps.

O presente invento inclui uma célula vegetal, planta ou tecido vegetal compreendendo o vector acima descrito, em que a expressão da sequência nucleotídica do vector resulta na produção de pelo menos um ácido gordo seleccionado de entre o grupo que consiste num ácido gordo mono-insaturado e um ácido

gordo poli-insaturado pela célula vegetal, planta ou tecido vegetal. O ácido gordo poli-insaturado pode ser, por exemplo, ácido di-homo- γ -linolénico (DGLA), 20:4n-3, e ácido adrénico (ADA). O presente invento inclui também um ou mais óleos vegetais ou ácidos gordos expressos pela célula vegetal, planta ou tecido vegetal. Adicionalmente, o presente invento engloba uma engloba uma planta transgénica compreendendo o vector acima descrito, em que a expressão da sequência nucleotídica do vector resulta na produção de um ácido gordo poli-insaturado em sementes da planta transgénica.

Para além disso, o presente invento inclui um mamífero transgénico não humano cujo genoma compreende uma sequência de ADN codificando uma alongase operativamente ligada a um promotor. A sequência de ADN pode ser representada por SEQ ID NO:1 (Figura 6).

Pode ser produzido pelo animal transgénico não humano um fluido (p.ex., leite) em que o fluido compreende um nível detectável de pelo menos uma alongase ou de produtos desta tais como, por exemplo, DGLA, ácido ω 6-docosapentaenóico, ADA e/ou 20:4n-3 (ver Figura 1).

Adicionalmente, o presente invento inclui um método para a produção de um ácido gordo poli-insaturado compreendendo os passos de: a) isolamento da referida sequência nucleotídica representada por SEQ ID NO:1 (Figura 6); b) construção de um vector compreendendo a sequência nucleotídica isolada; c) introdução do vector numa célula hospedeira durante o tempo e sob as condições suficientes para a expressão da enzima alongase codificada pela sequência nucleotídica isolada; e d) exposição da enzima alongase expressa a um ácido gordo poli-insaturado "substrato" para converter o substrato num ácido gordo poli-insaturado "produto". O ácido gordo poli-insaturado substrato pode ser seleccionado de entre o grupo que consiste em, por exemplo, ácido γ -linolénico (GLA), ácido estearidónico (STA) e ácido araquidónico (AA) e o ácido gordo poli-insaturado produto pode ser seleccionado de entre o grupo que consiste em, por exemplo, DGLA, 20:4n-3 e ADA respectivamente. O método pode ainda compreender o passo de exposição do ácido gordo poli-insaturado produto a pelo menos uma dessaturase para converter o ácido gordo poli-insaturado produto "noutro"

ácido gordo poli-insaturado. O ácido gordo poli-insaturado produto pode ser seleccionado a partir do grupo que consiste em, por exemplo, DGLA, $20;4n-3$ e ADA. O outro ácido gordo poli-insaturado pode ser seleccionado a partir do grupo que consiste em, por exemplo, AA, ácido eicosapentaenóico (EPA), ácido $\omega 6$ -docosapentaenóico, respectivamente, e pelo menos uma dessaturase é $\Delta 5$ -dessaturase, em relação à produção de AA ou EPA, e $\Delta 4$ -dessaturase, em relação à produção de ácido $\omega 6$ -docosapentaenóico. O método pode ainda compreender o passo de exposição de outro ácido gordo poli-insaturado a uma ou mais enzimas seleccionadas a partir o grupo que consiste em pelo menos uma alongase e pelo menos uma dessaturase adicional para converter o outro ácido gordo poli-insaturado num ácido gordo poli-insaturado "final". O ácido gordo poli-insaturado final pode ser, por exemplo, ácido docosa-hexaenóico (DHA), AA, ácido $\omega 6$ -docosapentaenóico ou ácido $\omega 3$ -docosapentaenóico.

O presente invento inclui também uma sequência nucleotídica isolada correspondente ou complementar a pelo menos cerca de 35% da sequência nucleotídica mostrada na Figura 22. A sequência codifica uma alongase funcionalmente activa que utiliza um ácido gordo poli-insaturado como substrato. Esta sequência pode também ser derivada, por exemplo, de um fungo do género *Mortierella*. Em particular, pode ser derivada de *M. alpina*.

Adicionalmente, o presente invento inclui uma proteína purificada codificada pela sequência nucleotídica de cima.

O presente invento inclui também um método de produção de uma enzima de alongamento tal como descrito acima. A sequência inserida no vector é representada pela Figura 22. A célula hospedeira pode ser procariótica ou eucariótica. Exemplos adequados são descritos acima.

O presente invento inclui também um vector compreendendo: a) uma sequência nucleotídica tal como representada pela Figura 22 operativamente ligada a b) um promotor, bem como uma célula hospedeira compreendendo este vector. Novamente, a célula hospedeira pode ser eucariótica ou procariótica. Exemplos adequados estão descritos acima.

O presente invento inclui também uma célula vegetal, planta ou tecido vegetal compreendendo o vector de cima, em que a expressão da sequência nucleotídica do vector resulta na produção de um ácido gordo poli-insaturado pela célula vegetal, planta ou tecido vegetal. O ácido gordo poli-insaturado pode ser, por exemplo, DGLA, 20:4n-3 ou ADA. Adicionalmente, o presente invento inclui um ou mais óleos vegetais ou ácidos gordos expressos pela célula vegetal, planta ou tecido vegetal.

Para além disso, o presente invento inclui também uma planta transgénica compreendendo o vector de cima, em que a expressão da sequência nucleotídica da Figura 22 do vector resulta na produção de um ácido gordo poli-insaturado em sementes da planta transgénica.

O presente invento inclui também um mamífero transgénico, não humano cujo genoma compreende uma sequência de ADN da Figura 22 codificando uma alongase operativamente ligada a um promotor. Pode ser produzido um fluido por este mamífero transgénico não humano em que o fluido compreende um nível detectável de pelo menos uma alongase ou de produtos desta.

O presente invento inclui também um método para produção de um ácido gordo poli-insaturado compreendendo os passos de: a) isolamento da sequência nucleotídica representada pela Figura 22; b) construção de um vector compreendendo a sequência nucleotídica isolada; c) introdução do vector numa célula hospedeira durante um tempo e sob condições suficientes para expressão de uma enzima alongase codificada pela sequência nucleotídica isolada; e d) exposição da enzima alongase expressa a um ácido gordo poli-insaturado substrato para converter o substrato num ácido gordo poli-insaturado produto. O ácido gordo poli-insaturado substrato pode ser, por exemplo, GLA, STA ou AA, o ácido gordo poli-insaturado produto pode ser, por exemplo, DGLA, 20:4n-3 ou ácido ω 6-docosapentaenóico, respectivamente. O método pode ainda compreender o passo de exposição da enzima alongase expressa a pelo menos uma dessaturase para converter o ácido gordo poli-insaturado produto noutro ácido gordo poli-insaturado. O ácido gordo poli-insaturado produto pode ser, por exemplo, DGLA, 20:4n-3 ou ADA, o outro ácido gordo poli-insaturado pode ser,

por exemplo, AA, EPA ou ácido ω 6-docosapentaenóico, respectivamente, e pelo menos uma dessaturase é Δ 5-dessaturase em relação à produção de AA ou EPA e Δ 4-dessaturase em relação à produção de ácido ω 6-docosapentaenóico. O método pode ainda compreender o passo de exposição do outro ácido gordo poli-insaturado a uma ou mais enzimas seleccionadas a partir do grupo que consiste em pelo menos uma alongase e pelo menos uma dessaturase adicional para converter o outro ácido gordo poli-insaturado num ácido gordo poli-insaturado final. O ácido gordo poli-insaturado final pode ser, por exemplo, ácido docosa-hexaenóico, AA, ácido ω 6-docosapentaenóico ou ácido ω 3-docosapentaenóico.

Para além disso, o presente invento inclui uma sequência nucleotídica isolada correspondente ou complementar à sequência nucleotídica mostrada na Figura 43.

Esta sequência codifica uma alongase funcionalmente activa que utiliza um ácido gordo poli-insaturado ou um ácido gordo mono-insaturado como substrato. A sequência é derivada de um mamífero tal como, por exemplo, um humano.

O presente invento inclui também uma proteína purificada codificada por esta sequência nucleotídica.

Adicionalmente, o presente invento inclui um método de produção de uma enzima alongase compreendendo os passos de: a) isolamento da sequência nucleotídica representada pela Figura 43; b) construção de um vector compreendendo: i) a sequência nucleotídica isolada operativamente ligada a ii) um promotor; e c) introdução do referido vector numa célula hospedeira durante um tempo e sob condições suficientes para a expressão da enzima alongase. A célula hospedeira pode ser a mesma que a descrita acima em relação aos métodos correspondentes utilizando as Figuras 6 ou 22. O presente invento inclui também um vector compreendendo: a) uma sequência nucleotídica tal como representada pela Figura 43 operativamente ligada a b) um promotor, bem como uma célula hospedeira compreendendo este vector. A célula hospedeira pode ser a mesma que a descrita acima.

O presente invento inclui também uma célula vegetal, planta ou tecido vegetal compreendendo o vector acima descrito compreendendo a sequência da Figura 43 em que a expressão da sequência nucleotídica do vector resulta na produção de pelo menos um ácido gordo seleccionado a partir do grupo que consiste num ácido gordo mono-insaturado e num ácido gordo poli-insaturado pela referida célula vegetal, planta ou tecido vegetal. O ácido gordo poli-insaturado pode ser, por exemplo, DGLA, 20:4n-3 ou ADA. O presente invento inclui também um ou mais óleos ou ácidos vegetais expressos pela célula vegetal, planta ou tecido vegetal.

O presente invento inclui também uma planta transgénica compreendendo o vector compreendendo a sequência da Figura 43, em que a expressão da sequência nucleotídica do vector resulta na produção de um ácido gordo poli-insaturado em sementes da planta transgénica.

Adicionalmente, o presente invento inclui um mamífero transgénico não humano cujo genoma compreende uma sequência de ADN humana codificando uma alongase operativamente ligada a um promotor. A sequência de ADN é representada pela Figura 43. Pode ser produzido um fluido pelo referido mamífero transgénico não humano em que o referido fluido compreende um nível detectável de pelo menos uma alongase ou de produtos desta.

O presente invento engloba também um método para produção de um ácido gordo poli-insaturado compreendendo os passos de: a) isolamento da sequência nucleotídica representada pela Figura 43; b) construção de um vector compreendendo a referida sequência nucleotídica; c) introdução do vector numa célula hospedeira durante um tempo e sob condições suficientes para expressão da enzima alongase codificada pela sequência nucleotídica isolada; e d) exposição da enzima alongase expressa a um ácido gordo poli-insaturado substrato para converter o substrato num ácido gordo poli-insaturado produto. O ácido gordo poli-insaturado substrato pode ser, por exemplo, GLA, STA ou AA e o ácido gordo poli-insaturado produto pode ser, por exemplo, DGLA, 20:4n-3 ou ADA, respectivamente. O método pode compreender ainda o passo de exposição do ácido gordo poli-insaturado produto a pelo menos uma dessaturase

para converter o ácido gordo poli-insaturado produto noutro ácido gordo poli-insaturado. O ácido gordo poli-insaturado produto pode ser, por exemplo, DGLA, 20:4n-3 e ADA, o outro ácido gordo poli-insaturado pode ser, por exemplo, AA, EPA e ácido ω 6-docosapentaenóico, respectivamente, e pelo menos uma dessaturase é Δ 5-dessaturase em relação à produção de AA ou EPA e Δ 4-dessaturase em relação à produção de ácido ω 6-docosapentaenóico. O método pode ainda compreender o passo de exposição do outro ácido gordo poli-insaturado a uma ou mais enzimas seleccionadas a partir do grupo que consiste em pelo menos uma alongase e pelo menos uma dessaturase adicional para converter o outro ácido gordo poli-insaturado num ácido gordo poli-insaturado final. O ácido gordo poli-insaturado final pode ser, por exemplo, DHA, ADA, ácido ω 6-docosapentaenóico e ácido ω 3-docosapentaenóico.

Adicionalmente, o presente invento inclui uma sequência nucleotídica isolada correspondendo ou complementar à sequência nucleotídica mostrada na Figura 46. Codifica uma alongase funcionalmente activa que utiliza um ácido gordo poli-insaturado como substrato. A sequência pode ser derivada ou isolada de um nematode do género *Caenorhabditis* e, em particular, pode ser isolada a partir de *C. elegans*.

Adicionalmente, o presente invento inclui um método de produção de uma enzima alongase compreendendo os passos de: a) isolamento da sequência nucleotídica representada pela Figura 46; b) construção de um vector compreendendo: i) a sequência nucleotídica isolada operativamente ligada a ii) um promotor; e c) introdução do vector numa célula hospedeira durante um tempo e sob condições suficientes para a expressão da enzima alongase. As propriedades da célula hospedeira são as mesmas que as descritas acima em relação às sequências das Figuras 6, 22 e 43.

O presente invento engloba também um vector compreendendo: a) uma sequência nucleotídica tal como representada pela Figura 46 operativamente ligada a b) um promotor, bem como uma célula hospedeira compreendendo este vector. A célula hospedeira tem as mesmas propriedades que as citadas acima em relação à célula hospedeira citada acima para as sequências das Figuras 6, 22 e 43.

Para além disso, o presente invento inclui uma célula vegetal, planta ou tecido vegetal compreendendo o vector de cima compreendendo a sequência da Figura 46, em que a expressão da referida sequência nucleotídica do vector resulta na produção de um ácido gordo poli-insaturado pela célula vegetal, planta ou tecido vegetal. O ácido gordo poli-insaturado pode ser, por exemplo, DGLA, 20:4n-3 ou ADA. O presente invento inclui também um ou mais óleos ácidos gordos vegetais ou expressos por esta célula vegetal, planta ou tecido vegetal.

O presente invento inclui também uma planta transgénica compreendendo o vector de cima incluindo a sequência nucleotídica correspondente à sequência da Figura 46, em que a expressão da sequência nucleotídica do vector resulta na produção de um ácido gordo poli-insaturado em sementes da planta transgénica.

Adicionalmente, o presente invento inclui um mamífero transgénico não humano cujo genoma compreende uma sequência de ADN de *C. elegans* codificando uma alongase operativamente ligada a um promotor. A sequência de ADN pode ser representada pela Figura 46. Pode ser produzido um fluido pelo referido mamífero transgénico não humano em que o referido fluido compreende um nível detectável de pelo menos uma alongase ou de produtos desta.

O presente invento inclui também um método para a produção de um ácido gordo poli-insaturado compreendendo os passos de: a) isolamento da sequência nucleotídica representada pela Figura 46; b) construção de um vector compreendendo a sequência nucleotídica isolada; c) introdução do vector numa célula hospedeira durante um tempo e sob condições suficientes para expressão de uma enzima alongase através da sequência nucleotídica isolada; e d) exposição da enzima alongase expressa a um ácido gordo poli-insaturado substrato para converter o substrato num ácido gordo poli-insaturado produto. O ácido gordo poli-insaturado substrato pode ser, por exemplo, GLA, STA ou AA, e o ácido gordo poli-insaturado produto pode ser, por exemplo, DGLA, 20:4n-3 ou ADA, respectivamente. O método pode ainda compreender o passo

de exposição da enzima alongase expressa a pelo menos uma dessaturase para converter o referido ácido gordo poli-insaturado produto noutro ácido gordo poli-insaturado. O ácido gordo poli-insaturado produto pode ser, por exemplo, DGLA, 20:4n-3 ou ADA, o outro ácido gordo poli-insaturado pode ser, por exemplo, AA, EPA ou ácido ω 6-docosapentaenóico, respectivamente, e a pelo menos uma dessaturase é Δ 5-dessaturase em relação à produção de AA ou EPA e Δ 4-dessaturase em relação à produção de ácido ω 6-docosapentaenóico. O método pode ainda compreender o passo de exposição do outro ácido gordo poli-insaturado a uma ou mais enzimas seleccionadas de entre o grupo que consiste em pelo menos uma alongase e pelo menos uma dessaturase adicional para converter o outro ácido gordo poli-insaturado num ácido gordo poli-insaturado final. O ácido gordo poli-insaturado final pode ser, por exemplo, DHA, ADA, ácido ω 6-docosapentaenóico ou ácido ω 3-docosapentaenóico.

BREVE DESCRIÇÃO DOS DESENHOS

A Figura 1 representa várias vias de biossíntese de ácidos gordos. Deve notar-se o papel da enzima alongase (elo).

A Figura 2 representa a percentagem de semelhança e a percentagem de identidade entres as sequências de aminoácidos de KCS de jojoba e ELO2.

A Figura 3 representa a sequência de ELO2 de *S. cerevisiae* homóloga à sequência de KCS de jojoba (sequência iniciadora sublinhada) da Figura 2.

A Figura 4A mostra o mapa físico de pRAE-2 contendo o ADNc de MAELO. A Figura 4B representa o mapa físico do vector de expressão constitutiva, pRAE-5, utilizado para a produção da enzima alongase em levedura.

A Figura 5 representa uma comparação das sequências nucleotídicas dos clones pRAE-5 e pRAE-6.

A Figura 6 ilustra a sequência nucleotídica completa da alongase de *Mortierella alpina* (MAELO).

A Figura 7 representa a sequência de aminoácidos da alongase de *Mortierella alpina* traduzida a partir de MAELO (ver Figura 6).

A Figura 8 representa um alinhamento de sequências de aminoácidos entre 3 alongases: EL02 de *S. cerevisiae* (GNS1), EL03 de *S. cerevisiae* (SUR4) e a sequência traduzida de MAELO tal como mostrado na Figura 7.

A Figura 9 representa uma comparação a sequência nucleotídica de MAELO e a sequência nucleotídica de EL02 de *S. cerevisiae*.

As Figuras 10A e 10B representam a actividade de alongase de AGPI de MAELO expressa em fermento de padeiro.

A Figura 11 ilustra a actividade de alongase de AGPI de MAELO quando co-expressa com o ADNc da $\Delta 5$ -dessaturase de *M. alpina* para produzir AA.

A Figura 12 compara a actividade de alongase de AGPI de MAELO com a sobre-expressão de EL02 de *S. cerevisiae* no fermento de padeiro.

As Figuras 13, 14 e 15 representam três comparações separadas de sequências de aminoácidos derivadas de sequências nucleotídicas de *C. elegans* na base de dados GenEMBL com a MAELO traduzida.

A Figura 16 mostra a comparação entre as traduções em aminoácidos de duas sequências de mamífero diferentes na base de dados GenEMBL e a MAELO traduzida.

A Figura 17 mostra a comparação de uma sequência de ADN traduzida (ver pedido PCT publicado WO 88/07577) com a sequência de aminoácidos derivada de MAELO, que foi detectada durante uma pesquisa na base de dados.

A Figura 18 mostra a sequência nucleotídica completa da $\Delta 5$ -dessaturase de *M. alpina*.

A Figura 19 representa a análise inicial GC-FAME do banco MAD708. Deve notar-se a detecção de um pico de DGLA (C20:3n-6).

A Figura 20 representa a actividade de alongase de AGPI dos cinco clones MAD708 em levedura com GLA como substrato. Todos os clones têm uma actividade de alongase aparente.

A Figura 21 representa a análise da sequenciação de ADN do plasmídeo pRPB2. A análise revela um enquadramento de leitura aberto de 957 pb de comprimento.

A Figura 22 mostra a sequência nucleotídica completa do ADNc de *M. alpina*, contida no plasmídeo pRPB2, que é designada GLELO pela sua actividade de alongase de GLA.

A Figura 23 representa a sequência de aminoácidos da alongase de *M. alpina* traduzida a partir de GLELO (ver Figura 22).

A Figura 24 ilustra a actividade de alongase de AGPI n-6 numa cultura induzida de 334(pRPB2) quando suplementada com GLA.

A Figura 25 representa a actividade de alongase de AGPI n-3 e n-6 numa cultura induzida de 334(pRPB2) quando suplementada com 25 µM de outros substratos de ácidos gordos.

A Figura 26A ilustra a actividade de alongase de GLELO com GLA como substrato quando co-expressa com o ADNc da $\Delta 5$ -dessaturase de *M. alpina* para produzir AA. A Figura 26B ilustra a actividade de alongase de GLELO com STA como substrato quando co-expressa com o ADNc de $\Delta 5$ -dessaturase de *M. alpina* para produzir EPA.

A Figura 27 ilustra a comparação entre a sequência de GLELO traduzida (ver Figura 23) e a sequência de MAELO traduzida (ver Figura 7).

A Figura 28 representa uma comparação da sequência de aminoácidos de 4 alongases: a sequência de aminoácidos traduzida de GLELO (ver Figura 23), MAELO (ver Figura 7), EL02

de *S. cerevisiae* (GNS1) e ELO3 de *S. cerevisiae* (SUR4). A caixa de histidinas está sublinhada.

A Figura 29 representa um alinhamento entre a sequência de MAELO traduzida e a sequência putativa homóloga humana traduzida de HS1.

A Figura 30 representa um alinhamento entre a sequência de MAELO traduzida e a sequência putativa homóloga humana traduzida de HS2.

A Figura 31 mostra um alinhamento entre a sequência de MAELO traduzida e a sequência putativa homóloga de ratinho traduzida de MM2.

A Figura 32 representa um alinhamento entre a sequência de MAELO traduzida e a sequência putativa homóloga de ratinho traduzida AI225632.

A Figura 33 ilustra um alinhamento entre a sequência de GLELO traduzida e a sequência putativa homóloga humana traduzida AI225632.

A Figura 34 mostra um alinhamento entre a sequência de GLELO traduzida e a sequência putativa homóloga humana traduzida de HS1.

A Figura 35 representa um alinhamento entre a sequência de GLELO traduzida e a sequência putativa homóloga humana traduzida AC004050.

A Figura 36 ilustra um alinhamento entre a sequência de GLELO traduzida e a sequência putativa homóloga de ratinho traduzida de MM2.

A Figura 37 representa um alinhamento da sequência de GLELO traduzida com uma sequência putativa homóloga de ratinho traduzida AI225632.

A Figura 38 ilustra um alinhamento da sequência de GLELO traduzida com uma sequência putativa homóloga de ratinho traduzida U97107.

A Figura 39 representa um alinhamento da sequência traduzida de GLELO com uma sequência putativa homóloga de *C. elegans* traduzida U68749 (F56H11.4).

A Figura 40 mostra um alinhamento entre a sequência traduzida de MAELO e uma sequência putativa homóloga de *C. elegans* traduzida U68749 (F56H11.4).

A Figura 41 representa um alinhamento entre a sequência de GLELO traduzida e uma sequência putativa homóloga de *Drosophila melanogaster* traduzida, DM1.

A Figura 42 ilustra um alinhamento entre a sequência de MAELO traduzida e uma sequência putativa homóloga de *Drosophila melanogaster* traduzida, DM1.

A Figura 43 ilustra a sequência nucleotídica completa de uma alongase humana HSEL01.

A Figura 44 representa a sequência de aminoácidos deduzida da alongase humana HSEL01.

A Figura 45 ilustra a actividade de alongase (AGPI e outros) de uma cultura induzida de 334(pRAE-58-A1) quando suplementada com GLA ou AA.

A Figura 46 mostra a sequência nucleotídica completa da alongase CELO de *C. elegans*.

A Figura 47 mostra os aminoácidos deduzidos da alongase CELO de *C. elegans*.

A Figura 48 ilustra a actividade de alongase de AGPI de uma cultura induzida de 334(pRET-21) e 334(pRET-22) quando suplementada com GLA e AA.

A Figura 49 representa a sequência nucleotídica completa do putativo gene da alongase humana HS3.

A Figura 50 ilustra a sequência de aminoácidos deduzida da enzima alongase putativa humana HS3.

DESCRIÇÃO DETALHADA DO INVENTO

O presente invento refere-se a sequências nucleotídicas e às correspondentes sequências de aminoácidos de dois ADNc de alongases derivados de *Mortierella alpina*, bem como às sequências nucleotídicas e às correspondentes sequências de aminoácidos de um ADNc de alongase derivado de um humano e a outro derivado de *C. elegans*. Para além disso, o presente invento inclui também utilizações dos ADNc das proteínas codificadas pelos genes. Por exemplo, os genes e as correspondentes enzimas podem ser utilizados na produção de ácidos gordos poli-insaturados e/ou ácidos gordos mono-insaturados tais como DGLA, AA, ADA, EPA e/ou DHA que podem ser adicionados a composições farmacêuticas, composições nutricionais e a outros produtos valorizados.

Os Genes de Alongases e Enzimas por Eles Codificados

Tal como observado acima, uma enzima alongase codificada por um ADNc de alongase é essencial na produção de vários ácidos gordos poli-insaturados, em particular, AGPI de 20-24 carbonos. Em relação ao presente invento, a sequência nucleotídica do ADNc da alongase de *M. alpina* isolado (MAELO) é mostrada na Figura 6 e a sequência de aminoácidos da correspondente proteína ou enzima purificada codificada por esta sequência nucleotídica é mostrada na Figura 7. Adicionalmente, a sequência nucleotídica do ADNc da alongase de GLA (GLELO) isolado é mostrada na Figura 22 e a sequência de aminoácidos da proteína ou enzima purificada correspondente codificada por esta sequência nucleotídica é mostrada na Figura 23. A sequência nucleotídica da alongase de sequência humana 1 (HSELO1) isolada é mostrada na Figura 43 e a sequência de aminoácidos da proteína ou enzima purificada correspondente codificada por esta sequência é mostrada na Figura 44. Para além disso, a sequência nucleotídica do ADNc da alongase de *C. elegans* (CEELO1) isolada é mostrada na Figura 46 e a sequência de aminoácidos da proteína ou enzima correspondente purificada codificada deste modo é mostrada na Figura 47.

Como exemplo, as alongases isoladas codificadas pelos ADNc do presente invento alongam GLA em DGLA ou alongam STA em 20:4n-3 ou alongam AA em ADA. A produção de ácido araquidônico a partir de DGLA, ou EPA a partir de 20:4n-3, é então catalisada por uma Δ^5 -dessaturase. Assim, nem AA (ou EPA), nem DGLA (ou 20:4n-3) nem ADA (ou ácido ω 3-docosapentaenóico), podem ser sintetizados sem pelo menos um ADNc de alongase e uma enzima codificada por este.

Deve notar-se que o presente invento engloba também sequências nucleotídicas (e as correspondentes proteínas codificadas) possuindo sequências correspondendo a (i.e., possuindo identidade com) ou complementares a pelo menos cerca de 50%, de preferência pelo menos cerca de 60% e de maior preferência pelo menos cerca de 70% dos nucleótidos na sequência nucleotídica do ADNc de MAELO aqui descrita (ver Figura 6). Para além disso, o presente invento inclui também sequências nucleotídicas (e as correspondentes proteínas codificadas) possuindo sequências correspondendo a (i.e., possuindo identidade com) ou complementares a pelo menos cerca de 35%, de preferência pelo menos cerca de 45% e de maior preferência pelo menos cerca de 55% dos nucleótidos da sequência nucleotídica do ADNc de GLELO aqui descrita (ver Figura 22). Adicionalmente, o presente invento inclui também sequências nucleotídicas (e as correspondentes proteínas codificadas) possuindo sequências correspondendo a (i.e., possuindo identidade com) ou complementares à sequência nucleotídica do ADNc da sequência humana 1 (HSEL01) aqui descrita (ver Figura 43). Adicionalmente, o presente invento inclui também sequências nucleotídicas possuindo sequências correspondendo a (i.e., possuindo identidade com) ou complementares aos nucleótidos na sequência nucleotídica do ADNc de *C. elegans*, CEEL01, aqui descrita (ver Figura 46). Tais sequências podem ser derivadas de fontes que não *Mortierella* (p.ex., um eucariota (p.ex., *Thraustochytrium* sps. (p.ex., *Thraustochytrium aureum* e *Thraustochytrium roseum*), *Schizochytrium* sps. (p.ex., *Schizochytrium aggregatum*), *Conidiobolus* sps. (p.ex., *Conidiobolus nanodes*), *Entomorphthora* sps. (p.ex., *Entomorphthora exitalis*), *Saprolegnia* sps. (p.ex., *Saprolegnia parasitica* e *Saprolegnia diclina*), *Leptomitus* sps. (p.ex., *Leptomitus lacteus*), *Entomorphthora* sps., *Pythium* sps., *Porphyridium* sps. (p.ex.,

Porphyridium cruentum), *Conidiobolus* sps., *Phytophthora* sps., *Penicillium* sps., *Coidosporium* sps., *Mucor* sps. (p.ex., *Mucor circinelloides* e *Mucor javanicus*), *Fusarium* sps., *Aspergillus* sps. e *Rhodotorula* sps.), uma levedura (p.ex., *Dipodascopsis uninucleata*), um organismo não mamífero tal como uma mosca (p.ex., *Drosophila melanogaster*) ou *Caenorhabditis* sps. (p.ex., *Caenorhabditis elegans*), ou um mamífero (p.ex., um humano ou um ratinho). Tais sequências podem ser derivadas de espécies dentro do género *Mortierella*, para além da espécie alpina, por exemplo, *Mortierella elongata*, *Mortierella exigua*, *Mortierella isabellina*, *Mortierella hygrophila* e *Mortierella ramanniana*, va. *angulispora*. Para além disso, o presente invento engloba também fragmentos e derivados das sequências nucleotídicas do presente invento (i.e. a sequência da Figura 6 (MAELO), a sequência da Figura 22 (GLELO), a sequência da Figura 43 (HSELO1) e a sequência da Figura 46 (CEELO1)), bem como das sequências derivadas de fontes que não *Mortierella* e possuindo a complementaridade ou correspondência/identidade acima descritas. Os equivalentes funcionais das sequências de cima (i.e., sequências possuindo actividade de alongase) são também englobados pelo presente invento.

Para fins do presente invento, "complementaridade" é definida como o grau de parentesco entre dois segmentos de ADN. É determinada pela medição da capacidade da cadeia com sentido de um segmento de ADN para hibridar com a cadeia anti-sentido do outro segmento de ADN, sob condições apropriadas, para formar uma dupla hélice. Na dupla hélice, sempre que aparece adenina numa cadeia aparece timina na outra cadeia. De modo semelhante, sempre que se verifica guanina numa cadeia verifica-se citosina na outra cadeia. Quanto maior for o parentesco entre as sequências nucleotídicas de dois segmentos de ADN, maior é a capacidade para formar cópias híbridas entre as cadeias de dois segmentos de ADN.

A "identidade" entre duas sequências nucleotídicas é definida como o grau de coincidência, correspondência ou equivalência entre as mesmas cadeias (com sentido ou anti-sentido) de dois segmentos de ADN. Quanto maior for a percentagem de identidade, maior é a correspondência, coincidência ou equivalência entre as cadeias.

A "semelhança" entre duas sequências de aminoácidos é definida como a presença de uma série de resíduos de aminoácidos idênticos bem como conservados em ambas as sequências. Quanto maior o grau de semelhança entre duas sequências de aminoácidos, maior é a correspondência, coincidência ou equivalência das duas sequências. ("Identidade" entre duas sequências de aminoácidos é definida como a presença de uma série de resíduos de aminoácidos exactamente iguais ou invariantes em ambas as sequências).

As definições de "complementaridade", "identidade" e "semelhança" são bem conhecidas dos vulgares peritos na arte.

O presente invento inclui também um polipéptido purificado que alonga os ácidos gordos poli-insaturados e mono-insaturados, em que os polipéptidos (ver, p.ex., a Figura 7 (MAELO)) são codificados pelas sequências nucleotídicas acima descritas. Adicionalmente, o presente invento inclui um polipéptido purificado que alonga ácidos gordos poli-insaturados, em que os polipéptidos (ver, p.ex., a Figura 23 (GLELO)) são codificados pelas sequências nucleotídicas acima descritas. Para além disso, o presente invento inclui também um polipéptido purificado que alonga ácidos gordos poli-insaturados e mono-insaturados, em que os polipéptidos (ver, p.ex., a Figura 44 (HSELO1)) são codificados pelas sequências nucleotídicas acima descritas.

O presente invento engloba também uma sequência nucleotídica isolada que codifica actividade de alongase de AGPI e que é hibridável, sob condições moderadamente restritivas, com um ácido nucleico possuindo uma sequência nucleotídica correspondendo ou complementar à sequência nucleotídica representada pela Figura 6 (MAELD) e/ou pela Figura 22 (GLELO) e/ou pela sequência (HSELO1) mostrada na Figura 43 e/ou a sequência (CEELO1) mostrada na Figura 46. Uma molécula de ácido nucleico é "hibridável" com outra molécula de ácido nucleico quando uma forma de cadeia simples da molécula de ácido nucleico se pode ligar à outra molécula de ácido nucleico sob as condições apropriadas de temperatura e força iónica (ver Sambrook et al., "Molecular Cloning. A Laboratory Manual", Segunda Edição, 1989, Cold Spring Harbor

Laboratory Press, Cold Spring Harbor, New York). As condições de temperatura e força iônica determinam a "severidade" da hibridação. A "hibridação" requer que dois ácidos nucleicos contenham sequências complementares. No entanto, dependendo da severidade da hibridação, podem ocorrer emparelhamentos errados entre as bases. A severidade apropriada para hibridar ácidos nucleicos depende do comprimento dos ácidos nucleicos e do grau de complementaridade. Tais variáveis são bem conhecidas na arte. Mais especificamente, quanto maior for o grau de semelhança ou homologia entre duas sequências nucleotídicas, maior é o valor da T_m , temperatura de fusão, para híbridos de ácidos nucleicos possuindo essas sequências. Para híbridos com mais de 100 nucleótidos de comprimento foram derivadas equações para o cálculo da T_m (ver Sambrook *et al.*, *supra*). Para hibridação com ácidos nucleicos mais curtos, a posição dos emparelhamentos errados torna-se importante e o comprimento do oligonucleótido determina a sua especificidade (ver Sambrook *et al.*, *supra*).

Produção da Enzima Alongase

Uma vez isolado o gene codificando a alongase pode então ser introduzido numa célula hospedeira procariótica ou eucariótica através da utilização de um vector, plasmídeo ou construção.

O vector, por exemplo, um bacteriófago, cosmídeo ou plasmídeo, pode compreender a sequência nucleotídica codificando a alongase bem como qualquer promotor que seja funcional na célula hospedeira e seja capaz de desencadear a expressão da alongase codificada pela sequência nucleotídica. O promotor fica em associação operativa com ou operativamente ligado à sequência nucleotídica. (Um promotor diz-se estar "operativamente ligado" a uma sequência de codificação se o promotor afectar a transcrição ou a expressão da sequência de codificação). Promotores adequados incluem, por exemplo, os dos genes codificando a álcool-desidrogenase, gliceraldeído-3-fosfato-desidrogenase, fosfogluco-isomerase, fosfoglicerato-quinase, fosfatase ácida, T7, TP1, lactase, metalotioneína, inicial imediato de citomegalovírus, proteína ácida do soro do leite, glucoamilase e promotores activados na presença de galactose, por exemplo, GAL1 e GAL10. Adicionalmente,

sequências nucleotídicas que codificam outras proteínas, oligossacáridos, lípidos, etc. podem também ser incluídas dentro do vector bem como outras sequências reguladoras tal como um sinal de poliadenilação (p.ex., o sinal poli-A do antigénio SV-40T, ovalbumina ou hormona de crescimento bovina). A escolha de sequências presentes na construção está dependente dos produtos de expressão desejados bem como da natureza da célula hospedeira.

Tal como observado acima, uma vez construído o vector, este pode ser introduzido na célula hospedeira de escolha através de métodos conhecidos dos vulgares peritos na arte incluindo, por exemplo, transfecção, transformação e electroporação (ver "Molecular Cloning: A Laboratory Manual", 2ª ed., Vol. 1-3, ed. Sambrook et al., Cold Spring Harbor Laboratory Press, 1989). A célula hospedeira é então cultivada sob condições adequadas permitindo a expressão do AGPI que é então recuperado e purificado.

Deve também notar-se que se pode desenhar um triglicérido ou óleo único se se utilizar uma construção ou vector compreendendo a sequência nucleotídica de dois ou mais ADNc (p.ex., MAELO, GLELO, HSEL01 e/ou CEEL01). Este vector pode então ser introduzido numa célula hospedeira. Alternativamente, cada uma das sequências pode ser introduzida num vector separado. Estes vectores podem então ser introduzidos em duas células hospedeiras, respectivamente, ou numa célula hospedeira.

Exemplos de células hospedeiras procarióticas adequadas incluem, por exemplo, bactéria tais como *Escherichia coli*, *Bacillus subtilis* bem como cianobactérias tais como *Spirulina* sps. (i.e., algas azuis-verdes). Exemplos de células hospedeiras eucarióticas adequadas incluem, por exemplo, células de mamífero, células vegetais, células de levedura tais como *Saccharomyces* sps., *Lipomyces* sps., *Candida* sps. tal como *Yarrowia* (Cândida) sps., *Kluyveromyces* sps., *Pichia* sps., *Trichoderma* sps. ou *Hansenula* sps., ou células fúngicas tais como células fúngicas filamentosas, por exemplo, *Aspergillus*, *Neurospora* e *Penicillium*. De preferência, é utilizada *Saccharomyces cerevisiae* (fermento de padeiro).

A expressão numa célula hospedeira pode ser alcançada de um modo transiente ou estável. A expressão transiente pode ocorrer a partir de construções introduzidas que contenham sinais de expressão funcionais na célula hospedeira, mas cujas construções não se repliquem e raramente se integrem na célula hospedeira, ou quando a célula hospedeira não está em proliferação. A expressão transiente também pode ser alcançada através da indução da actividade de um promotor regulável operativamente ligado ao gene de interesse, embora tais sistemas indutíveis exibam frequentemente um baixo nível basal de expressão. A expressão estável pode ser alcançada através da introdução de uma construção que se possa integrar no genoma do hospedeiro ou que se replique autonomamente na célula hospedeira. A expressão estável do gene de interesse pode ser seleccionada através da utilização de um marcador seleccionável situado ou transfectado com a construção de expressão, seguido de selecção de células expressando o marcador. Quando a expressão estável resulta de integração, o local da integração da construção pode ocorrer de modo aleatório dentro do genoma do hospedeiro ou pode ser direccionado através da utilização de construções contendo regiões de homologia com o genoma do hospedeiro suficientes para direccionar a recombinação com o *locus* hospedeiro. Quando as construções são direccionadas para um *locus* endógeno, todas ou algumas das regiões reguladoras da transcrição e da tradução podem ser proporcionadas pelo *locus* endógeno.

Pode também ser utilizado um mamífero transgénico para expressar a enzima de interesse (i.e., a alongase) codificada por uma ou ambas as sequências nucleotídicas acima descritas. Mais especificamente, uma vez criada a construção acima descrita, esta pode ser inserida no pronúcleo de um embrião. O embrião pode então ser implantado numa fêmea receptora. Alternativamente, pode também ser utilizado um método de transferência nuclear (Schnieke *et al.*, *Science* 278: 2130-2133, 1997). Deixa-se então que ocorra a gestação e o nascimento (ver, p.ex., Patente U.S. N.º 5750176 e Patente U.S. N.º 5700671). Leite, tecido ou outras amostras de fluidos da descendência devem então conter níveis alterados de AGPI em comparação com os níveis normalmente verificados no animal não transgénico. As gerações subsequentes podem ser monitorizadas quanto à produção dos níveis alterados ou aumentados de AGPI e

assim a incorporação do gene ou genes codificando a enzima alongase nos seus genomas. O mamífero utilizado como hospedeiro pode ser seleccionado a partir do grupo que consiste, por exemplo, num ratinho, um rato, um coelho, um porco, uma cabra, uma ovelha, um cavalo e uma vaca. No entanto, pode ser utilizado qualquer mamífero desde que tenha a capacidade para incorporar ADN codificando a enzima de interesse no seu genoma.

Para expressão de um polipéptido alongase, as regiões de iniciação e terminação da transcrição e da tradução estão operativamente ligadas ao ADN codificando o polipéptido alongase. As regiões de iniciação e terminação da transcrição e da tradução são derivadas de uma variedade de fontes não exclusivas, incluindo o ADN a expressar, genes conhecidos ou suspeitos de serem capazes de expressão no sistema desejado, vectores de expressão, síntese química ou de um *locus* endógeno numa célula hospedeira. A expressão num tecido vegetal e/ou parte de planta apresenta certas eficiências, particularmente quando o tecido ou parte é um que é colhido precocemente, tal como semente, folhas, frutos, flores, raízes, etc. A expressão pode ser dirigida a esse local na planta através da utilização de uma sequência reguladora específica tal como as das Patentes U.S. N^{os} 5463174, 4943674, 5106739, 5175095, 5420034, 5188958 e 5589379. Alternativamente, a proteína expressa pode ser uma enzima que produza um produto que possa ser incorporado, directamente ou após outras modificações, numa fracção líquida da planta hospedeira. A expressão de um gene ou genes de alongases ou transcritos anti-sentido de alongases, pode alterar os níveis de AGPI específicos, ou derivados destes, verificados em partes da planta e/ou tecidos vegetais. A região de codificação do polipéptido alongase pode ser expressa sozinha ou com outros genes, para produzir tecidos e/ou partes de planta contendo proporções maiores dos AGPI desejados ou em que a composição de AGPI se assemelha mais à do leite materno humano (Prieto *et al.*, publicação PCT WO 95/24494). A região de terminação pode ser derivada da região 3' do gene do qual a região de iniciação foi obtida ou de um gene diferente. É conhecido um grande número de regiões de terminação que se verificou serem satisfatórias numa variedade de hospedeiros do mesmo género e espécie ou de diferentes. A região de terminação é habitualmente

seleccionada por conveniência em vez de ser por qualquer propriedade em particular.

Tal como observado acima, uma planta (p.ex., *Glicine max* (soja) ou *Brassica napus* (canola)), um tecido vegetal, milho, batata, girassol, cártamo ou linho pode ser utilizado como hospedeiro ou célula hospedeira, respectivamente, para expressão da enzima ou enzimas alongases que podem, por sua vez, ser utilizadas na produção de ácidos gordos poli-insaturados. Mais especificamente, os AGPI desejados podem ser expressos na semente. Os métodos de isolamento de óleos de sementes são conhecidos na arte. Assim, para além de proporcionar uma fonte de AGPI, os componentes do óleo da semente podem ser manipulados através da expressão dos genes de alongases, bem como talvez genes de dessaturases, para proporcionar óleos de sementes que possam ser adicionados a composições nutricionais, composições farmacêuticas, rações animais e cosméticos. Novamente, um vector que compreende uma sequência de ADN codificando a alongase operativamente ligada a um promotor, será introduzido no tecido vegetal ou na planta durante um tempo e sob condições suficientes para a expressão do gene de alongase. O vector pode também compreender um ou mais genes que codifiquem outras enzimas, por exemplo, $\Delta 4$ -dessaturase, $\Delta 5$ -dessaturase $\Delta 6$ -dessaturase, $\Delta 8$ -dessaturase, $\Delta 9$ -dessaturase, $\Delta 10$ -dessaturase, $\Delta 12$ -dessaturase, $\Delta 13$ -dessaturase, $\Delta 15$ -dessaturase, $\Delta 17$ -dessaturase e/ou $\Delta 19$ -dessaturase. O tecido vegetal ou planta pode produzir o substrato relevante (p.ex., DGLA, GLA, STA, AA, ADA, EPA, 20:4n-3, etc.) sobre o qual as enzimas actuam ou pode ser introduzido um vector codificando enzimas que produzam tais substratos no tecido vegetal, célula vegetal, planta ou célula hospedeira de interesse. Adicionalmente, o substrato pode ser vaporizado sobre tecidos vegetais que expressem as enzimas apropriadas. Utilizando estas várias técnicas, pode-se produzir AGPI (p.ex., ácidos gordos insaturados n-6 tais como DGLA, AA ou ADA, ou ácidos gordos n-3 tais como EPA ou DHA) através da utilização de uma célula vegetal, tecido vegetal, planta, ou célula hospedeira de interesse. Deve também notar-se que o presente invento engloba também uma planta transgénica compreendendo o vector acima descrito, em que a expressão da sequência nucleotídica do vector resulta na

produção de um ácido gordo poli-insaturado, por exemplo, nas sementes da planta transgênica.

Os substratos que podem ser produzidos pela célula hospedeira naturalmente ou transgenicamente, bem como as enzimas que podem ser codificadas pelas sequências de ADN presentes no vector, que é subsequentemente introduzido na célula hospedeira, são mostrados na Figura 1.

Tendo em vista o de cima, o presente invento engloba também um método de produção de uma das enzimas alongases descritas acima compreendendo os passos de: 1) isolamento da sequência nucleotídica desejada do ADNc da alongase; 2) construção de um vector compreendendo a referida sequência nucleotídica; e 3) introdução do referido vector numa célula hospedeira durante um tempo e sob condições suficientes para a produção da enzima alongase.

O presente invento engloba também um método de produção de ácidos gordos poli-insaturados compreendendo a exposição de um ácido à alongase ou alongases produzidas tal como acima de modo que a alongase converta o ácido num ácido gordo poli-insaturado. Por exemplo, quando o GLA é exposto à alongase, é convertido em DGLA. O DGLA pode então ser exposto a $\Delta 5$ -dessaturase que converte o DGLA em AA. O AA pode então ser convertido em EPA através da utilização da $\Delta 17$ -dessaturase que pode ser, por sua vez, convertido em DHA através da utilização da alongase e uma $\Delta 4$ -dessaturase. Alternativamente, pode ser utilizada alongase para converter $18:4n-3$ em $20:4n-3$ que pode ser exposto a $\Delta 5$ -dessaturase e convertido em EPA. A alongase pode também ser utilizada para converter $18:3n-3$ em $20:3n-3$, que pode ser, por sua vez, convertido em $20:4n-3$ através de uma $\Delta 8$ -dessaturase. Assim, a alongase pode ser utilizada na produção de ácidos gordos poli-insaturados que, por sua vez, podem ser utilizados para determinados fins benéficos. (Ver Figura 1 para uma ilustração dos muitos papéis críticos que a alongase tem em várias vias biossintéticas).

Utilizações do Gene da Alongase e da Enzima Codificada por Ele

Tal como observado acima, os ADNc de alongase isolados e as correspondentes enzimas alongases (ou polipéptidos

purificados) codificadas por eles têm muitas utilizações. Por exemplo, cada ADNc e a correspondente enzima podem ser utilizados indirectamente ou directamente na produção de ácidos gordos poli-insaturados, por exemplo, DGLA, AA, ADA, $20:4n-3$ ou EPA. ("Directamente" pretende englobar a situação em que a enzima converte directamente o ácido noutro ácido, o último dos quais é utilizado numa composição (p.ex., a conversão de GLA em DGLA)). "Indirectamente" pretende englobar a situação em que um ácido gordo é convertido noutro ácido gordo (i.e., um intermediário da via) através da alongase (p.ex., GLA em DGLA) e depois o último ácido gordo é convertido noutro ácido gordo através da utilização de uma enzima não alongase (p.ex., DGLA em AA através da $\Delta 5$ -dessaturase). Estes ácidos gordos poli-insaturados (i.e., os produzidos directamente ou indirectamente através da actividade da enzima alongase) podem ser adicionados, por exemplo, a composições nutricionais, composições farmacêuticas, cosméticos e rações animais, todos os quais são englobados pelo presente invento. Estas utilizações estão descritas, em detalhe, abaixo.

Composições Nutricionais

Uma das utilizações do ácido gordo poli-insaturado produzido de acordo com os métodos do presente invento inclui composições nutricionais. Tais composições, para fins do presente invento, incluem qualquer alimento ou preparação para consumo humano incluindo para consumo entérico ou parentérico, que quando tomado pelo corpo (a) serve para alimentar ou construir tecidos ou fornecer energia e/ou (b) mantém, restaura ou suporta o estado nutricional ou a função metabólica adequada.

A composição nutricional compreende pelo menos um óleo ou ácido produzido através da utilização de pelo menos uma enzima alongase, produzida utilizando o respectivo gene de alongase, e pode estar numa forma sólida ou líquida. Adicionalmente, a composição pode incluir macronutrientes edíveis, vitaminas e minerais em quantidades desejadas para uma determinada utilização. A quantidade de tais ingredientes variará dependendo de se a composição é pretendida para utilização com bebés e crianças normais e saudáveis ou adultos possuindo

necessidades especiais tais como as que acompanham certas condições metabólicas (p.ex., distúrbios metabólicos).

Exemplos de macronutrientes que podem ser adicionados à composição incluem mas não se limitam a gorduras edíveis, hidratos de carbono e proteínas. Exemplos de tais gorduras edíveis incluem mas não se limitam a óleo de coco, óleo de soja e mono e diglicéridos. Exemplos de tais hidratos de carbono incluem mas não se limitam a glicose, lactose edível e amido hidrolisado. Adicionalmente, exemplos de proteínas que podem ser utilizadas na composição nutricional do presente invento incluem mas não se limitam a proteínas de soja, soro do leite electrodialisado, leite magro electrodialisado, soro do leite ou os hidrolisados destas proteínas.

Em relação às vitaminas e minerais, pode ser adicionado às composições nutricionais o seguinte: cálcio, fósforo, potássio, sódio, cloreto, magnésio, manganês, ferro, cobre, zinco; selénio, iodo e Vitaminas A, E, D, C e complexo B. Outras vitaminas e minerais como estes podem também ser adicionados.

Os componentes utilizados nas composições nutricionais serão de origem semi-purificada ou purificada. Por semi-purificado ou purificado entenda-se um material que foi preparado através de purificação de um material natural ou através de síntese.

Exemplos de composições nutricionais incluem mas não se limitam a fórmulas infantis, suplementos dietéticos, substitutos dietéticos e composições de re-hidratação. Composições nutricionais de particular interesse incluem mas não se limitam às utilizadas para suplementação entérica e parentérica de bebés, fórmulas infantis especiais, suplementos para idosos e suplementos para pessoas com dificuldades gastrointestinais e/ou mal-absorção.

A composição nutricional pode também ser adicionada à comida mesmo quando não é necessária suplementação da dieta. Por exemplo, a composição pode ser adicionada a comida de qualquer tipo incluindo mas não se limitando a margarinas, manteigas modificadas, queijos, leite, iogurte, chocolate,

doces, aperitivos, óleos de saladas, óleos de cozinha, gorduras de cozinha, carnes, peixe e bebidas.

Numa concretização preferida, a composição nutricional é um produto nutricional entérico, de preferência um produto nutricional entérico de adultos ou pediátrico. Esta composição pode ser administrada a adultos ou crianças sofrendo de stress ou possuindo necessidades especiais devido a estados de doença crónica ou aguda. A composição pode compreender, para além de ácidos gordos poli-insaturados produzidos de acordo com o presente invento, macronutrientes, vitaminas e minerais tal como descrito acima. Os macronutrientes podem estar presentes em quantidades equivalentes às presentes no leite humano ou numa base energética, i.e., com base nas calorias.

Os métodos para formulação de fórmulas nutricionais entéricas e parentéricas líquidas ou sólidas são bem conhecidos na arte (ver também os Exemplos abaixo).

A fórmula entérica, por exemplo, pode ser esterilizada e subsequentemente utilizada numa base de pronta-a-dar (RTF) ou armazenada num líquido ou pó concentrado. O pó pode ser preparado através de liofilização da fórmula preparada tal como indicado acima e reconstituindo-a através de re-hidratação do concentrado. As fórmulas de adulto e pediátricas são bem conhecidas na arte e estão comercialmente disponíveis (p.ex., Similac[®], Ensure[®], Jevity[®] e Alimentum[®] em Ross Products Division, Abbott Laboratories, Columbus, Ohio). Um óleo ou ácido gordo produzido de acordo com o presente invento pode ser adicionado a qualquer uma destas fórmulas.

A densidade energética das composições nutricionais do presente invento, quando na forma líquida, podem variar de cerca de 0,6 Kcal a cerca de 3 Kcal por ml. Quando na forma sólida ou em pó, os suplementos nutricionais podem conter de cerca de 1,2 a mais de 9 Kcal por grama, de preferência de cerca de 3 a 7 Kcal por g. Em geral, a osmolalidade de um produto líquido deve ser inferior a 700 mOsm e, de preferência, inferior a 660 mOsm.

A fórmula nutricional pode incluir macronutrientes, vitaminas e minerais, tal como observado acima, para além dos

AGPI produzidos de acordo com o presente invento. A presença destes componentes adicionais ajuda o indivíduo a ingerir os requisitos mínimos diários destes elementos. Para além de proporcionar AGPI, pode também ser desejado adicionar à composição zinco, cobre, ácido fólico e antioxidantes. Crê-se que esta substância reforce um sistema imunitário sob *stress* e proporcionará por essa razão mais benefícios ao indivíduo que receba a composição. Uma composição farmacêutica pode também ser suplementada com estes elementos.

A composição nutricional compreende, para além de antioxidantes e pelo menos um AGPI, uma fonte de hidratos de carbono em que pelo menos 5% do peso do hidrato de carbono é oligossacárido não digerível. A composição nutricional pode adicionalmente compreender proteína, taurina e carnitina.

Tal como observado acima, os AGPI produzidos de acordo com o presente invento, ou derivados destes, podem ser adicionados a um substituto ou suplemento dietético, particularmente uma fórmula infantil, para pacientes sujeitos a alimentação intravenosa ou para prevenção ou tratamento de mal-nutrição ou outras condições ou estados de doença. Como antecedente, deve notar-se que o leite materno humano tem um perfil de ácidos gordos compreendendo de cerca de 0,15% a cerca de 0,36% como DHA, de cerca de 0,03% a cerca de 0,13% como EPA, de cerca 0,30% a cerca de 0,88% como AA, de cerca de 0,22% a cerca de 0,67% como DGLA e de cerca de 0,27% a cerca de 1,04% como GLA. Assim, ácidos gordos como DGLA, AA, EPA e/ou ácido docosa-hexaenóico (DHA), produzidos de acordo com o presente invento, podem ser utilizados para alterar, por exemplo, a composição de fórmulas infantis para melhor replicar o teor de AGPI do leite materno humano ou para alterar a presença dos AGPI normalmente verificados no leite de um mamífero não humano. Em particular, uma composição para utilizar num suplemento farmacológico ou alimentar, particularmente um substituto do leite materno humano ou suplemento, compreenderá de preferência um ou mais de AA, DGLA e GLA. De maior preferência, a mistura de óleos compreenderá de cerca de 0,3 a 30% de AA, de cerca de 0,2 a 30% de DGLA, e/ou de cerca de 0,2 a cerca de 30% de GLA.

As composições nutricionais parentéricas compreendendo de cerca 2 a cerca de 30 por cento do peso de ácidos gordos calculado como triglicéridos são um exemplo de utilizações dos ácidos poli-insaturados do presente invento. A composição preferida tem cerca de 1 a cerca de 25 por cento do peso da composição total de AGPI como GLA (Patente U.S. N.º 5196198). Outras vitaminas, particularmente vitaminas solúveis em gordura tais como vitamina A, D, E e L-carnitina podem ser opcionalmente incluídas. Quando desejado, pode ser adicionado um conservante tal como alfa-tocoferol numa quantidade de cerca de 0,1% do peso.

Adicionalmente, as proporções de AA, DGLA e GLA podem ser adaptadas para uma determinada dada utilização final. Quando formulada como suplemento ou substituto do leite materno, será proporcionada uma composição que compreende um ou mais de AA, DGLA e GLA numa proporção de cerca de 1:19:30 a cerca de 6:1:0,2, respectivamente. Por exemplo, o leite materno dos animais pode variar nas proporções de AA:DGLA:GLA variando de 1:19:30 a 6:1:0,2, que inclui proporções intermédias que são de preferência cerca de 1:1:1, 1:2:1, 1:1:4. Quando produzidos juntos numa célula hospedeira, pode ser utilizado o ajustamento da taxa e percentagem de conversão de um substrato precursor tal como GLA e DGLA em AA para controlar de modo preciso as proporções dos AGPI. Por exemplo, pode ser utilizada uma taxa de conversão de 5% a 10% de DGLA para AA para produzir uma proporção de AA para DGLA de cerca de 1:19, enquanto pode ser utilizada uma taxa de conversão de cerca de 75% a 80% para produzir uma proporção de AA para DGLA de cerca de 6:1. Por essa razão, tanto num sistema de cultura de células como num animal hospedeiro, a regulação do momento, da extensão e da especificidade da expressão da alongase, bem como da expressão de outras dessaturases, pode ser utilizada para modular os níveis e as proporções de AGPI. Os AGPI/ácidos produzidos de acordo com o presente invento (p.ex., AA e DGLA) podem então ser combinados com outros AGPI/ácidos (p.ex., GLA) nas concentrações e proporções desejadas.

Adicionalmente, os AGPI produzidos de acordo com o presente invento ou as células hospedeiras contendo-os podem também ser utilizados como suplementos alimentares de animais para alterar a composição de ácidos gordos de um tecido ou do

leite de um animal para um mais desejável para consumo humano ou animal.

Composições Farmacêuticas

Outra utilização dos ácidos poli-insaturados do presente invento é na produção de uma composição farmacêutica compreendendo um ou mais dos ácidos gordos e/ou óleos resultantes produzidos utilizando pelo menos um dos ADNc de alongase (i.e., MAELO, GLELO, HSELO1 ou CEELO), de acordo com os métodos aqui descritos. Mais especificamente, uma tal composição farmacêutica pode compreender um ou mais dos ácidos e/ou óleos bem como um transportador, adjuvante ou veículo padrão, bem conhecido, não tóxico e farmacêuticamente aceitável tal como, por exemplo, solução salina tamponada com fosfato, água, etanol, polióis, óleos vegetais, um agente humidificante ou uma emulsão tal como uma emulsão de água/óleo. A composição pode estar numa forma líquida ou sólida. Por exemplo, a composição pode estar na forma de um comprimido, cápsula, líquido ou pó ingerível, injectável, ou unguento ou creme tópico. A fluidez apropriada pode ser mantida, por exemplo, através da manutenção do tamanho de partícula requerido no caso de dispersões e através da utilização de tensioactivos. Pode também ser desejável incluir agentes isotónicos, por exemplo, açúcares, cloreto de sódio e semelhantes. Para além de tais diluentes inertes, a composição pode também incluir adjuvantes, tais como agentes humidificantes, agentes emulsionantes e de suspensão, agentes adoçantes, agentes aromatizantes e agentes perfumantes.

As suspensões, para além dos compostos activos, podem compreender agentes de suspensão tais como, por exemplo, álcoois de isostearilo etoxilados, sorbitol de polioxietileno e ésteres de sorbitano, celulose microcristalina, meta-hidróxido de alumínio, bentonite, ágar-ágar e goma adragante ou misturas destas substâncias.

Formas de dosagem sólidas tais como comprimidos e cápsulas podem ser preparadas utilizando técnicas bem conhecidas na arte. Por exemplo, os AGPI produzidos de acordo com o presente invento podem ser comprimidos com bases de comprimidos convencionais tais como lactose, sacarose e amido

de milho em combinação com aglomerantes tais como goma-arábica, amido de milho ou gelatina, agentes desintegrantes tais como amido de batata ou ácido algínico e um lubrificante tal como ácido esteárico ou estearato de magnésio. As cápsulas podem ser preparadas através da incorporação destes excipientes numa cápsula de gelatina juntamente com antioxidantes e o AGPI ou AGPI relevantes. Os componentes antioxidantes e AGPI devem estar dentro das linhas orientadoras apresentadas acima.

Para administração intravenosa, os AGPI produzidos de acordo com o presente invento ou derivados destes podem ser incorporados em formulações comerciais tais como Intralipids[™]. O perfil plasmático típico normal de ácidos gordos de um adulto compreende 6,64 a 9,46% de AA, 1,45 a 3,11% de DGLA e 0,02 a 0,08% de GLA. Estes AGPI ou os seus precursores metabólicos podem ser administrados sozinhos ou em combinação com outros AGPI para alcançar um perfil de ácidos gordos normal num paciente. Quando desejado, os componentes individuais das formulações podem ser proporcionados individualmente, na forma de estojo, para uma única utilização ou múltipla. Uma dosagem típica de um determinado ácido gordo é de 0,1 mg a 20 g (até 100 g) diariamente e de preferência de 10 mg a 1, 2, 5 ou 10 g diariamente.

As possíveis vias de administração das composições farmacêuticas do presente invento incluem, por exemplo, entérica (p.ex., oral ou rectal) e parentérica. Por exemplo, uma preparação líquida pode ser administrada, por exemplo, oralmente ou rectalmente. Adicionalmente, pode ser completamente dispersa uma mistura homogénea em água, misturada sob condições estéreis com diluentes, conservantes, tampões ou propulsores fisiologicamente aceitáveis para formar um vapor ou inalante.

A via de administração dependerá, como é claro, do efeito desejado. Por exemplo, se a composição estiver a ser utilizada para tratar pele áspera, seca ou envelhecida, para tratar pele ferida ou queimada, ou para tratar pele ou cabelo afectado por uma doença ou condição, pode talvez ser aplicada topicamente.

A dosagem da composição a administrar ao paciente pode ser determinada por um vulgar perito na arte e depende de vários factores tais como o peso do paciente, a idade do paciente, o estado imunitário do paciente, etc.

Em relação à forma, a composição pode ser, por exemplo, uma solução, uma dispersão, uma suspensão, uma emulsão ou um pó estéril que é então reconstituído.

As composições farmacêuticas e/ou nutricionais de cima são úteis no tratamento de vários distúrbios. Em particular, estas composições podem ser utilizadas para tratar restenose após angioplastia. Para além disso, sintomas de inflamação, artrite reumatóide, asma e psoríase podem também ser tratados com estas composições. Provas indicam também que os AGPI podem estar envolvidos no metabolismo do cálcio; assim, as composições podem, talvez, ser utilizadas no tratamento ou prevenção de osteoporose e de pedras renais ou do tracto urinário.

Adicionalmente, estas composições podem também ser utilizadas no tratamento de cancro. Mostrou-se que as células malignas têm composições em ácidos gordos alteradas. Mostrou-se que a adição de ácidos gordos torna o seu crescimento mais lento, causa morte celular e aumenta a sua susceptibilidade a agentes quimioterapêuticos. Para além disso, as composições do presente invento podem também ser úteis para tratamento de caquexia associada a cancro.

Estas composições podem ser utilizadas para tratar diabetes (ver Patente U.S. N.º 4826877 e Horrobin *et al.*, *Am. J. Clin. Nutr.* Vol. 57(Supl.) 732S-737S). Foram demonstrados metabolismo e composição de ácidos gordos alterados em animais diabéticos.

Para além disso, estas composições, compreendendo AGPI produzidos directamente ou indirectamente através da utilização da enzima ou enzimas alongases, podem também ser utilizadas no tratamento de eczema, na redução da pressão sanguínea e no melhoramento dos valores matemáticos dos exames. Adicionalmente, estas composições podem ser utilizadas na inibição da agregação de plaquetas, indução de

vasodilatação, redução nos níveis de colesterol, inibição da proliferação de tecido de músculo liso e tecido fibroso na parede dos vasos (Brenner *et al.*, *Adv. Exp. Med. Siol.*, Vol. 83: p. 85-101, 1976), redução ou prevenção de hemorragia gastrointestinal e de outros efeitos secundários dos fármacos anti-inflamatórios não esteróides (ver Patente U.S. N.º 4666701), prevenção ou tratamento de endometriose e de síndrome pré-menstrual (ver Patente U.S. N.º 4758592) e tratamento de encefalomielite miálgica e de fadiga crónica após infecções virais (ver Patente U.S. N.º 5116871).

Outras utilizações das composições incluem a utilização no tratamento de SIDA, esclerose múltipla e distúrbios inflamatórios da pele, bem como para manutenção da saúde geral.

Adicionalmente, esta composição pode ser utilizada para fins cosméticos. Pode ser adicionada a composições cosméticas preexistentes de modo a que seja formada uma mistura ou pode ser utilizada como uma única composição.

Aplicações Veterinárias

Deve notar-se que as composições farmacêuticas e nutricionais acima descritas podem ser utilizadas em associação com animais (i.e., domésticos ou não domésticos), bem como com humanos, uma vez que os animais experimentam muitas das mesmas necessidades e condições que os humanos. Por exemplo, o óleo ou ácidos do presente invento podem ser utilizados em suplementos de rações animais, substitutos de rações animais, vitaminas animais ou em unguentos tópicos animais.

O presente invento pode ser ilustrado através da utilização dos seguintes exemplos não limitantes:

Exemplo I

Determinação da Utilização de Codões em *Mortierella alpina*

A extremidade 5' de 1000 clones de ADNc aleatórios foi sequenciada a partir de uma biblioteca de ADNc de *Mortierella alpina*. As sequências foram traduzidas em seis enquadramentos

de leitura utilizando GCG (Genetics Computer Group (Madison, Wisconsin)) com o algoritmo FastA (Pearson e Lipman, *Proc. Natl. Acad. Sci. USA* 85: 2444-2448, 1988) para pesquisar semelhanças entre uma sequência desconhecida e um grupo de sequências do mesmo tipo (ácido nucleico ou proteína), especificamente com a base de dados Swissprot (GeneBio, Genebra, Suíça). Muitos dos clones foram identificados como um gene putativo do metabolismo celular ("*housekeeping*") com base na homologia da sequência proteica com genes conhecidos. Vinte e uma sequências de ADNc de *M. alpina* que coincidiram com genes "*housekeeping*" conhecidos na base de dados foram seleccionadas (ver Tabela 1 abaixo). A tabela dos codões tendenciais de *M. alpina* (ver Tabela 2) foi gerada com base nestas 21 sequências bem como nas sequências inteiras da $\Delta 5$ -dessaturase (ver Figura 18), $\Delta 6$ -dessaturase e $\Delta 12$ -dessaturase de *M. alpina*. Uma vez que o alinhamento FastA entre a proteína putativa codificada pela sequência de ADNc de *M. alpina* e a sequência proteica conhecida era fraco nalgumas áreas, apenas foram utilizados os codões das áreas de forte homologia.

Tabela 1

N.º do Clone	Coincidência	n.º de pb	n.º de aa
193	Factor de alongamento 1-alfa	426	142
143	Proteína ribossómica 60S L17	417	139
235	Actina I	360	120
299	Proteína ribossómica 40S YS11	387	129
390	Proteína relacionada com Ras rab-1a	342	114
65	Proteína ribossómica 40S RP10	366	122
289	Enzima de conjugação com ubiquitina E2-16 KD	294	98
151	Ubiquinol-citocromo C-redutase	375	125
80	Factor de iniciação 5A-2	183	61
33	Proteína ribossómica 60S L15	252	84
132	Proteína ribossómica 60S L3-2	300	100
198	Histona H3	285	95
286	6-Fosfogluconato-desidrogenase, descarboxilante	363	121
283	Proteína ribossómica 40S S22	261	87
127	Factor de alongamento 2	231	77

197	Actina, gama	252	84
496	Proteína ribossômica 40S S16	270	90
336	Histona H4	219	73
262	Ubiquitina	228	76
188	Proteína semelhante à subunidade beta da proteína de ligação ao nucleótido guanina	213	71
81	Ubiquitina	228	76
21	TOTAL	6252	2084

Tabela 2

Aminoácido	Codão	%	Aminoácido	Codão	%
	Preferencial	utilizada		Preferencial	utilizada
Ala	GCC	63%	Lys	AAG	96%
Arg	CGC	50%	Met	ATG	100%
Asn	AAC	97%	Phe	TTC	78%
Asp	GAC	65%	Pro	CCC	68%
Cys	TGC	87%	Ser	TCC	46%
Gln	CAG	78%	Thr	ACC	78%
Glu	GAG	85%	Trp	TGG	100%
Gly	GGT	47%	Tyr	TAC	95%
His	CAC	91%	Val	GTC	72%
Ile	ATC	72%	Stop	TAA	50%
Leu	CTC	49%			

Exemplo IIClonagem de um ADNc Inteiro Semelhante a Alongase de *M. alpina*

A β -cetoacil-coenzima A-sintetase (KCS) de jojoba e a alongase de *Saccharomyces cerevisiae* (ELO2) foram alinhadas para determinar uma área de homologia de aminoácidos (ver Figura 2). O codão preferencial foi aplicado na área da sequência correspondente aos aminoácidos homólogos entre as duas alongases, e foram desenhados iniciadores com base nesta sequência preferencial (ver Figura 3). O ADNc foi excisado da biblioteca de ADNc de *M. alpina* M11 (Knutzon et al., *J. Biol. Chem.* 273: 29360-29366, 1998), que contém aproximadamente 6×10^5 clones com um tamanho de inserção médio de 1,1 Kb. O ADNc excisado foi amplificado com o iniciador interno R0339

(5'-TTG GAG AGG AGG AAG CGA CCA CCG AAG ATG ATG-3') e um iniciador directo no vector R0317 (5'-CAC ACA GGA AAC AGC TAT GAC CAT GAT TAC G-3'). A reacção em cadeia da polimerase (PCR) foi realizada num volume de 100 µl contendo: 300 ng de biblioteca de ADNc de *M. alpina* excisada, 50 pmol de cada iniciador, 10 µl de tampão 10×, 1 µl de Mistura de Nucleótidos de PCR 10 mM (Boehringer Mannheim Corp., Indianapolis, IN) e 1,0 U de Polimerase Taq. As condições dos ciclos térmicos em Perkin Elmer 9600 (Norwalk, CT) foram como se segue: 94°C durante 2 minutos, depois 30 ciclos de 94°C durante 1 min., 58°C durante 2 min. e 72°C durante 3 min. A PCR foi seguida por uma extensão adicional a 72°C durante 7 minutos.

O produto de PCR amplificado foi corrido num gel e o fragmento amplificado de aproximadamente 360 pb foi purificado em gel e o fragmento isolado foi directamente sequenciado utilizando um sequenciador de ADN ABI 373A (Perkin Elmer, Foster City, CA). O pacote de análise de sequências de GCG foi utilizado para comparar a sequência obtida com sequências conhecidas. A sequência foi traduzida em todos os seis enquadramentos de leitura no Programa de Análise de GCG utilizando o algoritmo FastA (Pearson e Lipman, *supra*). A base de dados Swissprot (GeneBio, Genebra, Suíça) de proteínas foi pesquisada. Este fragmento de ADNc traduzido foi identificado como parte de uma alongase putativa com base na homologia da sequência proteica putativa com a ELO2 de *S. cerevisiae* (GNS1), possuindo 41,3% de identidade em 63 aminoácidos.

Novos iniciadores foram desenhados com base na sequência de alongase putativa e a sequência do vector, pZL1 (Life Technologies, Inc., Gaithersburg, MD) utilizado para construir a biblioteca de ADNc de *M. alpina*. A biblioteca de ADNc excisado de *M. alpina* foi novamente amplificada por PCR utilizando os iniciadores R0350 (5'-CAT CTC ATG GAT CCG CCA TGG CCG CCG CAA TCT TG-3'), ao qual foi adicionado um local de restrição *Bam*HI (sublinhado), e o iniciador inverso do vector R0352 (5'-ACG CGT ACG TAA AGC TTG-3') para isolar o ADNc da alongase de *M. alpina* inteiro, utilizando as condições anteriormente descritas. Os terminais do fragmento de PCR amplificado de aproximadamente 1,5 kb foram preenchidos com ADN-polimerase T4 (Boehringer Mannheim Corp., Indianapolis, IN) para criar extremidades cegas e clonado no vector

pCR-blunt (Invitrogen Corp., Carlsbad, CA). Isto resultou em dois clones, pRAE-1 e pRAE-2 (ver Figura 4A). (O ADN plasmídico pRAE-2 foi depositado na American Type Culture Collection, 10801 University Boulevard, Manassas, VA 20110-2209, a 28 de Agosto de 1998, sob os termos do tratado de Budapeste e foi de acordo com o número de depósito ATCC 203166). Os ADNc de alongases destes vectores foram cortados como um fragmento *EcoRI* e clonados no vector pYX242 digerido com *EcoRI* (Novagen, Madison, WI). Os clones pRAE-5 e pRAE-6 (ver Figura 4B) têm os ADNc de alongases de pRAE-1 e pRAE-2, respectivamente. (O ADN plasmídico pRAE-5 foi depositado na American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia 20110-2209, a 28 de Agosto de 1998, sob os termos do Tratado de Budapeste, e foi-lhe atribuído o número de depósito 203167). A sequenciação de pRAE-5 e pRAE-6 revelou que a região 5' não traduzida do gene de alongase em pRAE-5 é 16 pb mais curta que em pRAE-6 (ver Figura 5). A sequência do ADNc da alongase de *M. alpina* completa, designada MAELO foi obtida a partir de pRAE-2 (ver Figura 6). A Figura 7 é a sequência de aminoácidos obtida a partir da tradução de MAELO. A base de dados Swissprot (GeneBio, Genebra, Suíça) foi novamente pesquisada, tal como anteriormente descrito, com a MAELO traduzida: MAELO tem uma identidade de 44,3% em 317 aminoácidos com GNS1(EL02) de *S. cerevisiae* e 44,7% de identidade em 318 aminoácidos com SUR4(EL03) de *S. cerevisiae*. O alinhamento FastA entre as três alongases é mostrado na Figura 8. Ao nível nucleotídico (ver Figura 9), MAELO tem uma identidade de 57,4% em 549 pb sobreponíveis com GNS1(EL02) de *S. cerevisiae* (N.º de Acesso GenBank S78624). No entanto, a identidade entre o gene MAELO completo de 954 pb e GNS1(EL02) de *S. cerevisiae* é de 33,0%.

Exemplo III

Expressão de ADNc de Alongase de *M. alpina* em Fermento de Padeiro

As construções pRAE-5 e pRAE-6 foram transformadas em *S. cerevisiae* 334 (Hoveland et al., Gene 83: 57-64, 1989) e pesquisadas quanto a actividade de alongase. O plasmídeo pCGN7875 (Calgene LLC, Davis, CA) contendo o gene de KCS de jojoba no vector pYES2 (Invitrogen Corp., Carlsbad, CA) foi utilizado como controlo positivo. O substrato utilizado para

detectar actividade de alongase na alongase de *M. alpina* (MAELO) foi GLA e em KCS de jojoba foi ácido oleico (OA). A estirpe de controlo negativo foi *S. cerevisiae* 334 contendo o vector pYX242. As culturas foram criadas durante 40-48 horas a 25°C, em meio selectivo (Ausubel et al., "Short Protocols in Molecular Biology", Cap. 13, pág. 3-5, 1992), na presença de um determinado substrato. A expressão do gene KCS de jojoba clonado em pYES2 estava sob o controlo do promotor GAL1, enquanto o promotor em pYX242 é TP1, que é constitutivo. Deste modo, as culturas de 334(pCGN7875) e 334(pYES2) foram induzidas com galactose. A análise de GC-FAME da fracção lipídica de cada sedimento celular foi efectuada tal como anteriormente descrito (Knutzon et al., supra).

Os resultados da actividade de alongase de diferentes experiências são proporcionados na Figura 10A e 10B. A KCS de jojoba alonga ácidos gordos mono-insaturados de cadeia longa 18:1n-9 a 20:1n-9. A homologia de aminoácidos entre a alongase de *M. alpina* (MAELO) e as ELO2 e ELO3 de *S. cerevisiae* sugeriu que as proteínas codificadas por estes genes podem ter uma especificidade de substrato semelhante. A actividade da alongase de *M. alpina*, alongamento (MAELO) de ácidos gordos mono-insaturados e saturados de cadeia longa, é observada na conversão de 18:1n-9 em 20:1n-9 e também na síntese de 24:0. A estirpe de controlo, 334(pYX242) tem muito pouca ou nenhuma quantidade detectável de 20:1 e 24:0 (ver Figura 10A). A alongase de *M. alpina* (MAELO) também actua pelo menos num AGPI, convertendo 18:3n-6(GLA) em 20:3n-6(DGLA). A percentagem de 20:3n-6 nos lípidos totais é maior na estirpe 334(pRAE-5) e 334(pRAE-6) com o ADNc da alongase de *M. alpina* (MAELO) quando comparada com a do controlo 334(pYX242). As percentagens de 20:3n-6 produzidas foram de 0,092% para 334(pYX242) vs. 0,324% para 334(pRAE-5) e 0,269% para 334(pRAE-6) (mostrado entre parêntesis nas Figuras 10A e 10B). Esta diferença no perfil de ácidos gordos é também observada na quantidade total de 20:3n-6 produzida. Apenas 0,226 µg de 20:3n-6 foram produzidos por 334(pYX242) enquanto que 334(pRAE-5) e 334(pRAE-6) produziram 2,504 µg de 20:3n-6 e 1,006 µg de 20:3n-6, respectivamente. Também, quando não é adicionado substrato, o nível de 20:3n-6 não é detectável.

Uma vez gerado 20:3n-6 pela alongase de *M. alpina* (MAELO), a $\Delta 5$ -dessaturase pode convertê-lo em AA no sistema de expressão desejado. Para testar esta hipótese, as construções pRAE-5 e pCGR-4 (um plasmídeo contendo $\Delta 5$ -dessaturase) foram co-transformados em *S. cerevisiae* 334 e pesquisados quanto à produção de AA. O substrato utilizado foi GLA (18:3n-6) 25 μ M. Se a alongase de *M. alpina* (MAELO) for activa em levedura, então o substrato será convertido em DGLA (20:3n-6), que a $\Delta 5$ -dessaturase converterá em AA (20:4n-6). Os resultados da Figura 11 confirmam a produção de AA e por essa razão, a actividade da alongase de *M. alpina* (MAELO).

A expressão da $\Delta 5$ -dessaturase, $\Delta 6$ -dessaturase e $\Delta 12$ -dessaturase, em levedura, juntamente com a alongase, deve resultar na produção de AA (ver Figura 1) sem necessidade de um fornecimento exógeno de ácidos gordos.

Exemplo IV

Uma Comparação da Expressão do ADNC da Alongase de *M. alpina* MAELO com a Alongase ELO2 de *S. cerevisiae* em Fermento de Padeiro

O gene ELO2 codificando para a alongase de levedura foi clonado a partir de uma biblioteca genómica de *S. cerevisiae* (Origene, Rockville, MD) utilizando os iniciadores R0514 (5'-GGC TAT GGA TCC ATG AAT TCA CTC GTT ACT CAA TAT G-3') e R0515 (5'-CCT GCC AAG CTT TTA CCT TTT TCT TCT GTG TTG AG-3') incorporando os locais de restrição (sublinhados) *Bam*HI e *Hind*III (respectivamente). O gene ELO2 foi clonado no vector pYX242 nos locais *Bam*HI e *Hind*III, designado pRELO, transformado no hospedeiro 334 de *S. cerevisiae* (Hoveland et al., supra) e pesquisado quanto à actividade de alongase de AGPI. O plasmídeo vector foi utilizado como controlo negativo e foi criada 334(pRAE-5) para comparar a actividade de alongase de AGPI. As culturas foram criadas tal como anteriormente descrito sem galactose no meio e foi adicionado GLA 25 μ M como substrato. A Figura 12 mostra que a quantidade de 20:3n-6 ou DGLA produzida (alongado a partir 18:3n-6 ou GLA) por 334(pRAE-5) era aproximadamente 4 vezes a do controlo negativo contendo o vector inalterado pYX242, enquanto que os dois clones individuais 334(pRELO-1) e 334(pRELO-2) foram apenas duas vezes o controlo negativo. Adicionalmente, quando

o DGLA produzido é expresso como uma percentagem dos lípidos totais (mostrado entre parêntesis, Figura 12), os clones 334(pRELO-1) e 334 (pRELO-2) produziram 0,153% e 0,2% de DGLA, respectivamente, enquanto que 334(pYX242) produziu 0,185% de DGLA. Deste modo todas estas estirpes produziram percentagens comparáveis de DGLA. A estirpe 334(pRAE-5), no entanto, produziu 0,279% de DGLA, um aumento de 50,8% em relação a 334(pYX242) (controlo negativo). Estes dados mostram que o gene ELO2 da alongase de *S. cerevisiae*, mesmo quando sobre-expresso em levedura, não alonga eficazmente GLA até DGLA. A actividade de alongase de AGPI de *M. alpina* é específica para esta conversão tal como evidenciado pela maior quantidade de DGLA produzida em comparação com o controlo, 334(pYX242).

Exemplo V

Identificação de Alongases a Partir de Outras Fontes Utilizando MAELO

O algoritmo TFastA (Pearson e Lipman, *supra*) é utilizado para pesquisar semelhanças entre uma sequência peptídica de busca e a sequência de ADN na base de dados traduzida em cada um dos seis enquadramentos de leitura. A MAELO traduzida foi utilizada como busca para uma pesquisa TFastA em GCG com a base de dados GenEMBL (6/98) de GCG para identificar outras potenciais sequências de alongases com base nas suas comparações de semelhança de aminoácidos com a MAELO traduzida. Por exemplo, nas Figuras 13 e 14, são mostrados dois alinhamentos entre as traduções de duas sequências diferentes de *C. elegans* do cromossoma III e MAELO. A sequência de ADN de *C. elegans* (número de acesso GenBank Z68749) foi anotada como denotando semelhança com GNS1 (ELO2), enquanto a sequência de ADN de *C. elegans* adicional (número de acesso GenBank U61954) foi anotada como semelhante tanto a GNS1 como a SUR4 (ELO3). Estes são fragmentos de ADN processados nos quais os intrões foram removidos da sequência genómica, e os exões montados e traduzidos. A quantidade de identidade de aminoácidos entre as alongases putativas de AGPI de *C. elegans* e a MAELO traduzida está à volta de 30%. Isto apontaria para uma função comum no metabolismo dos ácidos gordos, p.ex., uma alongase de AGPI. A Figura 15 é outro exemplo de uma sequência de *C. elegans* traduzida (número de acesso GenBank AF003134) do cromossoma III. Foi identificada

uma sequência de ADN que tinha homologia com o ADN para a EL02 de *S. cerevisiae*. Outras inspecções desta sequência de ADN e da sua tradução em aminoácidos determinou que havia homologia com a MAELO traduzida. *C. elegans*, pode, portanto, conter uma alongase de AGPI.

A Figura 16 mostra os alinhamentos das sequências de ADN traduzidas de ratinho e humano, respectivamente, com a MAELO traduzida. A sequência de ratinho CIG30, número de acesso GenBank U97107, foi isolada a partir de tecido adiposo castanho e relatada como sendo "semelhante à proteína SUR4 de levedura". Tal como mostrado na Figura 16, os aminoácidos numerados de 130 a 152 na tradução de U97107 contêm um elevado grau de semelhança com a MAELO traduzida. A sequência humana, número de acesso GenBank AC004050, do cromossoma 4 era de uma HTGS ("High Throughput Genome Sequence"). Não houve anotações contidas nesta sequência. No entanto, a AC004050 traduzida tinha 28,7% de identidade em 150 aminoácidos com a MAELO traduzida. Este fragmento génico podia ser um fragmento de uma alongase de AGPI humana com base na sua semelhança de aminoácidos com a MAELO traduzida.

A Figura 17 mostra o alinhamento de aminoácidos da MAELO traduzida com uma sequência de mamífero (número de acesso GenBank 105465, PCT n.º WO 88/07577) que reivindica que a proteína derivada da expressão desta sequência é um factor de inibição da glicosilação. Significando as identidades de aminoácidos entre as duas proteínas que podia haver uma função aparentada, tal como actividade de alongase de AGPI.

Estes exemplos de outras sequências de ADN traduzidas e a sua homologia com a MAELO traduzida ilustram que qualquer um dos exemplos de cima podia potencialmente ser uma alongase de AGPI. Estes exemplos não são inclusivos de todas as possíveis alongases. No entanto, a utilização de MAELO ou da sua tradução em aminoácidos como busca para pesquisas de bases de dados pode identificar outros genes que tenham actividades de alongase de AGPI.

Exemplo VI

Pesquisa da Biblioteca de ADNc de *M. alpina* Utilizando um Método de Hibridação de Placas

Num esforço para isolar genes de alongases de AGPI adicionais de *M. alpina*, foi utilizado um método de hibridação de placas convencional para pesquisar uma biblioteca de ADNc de *M. alpina* feita num vector lambda. A sonda de ADN foi gerada com base na sequência nucleotídica de MAELO e foi utilizada para pesquisar a biblioteca de ADNc de *M. alpina* M7+8 feita num vector λZiplox (Knutzon et al., *J. Biol. Chem.* 273: 29360-29366, 1998).

Para fazer a sonda de ADN para pesquisar a biblioteca, o ADNc de MAELO foi digerido com as endonucleases de restrição *NspI* e *PvuI*. Três pequenos fragmentos de ADN, com um tamanho médio de aproximadamente 300 pb, foram produzidos e utilizados como sondas. O racional para a utilização de uma mistura de ADNc de MAELO fragmentado baseou-se na assunção de que poderia haver uma região ou domínio comum na sequência de aminoácidos que fosse conservada entre as várias alongases de AGPI presentes em *M. alpina*. Utilizando sondas de ADN de MAELO, a biblioteca de ADNc foi pesquisada através de uma técnica de hibridação de placas de acordo com um protocolo padrão (Sambrook et al., "Molecular Cloning", 2ª Ed., Cold Spring Harbor, 1989).

Resumidamente, 50000 clones primários foram plaqueados e transferidos para membranas de *nylon*. As membranas foram desnaturadas e hibridadas com sondas de ADN de MAELO marcadas com alfa-³²P-dCTP de um dia para o outro no tampão de hibridação que continha formamida a 20%, PVP a 0,2%, BSA, Ficoll, SDS a 0,1% e NaCl 0,5 M. Os filtros foram lavados com SSC 0,5× a 37°C e expostos a película de raios X para autorradiografia. Este procedimento foi repetido três vezes. Quatro clones (designados como F1, F2, F3, e F4) que hibridaram repetidamente foram apanhados e suspensos em tampão SM (Sambrook et al., *supra*) contendo DMSO a 7%.

O maior enquadramento de leitura aberto de cada candidato foi subclonado no vector de expressão de levedura pYX242 (Novagen, Inc., Madison, Wisconsin). Os clones de ADNc F1 e F3

foram subclonados em pYX242 no local *EcoRI* enquanto F2 e F4 foram subclonados nos locais *NcoI/HindIII*. O pYX242 recombinante contendo cada candidato foi transformado em SC334 (Hoveland *et al.*, *supra*) para expressão em levedura. Para determinar a actividade de alongase, bem como a especificidade do substrato, SC334 contendo um clone de ADNc cada foi criada em meio mínimo sem leucina na presença de 25 μ M de substrato GLA tal como descrito no Exemplo III. A análise dos ácidos gordos foi efectuada tal como descrito em Knutzon *et al.* (*J. Biol. Chem.* 273:29360-29366, 1998). Os resultados indicam que nenhum destes quatro clones de ADNc mostrou qualquer actividade significativa na conversão de GLA em DGLA. Assim, a abordagem de hibridação pareceu não ter sido bem sucedida na identificação de mais alongases de AGPI.

Exemplo VII

Construção de uma Biblioteca de Expressão de ADNc Directa de *M. alpina* em Levedura

Para identificar genes de alongases de AGPI diferentes de MAELO, foi tomada uma abordagem diferente para pesquisar a biblioteca de ADNc de *M. alpina*. Em particular, uma vez que o fermento de padeiro é incapaz de produzir AGPI de cadeia longa devido à ausência das respectivas dessaturases e alongases, foi feita uma tentativa para construir uma biblioteca de ADNc de expressão de *M. alpina* em *Saccharomyces cerevisiae*. O vector pYES2 (Novagen, Inc., Madison, Wisconsin), contendo o promotor GALL, foi escolhido para a expressão da biblioteca de ADNc em *S. cerevisiae*.

O modo convencional através do qual uma biblioteca de ADNc é feita (i.e., transformação de uma mistura de ADN ligado ADNc/vector em células hospedeiras) é difícil em levedura porque a eficiência de transformação através de electroporação directa da mistura de ADN ligado é muito baixa em comparação com a eficiência de ADN plasmídico super-enrolado purificado. No entanto, a principal vantagem deste método é evitar a amplificação de clones primários que acontece quando a biblioteca é feita em *E. coli* como intermediária. Devido à limitação no número de colónias a pesquisar, foi decidido otimizar primeiro a eficiência de transformação em diferentes estirpes de *S. cerevisiae* utilizando uma mistura ligada de

ADNc/vector. Os melhores resultados foram obtidos com um rendimento de $4-5 \times 10^5$ transformantes por μg de ADN ligado na estirpe de *S. cerevisiae* SC334 (Hoveland et al., supra).

Para fazer uma biblioteca de expressão de ADNc de *M. alpina* em levedura o ARN total foi isolado do fungo. O fungo *M. alpina* (número ATCC 32221) foi plaqueado em ágar de farinha de milho (Difco Laboratories, Detroit, MI) e criado à temperatura ambiente durante 3-4 dias. Uma vez visível o crescimento do fungo, este foi inoculado em 50 ml de caldo de dextrose de batata e agitado à temperatura ambiente muito lentamente para formular esporos. Uma vez visíveis os esporos, a cultura de 50 ml foi inoculada numa cultura de 1 litro de dextrose de batata, e os esporos foram criados durante 72 horas. Após filtração através de gaze estéril, as células foram imediatamente congeladas em azoto líquido para futura extracção de ARN. O ARN total foi preparado a partir de 36 g de sedimento celular utilizando o método de extracção de fenol quente/LiCl (Sambrook et al., supra). Os sedimentos celulares foram homogeneizados numa solução de EDTA 10 mM, SDS 1% e acetato de sódio 200 mM, pH 4,8. Foram adicionados fenol e clorofórmio aos homogenatos e a camada aquosa foi extraída. A camada aquosa foi extraída de novo uma ou mais vezes com fenol e clorofórmio. Depois, foi adicionado um volume igual de cloreto de lítio 4 M. As amostras foram precipitadas em etanol em gelo durante 3 horas e os sedimentos foram obtidos através de centrifugação. Os sedimentos foram lavados com etanol a 70% e ressuspenso em água tratada com DEPC. O ARN total foi quantificado através de espectrofotometria e visualizado através de electroforese em gel de agarose para confirmar a presença das bandas ribossómicas 28S e 18S. Foram obtidos aproximadamente 15 mg de ARN total a partir de 36 gramas de sedimento celular.

A biblioteca foi construída de acordo com o protocolo padrão (Sambrook et al., "Molecular Cloning", 2ª Ed. Cold Spring Harbor, 1989). O ARN mensageiro foi preparado a partir do ARN total utilizando purificação por afinidade em celulose oligo dT. O ARN mensageiro foi transcrito de forma reversa com iniciador oligo dT contendo um local de restrição *Xho*I utilizando a transcriptase reversa AMV. Após a síntese da primeira cadeia de ADNc, a segunda cadeia de ADNc foi

sintetizada através da adição da ADN-polimerase de *E. coli*, ADN-ligase de *E. coli* e ARNase H.

O adaptador *EcoRI* foi ligado no ADNc terminado de forma cega através da ADN-ligase T4. A amostra de ADNc foi fosforilada utilizando a polinucleótido-quinase de T4 e digerida com *XhoI*, diluída com tampão de coluna e passada através de uma coluna Sephacryl S-400. As amostras de ADN foram eluídas através de tampão de elevada salinidade. As amostras contendo ADN de 400-5000 pb foram reunidas e utilizadas para ligação num vector pYES2 (Invitrogen Corp., Carlsbad, CA). O ADNc foi ligado no vector pYES2 digerido com *EcoRI/XhoI* utilizando ADN-ligase de T4. Foi realizada uma reacção de ligação em grande escala uma vez que é necessária uma grande quantidade de ADN ligado (2-3 µg) na transformação directa de levedura.

Para transformar células de levedura directamente com a mistura ligada de ADNc/pYES2, foram preparadas células competentes SC334 utilizando o método TRAF0 de LiAc (Gietz et al., *Mol. Cell. Biol.* 5: 255-269, 1995). Resumidamente, cultura fresca de SC334 da placa foi inoculada em 50 ml de meio YPD. A cultura foi criada a 30°C com agitação até a DO a 600 ter alcançado 1,0. Trinta ml deste material inicial foram inoculados em 300 ml de meio líquido YPD e incubados com agitação até o número de células da cultura alcançar $\approx 3-5 \times 10^6$ células/ml (aproximadamente 3-4 h). As células foram colhidas e lavadas com água estéril. O sedimento celular inteiro foi ressuspenso em 1,5 ml de TE/LiAc 1× preparado de fresco (LiAc 0,1 M). Estas células foram utilizadas imediatamente para as transformações.

Setecentos e cinquenta microlitros de células SC334 competentes foram divididos em alíquotas em tubos Falcon de 15 ml. Aproximadamente 2 µg de ADN ligado ADNc/pYES2 foram adicionados às células juntamente com ADN transportador e misturados suavemente. Três mililitros de PEG a 40%/LiAc estéril foram adicionados às células e misturados suavemente mas profundamente. As células foram incubadas a 30°C durante 30 min. com agitação e subsequentemente foi-lhes dado um choque térmico a 42°C durante 15 min. As células foram arrefecidas, sedimentadas e ressuspensas em 5 ml de TE 1×. Uma

alíquota de 100 µl das células de cima foi plaqueada sobre cinquenta placas de ágar selectivas de 150 mm sem uracilo (Ausubel et al., supra) e incubadas a 30°C durante 3 dias. Foi obtido um total de 8×10^5 clones primários. Foram reunidas cinco colónias em 1 ml de meio mínimo sem uracilo (Ausubel et al., supra) e foi adicionado glicerol para preparar reservas. Foi feito um total de 5000 bancos para pesquisa.

Exemplo VIII

Pesquisa MAD (*M. alpina* Directa) em Levedura

A qualidade da biblioteca foi analisada através da determinação do tamanho médio dos ADNc na biblioteca. Uma vez que a pesquisa da biblioteca foi baseada na expressão do ADNc, era importante determinar o tamanho médio do ADNc presente na biblioteca. A biblioteca de expressão contendo os ADNc mais longos seria a escolha mais apropriada para isolar ADNc inteiros de interesse. Para este fim, foram plaqueados bancos seleccionados de modo aleatório sobre placas de ágar selectivas, tal como descrito no Exemplo VII, para obter colónias individuais. Foram apanhadas aleatoriamente quarenta colónias diferentes e cada colónia foi inoculada em 5 ml de meio líquido selectivo sem uracilo (tal como descrito no Exemplo VII) e criada ao mesmo tempo que se agitava, durante 24 horas a 30°C. O ADN plasmídico foi extraído destas colónias através do método de agitação de contas (Hoffman et al., Gene 57: 267, 1987) adaptado como segue:

Os sedimentos de 5 ml de cultura foram lisados em 0,5 ml de uma solução de NaCl 100 mM, Tris 10 mM, pH 8,0, EDTA 1 mM e SDS a 0,1%. Contas de vidro de 0,5 mm estéreis em volume igual foram adicionadas e manualmente agitadas em vórtice durante 3 minutos. Foram adicionados duzentos microlitros do mesmo tampão e a mistura foi sujeita a vórtice durante mais um minuto. As amostras foram centrifugadas a alta velocidade durante 2 minutos e o extracto citoplasmático foi então transferido para um tubo novo. Foi adicionado à amostra um volume igual de fenol/ CHCl_3 , sujeitou-se a vórtice e centrifugou-se novamente durante 2 minutos. A camada aquosa foi novamente extraída duas vezes e precipitada com acetato de sódio 0,3 M e aproximadamente 2,5 volumes de etanol durante 30 minutos a -20°C. Os precipitados foram lavados com etanol a

70% e ressuspensos em água. Para eliminar ARN e qualquer contaminação proteica, os ADN plasmídicos isolados de 40 amostras diferentes foram ainda purificados utilizando o QIAprep Spin Miniprep Kit de acordo com o protocolo do fabricante (Qiagen Inc., Valencia, CA). As amostras de ADN plasmídico foram então restringidas com as endonucleases de restrição *EcoRI* e *XhoI* para libertar o fragmento de ADNc, e o digerido foi analisado em gel de agarose a 1%. Os resultados indicam que a maioria dos ADNc da biblioteca directa variaram em comprimento de 0,8 kb a 1,5 kb.

Para pesquisar a biblioteca, as reservas glicérol foram descongeladas e foram adicionados aproximadamente 0,5 ml de meio líquido selectivo sem uracilo (Ausubel *et al.*, *supra*) e criadas a 30°C durante 24 horas. A cultura foi então transferida para 50 ml de meio líquido selectivo sem uracilo com galactose a 2% e GLA 25 μ M (substrato para a enzima alongase) durante 24 horas a 25°C com agitação. A análise de GC-FAME do teor lipídico no sedimento celular de cada cultura induzida foi efectuada tal como anteriormente descrito (Knutzon *et al.*, *supra*). A MAELO (pRAE-5 em pYX242 criada em meio selectivo sem leucina) foi utilizada como controlo positivo em cada corrida. MAELO foi consistentemente capaz de converter 1,5% de GLA em DGLA (ver Exemplo III).

Exemplo IX

Identificação de um ADNc Codificando uma Potencial Alongase de AGPI

Após a pesquisa e análise de aproximadamente 750 bancos individuais através de análise GC-FAME, tal como descrito no Exemplo VIII, um banco de cinco colónias (i.e., MAD708) pareceu ter uma actividade enzimática significativa na conversão de GLA em DGLA. Verificou-se que esta actividade era aproximadamente 5 vezes maior que a actividade da alongase de *M. alpina* (MAELO) em termos de proporção DGLA/GLA (Figura 19). Este banco foi testado novamente sob condições de ensaio idênticas para confirmar as descobertas iniciais. A experiência repetida mostrou uma conversão de 9,5% de GLA em DGLA e foi novamente cerca de 5 vezes maior que a actividade de alongase de *M. alpina* (MAELO). Estes resultados indicam fortemente que o banco de MAD708 continha um candidato a

alongase que era específico para GLA como substrato. Uma vez que MAD708 era um banco de cinco clones diferentes, foi necessário isolar o clone de ADNc individual que codificava a actividade de alongase deste banco. Para o fazer, a reserva de glicerol de MAD708 original foi plaqueada sobre uma placa de ágar de meio selectivo sem uracilo (Ausubel *et al.*, *supra*). Trinta clones individuais foram apanhados e criados em meio líquido selectivo, sem uracilo com galactose a 2%, tal como anteriormente descrito no Exemplo VIII, na presença de GLA. O sedimento celular obtido a partir de cada cultura foi então sujeito a análise GC-FAME dos ácidos gordos (Knutzon *et al.*, *supra*) juntamente com um controlo positivo de 334(pRAE-5) (MAELO em pYX242). A análise dos ácidos gordos dos 30 clones individuais a partir do banco de expressão de MAD708 em levedura revelou que 5 dos 30 clones apresentavam actividade de alongase na conversão de GLA em DGLA. Os perfis de ácidos gordos dos clones activos MAD708-2, MAD708-10, MAD708-18, MAD708-19 e MAD708-30 são mostrados na Figura 20. Tal como mostrado nesta Figura, MAD708-2, 10, e 30 produziram a maioria do DGLA, aproximadamente 25 vezes mais que MAELO (pRAE-5). Estes 3 converteram no intervalo de 41% a 49% GLA em DGLA. Outros clones, MAD708-18 e MAD708-19, converteram 8% e 21% do GLA em DGLA, respectivamente. Todos os clones MAD708 converteram uma maior percentagem de GLA em DGLA em relação à alongase codificada por MAELO (3,4%).

Exemplo X

Caracterização dos ADNc Codificando a Alongase

O ADN plasmídico foi extraído dos clones de levedura SC334 (banco MAD708) que apresentaram uma actividade de alongase específica de GLA significativa através do método de agitação de contas, tal como descrito no Exemplo VIII. Para determinar o tamanho da inserção de ADNc, foi efectuada PCR utilizando cada ADN plasmídico obtido a partir dos clones positivos para alongase como molde. O iniciador directo R0541 (5'-GAC TAC TAG CAG CTG TAA TAC-3') e o iniciador reverso R0540 (5'-GTG AAT GTA AGC GTG ACA TAA-3') estão no local de clonagem múltipla do vector pYES2 vector e foram utilizados para amplificar a inserção de ADNc dentro dos locais *EcoRI* e *XhoI*. A reacção de PCR foi efectuada num volume de 50 µl contendo 4 µl de ADN plasmídico, 50 pmol de cada iniciador,

5 µl de tampão 10×, 1 µl de Mistura de Nucleótidos de PCR 10 µM (Boehringer Mannheim Corp., Indianapolis, IN) e 0,5 µl de polimerase Taq High Five (Boehringer Mannheim, Indianapolis, IN). A amplificação foi realizada como se segue: 2 min. de desnaturação a 94°C, depois 94°C durante 1 min, 55°C durante 2 min., e 72°C durante 3 min. durante 30 ciclos e 7 min. de extensão a 72°C no final da amplificação. A análise dos produtos amplificados por PCR num gel de agarose a 1% mostrou que os tamanhos dos ADNc de alongase eram de cerca de 1,0-1,2 kb. Os ADN plasmídicos, contendo os potenciais ADNc de alongase, foram designados como pRPB2, pRPB10, pRPB18, pRPB19 e pRPB30. Uma vez que a biblioteca de ADNc foi feita no vector pYES2 nos locais *EcoRI* e *XhoI*, o tamanho do ADNc presente em cada plasmídeo foi ainda confirmado através de digestão dos plasmídeos de cima com *EcoRI* e *XhoI*.

Os ADN plasmídicos isolados de levedura foram novamente amplificados em *E. coli* para armazenamento a longo prazo dos clones de ADNc bem como para sequenciação do ADN. As células de *E. coli* TOP10 (Invitrogen Corp., Carlsbad, CA) foram transformadas com os plasmídeos recombinantes pRPB de acordo com o protocolo do fabricante. Os transformantes obtidos a partir de cada ADN plasmídico foram inoculados em LB contendo ampicilina (50 µg/ml) e criados de um dia para o outro a 37°C com agitação. Os ADN plasmídicos foram isolados a partir destas culturas através da utilização de QIAprep Spin Miniprep (Qiagen Inc., Valencia, CA) de acordo com o protocolo do fabricante. Os ADN plasmídicos purificados foram então utilizados para sequenciação a partir de ambas as extremidades 5' e 3'. A sequenciação de ADN foi efectuada através da utilização de um sequenciador de ADN automático 373A Stretch ABI (Perkin Elmer, Foster City, CA) de acordo com o protocolo do fabricante. Os iniciadores utilizados para sequenciação foram o iniciador directo R0541 (5'-GAT TAC TAG CAG CTG TAA TAC-3') e o iniciador reverso R0540 (5'-GTG AAT GTA AGC GTG ACA TAA-3') contidos nos locais de clonagem múltipla o vector pYES2. As sequências nucleotídicas obtidas foram transferidas para o programa de suporte lógico Sequencher (Gene Codes Corporation, Ann Arbor, MI) para análise. A análise da sequência de ADN revelou que todos os cinco ADNc de alongase continham a sequência nucleotídica idêntica a uma sobreposição comum de 301 nucleótidos. Cada

sequência de ADN contém um putativo local de início no começo da extremidade 5' e um codão de terminação com uma cauda poli-A no final do local 3'. Para confirmar melhor a sequência de ADN, os iniciadores directos internos R0728 (5'-GAG ACT TTG AGC GGT TCG-3') e R0730 (5'-TCT CTG CTG CGT TGA ACT CG-3'), juntamente com os iniciadores reversos R0729 (5'-AAA GCT CTT GAC CTC GAA C-3') e R0731 (5'-AAC TTG ATG AAC GAC ACG TG-3') foram desenhados dentro do ADNc e utilizados para sequenciação de pRPB2, uma vez que este candidato possuía a actividade de alongase mais elevada. A sequência nucleotídica inteira foi analisada através do programa Sequencher (Figura 21) e o maior enquadramento de leitura aberto deduzido da sequência de ADNc inteira de pRPB2 pareceu ter 957 pb de comprimento (Figura 22). O enquadramento de leitura aberto deduzido foi então traduzido na sequência de aminoácidos correspondente e a sequência prevista é mostrada na Figura 23. A alongase codificada pelo ADNc (pRPB2) identificado a partir de *M. alpina* parece ser uma proteína de 318 aminoácido de comprimento que é quase idêntica em tamanho à MAELO traduzida. Este novo ADNc de alongase foi designado como "GLELO" e a sua proteína codificada foi designada "alongase de GLA".

O ADN plasmídico pRPB2 foi depositado na American Type Culture Collection, 10801 University Boulevard, Manassas, Virginia 20110-2209 a 22 de Julho de 1999 sob os termos do Tratado de Budapeste. Foi-lhe atribuído o depósito ATCC n.º PTA-402.

Exemplo XI

Caracterização Bioquímica da Alongase de GLA (GLELO)

A. Confirmação da Actividade de Alongase de GLA

Para confirmar melhor a actividade da alongase de GLA codificada pelo plasmídeo recombinante pRPB2, a pesquisa da actividade de alongase foi repetida no clone de levedura SC334 contendo o plasmídeo pRPB2. Esta experiência foi também conduzida para assegurar uma extracção de lípidos consistente e para detectar a actividade de alongase de GLA através do cálculo da média de quatro experiências independentes. A reserva em glicerol de *S. cerevisiae* 334 contendo pRPB2 foi plaqueada em placas de ágar de meio mínimo sem uracilo. Colónias individuais foram apanhadas aleatoriamente e criadas

em meio mínimo sem uracilo, tal como descrito no Exemplo VIII. As quatro culturas independentes foram combinadas e uma alíquota de 5 ml foi utilizada como inóculo para quatro culturas de 50 ml separadas. As culturas foram então criadas na presença de GLA e foram sujeitas a análise de ácidos gordos juntamente com um controlo negativo de *S. cerevisiae* 334 contendo pYES2, tal como descrito no Exemplo VIII. A actividade de alongase média de quatro culturas independentes de 334(pRPB2) com 25 μ M de GLA é mostrada na Figura 24. A actividade de alongase de GLA de cada uma das quatro amostras independentes de 334(pRPB2) pareceu ser consistente com uma conversão média de 62% de GLA em DGLA.

B. Determinação da Especificidade do Substrato de GLELO para a Alongase de GLA

Para analisar a especificidade de substrato da alongase de GLA, a cultura de 334(pRPB2) foi testada com diferentes substratos de ácidos gordos para além de GLA (p.ex., SA(18:0), OA(18:1), LA(18:2n-6), AA(20:4n-6), ADA(22:4n-6), ALA(18:3n-3), STA(18:4n-3) e EPA(20:5n-3). Sob condições de ensaio idênticas, o único outro substrato utilizado pela enzima alongase foi STA, um ácido gordo da via n-3. A alongase de GLA foi capaz de converter 73% de STA em 20:4n-3 (Figura 25). Destas experiências, pode-se concluir que a alongase de GLA tem especificidade de substrato tanto para GLA como para STA, indicando que possui actividade de alongase tanto na via n-6 como na n-3.

C. Co-expressão dos Genes de GLELO e da Δ 5-Dessaturase Fúngicos em Levedura

Uma vez produzido DGLA (20:3n-6) através da alongase de GLA, a Δ 5-dessaturase pode convertê-lo em AA (20:4n-6) num sistema de co-expressão desejado. Este esquema, tal como representado na Figura 1, pode ser testado através da co-transformação de *S. cerevisiae* 334 com os plasmídeos pRPB2 e pRPE31 (o plasmídeo recombinante pYX242 contendo um ADNc de Δ 5-dessaturase (Figura 18) clonado no local *Eco*RI. As culturas de levedura co-transformadas foram suplementadas com GLA 25 μ M e analisadas quanto à síntese de AA. Se ambas as enzimas alongase e Δ 5-dessaturase forem expressas, o substrato GLA será convertido em DGLA, que será então convertido em AA. Os

resultados da Figura 26A indicam que a acção sequencial da alongase de GLA e da $\Delta 5$ -dessaturase sobre o substrato GLA resultou numa conversão média de 27% de GLA em AA. Por essa razão, a alongase de GLA tem a capacidade de funcionar com outras enzimas na via sintética de AGPI n-6 para produzir ácidos gordos desejáveis.

Para determinar se a conversão de cima é também verdadeira nas vias n-3, foram realizadas as experiências de co-expressão semelhantes na presença de STA 25 μ M. Novamente, se ambas as enzimas forem expressas, o substrato STA será convertido em 20:4n-3 que será então convertido em EPA (20:5n-3) através da $\Delta 5$ -dessaturase. A Figura 26B mostra os resultados nos quais é observada a produção de EPA (aprox. 40%). Uma vez mais, a alongase de GLA demonstra a sua capacidade para funcionar com a $\Delta 5$ -dessaturase na via n-3 para produzir ácidos gordos desejáveis.

Exemplo XII

Comparação de Sequências entre GLELO e Outras Alongases Fúngicas

O pacote de análise de sequências de GCG (ver Exemplo I) foi utilizado para comparar a sequência GLELO com sequências proteicas conhecidas. A sequência nucleotídica do enquadramento de leitura aberto de GLELO foi primeiro traduzida na sequência de aminoácidos que foi utilizada como sequência de busca para pesquisar a base de dados Swissprot (ver Exemplo I) utilizando o algoritmo FastA (ver Exemplo I). Com base na semelhança da sequência de aminoácidos, as melhores coincidências foram verificadas com YJT6 de *S. cerevisiae* (uma EST com anotação desconhecida) com 33,9% de identidade numa sobreposição de 189 aminoácidos, EL02 de *S. cerevisiae* (GNS1) com 25,8% de identidade numa sobreposição de 295 aminoácidos e EL03 de *S. cerevisiae* (SUR4) com 25,2% de identidade numa sobreposição de 313 aminoácidos. O alinhamento FastA de GLELO com MAELO mostrou 30,9% de identidade em 275 aminoácidos (Figura 27). O programa Pileup de GCG cria um alinhamento de múltiplas sequências a partir de um grupo de sequências aparentadas utilizando alinhamentos aos pares progressivos (ver Exemplo I) e foi utilizado com as alongases descritas acima. Os resultados de Pileup indicam que existem

muitas regiões conservadas entre as alongases incluindo uma caixa de histidinas putativa que está sublinhada (Knutzon *et al.*, *J. Biol. Chem.* 273: 29360-29366, 1998) (Figura 28). Assim, embora GLELO tenha semelhança com MAELO, a diferença nas suas alongases codificadas pode ser presumivelmente devida à sua preferência de substrato. A alongase de GLA pode converter uma maior percentagem de GLA em DGLA que a alongase de *M. alpina*. Adicionalmente, a expressão de MAELO em *S. cerevisiae* mostrou alongamento de ácidos gordos saturados e mono-insaturados para além do alongamento de GLA até DGLA (ver Exemplo III).

Exemplo XIII

Identificação de Homólogos de MAELO de *M. alpina* em Mamíferos

A sequência traduzida de MAELO foi utilizada para pesquisar a Unified Human Transcript Database de Abbott Laboratories, 100 Abbott Park Rd., Abbott Park, Illinois 60064. Esta base de dados foi pesquisada utilizando Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, *Nuc. Acids Res.* 25: 3389-3402, 1997) que "é um conjunto de programas de pesquisa de semelhanças desenhado para explorar todas as bases de dados de sequências disponíveis independentemente da busca ser uma proteína ou ADN". Especificamente foi utilizado o algoritmo tblastn (i.e., uma pesquisa de busca de proteína numa base de dados de nucleótidos traduzidos em seis enquadramentos de leitura). As sequências contíguas (CC) na Unified Human Transcript Database são sequências de consenso representando grupos de ADNc de sequências etiquetas expressas (EST) derivados do domínio público e da base de dados de EST Incyte LIFESEQ[™] (Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304) que estão agrupados com base numa homologia de sequência definida, e montados com base na sobreposição das sequências. Duas sequências desta base de dados, CC067284R1 e CC1484548T1 tinham 28% de identidade numa sobreposição de 242 aminoácidos e 28,6% de identidade numa sobreposição de 266 aminoácidos, respectivamente, com a sequência MAELO traduzida. As duas sequências derivadas e editadas foram designadas como hs1 e hs2, respectivamente, e copiadas para o pacote de suporte lógico de GCG de análise de sequências (ver Exemplo I). A sequência MAELO traduzida foi alinhada com as sequências de

ADNc traduzidas de HS1 (28,5% de identidade em 242 aminoácidos) e HS2 (28,2% de identidade em 266 aminoácidos) utilizando o algoritmo utilizando o algoritmo FastA, tal como mostrado nas Figuras 29 e 30, respectivamente. A sequência nucleotídica do ADNc de HS1 tinha 86,9% de identidade em 844 pb com a sequência nucleotídica I05465 (ver Exemplo V). A sequência de ADNc de HS2 tinha 100% de identidade com a sequência de aminoácidos com o número de acesso GenBank W74824 (ver pedido PCT publicado WO9839448).

O National Center for Biotechnology Information (NCBI em <http://www.ncbi.nlm.nih.gov/>) foi utilizado para conduzir pesquisas de bases de dados utilizando tblastn com a sequência de 28 aminoácidos (DTIFIILRKQKLIFLHWYHHITVLLYSW) traduzida a partir de AC004050 (uma sequência humana identificada numa pesquisa TFastA, ver Exemplo V). Esta sequência de aminoácidos contém uma caixa de histidinas (sublinhada), que tem um motivo anotado de dessaturases (Knutzon *et al.*, *supra*), e ambas as alongases de AGPI, MAELO e GLELO (ver Figura 28). Uma sequência de ratinho traduzida mostrada anteriormente no Exemplo V (número de acesso GenBank U97107) e uma sequência de *C. elegans* traduzida (número de acesso GenBank # U41011) tiveram as maiores coincidências com esta busca de 28 aminoácidos. A base de dados de EST de ratinho de NCBI foi pesquisada novamente com tblastn, utilizando U41011 traduzida como busca. Foi identificada uma sequência de ratinho adicional (número de acesso GenBank AF014033.1), anotada como "putativo envolvimento no alongamento de ácidos gordos". Três sequências maiores (números de acesso GenBank AA591034, AA189549 e AA839346) foram identificadas através de uma pesquisa tblastn da base de dados de EST de ratinho com AF014033.1 traduzida e combinadas numa sequência designada como mm2. O alinhamento FastA (ver Exemplo I) de mm2 e MAELO traduzidas é mostrado na Figura 31. Outra sequência de ratinho aparentada mas não idêntica (número de acesso GenBank AI225632), foi também identificada numa pesquisa de tblastn da base de dados de EST de ratinho com AF014033.1. O alinhamento FastA da AI225632 com MAELO traduzidas é mostrado na Figura 32. A percentagem de identidade para ambas MM2 e AI225632 traduzidas com a MAELO traduzida é de 30,4% numa sobreposição de 191 e 115 aminoácidos, respectivamente. O nível de identidade de aminoácidos com a MAELO traduzida com

estas duas sequências de ratinho traduzidas identifica-as como putativos homólogos das alongases de AGPI.

Exemplo XIV

Identificação de Homólogos de GLELO de *M. alpina* em Mamíferos

O algoritmo TFastA, que compara uma sequência proteica com a sequência de ADN da base de dados traduzida em cada um dos seis enquadramentos de leitura, foi utilizado com GLELO como busca. A base de dados GenEMBL de GCG foi utilizada para identificar outras potenciais sequências de alongases com base na sua semelhança de aminoácidos com a GLELO traduzida. Verificou-se que três sequências humanas tinham coincidências com a sequência de aminoácidos de GLELO. Estas sequências tinham os números de acesso GenBank 1) AI815960, 2) AL034374, e 3) AC004050. AI815960, uma sequência EST de *Homo sapiens*, tem 40,3% de identidade numa sobreposição de 144 aminoácidos com a GLELO traduzida (ver Figura 33). Uma região traduzida da sequência genómica humana AL034374, derivada do cromossoma VI tem 46,7% de identidade numa sobreposição de 60 aminoácidos com a GLELO traduzida. Esta região homóloga em AL034374 pareceu ser parte da sequência de aminoácidos de HS1 que se mostrou ter homologia com a MAELO traduzida (ver Exemplo XIII). Por essa razão, a sequência de HS1 tem semelhança tanto com MAELO (ver Figura 29) como com GLELO (ver Figura 34). Uma região traduzida de uma sequência genómica humana AC004050 do cromossoma IV tem 34,8% de identidade numa sobreposição de 89 aminoácidos com a GLELO traduzida (ver Figura 35). As identidades de aminoácidos entre GLELO e estas sequências humanas indicam que as proteínas derivadas destas sequências humanas podiam ter uma função aparentada, tal como actividade de alongase de AGPI.

Para identificar um ADNc de ratinho semelhante a GLELO, foram efectuadas pesquisas TFastA com a base de dados GenEMBL utilizando GLELO traduzida como busca. Das pesquisas TFastA, foram identificadas as três sequências de ratinho com as maiores coincidências com a GLELO traduzida: (números de acesso GenBank 1) AF104033, 2) AI595258, e 3) U97107). AF104033 está anotada como "proteína MUEL possuindo alongase putativa de ácidos gordos com homologia com a ELO3 de levedura (SUR4)" e é parte da sequência de MM2. A sequência de MM2 foi

inicialmente derivada da sequência de ratinho AF104033, mas sequência de MM2 inteira foi finalmente obtida através de mais pesquisas da base de dados de EST de ratinho e mostrou-se também que tinha homologia com a MAELO traduzida (ver Exemplo XIII e Figura 31). Quando esta sequência de aminoácidos de MM2 foi alinhada com a sequência GLELO traduzida utilizando FastA, verificou-se uma identidade de 34,6% numa sobreposição de 211 aminoácidos (ver Figura 36) indicando que a MM2 tinha também homologia com GLELO. AI595258 é um clone de ADNc de ratinho possuindo semelhança a 5' com a alongase ELO3 de levedura e é parte do ADNc da EST de ratinho AI225632. Mostrou-se que a sequência de ratinho AI225632, que é uma sequência maior que AI595258, tem semelhança com a MAELO traduzida (ver Figura 32). A AI225632 foi também alinhada com a GLELO traduzida e o alinhamento FastA é mostrado na Figura 37. Verificou-se uma identidade de 35,3% numa sobreposição de 199 aminoácidos. A terceira sequência, U97107, uma sequência de ratinho, foi anotada como "semelhante ao gene de ELO3 (SUR4) de levedura". O alinhamento FastA da GLELO traduzida com U97107 é mostrado na Figura 38 onde se verificou uma identidade de 23,7% numa sobreposição de 279 aminoácidos. Anteriormente, verificou-se que uma região de U97107 tinha também um elevado grau de homologia com MAELO com base num alinhamento FastA (ver Exemplo V e Figura 16).

As pesquisas de cima indicam claramente que foram obtidas as mesmas sequências humanas e de ratinho através da utilização de MAELO ou GLELO como busca.

Exemplo XV

Identificação de Homólogos de GLELO e MAELO de *M. alpina* Noutros Organismos Produtores de AGPI

A) *Caenorhabditis elegans*:

Uma sequência putativa de aminoácidos deduzida de uma sequência cromossômica de *C. elegans* (número de acesso GenBank U41011) foi capaz de identificar uma sequência parcial contida na alongase putativa de AGPI MM2 de ratinho que tem semelhança de aminoácidos tanto com a alongase de GLA (GLELO) como com a alongase de *M. alpina* (MAELO). Era por isso concebível que homólogos de *C. elegans* de GLELO ou MAELO pudessem estar

presentes na base de dados de nematodes. As sequências de aminoácidos putativas derivadas das sequências GLELO e MAELO foram utilizadas como buscas independentemente para pesquisar as bases de dados do nematode. Uma pesquisa BLAST (ver Exemplo XIII) foi efectuada nas bases de dados wormpep16 (blastp compara uma sequência de aminoácidos de busca contra uma base de dados de sequências nucleotídicas) e wormpep16cDNA (tblastn) que são proteínas previstas e ADNc obtidos a partir do projecto de sequenciação do genoma de *C. elegans* ou EST e as suas correspondentes sequências de ADNc, respectivamente. Estes dados das sequências foram produzidos pelo grupo de sequenciação de *C. elegans*, realizada em conjunto com o Sanger Centre and Genome Sequencing Center, e podem ser obtidos em <ftp://ftp.sanger.ac.uk/pub/databases/wormpep/>. Foram identificadas pelo menos sete sequências putativas traduzidas de *C. elegans* através da sua homologia da sequência de aminoácidos com a sequência de aminoácidos traduzida tanto de GLELO como de MAELO. Os números de acesso GenBank das sequências genómicas contendo os aminoácidos deduzidos foram identificados como Z19154, U68749 (2 proteínas deduzidas (F56H11.4 e F56H11.3 (números de acesso wormpep)), U41011, U61954 (2 proteínas deduzidas (F41H10.7 e F41H10.8 (números de acesso wormpep)), e Z81058. Os sublinhados foram identificados numa pesquisa anterior utilizando a MAELO traduzida como busca (ver Exemplo V). Como exemplo, os alinhamentos de aminoácidos FastA da U68749 (F56H11.4) traduzida com a GLELO e MAELO traduzidas são mostrados nas Figuras 39 e 40. A U68749 (F56H11.4) traduzida tem 25-30% de identidade com a alongase de *M. alpina* e a alongase de GLA numa sobreposição de aproximadamente 200 aminoácidos (ver Figuras 39 e 40). Para todos os sete putativos ADNc de *C. elegans* traduzidos, os alinhamentos FastA com a GLELO traduzida estiveram entre 25-30% de identidade numa sobreposição de 200 aminoácidos, enquanto que a identidade foi 26-34% numa sobreposição de pelo menos 188 aminoácidos para a MAELO traduzida. As semelhanças de alinhamento indicam que a GLELO ou a MAELO traduzidas podem ser utilizadas para identificar potenciais genes de *C. elegans* com actividade de alongase.

B) *Drosophila melanogaster*:

O ADNc deduzido traduzido da sequência genómica U41011 (*C. elegans*) tinha a sua maior coincidência com uma EST de *Drosophila melanogaster*, número de acesso AI134173 numa pesquisa blastn (compara uma sequência nucleotídica de busca contra uma base de dados nucleotídica) da base de dados das "outras EST" através de NCBI (ver Exemplo XIII) e foi montado com um fragmento de ADN EST sobreponível, número de acesso AI517255. O fragmento de ADN traduzido DM1, derivado das duas sequências sobreponíveis foi alinhado com a GLELO traduzida bem como com a MAELO (ver Figuras 41 e 42) utilizando FastA em GCG (ver Exemplo I). Os alinhamentos mostraram 27,2% de identidade com a alongase de GLA numa sobreposição de 206 de aminoácidos e 30% de identidade com a alongase de *M. alpina* numa sobreposição de 237 aminoácidos. Assim, com base na semelhança de aminoácidos, a DM1 podia ser um potencial homólogo de GLELO ou MAELO possuindo actividade semelhante a alongase de AGPI. Para além disso, utilizando sequências de ADN de GLELO e MAELO como busca para as pesquisas nas bases de dados, podem ser identificados homólogos com actividade de alongase de AGPI a partir de *Drosophila*.

Exemplo XVI

Clonagem e Expressão de um Homólogo de Alongase de AGPI Humano

Muitas potenciais sequências de alongases de AGPI foram identificadas com base nas suas semelhanças de aminoácidos com a GLELO e/ou MAELO traduzidas. Para determinar as potenciais actividades de alongase destas sequências, o ADNc codificando a proteína inteira é então identificado, clonado e expresso, tal como demonstrado no presente exemplo.

Os iniciadores R0719 (5'-GGT TCT CCC ATG GAA CAT TTT GAT GCA TC-3') e R0720 (5'-GGT TTC AAA GCT TTG ACT TCA ATC CTT CCG-3') foram desenhados com base na sequência putativa de HS1, e utilizados para amplificar o ADNc Marathon-Ready de fígado humano (Clontech Laboratories, Inc., Palo Alto, California). A reacção em cadeia da polimerase (PCR) foi realizada num volume de 50 µl contendo: 5 µl de ADNc Marathon-Ready de fígado humano, 50 pmol de cada iniciador, 1 µl de Mistura de Nucleótidos de PCR 10 mM (Boehringer Mannheim Corp., Indianapolis, IN), 5 µl de tampão 10× e 1,0 U de

Mistura de Polimerase Advantage KlenTaq (Clontech Laboratories, Inc., Palo Alto, CA). As condições de ciclo térmico em Perkin Elmer 9600 (Norwalk, CT) foram como se segue: 94°C durante 2 min, depois 30 ciclos de 94°C durante 1 min., 58°C durante 2 min e 72°C durante 3 min. A PCR foi seguida de um ciclo de extensão adicional a 72°C durante 7 minutos.

O produto de PCR amplificado foi corrido num gel, um fragmento amplificado de aproximadamente 960 pb foi purificado em gel, os terminais do fragmento foram preenchidos com ADN-polimerase T4 (Boehringer Mannheim, Corp., Indianapolis, IN) e clonados no vector pCR-Blunt (Invitrogen Corp., Carlsbad, CA) seguindo o protocolo do fabricante. O novo plasmídeo foi designado pRAE-52 e o putativo ADNc de alongase de AGPI neste clone foi sequenciado utilizando um ABI 373A Stretch DNA Sequencer (Perkin Elmer, Foster City, CA). A sequência putativa do ADNc de alongase de AGPI no plasmídeo pRAE-52 é mostrada na Figura 43 e a sequência traduzida é mostrada na Figura 44.

O putativo ADNc de alongase de AGPI do plasmídeo pRAE-52 foi então digerido com *NcoI/HindIII*, purificado em gel e ligado em pYX242(*NcoI/HindIII*). O novo plasmídeo foi designado pRAE-58-A1. (O plasmídeo 58-A1 foi depositado na American Type Culture Collection, 10801 University Boulevard, Manassas, VA 20110-2209 a 19 de Agosto de 1999, sob os termos do Tratados de Budapeste e foi-lhe atribuído o número de depósito ____.)

A construção pRAE-58-A1 foi transformada em *S. cerevisiae* 334 (Hoveland *et al.*, *supra*) e pesquisada quanto à actividade de alongase. A estirpe de controlo negativo foi *S. cerevisiae* 334 contendo o vector pYX242. As culturas foram criadas durante 24 horas a 30°C, em meio selectivo (Ausubel *et al.*, *supra*), na presença de 25 µM de GLA ou AA. Neste estudo, DGLA ou ácido adrénico (ADA, 22:4n-6), respectivamente, foi o produto previsto da actividade de alongase humana. Quando foi utilizado GLA como substrato, as células de levedura contendo o ADNc da alongase humana continham níveis elevados de DGLA em comparação com as células de controlo, 2,75% vs. 0,09% de ácidos gordos totais, respectivamente (ver Figura 45). Quando foi utilizada AA como substrato, as células de levedura

contendo o ADNc da alongase humana continham níveis elevados de ADA em comparação com as células de controlo, nenhum detectado vs. 1,21% de ácidos gordos totais, respectivamente. Assim, a alongase humana converte ambos os AGPI com uma cadeia de 18 e 20 carbonos de comprimento nos seus respectivos ácidos gordos alongados.

As células de levedura contendo o ADNc da alongase humana tiveram também níveis elevados de ácidos gordos mono-insaturados incluindo 18:1n-7, 20:1n-7, 20:1n-9 e 18:1n-5, em comparação com a estirpe de controlo. Por essa razão, estes resultados indicam que a alongase humana identificada é capaz de utilizar AGPI bem como ácidos gordos mono-insaturados como substratos. Assim, esta sequência humana HSELO1, e a sua proteína codificada, possuem actividade de alongase independente da especificidade do substrato.

Exemplo XVII

Clonagem e Expressão de uma Alongase de AGPI de *C. elegans*

Foram identificadas várias alongases putativas de *C. elegans* com homologia de aminoácidos com ambas GLELO e MAELO traduzidas. Tal como com a sequência de ADNc humana, foi utilizada a clonagem de um ADNc e a expressão em levedura para determinar se de facto era uma alongase de AGPI. Os iniciadores R0738 (5'-AAT CAG GAA TTC ATG GCT CAG CAT CCG CTC GTT CAA C-3') e R0739 (5'-CCG CTT GTC GAC TTA GTT GTT CTT CTT CTT TGG CAC-3') com os locais de restrição *EcoRI* e *SalI* (sublinhados), respectivamente, foram baseados na sequência putativa de ADNc contida na sequência genómica U68749 (número de acesso wormpep do ADNc F56H11.4). Foi efectuada uma amplificação por PCR num volume de 100 µl contendo: 250 ng de ADNc da biblioteca de *C. elegans* excisado (OriGene Technologies Inc., Rockville, MD), 50 pmol de cada iniciador, 10 µl de tampão reaccional 10× (Boehringer Mannheim Corp., Indianapolis, IN), 1 µl de mistura de Nucleótidos de PCR 10 mM (Boehringer Mannheim Corp., Indianapolis, IN) e 2,5 U de polimerase Taq (Boehringer Mannheim Corp., Indianapolis, IN). As condições dos ciclos térmicos num Perkin Elmer 9600 (Norwalk, CT) foram como se segue: 95°C durante 5 min., depois 25 ciclos de 94°C durante 30 seg., 55°C durante 2 min. e 72°C

durante 2 min. A PCR foi seguida de um ciclo adicional de 72°C durante 7 minutos.

O produto amplificado por PCR foi purificado num gel de agarose, cortado com *EcoRI* e *SalI*, ligado a pYX242 (Invitrogen Corp., Carlsbad, CA) (linearizado com *EcoRI* e *SalI*) utilizando o estojó Rapid Ligation (Boehringer Mannheim Corp., Indianapolis, IN), de acordo com o protocolo do fabricante e transformado em células Top10 de *E. coli* (Invitrogen Corp., Carlsbad, CA). Os novos plasmídeos, designados pRET-21 e pRET-22 (dois clones individuais da ligação), foram sequenciados com o 373A Stretch DNA sequencer ABI (Perkin Elmer, Foster City, CA), e as sequências de ADNc eram idênticas. A sequência nucleotídica de ADNc de 867 bases do plasmídeo pRET-22 contendo a alongase putativa é mostrada Figura 46 e a sequência traduzida de 288 aminoácidos é mostrada na Figura 47. (O plasmídeo pRET-22 foi depositado na American Type Culture Collection, 10801 University Boulevard, Manassas, VA 20110-2209 a 19 de Agosto de 1999 sob os termos do Tratado de Budapeste e foi-lhe atribuído o número de depósito ____).

Os plasmídeos pRET-21 e pRET-22 foram transformados em *S. cerevisiae* 334 tal como anteriormente descrito (ver Exemplo III) e as culturas de levedura resultantes (334(pRET-21) e 334(pRET-22) criadas em 100 ml de meio selectivo sem leucina (Ausubel et al., *supra*) durante 48 horas a 20°C na presença de 50 µM de GLA e AA. Os sedimentos celulares foram recolhidos e sujeitos a análise de ácidos gordos e os resultados mostrados na Figura 48. Verificou-se que DGLA, o produto previsto do alongamento de GLA, era uma média de 1,79% dos lípidos totais nas duas amostras, *versus* 0,13% para o controlo negativo (334 contendo o plasmídeo pYX242) indicando que a enzima codificada tanto por pRET-21 como por pRET-22 possuía actividade de alongase de GLA. A percentagem de conversão de GLA em DGLA por 334(pRET-21) e 334(pRET-22) foi de 11,1% e 19,4%, respectivamente, com uma média de 15,25%. Interessantemente, quase não foi observado alongamento de AA ou de qualquer ácido gordo endógeno (Fig. 48). Estes resultados indicam que a alongase codificada por este ADNc de *C. elegans* recentemente identificado, CEEL01, é capaz de alongar especificamente GLA em DGLA, sugerindo que pode ser um homólogo de *C. elegans* da alongase de GLA.

Exemplo XVIII

Isolamento de um ADNc de uma Alongase Putativa Humana com Base na Sequência AC004050

Para isolar o ADNc inteiro da alongase putativa com base na sequência AC004050, os iniciadores RP735 (5'-CCT CCT GAA TTC CAQA CAC TAT TCA GCT TTC-3') e R073 (5'-TAA TAC GAC TCA CTA TAG GG-3') foram utilizados para amplificar por PCR o ADNc Marathon-Ready de fígado humano (Clontech Laboratories, Inc., Palo Alto, CA). A PCR foi realizada utilizando o estojo de PCR de ADNc Advantage™ (Clontech Laboratories, Inc., Palo Alto, CA) com 5 µl de ADNc Marathon-Ready de fígado humano e 50 pmol de cada iniciador seguindo as instruções do fabricante. As condições dos ciclos térmicos em Perkin Elmer 9600 (Norwalk, CT) foram como se segue: 94°C durante 2 min, depois 30 ciclos de 94°C durante 1 min., 58°C durante 2 min. e 72°C durante 3 min. A PCR foi seguida de uma extensão adicional a 72°C durante 7 min.

O produto de PCR amplificado foi corrido num gel, um fragmento amplificado de aproximadamente 1 kb foi purificado em gel, os terminais do fragmento foram preenchidos com ADN-polimerase T4 (Boehringer Mannheim, Corp., Carlsbad, CA) seguindo as instruções do fabricante. O novo plasmídeo foi designado como pRAE-59 e o ADNc da alongase putativa de AGPI neste plasmídeo, designado como HS3, foi sequenciado utilizando o ABI 373A Stretch Sequencer (Perkin Elmer, Foster City, CA). A sequência de ADNc da alongase putativa de AGPI HS3 é mostrada na Figura 49, e a sequência traduzida é mostrada na Figura 50.

Composições Nutricionais

Os AGPI descritos nas Descrição Detalhada podem ser utilizados em vários suplementos nutricionais, formulações infantis, substitutos nutricionais e outras soluções nutricionais.

I. FORMULAÇÕES INFANTIS

A. Fórmula de Soja Isomil® com Ferro:

Utilização: Como bebida para bebês, crianças e adultos com uma alergia ou sensibilidade a leite de vaca. Um alimento para pacientes com distúrbios para os quais a lactose deve ser evitada: deficiência de lactase, intolerância à lactose e galactosemia.

Características:

- Isolado de proteínas de soja para evitar sintomas de alergia ou sensibilidade a leite de vaca;
- Formulação sem lactose para evitar diarreia associada a lactose;
- Baixa osmolalidade (240 mOs/kg de água) para reduzir o risco de diarreia osmótica;
- Hidratos de carbono duplos (xarope de milho e sacarose) concebidos para aumentar a absorção de hidratos de carbono e reduzir o risco de exceder a capacidade de absorção do intestino lesado;
- 1,8 mg de Ferro (como sulfato ferroso) por 100 Calorias para ajudar a prevenir a deficiência em ferro;
- Níveis recomendados de vitaminas e minerais;
- Óleos vegetais para proporcionar os níveis recomendados de ácidos gordos essenciais;
- Cor branca do leite, consistência semelhante à do leite e aroma agradável.

Ingredientes: (Pareve) 85% de água, 4,9% de xarope de milho, 2,6% de açúcar (sacarose), 2,1% de óleo de soja, 1,9% de isolado de proteínas de soja, 1,4% de óleo de coco, 0,15% de citrato de cálcio, 0,11% de fosfato de cálcio tribásico, citrato de potássio, fosfato de potássio monobásico, cloreto de potássio, monoglicéridos e diglicéridos, lecitina de soja, carragenano, ácido ascórbico, L-metionina, cloreto de magnésio, fosfato de potássio dibásico, cloreto de sódio, cloreto de colina, taurina, sulfato ferroso, m-inositol, acetato de alfa-tocoferilo, sulfato de zinco, L-carnitina, niacinamida, pantotenato de cálcio, sulfato de cobre,

palmitato de vitamina A, cloridrato de cloreto de tiamina, riboflavina, cloridrato de piridoxina, ácido fólico, sulfato de manganês, iodeto de potássio, filoquinona, biotina, selenito de sódio, vitamina D3 e cianocobalamina.

B. Fórmula de Soja Isomil[®] DF para Diarreia:

Utilização: Como alimentação de curto prazo para a gestão dietética da diarreia em crianças e bebês.

Características:

- Primeira fórmula infantil a conter fibra dietética adicionada a partir de fibra de soja especificamente para a gestão da diarreia;
- Mostrou-se clinicamente que reduz a duração da soltura e das fezes líquidas durante diarreia moderada a grave em crianças;
- Nutricionalmente completa para preencher as necessidades nutricionais da criança;
- Isolado de proteínas de soja com adição de L-metionina preenche ou excede os requisitos de uma criança quanto a todos os aminoácidos essenciais;
- Formulação sem lactose para evitar diarreia associada à lactose;
- Baixa osmolalidade (240 mOsm/kg de água) para reduzir o risco de diarreia osmótica;
- Hidratos de carbono duplos (xarope de milho e sacarose) concebidos para aumentar a absorção de hidratos de carbono e reduzir o risco de exceder a capacidade de absorção do intestino lesado;
- Preenche ou excede os níveis de vitaminas e minerais recomendados pelo Comité de Nutrição da Academia Americana de Pediatria e requeridos pelo Infant Formula Act;
- 1,8 mg de ferro (como sulfato ferroso) por 100 Calorias para ajudar a prevenir a deficiência em ferro;
- Óleos vegetais para proporcionar os níveis recomendados de ácidos gordos essenciais.

Ingredientes: (Pareve) 86% de água, 4,8% de xarope de milho, 2,5% de açúcar (sacarose), 2,1% de óleo de soja, 2,0% de isolado de proteínas de soja, 1,4% de óleo de coco, 0,77% de fibra de soja, 0,12% de citrato de cálcio, 0,11% de fosfato de cálcio tribásico, 0,10% de citrato de potássio, cloreto de potássio, fosfato de potássio monobásico, monoglicéridos e diglicéridos, lecitina de soja, carragenano, cloreto de magnésio, ácido ascórbico, L-metionina, fosfato de potássio dibásico, cloreto de sódio, cloreto de colina, taurina, sulfato ferroso, m-inositol, acetato de alfa-tocoferilo, sulfato de zinco, L-carnitina, niacinamida, pantotenato de cálcio, sulfato de cobre, palmitato de vitamina A, cloridrato de cloreto de tiamina, riboflavina, cloridrato de piridoxina, ácido fólico, sulfato de manganês, iodeto de potássio, filoquinona, biotina, selenito de sódio, vitamina D3 e cianocobalamina.

C. Fórmula de Soja sem Sacarose Isomil® SF com Ferro:

Utilização: Como bebida para bebês, crianças e adultos com uma alergia ou sensibilidade a proteínas do leite de vaca ou uma intolerância à sacarose. Uma alimentação para pacientes com distúrbios para os quais a sacarose deve ser evitada.

Características:

- Isolado de proteínas de soja para evitar sintomas de alergia ou sensibilidade a proteínas do leite de vaca;
- Formulação sem lactose para evitar diarreia associada à lactose (a fonte de hidratos de carbono são Polímeros de Glicose Polycose®);
- Sem sacarose para o paciente que não consegue tolerar sacarose;
- Baixa osmolalidade (180 mOs/kg de água) para reduzir o risco de diarreia osmótica;
- 1,8 mg de ferro (como sulfato ferroso) por 100 Calorias para ajudar a prevenir a deficiência em ferro;
- Níveis recomendados de vitaminas e minerais;
- Óleos vegetais para proporcionar os níveis recomendados de ácidos gordos essenciais;

- Cor branca do leite, consistência semelhante à do leite e aroma agradável.

Ingredientes: (Pareve) 75% de água, 11,8% de amido de milho hidrolisado, 4,1% de óleo de soja, 4,1% de isolado de proteínas de soja, 2,8% de óleo de coco, 1,0% de amido de milho modificado, 0,38% de fosfato de cálcio tribásico, 0,17% de citrato de potássio, 0,13% de cloreto de potássio, monoglicéridos e diglicéridos, lecitina de soja, cloreto de magnésio, ácido ascórbico, L-metionina, carbonato de cálcio, cloreto de sódio, cloreto de colina, carragenano, taurina, sulfato ferroso, m-inositol, acetato de alfa-tocoferilo, sulfato de zinco, L-carnitina, niacinamida, pantotenato de cálcio, sulfato de cobre, palmitato de vitamina A, cloridrato de cloreto de tiamina, riboflavina, cloridrato de piridoxina, ácido fólico, sulfato de manganês, iodeto de potássio, filoquinona, biotina, selenito de sódio, vitamina D3 e cianocobalamina.

D. Fórmula de Soja Isomil® 20 com Ferro Pronta a Comer, 20 Cal/fl oz.:

Utilização: Quando é desejada uma alimentação de soja.

Ingredientes: (Pareve) 85% de água, 4,9% de xarope de milho, 2,6% de açúcar (sacarose), 2,1% de óleo de soja, 1,9% de isolado de proteínas de soja, 1,4% de óleo de coco, 0,15% de citrato de cálcio, 0,11% de fosfato de cálcio tribásico, citrato de potássio, fosfato de potássio monobásico, cloreto de potássio, monoglicéridos e diglicéridos, lecitina de soja, carragenano, ácido ascórbico, L-metionina, cloreto de magnésio, fosfato de potássio dibásico, cloreto de sódio, cloreto de colina, taurina, sulfato ferroso, m-inositol, acetato de alfa-tocoferilo, sulfato de zinco, L-carnitina, niacinamida, pantotenato de cálcio, sulfato de cobre, palmitato de vitamina A, cloridrato de cloreto de tiamina, riboflavina, cloridrato de piridoxina, ácido fólico, sulfato de manganês, iodeto de potássio, filoquinona, biotina, selenito de sódio, vitamina D3 e cianocobalamina.

E. Fórmula Infantil Similac®:

Utilização: Quando é necessária uma fórmula infantil: se for tomada a decisão de descontinuar a amamentação antes da idade de 1 ano, se for necessário um suplemento à amamentação ou como alimentação de rotina se não for adoptada a amamentação.

Características:

- Proteína de qualidade e quantidade apropriadas para um bom crescimento; desnaturada pelo calor, o que reduz o risco de perda de sangue entérica associada ao leite;
- Gordura de uma mistura de óleos vegetais (duplamente homogeneizados), proporcionando ácido linoleico essencial que é facilmente absorvido;
- Hidratos de carbono como lactose em proporção semelhante à do leite humano;
- Baixa carga de solutos renais para minimizar o stress no desenvolvimento dos órgãos;
- Formas em Pó, Líquido Concentrado e Pronta a Comer.

Ingredientes: (-D) Água, leite magro, lactose, óleo de soja, óleo de coco, monoglicéridos e diglicéridos, lecitina de soja, ácido ascórbico, carragenano, cloreto de colina, taurina, m-inositol, acetato de alfa-tocoferilo, sulfato de zinco, niacinamida, sulfato ferroso, pantotenato de cálcio, sulfato de cobre, palmitato de vitamina A, cloridrato de cloreto de tiamina, riboflavina, cloridrato de piridoxina, ácido fólico, sulfato de manganês, filoquinona, biotina, selenito de sódio, vitamina D3 e cianocobalamina.

F. Fórmula para Bebés Prematuros Similac® NeoCare com Ferro:

Utilização: Para as necessidades nutricionais especiais dos bebés prematuros após alta hospitalar. Similac NeoCare é uma fórmula nutricionalmente completa desenvolvida para proporcionar aos bebés prematuros calorias, proteína, vitaminas e minerais extra necessários para promover o crescimento acelerado ("catch-up growth") e suportar o desenvolvimento.

Características:

- Reduz a necessidade de suplementação calórica e vitamínica. Mais calorias (22 Cal/fl oz) que as fórmulas padrão de termo (20 Cal/fl oz);
- Mistura de gorduras altamente absorvível, com triglicéridos de cadeia média (óleo MCT) para ajudar a preencher as necessidades digestivas especiais dos bebés prematuros;
- Níveis mais elevados de proteína, vitaminas e minerais por 100 calorias para prolongar o suporte nutricional iniciado no hospital;
- Mais cálcio e fósforo para melhor mineralização óssea.

Ingredientes: -D Sólidos de xarope de milho, leite magro, lactose, concentrado de proteínas do soro do leite, óleo de soja, óleo de cártamo de elevado teor oleico, óleo de coco fraccionado (triglicéridos de cadeia média), óleo de coco, citrato de potássio, fosfato de cálcio tribásico, carbonato de cálcio, ácido ascórbico, cloreto de magnésio, cloreto de potássio, cloreto de sódio, taurina, sulfato ferroso, m-inositol, cloreto de colina, palmitato de ascorbilo, L-carnitina, acetato de alfa-tocoferilo, sulfato de zinco, niacinamida, tocoferóis mistos, citrato de sódio, pantotenato de cálcio, sulfato de cobre, cloridrato de cloreto de tiamina, palmitato de vitamina A, betacaroteno, riboflavina, cloridrato de piridoxina, ácido fólico, sulfato de manganês, filoquinona, biotina, selenito de sódio, vitamina D3 e cianocobalamina.

G. Fortificante do Leite Humano Similac Cuidado Natural Pobre em Ferro Pronto a Utilizar, 24 Cal/fl oz.:

Utilização: Concebida para ser misturada com o leite humano ou para ser dada alternadamente com o leite humano a bebés com baixo peso à nascença.

Ingredientes: -D Água, leite magro, amido de milho hidrolisado, lactose, óleo de coco fraccionado (triglicéridos de cadeia média), concentrado de proteínas do soro do leite, óleo de soja, óleo de coco, fosfato de cálcio tribásico, citrato de potássio, cloreto de magnésio, citrato de sódio, ácido ascórbico, carbonato de cálcio, monoglicéridos e diglicéridos, lecitina de soja, carragenano, cloreto de

colina, m-inositol, taurina, niacinamida, L-carnitina, acetato de alfa-tocoferilo, sulfato de zinco, cloreto de potássio, pantotenato de cálcio, sulfato ferroso, sulfato de cobre, riboflavina, palmitato de vitamina A, cloridrato de cloreto de tiamina, cloridrato de piridoxina, biotina, ácido fólico, sulfato de manganês, filoquinona, vitamina D3, selenito de sódio e cianocobalamina.

Vários AGPI deste invento podem ser substituídos e/ou adicionados às fórmulas de bebês descritas acima e a outras fórmulas de bebês conhecidas dos peritos na arte.

II. FORMULAÇÕES NUTRICIONAIS

A. ENSURE®

Utilização: ENSURE é um alimento líquido de baixo resíduo concebido primeiramente como suplemento nutricional oral para ser utilizado com ou entre as refeições ou, em quantidades apropriadas, como substituto das refeições. ENSURE é isento de lactose e glúten e é adequado para utilizar em dietas modificadas, incluindo dietas de baixo colesterol. Embora seja primeiramente um suplemento oral, pode ser dado através de tubo.

Condições dos Pacientes:

- Para pacientes a dietas modificadas;
- Para pacientes idosos em risco de nutrição;
- Para pacientes com perda involuntária de peso;
- Para pacientes a recuperar de doença ou cirurgia;
- Para pacientes que necessitem de uma dieta de baixo resíduo.

Ingredientes: -D Água, Açúcar (Sacarose), Maltodextrina (Milho), Caseinatos de Cálcio e Sódio, Óleo de Cártamo De elevado teor oleico, Isolado de Proteínas de Soja, Óleo de Soja, Óleo de Canola, Citrato de Potássio, Fosfato de Cálcio Tribásico, Citrato de Sódio, Cloreto de Magnésio, Fosfato de Magnésio Dibásico, Aroma Artificial, Cloreto de Sódio, Lecitina de Soja, Cloreto de Colina, Ácido Ascórbico, Carragenano, Sulfato de Zinco, Sulfato Ferroso, Acetato de Alfa-tocoferilo, Goma Gelana, Niacinamida, Pantotenato de Cálcio, Sulfato de Manganês, Sulfato de Cobre, Palmitato de Vitamina A, Cloridrato de Cloreto de Tiamina, Cloridrato de

Piridoxina, Riboflavina, Ácido Fólico, Molibdato de Sódio, Cloreto de Crômio, Biotina, Iodeto de Potássio, Selenato de Sódio.

B. BARRAS ENSURE®:

Utilização: As BARRAS ENSURE são uma nutrição completa e equilibrada para utilização suplementar entre ou com as refeições. Elas proporcionam uma alternativa deliciosa rica em nutrientes a outras refeições ligeiras ("snacks"). BARRAS ENSURE contêm <1 g de lactose/barra e o aroma "Chocolate Fudge Brownie" é isento de glúten (o aroma "Honey Graham Crunch" contém glúten).

Condições dos Pacientes:

- Para pacientes que necessitam de calorias, proteína, vitaminas e minerais extra;
- Especialmente útil para pessoas que não tomam calorias nem nutrientes suficientes;
- Para pessoas que têm a capacidade de mastigar e engolir;
- Não deve ser utilizado por ninguém com alergia a amendoim ou qualquer tipo de alergia a frutos secos.

Ingredientes: "Honey Graham Crunch" - Xarope de Milho Rico em Frutose, Isolado de Proteínas de Soja, Açúcar Mascavado, Mel, Maltodextrina (Milho), Arroz Crocante (Arroz Moído, Açúcar [Sacarose], Sal [Cloreto de Sódio] e Malte), Farelo de Aveia, Semente de Algodoeiro Parcialmente Hidrogenada e Óleos de Soja, Polissacárido de Soja, Glicerina, Concentrado de Proteínas do Soro do Leite, Polidextrose, Frutose, Caseinato de Cálcio, Cacau em Pó, Aromas Artificiais, Óleo de Canola, Óleo de Cártamo de elevado teor oleico, Leite em Pó Magro, Soro de Leite em Pó, Lecitina de Soja e Óleo de Milho. Fabricado numa instalação que processa frutos secos.

Vitaminas e Minerais: Fosfato de Cálcio Tribásico, Fosfato de Potássio Dibásico, Óxido de Magnésio, Sal (Cloreto de Sódio), Cloreto de Potássio, Ácido Ascórbico, Ortofosfato Férrico, Acetato de Alfa-tocoferilo, Niacinamida, Óxido de Zinco, Pantotenato de Cálcio, Gluconato de Cobre, Sulfato de Manganês, Riboflavina, Betacaroteno, Cloridrato de Piridoxina, Mononitrato de Tiamina, Ácido Fólico, Biotina, Cloreto de

Crómio, Iodeto de Potássio, Selenato de Sódio, Molibdato de Sódio, Filoquinona, Vitamina D3 e Cianocobalamina.

Proteína: “Honey Graham Crunch” - A fonte de proteínas é uma mistura de isolado de proteínas de soja e de proteínas do leite.

Isolado de proteína de soja	74%
Proteínas do leite	26%

Gordura: “Honey Graham Crunch” - A fonte de gordura é uma mistura de óleos de semente de algodoeiro parcialmente hidrogenada e soja, canola, cártamo de elevado teor oleico, e lecitina de soja.

Óleo de semente de algodoeiro parcialmente hidrogenado e soja	76%
Óleo de canola	8%
Óleo de cártamo de elevado teor oleico	8%
Óleo de milho	4%
Lecitina de soja	4%

Hidratos de carbono: “Honey Graham Crunch” - A fonte de hidratos de carbono é uma combinação de xarope de milho rico em frutose, açúcar mascavado, maltodextrina, mel, arroz crocante, glicerina, polissacáridos de soja e farelo de aveia.

Xarope de milho rico em frutose	24%
Açúcar mascavado	21%
Maltodextrina	12%
Mel	11%
Arroz crocante	9%
Glicerina	9%
Polissacáridos de soja	7%
Farelo de aveia	7%

C. ENSURE® RICO EM PROTEÍNA:

Utilização: ENSURE RICO EM PROTEÍNA é um alimento líquido concentrado, rico em proteína concebido para pessoas que requeiram calorias, proteína, vitaminas e minerais adicionais nas suas dietas. Pode ser utilizado como um suplemento

nutricional oral com ou entre as refeições ou, nas quantidades apropriadas, como um substituto da refeição. ENSURE RICO EM PROTEÍNA é isento de lactose e glúten e é adequado para utilização por pessoas a recuperar de cirurgia geral ou de fracturas da anca e por pacientes em risco de úlceras de pressão.

Condições dos Pacientes:

- Para pacientes que requeiram calorias, proteína, vitaminas e minerais adicionais, tais como pacientes a recuperar de cirurgia geral ou de fracturas da anca, pacientes em risco de úlceras de pressão e pacientes a dietas de baixo colesterol.

Características:

- Pobre em gorduras saturadas;
- Contém 6 g de gordura total e <5 mg de colesterol por dose;
- Sabor rico e cremoso;
- Excelente fonte de proteína, cálcio e outras vitaminas e minerais essenciais;
- Para dietas de baixo colesterol;
- Sem lactose e facilmente digerido.

Ingredientes:

“Vanilla Supreme”: -D Água, Açúcar (Sacarose), Maltodextrina (Milho), Caseinatos de Cálcio e Sódio, Óleo de Cártamo De elevado teor oleico, Isolado de Proteínas de Soja, Óleo de Soja, Óleo de Canola, Citrato de Potássio, Fosfato de Cálcio Tribásico, Citrato de Sódio, Cloreto de Magnésio, Fosfato de Magnésio Dibásico, Aroma Artificial, Cloreto de Sódio, Lecitina de Soja, Cloreto de Colina, Ácido Ascórbico, Carragenano, Sulfato de Zinco, Sulfato Ferroso, Acetato de Alfa-tocoferilo, Goma Gelana, Niacinamida, Pantotenato de Cálcio, Sulfato de Manganês, Sulfato de Cobre, Palmitato de Vitamina A, Cloridrato de Cloreto de Tiamina, Cloridrato de Piridoxina, Riboflavina, Ácido Fólico, Molibdato de Sódio, Cloreto de Crómio, Biotina, Iodeto de Potássio, Selenato de Sódio, Filoquinona, Vitamina D3 e Cianocobalamina.

Proteína:

A fonte de proteínas é uma mistura de duas proteínas de elevado valor biológico: caseína e soja.

Caseinatos de sódio e cálcio	85%
Isolado de proteínas de soja	15%

Gordura:

A fonte de gordura é uma mistura de três óleos: cártamo de elevado teor oleico, canola e soja.

Óleo de cártamo de elevado teor oleico	40%
Óleo de canola	30%
Óleo de soja	30%

O nível de gordura em ENSURE RICO EM PROTEÍNA verifica as linhas orientadoras da American Heart Association (AHA). Os 6 gramas de gordura em ENSURE RICO EM PROTEÍNA representam 24% das calorias totais, sendo 2,6% da gordura de ácidos gordos saturados e 7,9% de ácidos gordos poli-insaturados. Estes valores estão dentro das linhas orientadoras da AHA de <30% das calorias totais provenientes de gordura, <10% das calorias de ácidos gordos saturados, e <10% das calorias totais de ácidos gordos poli-insaturados.

Hidratos de carbono:

ENSURE RICO EM PROTEÍNA contém uma combinação de maltodextrina e sacarose. A doçura moderada e a variedade de aromas ("vanilla supreme", "chocolate royal", frutos silvestres e banana), mais os Pacotes de Aromas VARI-FLAVORS® a noz pecã, cereja, morango, limão e laranja, ajudam a prevenir a fadiga do aroma e contribuem para a adesão do paciente.

Baunilha e outros aromas que não chocolate:

Sacarose	60%
Maltodextrina	40%

Chocolate:

Sacarose	70%
Maltodextrina	30%

D. ENSURE® LIGHT

Utilização: ENSURE LIGHT é um alimento líquido pobre em gordura concebido para utilizar como suplemento nutricional oral com ou entre as refeições. ENSURE LIGHT é isento de lactose e glúten e é adequado para utilizar em dietas modificadas, incluindo dietas de baixo colesterol.

Condições dos Pacientes:

- Para pacientes de peso normal ou com excesso de peso que necessitem de nutrição extra num suplemento que contém 50% menos gordura e 20% menos calorias que ENSURE;
- Para adultos saudáveis que não se alimentem bem e necessitem de nutrição extra.

Características:

- Pobre em gordura e em gordura saturada;
- Contém 3 g de gordura total por dose e <5 mg de colesterol;
- Sabor rico e cremoso;
- Excelente fonte de cálcio e outras vitaminas e minerais essenciais;
- Para dietas de baixo colesterol;
- Sem lactose e de fácil digestão.

Ingredientes:

“French Vanilla”: -D Água, Maltodextrina (Milho), Açúcar (Sacarose), Caseinato de Cálcio, Óleo de Cártamo De elevado teor oleico, Óleo de Canola, Cloreto de Magnésio, Citrato de Sódio, Citrato de Potássio, Fosfato de Potássio Dibásico, Fosfato de Magnésio Dibásico, Aroma Natural e Artificial, Fosfato de Cálcio Tribásico, Gel de Celulose, Cloreto de Colina, Lecitina de Soja, Carragenano, Sal (Cloreto de Sódio), Ácido Ascórbico, Goma Celulósica, Sulfato Ferroso, Acetato de Alfa-tocoferilo, Sulfato de Zinco, Niacinamida, Sulfato de Manganês, Pantotenato de Cálcio, Sulfato de Cobre, Cloridrato de Cloreto de Tiamina, Palmitato de Vitamina A, Cloridrato de Piridoxina, Riboflavina, Cloreto de Crómio, Ácido Fólico, Molibdato de Sódio, Biotina, Iodeto de Potássio, Selenato de Sódio, Filoquinona, Vitamina D3 e Cianocobalamina.

Proteína:

A fonte proteica é caseinato de cálcio.

Caseinato de cálcio	100%
---------------------	------

Gordura:

A fonte de gordura é uma mistura de dois óleos: cártamo de elevado teor oleico e canola.

Óleo de cártamo de elevado teor oleico	70%
Óleo de canola	30%

O nível de gordura em ENSURE LIGHT verifica as linhas orientadoras da American Heart Association (AHA). Os 3 gramas de gordura em ENSURE LIGHT representam 13,5% das calorias totais, sendo 1,4% da gordura proveniente de ácidos gordos saturados e 2,6% de ácidos gordos poli-insaturados. Estes valores estão dentro das linhas orientadoras da AHA de <30% das calorias totais provenientes de gordura, <10% das calorias de ácidos gordos saturados, e <10% das calorias totais de ácidos gordos poli-insaturados.

Hidratos de carbono:

ENSURE LIGHT contém uma combinação de maltodextrina e sacarose. O aroma a chocolate também contém xarope de milho. A doçura moderada e a variedade de aromas ("French Vanilla", "Chocolate Supreme", "Strawberry Swirl"), mais os pacotes de aromas VARI-FLAVORS® a noz pecã, cereja, morango, limão e laranja, ajudam a prevenir a fadiga do aroma e contribuem para a adesão do paciente.

Baunilha e outros aromas que não chocolate:

Sacarose	51%
Maltodextrina	49%

Chocolate:

Sacarose	47,0%
Xarope de milho	26,5%
Maltodextrina	26,5%

Vitaminas e Minerais:

Uma dose de 8-fl-oz de ENSURE LIGHT proporciona pelo menos 25% da DDR para 24 vitaminas e minerais chave.

Cafeína:

O aroma de chocolate contém 2,1 mg de cafeína/8 fl oz.

E. ENSURE PLUS®

Utilização: ENSURE PLUS é um alimento líquido rico em calorias e de resíduo reduzido para utilizar quando são necessárias calorias e nutrientes extra, mas uma concentração de proteína normal. Está concebido primeiramente como suplemento nutricional oral a utilizar com ou entre as refeições ou, nas quantidades apropriadas, como um substituto de refeições. ENSURE PLUS é isento de lactose e glúten. Embora seja primeiramente um suplemento nutricional oral, pode ser dado através de um tubo.

Condições dos Pacientes:

- Para pacientes que necessitem de calorias e nutrientes extra mas de uma concentração normal de proteína, num volume limitado;
- Para pacientes que necessitem de ganhar ou manter um peso saudável.

Características:

- Sabor rico e cremoso;
- Boa fonte de vitaminas e minerais essenciais.

Ingredientes:

"Vanilla": -D Água, Xarope de milho, Maltodextrina (Milho), Óleo de Milho, Caseinatos de Sódio e Cálcio, Açúcar (Sacarose), Isolado de Proteínas de Soja, Cloreto de Magnésio, Citrato de Potássio, Fosfato de Cálcio Tribásico, Lecitina de Soja, Aroma Natural e Artificial, Citrato de Sódio, Cloreto de Potássio, Cloreto de Colina, Ácido Ascórbico, Carragenano, Sulfato de Zinco, Sulfato Ferroso, Acetato de Alfa-tocoferilo, Niacinamida, Pantotenato de Cálcio, Sulfato de Manganês, Sulfato de Cobre, Cloridrato de Cloreto de Tiamina, Cloridrato de Piridoxina, Riboflavina, Palmitato de Vitamina A, Ácido Fólico, Biotina, Cloreto de Crómio, Molibdato de Sódio, Iodeto de Potássio, Selenito de sódio, Filoquinona, Cianocobalamina e Vitamina D3.

Proteína:

A fonte proteica é uma mistura de duas proteínas de elevado valor biológico: caseína e soja.

Caseinatos de sódio e cálcio	84%
Isolado de proteínas de soja	16%

Gordura:

A fonte de gordura é óleo de milho.

Óleo de milho	100%
---------------	------

Hidratos de carbono:

ENSURE PLUS contém uma combinação de maltodextrina e sacarose. A doçura moderada e a variedade de aromas (baunilha, chocolate, morango, café, manteiga de noz pecã e "eggnog"), mais os Pacotes de Aromas VARI-FLAVORS® a noz pecã, cereja, morango, limão e laranja, ajudam a prevenir a fadiga do aroma e contribuem para a adesão do paciente.

Aromas de baunilha, morango, manteiga de noz pecã e café:

Xarope de Milho	39%
Maltodextrina	38%
Sacarose	23%

Aromas de chocolate e "eggnog":

Xarope de Milho	36%
Maltodextrina	34%
Sacarose	30%

Vitaminas e Minerais:

Uma dose de 8-fl-oz de ENSURE PLUS proporciona pelo menos 15% da DDR para 25 Vitaminas e minerais chave.

Cafeína:

O aroma de chocolate contém 3,1 mg de Cafeína/8 fl oz. O aroma de café contém uma quantidade vestigiária de cafeína.

F. ENSURE PLUS® HN

Utilização: ENSURE PLUS HN é um alimento líquido nutricionalmente completo rico em calorias e rico em azoto, concebido para pessoas com elevadas necessidades calóricas e proteicas ou uma tolerância de volume limitada. Pode ser utilizado para suplementação oral ou para suporte nutricional total através de tubo. ENSURE PLUS HN é isento de lactose e glúten.

Condições dos Pacientes:

- Para pacientes com maiores necessidades calóricas e proteicas, tais como após cirurgia ou lesão;
- Para pacientes com tolerância de volume limitada e saciedade precoce.

Características:

- Para nutrição suplementar ou total;
- Para alimentação oral ou por tubo;
- 1,5 CaVmL;
- Rico em azoto;
- Caloricamente denso.

Ingredientes:

“Vanilla”: -D Água, Maltodextrina (Milho), Caseinatos de Sódio e Cálcio, Óleo de Milho, Açúcar (Sacarose), Isolado de Proteínas de Soja, Cloreto de Magnésio, Citrato de Potássio, Fosfato de Cálcio Tribásico, Lecitina de Soja, Aroma Natural e Artificial, Citrato de Sódio, Cloreto de Colina, Ácido Ascórbico, Taurina, L-carnitina, Sulfato de Zinco, Sulfato Ferroso, Acetato de Alfa-tocoferilo, Niacinamida, Carragenano, Pantotenato de Cálcio, Sulfato de Manganês, Sulfato de Cobre, Cloridrato de Cloreto de Tiamina, Cloridrato de Piridoxina, Riboflavina, Palmitato de Vitamina A, Ácido Fólico, Biotina, Cloreto de Crómio, Molibdato de Sódio, Iodeto de Potássio, Selenito de Sódio, Filoquinona, Cianocobalamina e Vitamina D3.

G. ENSURE® PÓ:

Utilização: ENSURE PÓ (reconstituído com água) é um alimento líquido de resíduo reduzido concebido primeiramente como suplemento nutricional oral a utilizar com ou entre as

refeições. ENSURE PÓ é isento de lactose e glúten e é adequado para utilizar em dietas modificadas, incluindo dietas de baixo colesterol.

Condições dos Pacientes:

- Para pacientes em dietas modificadas;
- Para pacientes idosos em risco de nutrição;
- Para pacientes a recuperar de doença/cirurgia;
- Para pacientes que necessitem de uma dieta de resíduo reduzido.

Características:

- Conveniente, fácil de misturar;
- Pobre em gordura saturada;
- Contém 9 g de gordura total e <5 mg de colesterol por dose;
- Rico em vitaminas e minerais;
- Para dietas pobres em colesterol;
- Sem colesterol e de fácil digestão.

Ingredientes: -D Xarope de Milho, Maltodextrina (Milho), Açúcar (Sacarose), Óleo de Milho, Caseinatos de Sódio e Cálcio, Isolado de Proteínas de Soja, Aroma Artificial, Citrato de Potássio, Cloreto de Magnésio, Citrato de Sódio, Fosfato de Cálcio Tribásico, Cloreto de Potássio, Lecitina de Soja, Ácido Ascórbico, Cloreto de Colina, Sulfato de Zinco, Sulfato Ferroso, Acetato de Alfa-tocoferilo, Niacinamida, Pantotenato de Cálcio, Sulfato de Manganês, Cloridrato de Cloreto de Tiamina, Sulfato de Cobre, Cloridrato de Piridoxina, Riboflavina, Palmitato de Vitamina A, Ácido Fólico, Biotina, Molibdato de Sódio, Cloreto de Crómio, Iodeto de Potássio, Selenato de Sódio, Filoquinona, Vitamina D3 e Cianocobalamina.

Proteína:

A fonte proteica é uma mistura de duas proteínas de elevado valor biológico: caseína e soja.

Caseinatos de sódio e cálcio	84%
Isolado de proteínas de soja	16%

Gordura:

A fonte de gordura é o óleo de milho.

Óleo de milho	100%
---------------	------

Hidratos de carbono:

ENSURE PÓ contém uma combinação de xarope de milho, maltodextrina e sacarose. A doçura moderada de ENSURE PÓ, mais os Pacotes de Aromas VARI-FLAVORS® a noz pecã, cereja, morango, limão e laranja, ajudam a prevenir a fadiga do aroma e contribuem para a adesão do paciente.

Baunilha:

Xarope de Milho	35%
Maltodextrina	35%
Sacarose	30%

H. ENSURE® PUDIM

Utilização: ENSURE® PUDIM é um suplemento denso em nutrientes que proporciona uma nutrição equilibrada numa forma não líquida a utilizar com ou entre as refeições. É apropriado para dietas de consistência modificada (p.ex., mole, em puré ou completamente líquida) ou para pessoas com dificuldades de deglutição. ENSURE PUDIM é isento de glúten.

Condições dos Pacientes:

- Para pacientes a dietas de consistência modificada (p.ex., mole, em puré ou completamente líquida);
- Para pacientes com dificuldades de deglutição.

Características:

- Rico e cremoso, bom paladar;
- Boa fonte de vitaminas e minerais essenciais;
- Conveniente, não necessita de refrigeração;
- Sem glúten.

Perfil de nutrientes por 5 oz: Calorias 250, Proteína 10,9%, Gordura Total 34,9%, Hidratos de Carbono 54,2%.

Ingredientes:

"Vanilla": -D Leite Magro, Água, Açúcar (Sacarose), Óleo de Soja Parcialmente Hidrogenado, Amido Alimentar Modificado, Sulfate de Magnésio, Estearoil-lactilato de Sódio, Fosfato de Sódio Dibásico, Aroma Artificial, Ácido Ascórbico, Sulfato de Zinco, Sulfato Ferroso, Acetato de Alfa-tocoferilo, Cloreto de Colina, Niacinamida, Sulfato de Manganês, Pantotenato de Cálcio, FD&C Amarelo n.º 5, Citrato de Potássio, Sulfato de Cobre, Palmitato de Vitamina A, Cloridrato de Cloreto de Tiamina, Cloridrato de Piridoxina, Riboflavina, FD&C Amarelo n.º 6, Ácido Fólico, Biotina, Filoquinona, Vitamina D3 e Cianocobalamina.

Proteína:

A fonte proteica é leite magro.

Leite magro	100%
-------------	------

Gordura:

A fonte de gordura é óleo de soja hidrogenado:

Óleo de soja hidrogenado	100%
--------------------------	------

Hidratos de Carbono:

ENSURE PUDIM contém uma combinação de sacarose e amido alimentar modificado. A doçura moderada e a variedade de aromas (baunilha, chocolate, "butterscotch", e tapioca) ajudam a prevenir a fadiga do aroma. O produto contém 9,2 gramas de lactose por dose.

Baunilha e outros aromas que não chocolate:

Sacarose	56%
Lactose	27%
Amido alimentar modificado	17%

Chocolate:

Sacarose	58%
Lactose	26%
Amido alimentar modificado	16%

I. ENSURE® COM FIBRA:

Utilização: ENSURE COM FIBRA é um alimento líquido nutricionalmente completo contendo fibra concebido para pessoas que possam beneficiar do aumento de fibra e nutrientes dietéticos. ENSURE COM FIBRA é adequado para pessoas que não requeiram uma dieta de resíduo reduzido. Pode ser dado oralmente ou através de tubo e pode ser utilizado como suplemento nutricional a uma dieta regular ou, nas quantidades apropriadas, como um substituto de refeição. ENSURE COM FIBRA é isento de lactose e glúten e é adequado para utilizar em dietas modificadas, incluindo dietas de baixo colesterol.

Condições dos Pacientes:

- Para pacientes que possam beneficiar de aumento de fibra e nutrientes dietéticos.

Características:

- Nova fórmula avançada pobre em gordura saturada, rica em vitaminas e minerais;
- Contém 6 g de gordura total e <5 mg de colesterol por dose;
- Paladar rico e cremoso;
- Boa fonte de fibras;
- Excelente fonte de vitaminas e minerais essenciais;
- Para dietas de baixo colesterol;
- Sem lactose nem glúten.

Ingredientes:

“Vanilla”: -D Água; Maltodextrina (Milho), Açúcar (Sacarose), Caseinatos de Sódio e Cálcio, Fibra de Aveia, Óleo de Cártamo De elevado teor oleico, Óleo de Canola, Isolado de Proteínas de Soja, Óleo de Milho, Fibra de Soja, Fosfato de Cálcio Tribásico, Cloreto de Magnésio, Citrato de Potássio, Gel de Celulose, Lecitina de Soja, Fosfato de Potássio Dibásico, Citrato de Sódio, Aromas Naturais e Artificiais, Cloreto de Colina, Fosfato de Magnésio, Ácido Ascórbico, Goma Celulósica, Cloreto de Potássio, Carragenano, Sulfato Ferroso, Acetato de Alfa-tocoferilo, Sulfato de Zinco, Niacinamida, Sulfato de Manganês, Pantotenato de Cálcio, Sulfato de Cobre, Palmitato de Vitamina A, Cloridrato de Cloreto de Tiamina,

Cloridrato de Piridoxina, Riboflavina, Ácido Fólico, Cloreto de Crômio, Biotina, Molibdato de Sódio, Iodeto de Potássio, Selenato de Sódio, Filoquinona, Vitamina D3 e Cianocobalamina.

Proteína:

A fonte proteica é uma mistura de duas proteínas de elevado valor biológico: caseína e soja.

Caseinatos de sódio e cálcio	80%
Isolado de proteínas de soja	20%

Gordura:

A fonte de gordura é uma mistura de três óleos: cártamo de elevado teor oleico, canola e milho.

Óleo de cártamo de elevado teor oleico	40%
Óleo de canola	40%
Óleo de milho	20%

O nível de gordura em ENSURE COM FIBRA verifica as linhas orientadoras da American Heart Association (AHA). Os 6 gramas de gordura em ENSURE COM FIBRA representam 22% das calorias totais, sendo 2,01% da gordura de ácidos gordos saturados e 6,7% de ácidos gordos poli-insaturados. Estes valores estão dentro das linhas orientadoras de AHA de $\leq 30\%$ das calorias totais serem provenientes de gordura, $< 10\%$ das calorias de ácidos gordos saturados, e $\leq 10\%$ das calorias totais de ácidos gordos poli-insaturados.

Hidratos de carbono:

ENSURE COM FIBRA contém uma combinação de maltodextrina e sacarose. A doçura moderada e a variedade de aromas (baunilha, chocolate e manteiga de noz pecã), mais os Pacotes de Aromas VARI-FLAVORS® a noz pecã, cereja, morango, limão e laranja, ajudam a prevenir a fadiga do aroma e contribuem para a adesão do paciente.

Baunilha e outros aromas que não chocolate:

Maltodextrina	66%
Sacarose	25%
Fibra de Aveia	7%
Fibra de Soja	2%

Chocolate:

Maltodextrina	55%
Sacarose	36%
Fibra de Aveia	7%
Fibra de Soja	2%

Fibra:

A mistura de fibras utilizada em ENSURE COM FIBRA consiste em fibra de aveia e polissacáridos de soja. Esta mistura resulta em aproximadamente 4 gramas de fibra dietética total por lata de 8-fl. oz. A proporção de fibra insolúvel para fibra solúvel é de 95:5.

Os vários suplementos nutricionais descritos acima e conhecidos de outros peritos na arte podem ser substituídos e/ou suplementados com os AGPI produzidos de acordo com o presente invento.

J. Produto Nutricional Oxepa™

Oxepa é um produto é um produto nutricional entérico pobre em hidratos de carbono e caloricamente denso concebido para a gestão dietética de pacientes com ou em risco de SIRA. Tem uma combinação única de ingredientes, incluindo uma mistura patenteada de óleos contendo ácido eicosapentaenóico (EPA de óleo de peixe), ácido γ -linolénico (GLA de óleo de borragem) e níveis elevados de antioxidantes.

Distribuição Calórica:

A densidade calórica é elevada a 1,5 Cal/ml (355 Cal/8 fl oz), para minimizar o volume requerido para preencher as necessidades energéticas. A distribuição de calorias em Oxepa é mostrada na Tabela IV.

Tabela IV. Distribuição Calórica de Oxepa

	por 8 fl oz.	por litro	% de Cal
Calorias	355	1,500	---
Gordura (g)	22,2	93,7	55,2
Hidratos de carbono (g)	25	105,5	28,1
Proteína (g)	14,8	62,5	16,7
Água (g)	186	785	---

Gordura:

- Oxepa contém 22,2 g de gordura por dose de 8-fl oz (93,7 g/L);
- A fonte de gordura é uma mistura de óleos de 31,8% de óleo de canola, 25% de triglicéridos de cadeia média (MCT), 20% de óleo de borragem, 20% de óleo de peixe e 3,2% de lecitina de soja. O perfil típico de ácidos gordos de Oxepa é mostrado na Tabela V;
- Oxepa proporciona uma quantidade equilibrada de ácidos gordos poli-insaturados, mono-insaturados e saturados, tal como mostrado na Tabela VI;
- Os triglicéridos de cadeia média (MCT) - 25% da mistura de gorduras - ajudam ao esvaziamento gástrico porque são absorvidos pelo tracto intestinal sem emulsão pelos ácidos biliares.

Os vários componentes ácidos gordos do produto nutricional Oxepa™ podem ser substituídos e/ou suplementados com os AGPI produzidos de acordo com este invento.

Tabela V. Perfil Típico de Ácidos Gordos

	% Total de Ácidos Gordos	g/8 fl oz*	g/l*
Capróico (6:0)	0,2	0,04	0,18
Caprílico (8:0)	14,69	3,1	13,07
Cáprico (10:0)	11,06	2,33	9,87
Palmítico (16:0)	5,59	1,18	4,98
Palmitoleico	1,82	0,38	1,62
Esteárico	1,94	0,39	1,64
Oleico	24,44	5,16	21,75
Linoleico	16,28	3,44	14,49
α-Linolénico	3,47	0,73	3,09
γ-Linolénico	4,82	1,02	4,29
Eicosapentaenóico	5,11	1,08	4,55
n-3-Docosapentaenóico	0,55	0,12	0,49
Docosa-hexaenóico	2,27	0,48	2,02
Outros	7,55	1,52	6,72

Os ácidos gordos são aproximadamente iguais a 95% da gordura total.

Tabela VI. Perfil de Gorduras de Oxepa

% de calorias totais provenientes de gordura	55,2
Ácidos gordos poli-insaturados	31,44 g/l
Ácidos gordos mono-insaturados	25,53 g/l
Ácidos gordos saturados	32,38 g/l
Proporção de n-6 para n-3	1,75:1
Colesterol	9,49 mg/8 fl oz (40,1 mg/l)

Hidratos de carbono:

- O teor em hidratos de carbono é de 25,0 g por dose de 8 fl oz (105,5 g/l);
- As fontes de hidratos de carbono são 45% de maltodextrina (um hidrato de carbono complexo) e 55% de sacarose (um açúcar simples), ambas as quais são de fácil digestão e absorção;
- O teor rico em gordura e pobre em hidratos de carbono de Oxepa está concebido para minimizar a produção de dióxido de carbono (CO₂). Níveis elevados de CO₂ podem complicar o desmame de pacientes dependentes de ventilador. O baixo nível de hidratos de carbono pode também ser útil para os pacientes que tenham desenvolvido hiperglicemia induzida por *stress*;
- Oxepa é isento de lactose.

Os hidratos de carbono dietéticos, os aminoácidos das proteínas e a porção glicerol das gorduras podem ser convertidos em glicose no corpo. Através deste processo, são preenchidos os requisitos de hidratos de carbono dos tecidos dependentes de glicose (tais como o sistema nervoso central e os glóbulos vermelhos). No entanto, uma dieta isenta de hidratos de carbono pode conduzir a cetose, catabolismo excessivo das proteínas dos tecidos e perda de líquido e electrólitos. Estes efeitos podem ser prevenidos através da ingestão diária de 50 a 100 g de hidratos de carbono digeríveis, se a toma calórica for adequada. O nível de

hidratos de carbono em Oxepa é também suficiente para minimizar a gliconeogénese, se as necessidades energéticas estiverem a ser preenchidas.

Proteína:

- Oxepa contém 14,8 g de proteína por dose de 8-fl-oz (62,5 g/l);
- A proporção de calorías totais/azoto (150:1) preenche a necessidade de pacientes sob *stress*;
- Oxepa proporciona proteína suficiente para promover anabolismo e a manutenção de uma massa corporal magra sem precipitação de problemas respiratórios. Tomas elevadas de proteína são uma preocupação em pacientes com insuficiência respiratória. Embora a proteína tenha pouco efeito na produção de CO₂, uma dieta rica em proteínas aumentará a condução ventilatória;
- As fontes proteicas de Oxepa são 86,8% de caseinato de sódio e 13,2% de caseinato de cálcio;
- O perfil de aminoácidos do sistema proteico em Oxepa preenche ou excede o padrão para proteína de elevada qualidade estabelecido pela National Academy of Sciences.

* Oxepa é isento de glúten.

Definições por defeito para os programas de análise

Programas de GCG

Pesquisa FastA

Parâmetros por defeito:

intervalo de interesse	"Begin"= 1	"END"= última proteína ou ácido nucleico
conjunto de pesquisa	toda a SwissProt (proteína) ou GenEMBL (ácido nucleico)	
tamanho da palavra	= (2) para proteína	= (6) para ácido nucleico

Os registos esperados lista os registos até o valor de E() alcançar 2,0.

Pesquisa TFastA

Parâmetros por defeito:

intervalo de interesse "Begin"=1 "END"=último ácido nucleico
 conjunto de pesquisa Toda a GenEMBL
 tamanho da palavra =(2) para proteína

Os registos esperados lista os registos até o valor de E() alcançar 2,0.

Pileup

Parâmetros por defeito:

penalidade por criação de intervalo "gap weight" = 5
 penalidade por prolongamento de intervalo "gap length weight" = 12
 aspecto do gráfico "one page plot density" = 2,7

Programa Sequencher

Parâmetros por defeito:

Montagem Automática Algoritmo de dados sujos=montagem de
 contíguos mais lenta mas comparações mais
 rigorosas entre as sequências

 coincidências mínimas=85%
 sobreposição mínima=20

BLAST 2 (blastp, tblastn)

Parâmetros por defeito:	V=50	Lambda=,329	W=3
	B=50	K=0,140	X=22
	E=10	H=0,427	
blast n			
Parâmetros por defeito:	V=100	Lambda=1,37	W=11
	B=250	K=0,171	X1=22
	E=10	H=1,31	X2=25

BLAST 2 Argumentos da Linha de Comandos

-v Acertos	número de melhores registos para mostrar
-b Alinhamentos	número de melhores alinhamentos para mostrar
-e Valor de Expectativa (E)	[Real] por defeito=10,0
-m Opções de vista do alinhamento:	0=aos pares, 1=principal-secundário mostrando identidades, 2=principal-secundário, sem identidades, 3=principal-secundário plano, mostra identidades, 4=principal-secundário plano, sem identidades, 5=principal-secundário, sem identidades e extremidades cegas, 6= principal-secundário plano, sem identidades e extremidades cegas, [Inteiro] por defeito=0
-F Seq. do filtro de busca	(DUST com blastn, SEG com outros) [T/F] defeito=T
-G Custo de abrir um intervalo	(zero invoca comportamento por defeito) [Inteiro] por defeito=0
-E Custo de prolongar um intervalo	(zero invoca comportamento por defeito) [Inteiro] por defeito=0
-X valor da queda do X para alinhamento com intervalos	(em bits) (zero invoca comportamento por defeito) [Inteiro] por defeito=0
-I Mostra GI em linhas	[T/F] por defeito=F
-q Penalidade para um nucleótido errado	(apenas blastn) [Inteiro] por defeito=-3
-r Recompensa para uma coincidência de nucleótidos	(apenas blastn) [Inteiro] por defeito=1
-f Limiar para prolongamento de acertos	por defeito se zero [Inteiro] por defeito=0
-g Efectua alinhamento com intervalos	(não disponível com tblastx) [T/F] por defeito=T
-q Busca o código genético a utilizar	[Inteiro] por defeito=1
-D BD do código genético	(apenas para tblast[nx]) [Inteiro] por defeito=1
-J Crê na linha da busca	[T/F] por defeito=F
-M Matriz	[String] por defeito=BLOSUM62
-W Tamanho da palavra	por defeito se zero [Inteiro] por defeito=0
-z Comprimento eficaz da base de dados	(utilizar zero para o tamanho real) [Inteiro] por defeito=0
-a Número de processadores a utilizar	[Inteiro] por defeito=sítio configurável (SeqServer.conf)

Os valores permitidos e por defeito para os parâmetros do custo de abertura de intervalo/prolongamento de intervalo(-G/E):

BLOSUM62

-G	9	8	7	12	11	10	
-E	2	2	2		1	1	1

BLOSUM50

-G	12	11	10	9	15	14	13	12	18	17	16	15
-E	3	3	3	3	2	2	2	2	1	1	1	1

PAM250

-G	13	12	11	10	15	14	13	12	19	18	17	16
-E	3	3	3	3	2	2	2	2	1	1	1	1

BLOSUM90

-G	8	7	6	11	10	9
-E	2	2	2	1	1	1

PAM30

-G	5	4	3	7	6	5	10	9	8
-E	3	3	3	2	2	2	1	1	1

PAM70

-G	6	5	4	8	7	6	11	10	9
-E	3	3	3	2	2	2	1	1	1

LISTAGEM DE SEQUÊNCIAS

<110> Pradip MUKERJI,
 Tapas DAS,
 Yung-Sheng HUANG,
 Jennifer M. PARKER-BARNES,
 Amanda Eun-Yeong LEONARD,
 Jennifer THURMOND,

<120> GENES DE ALONGASE E SUAS UTILIZAÇÕES

<130> 6407.US.P1

<140> US 09/145,828

<141> 1998-09-02

<160> 78

<170> FastSEQ for Windows Versão 3.0

<210> 1

<211> 954

<212> ADN

<213> *Mortierella alpin*

<400> 1

atggccgccc	caatcttga	caaggccaac	ttcggcattg	atcagccctt	cggaatcaag	60
ctcgacacct	actttgtca	ggcctatgaa	ctcgtcaccg	gaaagtccat	cgactccttc	120
gtcttccagg	agggcgctac	gcctctctcg	accagagag	aggctgceat	gtggactatc	180
acttacttcg	tcgtcatctt	tgggtgctcg	cagatcatga	agagccagga	cgcttcaag	240
ctcaagcccc	tcttcatect	ccacaacttc	ctcctgacga	tcgcgtccgg	atcgctgttg	300
ctcctgttca	tcgagaacct	ggtcccccac	ctcgccagaa	acggactttt	ctacgccatc	360
tgcgacgacg	gtgcctggac	ccagcgcttc	gagctcctct	actacctcaa	ctacctggtc	420
aagtactggg	agttggccga	caccgtcttt	ttggtcctca	agaagaagcc	tcttgagttc	480
ctgcactact	tccaccactc	gatgaccatg	gttctctgct	ttgtccagct	tggaggatac	540
acttcagtgt	cctgggtccc	tattaccctc	aacttgactg	tccacgtctt	catgtactac	600
tactacatgc	gctccgtgcg	cggtgttcgc	atctggtgga	agcagtactt	gaccactctc	660
cagatcgtcc	agttcgttct	tgacctcgga	ttcatctact	tctgcgccta	cacctacttc	720
gccttcacct	acttcccctg	ggctcccaac	gtcgccaagt	gcgcgggtac	cgagggtgct	780
gctctctttg	gctgcgact	cctctccagc	tatctcttgc	tctttatcaa	cttctaccgc	840
attacctaca	atgccaaagg	caaggcagcc	aaggagcgtg	gaagcaactt	tacccccaag	900
actgtcaagt	ccggcgatc	gccaagaag	ccctccaaga	gcaagcaca	ctaa	954

<210> 2

<211> 955

<212> ADN

<213> *Mortierella alpin*

<400> 2

atggagtcga	ttgcgccatt	cctcccatca	aagatgccgc	aagatctggt	tatggacctt	60
gccaccgcta	tcgggtgccg	ggccgcgccc	tatgtcgatc	ctctcgaggg	cgcgctgggtg	120
gcccaggccg	agaagtacat	ccccacgatt	gtccatcaca	cgcgctgggtt	cctggctcgcg	180
gtggagtcgc	ctttggcccg	tgagctgccg	ttgatgaacc	cgttccacgt	gctgttgatc	240
gtgctcgctt	atttggtcac	ggtctttgtg	ggcatgcaga	tcatagaaga	ctttgagcgg	300
ttcgaggcca	agacgttttc	gctcctgcac	aacttttgtc	tggctcgcgt	cagcgccctac	360
atgtgcgggtg	ggatcctgta	aggcttatca	ggccaactat	ggactgtttg	agaacgctgc	420
tgatcatacc	ttcaagggtc	ttcctatggc	caagatgate	tggctcttct	acttctccaa	480
gatcatggag	tttgtcgaca	ccatgatcat	ggtcctcaag	aagaacaacc	gccagatctc	540
cttcttgcac	gtttaccacc	acagctccat	cttcaccatc	tgggtggttg	tcacctttgt	600
tgcacccaac	ggtgaagcct	acttctctgc	tgcgttgaac	tcgttcatcc	atgtgatcat	660
gtacggctac	tacttcttgt	cgcccttggg	cttcaagcag	gtgtcggttc	tcaagttcta	720
catcacgcgc	tcgcagatga	cacagttctg	catgatgtcg	gtccagtcct	cctgggacat	780
gtacgccatg	aaggctcttg	gcgcgcccg	atacccttct	ttcatcacgg	ctctgctttg	840
gttctacatg	tggaccatgc	tcggctctct	ctacaacttt	tacagaaaga	acgccaaagt	900
ggccaagcag	gccaaggccg	acgctgcca	ggagaaggca	aggaaagttg	agtaa	955

<210> 3

<211> 914

<212> ADN

<213> *Homo sapiens*

<400> 3

atggaacatt	ttgatgcac	acttagtacc	tatttcaagg	cattgctagg	ccctcgagat	60
actagagtaa	aaggatgggt	tcttctggac	aattatatac	ccacatttat	ctgctctgtc	120
atataattac	taattgtatg	gctgggacca	aaatacatga	ggaataaaca	gccattctct	180
tgccggggga	ttttagtgg	gtataacctt	ggactcacac	tgtgtctctc	gtatatgttc	240
tgtgagttag	taacaggagt	atgggaaggc	aaatacaact	tcttctgtca	gggcacacgc	300
accgcaggag	aatcagatat	gaagattatc	cgtgtcctct	ggtggtacta	cttctccaaa	360
ctcatagaat	ttatggacac	tttcttcttc	atcctgcgca	agaacaacca	ccagatcacg	420
gtcctgcacg	tctaccacca	tgccctgatg	ctgaacatct	ggtggtttgt	gatgaactgg	480
gtccctcgcg	gccactctta	ttttggtgcc	acacttaata	gcttcatcca	cgtcctcatg	540
tactcttact	atggtttgtc	gtcagtcctc	tccatgcgtc	catacctctg	gtggaagaag	600
tacatcactc	aggggcagct	gcttcagttt	gtgctgacaa	tcattccagac	cagctgcggg	660
gtcatctggc	cgtgcacatt	ccctcttgg	tgggtgtatt	tccagattgg	atacatgatt	720
tccttgattg	ctctcttcac	aaacttctac	attcagacct	acaacaagaa	aggggcctcc	780
cgaaggaaag	accacctgaa	ggaccaccag	aatgggtcca	tggctgctgt	gaatggacac	840
accaacagct	tttcacccct	ggaaaacaat	gtgaagccaa	ggaagctgcg	gaaggattga	900
agtcaaagaa	ttga					914

<210> 4
 <211> 867
 <212> ADN
 <213> *Caenorhabditis elegans*

<400> 4
 atggctcagc atccgctcgt tcaacggctt ctcgatgtca aattcgacac gaaacgattt 60
 gtggctattg ctactcatgg gccaaagaat ttccctgaag cagaaggtag caagttcttt 120
 gctgatcact ttgatgttac tattcaggct tcaatcctgt acatggtagt tgtgttcgga 180
 acaaaatggg tcatgcgtaa tegtcaacca ttccaattga ctattccact caacatctgg 240
 aatttcaccc tgcgcgcatt ttccatcgca ggagctgtca aaatgacccc agagttcttt 300
 ggaaccattg ccaacaaagg aattgtcgca tctactgca aagtgtttga ttccacgaaa 360
 ggagagaatg gatactgggt gtggctcttc atggcttcca aacttttoga acttgttgac 420
 accatcttct tggttctccg taaacgtcca ctcatgttcc ttactggta toaccatatt 480
 ctccaccatga tctacgctg gtactctcat ccattgaccc caggattcaa cagatacgga 540
 atttatctta actttgtcgt ccacgccttc atgtactctt actacttctt tgcgtcgatg 600
 aagattcgcg tgccaggatt catcgcccaa gctatcacat ctcttcaa atcggtcaattc 660
 atcatctctt gcgccgttct tgcctatctt gggtatctca tgcacttcac caatgccaac 720
 tgtgatttgc agccatcagt attcaagctc gcagttttca tggacacaac atacttggtc 780
 cttttcgtca acttcttctt ccaatcatat gttctccgcg gaggaaaaga caagtacaag 840
 gcagtgccaa agaagaagaa caactaa 867

<210> 5
 <211> 33
 <212> ADN
 <213> *Mortierella alpin*

<400> 5
gaattcaggc atggccgccc caatcttgga caa 33

<210> 6
 <211> 49
 <212> ADN
 <213> *Mortierella alpin*

<400> 6
gaattcaggc atctcatgga tccgccatgg ccgcccgaat cttggacaa 49

<210> 7
 <211> 317
 <212> PRT
 <213> *Mortierella alpin*

<400> 7

```

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

```

<210> 8

<211> 347

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 8

Met Asn Ser Leu Val Thr Gln Tyr Ala Ala Pro Leu Phe Glu Arg Tyr

```

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

```

<210> 9

<211> 345

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 9

```

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

```


<400> 12

```

cattaagcac ttgccccct gtgtatacgc ccatcactgc ctattacggtt attatttttg      60
gtggcagggtt ttgtttaagt aagtcgaaac catttaaatt aaatggcctt ttccaattgc      120
ataatttgggt ttttaacttca ctttcattga cgcttttatt gcttatgggt gaacaattag      180
tgccaattat tggtcagcac ggggtatact tcgctatctg taatattgggt gcttggactc      240
aaccgctcgt tacattatat tacatgaatt acattgtcaa gtttattgaa tttatagaca      300
cctttttctt ggtgctaaaa cataaaaaat tgacattttt gcatacttat caccatggcg      360
ctactgcctt attatgttac acccaattga tgggcaccac atctatttct tgggtcccta      420
tttcattgaa ccttgggtgtt cacgtgggta tgtattggta ctatttcttg gctgccagag      480
gcacaggggt ctgggtggaag gaatgggta ccagatttca aattatccaa tttgttttgg      540
atatcggttt catatatttt gctgtctacc aaaaagcagt tcacttgtat      590

```

<210> 13

<211> 278

<212> PRT

<213> *Caenorhabditis elegans*

<400> 13

```

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

```

<210> 14

<211> 293

<212> PRT

<213> Sequência Artificial

<220>

<223> Posição 293 é artificial ou desconhecida

<400> 14

```

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

```

<210> 15

<211> 238

<212> PRT

<213> *Caenorhabditis elegans*

<400> 15

```

Ser Leu Leu Thr Asn Gln Asp Glu Val Phe Pro His Ile Arg Ala Arg
 1          5          10          15
Arg Phe Ile Gln Glu His Phe Gly Leu Phe Val Gln Met Ala Ile Ala
 20          25          30
Tyr Val Ile Leu Val Phe Ser Ile Lys Arg Phe Met Arg Asp Arg Glu
 35          40          45
Pro Phe Gln Leu Thr Thr Ala Leu Arg Leu Trp Asn Phe Phe Leu Ser
 50          55          60
Val Phe Ser Ile Tyr Gly Ser Trp Thr Met Phe Pro Phe Met Val Gln
 65          70          75          80

```

```

Gln Ile Arg Leu Tyr Gly Leu Tyr Gly Cys Gly Cys Glu Ala Leu Ser
      85                      90                      95
Asn Leu Pro Ser Gln Ala Glu Tyr Trp Leu Phe Leu Thr Ile Leu Ser
      100                    105                    110
Lys Ala Val Glu Phe Val Asp Thr Phe Phe Leu Val Leu Arg Lys Lys
      115                    120                    125
Pro Leu Ile Phe Leu His Trp Tyr His His Met Ala Thr Phe Val Phe
      130                    135                    140
Phe Cys Ser Asn Tyr Pro Thr Pro Ser Ser Gln Ser Arg Val Gly Val
      145                    150                    155                    160
Ile Val Asn Leu Phe Val His Ala Phe Met Tyr Pro Tyr Tyr Phe Thr
      165                    170                    175
Arg Ser Met Asn Ile Lys Val Pro Ala Lys Ile Ser Met Ala Val Thr
      180                    185                    190
Val Leu Gln Leu Thr Gln Phe Met Cys Phe Ile Tyr Gly Cys Thr Leu
      195                    200                    205
Met Tyr Tyr Ser Leu Ala Thr Asn Gln Ala Arg Tyr Pro Ser Asn Thr
      210                    215                    220
Pro Ala Thr Leu Gln Cys Leu Ser Tyr Thr Leu His Leu Leu
      225                    230                    235

```

<210> 16

<211> 293

<212> PRT

<213> Sequência Artificial

<220>

<223> Posição 293 é artificial ou desconhecida

<400> 16

```

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

```

```

                245                250                255
Thr Tyr Asn Ala Lys Ala Lys Ala Ala Lys Glu Arg Gly Ser Asn Phe
                260                265                270
Thr Pro Lys Thr Val Lys Ser Gly Gly Ser Pro Lys Lys Pro Ser Lys
                275                280                285
Ser Lys His Ile Xaa
                290

```

<210> 17

<211> 101

<212> PRT

<213> *Caenorhabditis elegans*

<400> 17

```

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

```

<210> 18

<211> 115

<212> PRT

<213> *Caenorhabditis elegans*

<400> 18

```

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

```

<210> 19

<211> 272

<212> PRT

<213> Sequência Artificial

<220>

<223> Posição número 272 é artificial ou desconhecida

<400> 19

```

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

```

<210> 20

<211> 318

<212> PRT

<213> Sequência Artificial

<220>

<223> Posição Número 318 é artificial ou desconhecida

<400> 20

```

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

```

130	135	140
Leu Ala Asp Thr Val Phe	Leu Val Leu Lys Lys Lys	Pro Leu Glu Phe
145	150	155
Leu His Tyr Phe His His	Ser Met Thr Met Val Leu Cys Phe	Val Gln
165	170	175
Leu Gly Gly Tyr Thr Ser	Val Ser Trp Val Pro Ile Thr	Leu Asn Leu
180	185	190
Thr Val His Val Phe Met	Tyr Tyr Tyr Met Arg Ser	Ala Ala Gly
195	200	205
Val Arg Ile Trp Trp Lys	Gln Tyr Leu Thr Thr Leu	Gln Ile Val Gln
210	215	220
Phe Val Leu Asp Leu Gly	Phe Ile Tyr Phe Cys Ala	Tyr Thr Tyr Phe
225	230	235
Ala Phe Thr Tyr Phe Pro	Trp Ala Pro Asn Val Gly	Lys Cys Ala Gly
245	250	255
Thr Glu Gly Ala Ala Leu	Phe Gly Cys Gly Leu Leu	Ser Ser Tyr Leu
260	265	270
Leu Leu Phe Ile Asn Phe	Tyr Arg Ile Thr Tyr Asn	Ala Lys Ala Lys
275	280	285
Ala Ala Lys Glu Arg Gly	Ser Asn Phe Thr Pro Lys	Thr Val Lys Ser
290	295	300
Gly Gly Ser Pro Lys Lys	Pro Ser Lys Ser Lys	His Ile Xaa
305	310	315

<210> 21

<211> 178

<212> PRT

<213> *Homo sapiens*

<400> 21

Asn Leu Val Pro Ile Leu	Ala Arg Asn Gly Leu Phe	Tyr Ala Ile Cys
1	5	10
Asp Asp Gly Ala Trp Thr	Gln Arg Leu Glu Leu	Tyr Tyr Leu Asn
20	25	30
Tyr Leu Val Lys Tyr Trp	Glu Leu Ala Asp Thr	Val Phe Leu Val Leu
35	40	45
Lys Lys Lys Pro Leu Glu	Phe Leu His Tyr Phe	His His Ser Met Thr
50	55	60
Met Val Leu Cys Phe Val	Gln Leu Gly Gly Tyr	Thr Ser Val Ser Trp
65	70	75
Val Pro Ile Thr Leu Asn	Leu Thr Val His Val	Phe Met Tyr Tyr Tyr
85	90	95
Tyr Met Arg Ser Ala Ala	Gly Val Arg Ile Trp	Trp Lys Gln Tyr Leu
100	105	110
Thr Thr Leu Gln Ile Val	Gln Phe Val Leu Asp	Leu Gly Phe Ile Tyr
115	120	125
Phe Cys Ala Tyr Thr Tyr	Phe Ala Phe Thr Tyr	Phe Pro Trp Ala Pro
130	135	140
Asn Val Gly Lys Cys Ala	Gly Thr Glu Gly Ala	Ala Leu Phe Gly Cys
145	150	155
Gly Leu Leu Ser Ser Tyr	Leu Leu Leu Phe Ile	Asn Phe Tyr Arg Ile
165	170	175
Thr Tyr		

<210> 22

<211> 147

<212> PRT

<213> *Homo sapiens*

<400> 22

```

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

```

<210> 23

<211> 280

<212> PRT

<213> Sequência Artificial

<220>

<223> Posição 280 é artificial ou desconhecida

<400> 23

```

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

```

				245					250					255		
Ser	Asn	Phe	Thr	Pro	Lys	Thr	Val	Lys	Ser	Gly	Gly	Ser	Pro	Lys	Lys	
			260					265					270			
Pro	Ser	Lys	Ser	Lys	His	Ile	Xaa									
		275					280									

<210> 24

<211> 283

<212> PRT

<213> Sequência Artificial

 $\langle 220 \rangle$

<223> Posição 282 é artificial ou desconhecida

<400> 24

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

<210> 25

<211> 446

<212> PRT

<213> *Mortierella alpin*

<400> 25

Met Gly Thr Asp Gln Gly Lys Thr Phe Thr Trp Glu Glu Leu Ala Ala
1 5 10 15
His Asn Thr Lys Asp Asp Leu Leu Leu Ala Ile Arg Gly Arg Val Tyr

```
<210> 26
<211> 318
<212> PRT
<213> Mortierella alpin
```

<400> 26
Met Glu Ser Ile Ala Pro Phe Leu Pro Ser Lys Met Pro Gln Asp Leu

```

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

```

<210> 27

<211> 279

<212> PRT

<213> *Mortierella alpin*

<400> 27

```

Val Ala Gln Ala Glu Lys Tyr Ile Pro Thr Ile Val His His Thr Arg
      1           5           10           15
Gly Phe Leu Val Ala Val Glu Ser Pro Leu Ala Arg Glu Leu Pro Leu
      20           25           30
Met Asn Pro Phe His Val Leu Leu Ile Val Leu Ala Tyr Leu Val Thr
      35           40           45
Val Phe Val Gly Met Gln Ile Met Lys Asn Phe Glu Arg Phe Glu Val
      50           55           60
Lys Thr Phe Ser Leu Leu His Asn Phe Cys Leu Val Ser Ile Ser Ala
65           70           75           80
Tyr Met Cys Gly Gly Ile Leu Tyr Glu Ala Tyr Gln Ala Asn Tyr Gly
      85           90           95
Leu Phe Glu Asn Ala Ala Asp His Thr Phe Lys Gly Leu Pro Met Ala
      100          105          110
Lys Met Ile Trp Leu Phe Tyr Phe Ser Lys Ile Met Glu Phe Val Asp

```

```

      115      120      125
Thr Met Ile Met Val Leu Lys Lys Asn Asn Arg Gln Ile Ser Phe Leu
130      135      140
His Val Tyr His His Ser Ser Ile Phe Thr Ile Trp Trp Leu Val Thr
145      150      155      160
Phe Val Ala Pro Asn Gly Glu Ala Tyr Phe Ser Ala Ala Leu Asn Ser
      165      170      175
Phe Ile His Val Ile Met Tyr Gly Tyr Tyr Phe Leu Ser Ala Leu Gly
      180      185      190
Phe Lys Gln Val Ser Phe Ile Lys Phe Tyr Ile Thr Arg Ser Gln Met
195      200      205
Thr Gln Phe Cys Met Met Ser Val Gln Ser Ser Trp Asp Met Tyr Ala
210      215      220
Met Lys Val Leu Gly Arg Pro Gly Tyr Pro Phe Ile Thr Ala Leu
225      230      235      240
Leu Trp Phe Tyr Met Trp Thr Met Leu Gly Leu Phe Tyr Asn Phe Tyr
      245      250      255
Arg Lys Asn Ala Lys Leu Ala Lys Gln Ala Lys Ala Asp Ala Ala Lys
260      265      270
Glu Lys Ala Arg Lys Leu Gln
275

```

<210> 28

<211> 301

<212> PRT

<213> Sequência Artificial

<220>

<223> Posição número 301 é artificial ou desconhecida

<400> 28

```

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

```

```

Glu Gly Ala Ala Leu Phe Gly Cys Gly Leu Leu Ser Ser Tyr Leu Leu
      245      250      255
Leu Phe Ile Asn Phe Tyr Arg Ile Thr Tyr Asn Ala Lys Ala Lys Ala
      260      265      270
Ala Lys Glu Arg Gly Ser Asn Phe Thr Pro Lys Thr Val Lys Ser Gly
      275      280      285
Gly Ser Pro Lys Lys Pro Ser Lys Ser Lys His Ile Xaa
      290      295      300

```

<210> 29

<211> 289

<212> PRT

<213> *Mortierella alpin*

<400> 29

```

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

```

<210> 30

<211> 292

<212> PRT

<213> Sequência Artificial

<220>

<223> Posição 292 é artificial ou desconhecida

<400> 30

```

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

```

<210> 31

<211> 291

<212> PRT

<213> *Mortierella alpin*

<400> 31

```

Gln Ala Tyr Glu Leu Val Thr Gly Lys Ser Ile Asp Ser Phe Val Phe
 1          5          10          15
Gln Glu Gly Val Thr Pro Leu Ser Thr Gln Arg Glu Val Ala Met Trp
 20          25          30
Thr Ile Thr Tyr Phe Val Val Ile Phe Gly Gly Arg Gln Ile Met Lys
 35          40          45
Ser Gln Asp Ala Phe Lys Leu Lys Pro Leu Phe Ile Leu His Asn Phe
 50          55          60
Leu Leu Thr Ile Ala Ser Gly Ser Leu Leu Leu Leu Phe Ile Glu Asn
 65          70          75          80
Leu Val Pro Ile Leu Ala Arg Asn Gly Leu Phe Tyr Ala Ile Cys Asp
 85          90          95
Asp Gly Ala Trp Thr Gln Arg Leu Glu Leu Leu Tyr Tyr Leu Asn Tyr
100          105          110

```

```

Leu Val Lys Tyr Trp Glu Leu Ala Asp Thr Val Phe Leu Val Leu Lys
    115                      120                      125
Lys Lys Pro Leu Glu Phe Leu His Tyr Phe His His Ser Met Thr Met
    130                      135                      140
Val Leu Cys Phe Val Gln Leu Gly Gly Tyr Thr Ser Val Ser Trp Val
    145                      150                      155                      160
Pro Ile Thr Leu Asn Leu Thr Val His Val Phe Met Tyr Tyr Tyr Tyr
    165                      170                      175
Met Arg Ser Ala Ala Gly Val Arg Ile Trp Trp Lys Gln Tyr Leu Thr
    180                      185                      190
Thr Leu Gln Ile Val Gln Phe Val Leu Asp Leu Gly Phe Ile Tyr Phe
    195                      200                      205
Cys Ala Tyr Thr Tyr Phe Ala Phe Thr Tyr Phe Pro Trp Ala Pro Asn
    210                      215                      220
Val Gly Lys Cys Ala Gly Thr Glu Gly Ala Ala Leu Phe Gly Cys Gly
    225                      230                      235                      240
Leu Leu Ser Ser Tyr Leu Leu Leu Phe Ile Asn Phe Tyr Arg Ile Thr
    245                      250                      255
Tyr Asn Ala Lys Ala Lys Ala Ala Lys Glu Arg Gly Ser Asn Phe Thr
    260                      265                      270
Pro Lys Thr Val Lys Ser Gly Gly Ser Pro Lys Lys Pro Ser Lys Ser
    275                      280                      285
Lys His Ile
    290

```

<210> 32

<211> 276

<212> PRT

<213> Sequência Artificial

<220>

<223> Posição 276 é artificial ou desconhecida

<400> 32

```

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

```

```

      210              215              220
Tyr Pro Val Ile Ile His Leu Ile Trp Met Tyr Gly Thr Ile Phe Phe
225              230              235              240
Met Leu Phe Ser Asn Phe Trp Tyr His Ser Tyr Thr Lys Gly Lys Arg
      245              250              255
Leu Pro Arg Ala Leu Gln Gln Asn Gly Ala Pro Gly Ile Ala Lys Val
      260              265              270
Lys Ala Asn Xaa
      275

```

<210> 33

<211> 219

<212> PRT

<213> *Mortierella alpin*

<400> 33

```

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

```

<210> 34

<211> 202

<212> PRT

<213> *Mortierella alpin*

<400> 34

```

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

```

```

Ile Lys Ile Ala Pro Gly Gly Met Gly Ser Phe His Ala Met Ile Asn
      85      90      95
Ser Ser Val His Val Val Met Tyr Leu Tyr Tyr Gly Leu Ser Ala Leu
      100      105      110
Gly Pro Val Ala Gln Pro Tyr Leu Trp Trp Lys Lys His Met Thr Ala
      115      120      125
Ile Gln Leu Ile Gln Phe Val Leu Val Ser Leu His Ile Ser Gln Tyr
      130      135      140
Tyr Phe Met Pro Ser Cys Asn Tyr Gln Tyr Pro Val Ile Ile His Leu
145      150      155      160
Ile Trp Met Thr Ile Phe Phe Ile Leu Phe Ser Asn Phe Trp Tyr His
      165      170      175
Ser Tyr Thr Lys Gly Lys Arg Leu Pro Arg Ala Val Gln Gln Asn Gly
      180      185      190
Ala Pro Ala Thr Thr Lys Val Lys Ala Asn
      195      200

```

<210> 35

<211> 174

<212> PRT

<213> *Mortierella alpin*

<400> 35

```

Tyr Glu Leu Val Thr Gly Lys Ser Ile Asp Ser Phe Val Phe Gln Glu
 1      5      10      15
Gly Val Thr Pro Leu Ser Thr Gln Arg Glu Val Ala Met Trp Thr Ile
      20      25      30
Thr Tyr Phe Val Val Ile Phe Gly Gly Arg Gln Ile Met Lys Ser Gln
      35      40      45
Asp Ala Phe Lys Leu Lys Pro Leu Phe Ile Leu His Asn Phe Leu Leu
50      55      60
Thr Ile Ala Ser Gly Ser Leu Leu Leu Phe Ile Glu Asn Leu Val
65      70      75      80
Pro Ile Leu Ala Arg Asn Gly Leu Phe Tyr Ala Ile Cys Asp Asp Gly
      85      90      95
Ala Trp Thr Gln Arg Leu Glu Leu Leu Tyr Tyr Leu Asn Tyr Leu Val
      100      105      110
Lys Tyr Trp Glu Leu Ala Asp Thr Val Phe Leu Val Leu Lys Lys Lys
      115      120      125
Pro Leu Glu Phe Leu His Tyr Phe His His Ser Met Thr Met Val Leu
130      135      140
Cys Phe Val Gln Leu Gly Gly Tyr Thr Ser Val Ser Trp Val Pro Ile
145      150      155      160
Thr Leu Asn Leu Thr Val His Val Phe Met Tyr Tyr Tyr Tyr
      165      170

```

<210> 36

<211> 145

<212> PRT

<213> *Mortierella alpin*

<400> 36

```

Asn Ala Phe Leu Asp Asn Met Phe Gly Pro Arg Asp Ser Arg Val Arg
 1      5      10      15
Gly Trp Phe Leu Leu Asp Ser Tyr Leu Pro Thr Phe Ile Leu Thr Ile
      20      25      30
Thr Tyr Leu Leu Ser Ile Trp Leu Gly Asn Lys Tyr Met Lys Asn Arg
      35      40      45
Pro Ala Leu Ser Leu Arg Gly Ile Leu Thr Leu Tyr Asn Leu Ala Ile
50      55      60
Thr Leu Leu Ser Ala Tyr Met Leu Val Glu Leu Ile Leu Ser Ser Trp

```

65					70					75					80
Glu	Gly	Gly	Tyr	Asn	Leu	Gln	Cys	Gln	Asn	Leu	Asp	Ser	Ala	Gly	Glu
				85					90					95	
Gly	Asp	Val	Arg	Val	Ala	Lys	Val	Leu	Val	Trp	Tyr	Tyr	Phe	Ser	Lys
			100					105					110		
Leu	Val	Glu	Phe	Leu	Asp	Thr	Ile	Phe	Phe	Val	Leu	Arg	Lys	Lys	Ala
		115					120					125			
Asn	Gln	Ile	Thr	Phe	Leu	His	Val	Tyr	His	His	Ala	Ser	Met	Phe	Asn
	130					135					140				
Ile															
145															

<210> 37

<211> 278

<212> PRT

<213> *Caenorhabditis elegans*

<400> 37

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

<210> 38

<211> 293

<212> PRT

<213> *Caenorhabditis elegans*

<400> 38

```

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

```

<210> 39

<211> 174

<212> PRT

<213> *Caenorhabditis elegans*

<400> 39

```

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

```

```

Pro Leu Glu Phe Leu His Tyr Phe His His Ser Met Thr Met Val Leu
 130                      135                      140
Cys Phe Val Gln Leu Gly Gly Tyr Thr Ser Val Ser Trp Val Pro Ile
145                      150                      155                      160
Thr Leu Asn Leu Thr Val His Val Phe Met Tyr Tyr Tyr Tyr
                      165                      170

```

<210> 40

<211> 141

<212> PRT

<213> Sequência Artificial

<220>

<223> Posição Número 141 é artificial ou desconhecida

<400> 40

```

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

```

<210> 41

<211> 238

<212> PRT

<213> *Caenorhabditis elegans*

<400> 41

```

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

```

```

Lys Phe Tyr Ile Thr Arg Ser Gln Met Thr Gln Phe Cys Met Met Ser
      165      170      175
Val Gln Ser Ser Trp Asp Met Tyr Ala Met Lys Val Leu Gly Arg Pro
      180      185      190
Gly Tyr Pro Phe Phe Ile Thr Ala Leu Leu Trp Phe Tyr Met Trp Thr
      195      200      205
Met Leu Gly Leu Phe Tyr Asn Phe Tyr Arg Lys Asn Ala Lys Leu Ala
      210      215      220
Lys Gln Ala Lys Ala Asp Ala Ala Lys Glu Lys Ala Arg Lys
      225      230      235

```

<210> 42

<211> 144

<212> PRT

<213> *Caenorhabditis elegans*

<400> 42

```

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

```

<210> 43

<211> 241

<212> PRT

<213> *Caenorhabditis elegans*

<400> 43

```

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

```

```

145          150          155          160
Ile Lys Phe Tyr Ile Thr Arg Ser Gln Met Thr Gln Phe Cys Met Met
          165          170          175
Ser Val Gln Ser Ser Trp Asp Met Tyr Ala Met Lys Val Leu Gly Arg
          180          185          190
Pro Gly Tyr Pro Phe Phe Ile Thr Ala Leu Leu Trp Phe Tyr Met Trp
          195          200          205
Thr Met Leu Gly Leu Phe Tyr Asn Phe Tyr Arg Lys Asn Ala Lys Leu
          210          215          220
Ala Lys Gln Ala Lys Ala Asp Ala Ala Lys Glu Lys Ala Arg Lys Leu
225          230          235          240
Gln

```

<210> 44

<211> 217

<212> PRT

<213> *Caenorhabditis elegans*

<400> 44

```

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

```

<210> 45

<211> 178

<212> PRT

<213> *Caenorhabditis elegans*

<400> 45

```

Pro Thr Ile Val His His Thr Arg Gly Phe Leu Val Ala Val Glu Ser
1          5          10          15
Pro Leu Ala Arg Glu Leu Pro Leu Met Asn Pro Phe His Val Leu Leu
          20          25          30
Ile Val Leu Ala Tyr Leu Val Thr Val Phe Val Gly Met Gln Ile Met
          35          40          45

```

```

Lys Asn Phe Glu Arg Phe Glu Val Lys Thr Phe Ser Leu Leu His Asn
 50          55          60
Phe Cys Leu Val Ser Ile Ser Ala Tyr Met Cys Gly Gly Ile Leu Tyr
65          70          75          80
Glu Ala Tyr Gln Ala Asn Tyr Gly Leu Phe Glu Asn Ala Ala Asp His
          85          90          95
Thr Phe Lys Gly Leu Pro Met Ala Lys Met Ile Trp Leu Phe Tyr Phe
          100          105          110
Ser Lys Ile Met Glu Phe Val Asp Thr Met Ile Met Val Leu Lys Lys
          115          120          125
Asn Asn Arg Gln Ile Ser Phe Leu His Val Tyr His His Ser Ser Ile
130          135          140
Phe Thr Ile Trp Trp Leu Val Thr Phe Val Ala Pro Asn Gly Glu Ala
145          150          155          160
Tyr Phe Ser Ala Ala Leu Asn Ser Phe Ile His Val Ile Met Tyr Gly
          165          170          175
Tyr Tyr

```

<210> 46

<211> 148

<212> PRT

<213> *Caenorhabditis elegans*

<400> 46

```

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

```

<210> 47

<211> 302

<212> PRT

<213> *Caenorhabditis elegans*

<400> 47

```

Phe Met Asp Leu Ala Thr Ala Ile Gly Val Arg Ala Ala Pro Tyr Val
 1          5          10          15
Asp Pro Leu Glu Ala Ala Leu Val Ala Gln Ala Glu Lys Tyr Ile Pro
          20          25          30
Thr Ile Val His His Thr Arg Gly Phe Leu Val Ala Val Glu Ser Pro
          35          40          45
Leu Ala Arg Glu Leu Pro Leu Met Asn Pro Phe His Val Leu Leu Ile
          50          55          60
Val Leu Ala Tyr Leu Val Thr Val Phe Val Gly Met Gln Ile Met Lys

```

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

<210> 48

<211> 271

<212> PRT

<213> *Caenorhabditis elegans*

<400> 48

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

195					200					205					
Leu	Gln	Met	Val	Leu	Gly	Thr	Ile	Phe	Gly	Ile	Leu	Asn	Tyr	Ile	Trp
210					215					220					
Arg	Gln	Glu	Lys	Gly	Cys	His	Thr	Thr	Thr	Glu	His	Phe	Phe	Trp	Ser
225					230					235					240
Phe	Met	Leu	Tyr	Gly	Thr	Tyr	Phe	Ile	Leu	Phe	Ala	His	Phe	Phe	His
245					250					255					
Arg	Ala	Tyr	Leu	Arg	Pro	Lys	Gly	Lys	Val	Ala	Ser	Lys	Ser	Gln	
260					265					270					

<210> 49

<211> 265

<212> PRT

<213> *Caenorhabditis elegans*

<400> 49

Thr 1	Arg	Gly	Phe	Leu 5	Val	Ala	Val	Glu	Ser 10	Pro	Leu	Ala	Arg	Glu 15	Leu
Pro	Leu	Met	Asn 20	Pro	Phe	His	Val	Leu 25	Leu	Ile	Val	Leu	Ala 30	Tyr	Leu
Val	Thr	Val 35	Phe	Val	Gly	Met	Gln 40	Ile	Met	Lys	Asn	Phe 45	Glu	Arg	Phe
Glu	Val	Lys	Thr	Phe	Ser	Leu 55	Leu	His	Asn	Phe	Cys 60	Leu	Val	Ser	Ile
Ser 65	Ala	Tyr	Met	Cys	Gly 70	Ile	Leu	Tyr	Glu	Ala	Tyr	Gln	Ala 80	Asn	
Tyr	Gly	Leu	Phe 85	Glu	Asn	Ala	Ala	Asp 90	His	Thr	Phe	Lys	Gly 95	Leu	Pro
Met	Ala	Lys	Met 100	Ile	Trp	Leu	Phe	Tyr 105	Phe	Ser	Lys	Ile	Met 110	Glu	Phe
Val	Asp	Thr 115	Met	Ile	Met	Val	Leu 120	Lys	Lys	Asn	Asn	Arg 125	Gln	Ile	Ser
Phe	Leu	His	Val	Tyr	His	His 135	Ser	Ser	Ile	Phe	Thr 140	Ile	Trp	Trp	Leu
Val 145	Thr	Phe	Val	Ala	Pro	Asn 150	Gly	Glu	Ala	Tyr	Phe	Ser	Ala	Ala	Leu
Asn	Ser	Phe	Ile 165	His	Val	Ile	Met	Tyr	Gly 170	Tyr	Tyr	Phe	Leu 175	Ser	Ala
Leu	Gly	Phe	Lys 180	Gln	Val	Ser	Phe	Ile 185	Lys	Phe	Tyr	Ile	Thr 190	Arg	Ser
Gln	Met	Thr 195	Gln	Phe	Cys	Met	Met 200	Ser	Val	Gln	Ser	Ser 205	Trp	Asp	Met
Tyr	Ala 210	Met	Lys	Val	Leu	Gly 215	Arg	Pro	Gly	Tyr	Pro	Phe 220	Phe	Ile	Thr
Ala 225	Leu	Leu	Trp	Phe	Tyr	Met 230	Trp	Thr	Met	Leu	Gly	Leu	Phe	Tyr	Asn
Phe	Tyr	Arg	Lys 245	Asn	Ala	Lys	Leu	Ala	Lys 250	Gln	Ala	Lys	Ala 255	Asp	Ala
Ala	Lys	Glu	Lys 260	Ala	Arg	Lys	Leu	Gln 265							

<210> 50

<211> 265

<212> PRT

<213> *Caenorhabditis elegans*

<400> 50

Ala Thr His Gly Pro Lys Asn Phe Pro Asp Ala Glu Gly Arg Lys Phe
1 5 10 15
Phe Ala Asp His Phe Asp Val Thr Ile Gln Ala Ser Ile Leu Tyr Met
20 25 30

```

Val Val Val Phe Gly Thr Lys Trp Phe Met Arg Asn Arg Gln Pro Phe
   35           40           45
Gln Leu Thr Ile Pro Leu Asn Ile Trp Asn Phe Ile Leu Ala Ala Phe
   50           55           60
Ser Ile Ala Gly Ala Val Lys Met Thr Pro Glu Phe Phe Gly Thr Ile
  65           70           75           80
Ala Asn Lys Gly Ile Val Ala Ser Tyr Cys Lys Val Phe Asp Phe Thr
           85           90           95
Lys Gly Glu Asn Gly Tyr Trp Val Trp Leu Phe Met Ala Ser Lys Leu
   100           105           110
Phe Glu Leu Val Asp Thr Ile Phe Leu Val Leu Arg Lys Arg Pro Leu
   115           120           125
Met Phe Leu His Trp Tyr His His Ile Leu Thr Met Ile Tyr Ala Trp
   130           135           140
Tyr Ser His Pro Leu Thr Pro Gly Phe Asn Arg Tyr Gly Ile Tyr Leu
  145           150           155           160
Asn Phe Val Val His Ala Phe Met Tyr Ser Tyr Tyr Phe Leu Arg Ser
           165           170           175
Met Lys Ile Arg Val Pro Gly Phe Ile Ala Gln Ala Ile Thr Ser Leu
   180           185           190
Gln Ile Val Gln Phe Ile Ile Ser Cys Ala Val Leu Ala His Leu Gly
   195           200           205
Tyr Leu Met His Phe Thr Asn Ala Asn Cys Asp Phe Glu Pro Ser Val
   210           215           220
Phe Lys Leu Ala Val Phe Met Asp Thr Thr Tyr Leu Ala Leu Phe Val
  225           230           235           240
Asn Phe Phe Leu Gln Ser Tyr Val Leu Arg Gly Gly Lys Asp Lys Tyr
           245           250           255
Lys Ala Val Pro Lys Lys Lys Asn Asn
           260           265

```

<210> 51

<211> 317

<212> PRT

<213> *Caenorhabditis elegans*

<400> 51

```

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

```

```

Thr Val His Val Phe Met Tyr Tyr Tyr Tyr Met Arg Ser Ala Ala Gly
      195      200      205
Val Arg Ile Trp Trp Lys Gln Tyr Leu Thr Thr Leu Gln Ile Val Gln
      210      215      220
Phe Val Leu Asp Leu Gly Phe Ile Tyr Phe Cys Ala Tyr Thr Tyr Phe
      225      230      235      240
Ala Phe Thr Tyr Phe Pro Trp Ala Pro Asn Val Gly Lys Cys Ala Gly
      245      250      255
Thr Glu Gly Ala Ala Leu Phe Gly Cys Gly Leu Leu Ser Ser Tyr Leu
      260      265      270
Leu Leu Phe Ile Asn Phe Tyr Arg Ile Thr Tyr Asn Ala Lys Ala Lys
      275      280      285
Ala Ala Lys Glu Arg Gly Ser Asn Phe Thr Pro Lys Thr Val Lys Ser
      290      295      300
Gly Gly Ser Pro Lys Lys Pro Ser Lys Ser Lys His Ile
      305      310      315

```

<210> 52

<211> 288

<212> PRT

<213> *Caenorhabditis elegans*

<400> 52

```

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

```

<210> 53

<211> 282

<212> PRT

<213> *Drosophila melanogaster*

<400> 53

```

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

```

<210> 54

<211> 278

<212> PRT

<213> *Drosophila melanogaster*

<400> 54

```

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

```

```

      100      105      110
Glu Gln Asp Arg Val Cys Gly Phe Trp Thr Trp Leu Phe Val Leu Ser
      115      120      125
Lys Leu Pro Glu Leu Gly Asp Thr Ile Phe Ile Val Leu Arg Lys Gln
      130      135      140
Pro Leu Ile Phe Leu His Trp Tyr His His Ile Thr Val Leu Ile Tyr
145      150      155      160
Ser Trp Phe Ser Tyr Thr Glu Tyr Thr Ser Ser Ala Arg Trp Phe Ile
      165      170      175
Val Met Asn Tyr Cys Val His Ser Val Met Tyr Ser Tyr Tyr Ala Leu
      180      185      190
Lys Ala Ala Arg Phe Asn Pro Pro Arg Phe Ile Ser Met Ile Ile Thr
      195      200      205
Ser Leu Gln Leu Ala Gln Met Ile Ile Gly Cys Ala Ile Asn Val Trp
      210      215      220
Ala Asn Gly Phe Leu Lys Thr His Gly Thr Xaa Ser Cys His Ile Ser
225      230      235      240
Gln Arg Asn Ile Asn Leu Ser Ile Ala Met Tyr Ser Ser Tyr Phe Val
      245      250      255
Leu Phe Ala Arg Phe Phe Tyr Lys Ala Tyr Leu Ala Pro Gly Gly His
      260      265      270
Lys Ser Arg Arg Met Ala
      275

```

<210> 55

<211> 286

<212> PRT

<213> *Drosophila melanogaster*

<400> 55

```

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

```

				245					250				255				
Lys	Ala	Ala	Lys	Glu	Arg	Gly	Ser	Asn	Phe	Thr	Pro	Lys	Thr	Val	Lys		
			260					265					270				
Ser	Gly	Gly	Ser	Pro	Lys	Lys	Pro	Ser	Lys	Ser	Lys	His	Ile				
		275					280					285					

<210> 56

<211> 261

<212> PRT

<213> *Drosophila melanogaster*

<400> 56

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

<210> 57

<211> 299

<212> PRT

<213> Alongase Humana

<400> 57

Met	Glu	His	Phe	Asp	Ala	Ser	Leu	Ser	Thr	Tyr	Phe	Lys	Ala	Leu	Leu		
1				5					10				15				
Gly	Pro	Arg	Asp	Thr	Arg	Val	Lys	Gly	Trp	Phe	Leu	Leu	Asp	Asn	Tyr		
			20					25					30				
Ile	Pro	Thr	Phe	Ile	Cys	Ser	Val	Ile	Tyr	Leu	Leu	Ile	Val	Trp	Leu		
	35					40						45					
Gly	Pro	Lys	Tyr	Met	Arg	Asn	Lys	Gln	Pro	Phe	Ser	Cys	Arg	Gly	Ile		
	50					55					60						

```

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

```

<210> 58

<211> 288

<212> PRT

<213> *Caenorhabditis elegans*

<400> 58

```

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

```

```

Ser Tyr Tyr Phe Leu Arg Ser Met Lys Ile Arg Val Pro Gly Phe Ile
    195                200                205
Ala Gln Ala Ile Thr Ser Leu Gln Ile Val Gln Phe Ile Ile Ser Cys
    210                215                220
Ala Val Leu Ala His Leu Gly Tyr Leu Met His Phe Thr Asn Ala Asn
    225                230                235                240
Cys Asp Phe Glu Pro Ser Val Phe Lys Leu Ala Val Phe Met Asp Thr
    245                250                255
Thr Tyr Leu Ala Leu Phe Val Asn Phe Phe Leu Gln Ser Tyr Val Leu
    260                265                270
Arg Gly Gly Lys Asp Lys Tyr Lys Ala Val Pro Lys Lys Lys Asn Asn
    275                280                285

```

<210> 59

<211> 798

<212> ADN

<213> Alongase Humana

<400> 59

```

atgaacatgt cagtgttgac tttaacaagaa tatgaattcg aaaagcagtt caacgagaat      60
gaagccatcc aatggatgca ggaaaactgg aagaaatctt tcctgttttc tgctctgtat      120
gctgccttta tattcgggtgg tcggcaccta atgaataaac gagcaaagtt tgaactgagg      180
aagccattag tgctctggtc tctgaccctt gcagtcctca gtatattcgg tgctcttcga      240
actggtgctt atatggtgta cattttgatg accaaaaggcc tgaagcagtc agtttgtgac      300
caggggtttt acaatggacc tgtcagcaaa ttctgggctt atgcatttgt gctaagcaaa      360
gcacccgaac taggagatac aatattcatt attctgagga agcagaagct gatcttcctg      420
cactgggtatc accacatcac tgtgctcctg tactcttggt actoctacaa agacatgggt      480
gccgggggag gttggttcat gactatgaac tatggcgtgc acgccgtgat gtactcttac      540
tatgccttgc gggcggcagg tttccgagtc tcccgggaagt ttgccatgtt catcaccttg      600
tcccagatca ctcagatgct gatgggctgt gtgggttaact acctgggtct ctgctggatg      660
cagcatgacc agtgtcactc tcactttcag aacatcttct ggctctcact catgtacctc      720
agctaccttg tgctcttctg ccattttctt tttgaggcct acatcggcaa aatgaggaaa      780
acaacgaaag ctgaatag

```

<210> 60

<211> 265

<212> PRT

<213> Alongase Humana

<400> 60

```

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

```

```

          165          170          175
Met Tyr Ser Tyr Tyr Ala Leu Arg Ala Ala Gly Phe Arg Val Ser Arg
          180          185          190
Lys Phe Ala Met Phe Ile Thr Leu Ser Gln Ile Thr Gln Met Leu Met
          195          200          205
Gly Cys Val Val Asn Tyr Leu Val Phe Cys Trp Met Gln His Asp Gln
          210          215          220
Cys His Ser His Phe Gln Asn Ile Phe Trp Ser Ser Leu Met Tyr Leu
          225          230          235          240
Ser Tyr Leu Val Leu Phe Cys His Phe Phe Phe Glu Ala Tyr Ile Gly
          245          250          255
Lys Met Arg Lys Thr Thr Lys Ala Glu
          260          265

```

<210> 61

<211> 82

<212> PRT

<213> Desconhecido

<220>

<223> Descrição de sequência desconhecida: KCS de jojoba; Acesso no Genbank N.º U37088

<400> 61

```

Ala Thr Leu Pro Asn Phe Lys Ser Ser Ile Asn Leu His His Val Lys
 1          5          10          15
Leu Gly Tyr His Tyr Leu Ile Ser Asn Ala Leu Phe Leu Val Phe Ile
          20          25          30
Pro Leu Leu Gly Leu Ala Ser Ala His Leu Ser Ser Phe Ser Ala His
          35          40          45
Asp Ser Leu Leu Phe Asp Leu Arg Arg Asn Leu Leu Pro Val Val
          50          55          60
Val Cys Ser Phe Leu Phe Val Leu Leu Ala Thr Leu His Phe Leu Thr
          65          70          75          80
Arg Pro

```

<210> 62

<211> 80

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 62

```

Ser Thr Leu Pro Pro Val Leu Tyr Ala Ile Thr Ala Tyr Tyr Val Ile
 1          5          10          15
Ile Phe Gly Gly Arg Phe Leu Leu Ser Lys Ser Lys Pro Phe Lys Leu
          20          25          30
Asn Gly Leu Phe Gln Leu His Asn Leu Val Leu Thr Ser Leu Ser Leu
          35          40          45
Thr Leu Leu Leu Leu Met Val Glu Gln Leu Val Pro Ile Ile Val Gln
          50          55          60
His Gly Leu Tyr Phe Ala Ile Cys Asn Ile Gly Ala Trp Thr Gln Pro
          65          70          75          80

```

<210> 63

<211> 18

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador R0352

<400> 63

acgcgtacgt aaagcttg 18

<210> 64

<211> 36

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador R0514

<400> 64

ggctatggat ccatgaattc actcgttact caatat 36

<210> 65

<211> 33

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador R0515

<400> 65

cctgccaaagc ttttaccttt ttctttctgtg ttg 33

<210> 66

<211> 21

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador R0541

<400> 66

gattactagc agctgtaata c 21

<210> 67

<211> 21

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador R0540

<400> 67

gtgaatgtaa cggtgacata a 21

<210> 68

<211> 28

<212> PRT

<213> *Homo sapiens*

<400> 68

Asp	Thr	Ile	Phe	Ile	Ile	Leu	Arg	Lys	Gln	Lys	Leu	Ile	Phe	Leu	His
1				5					10					15	
Trp	Tyr	His	His	Ile	Thr	Val	Leu	Leu	Tyr	Ser	Trp				
			20					25							

<210> 69

<211> 29

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador R0719

<400> 69

ggttctccca tggaacattt tgatgcatc 29

<210> 70

<211> 30

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador R0720

<400> 70

ggtttcaaag ctttgacttc aatccttccg 30

<210> 71

<211> 36

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador R0738

<400> 71

aatcaggaat tcatggctca gcatccgctc gttcaa 36

<210> 72

<211> 36

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador R0739

<400> 72

ccgcttgctcg acttagttgt tcttcttctt tggcac 36

<210> 73

<211> 30

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador RP735

<400> 73

cctcctgaat tccaacacta ttcagctttc 30

<210> 74

<211> 20

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador R073

<400> 74

taatacgact cactataggg 20

<210> 75

<211> 18

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador R0728

<400> 75

gagactttga gcggttcg 18

<210> 76

<211> 20

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador R0730

<400> 76

tctctgctgc gttgaactcg 20

<210> 77

<211> 19

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador R0729

<400> 77

aaagctcttg acctcgaac 19

<210> 78

<211> 20

<212> ADN

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: iniciador R0731

<400> 78

aacttgatga acgacacgtg 20

Lisboa, 2008-06-27

REIVINDICAÇÕES

1. Sequência nucleotídica isolada codificando uma alongase funcionalmente activa possuindo identidade com pelo menos cerca de 50% da sequência nucleotídica mostrada na Figura 6, ou o seu complemento.

2. Sequência nucleotídica isolada de acordo com a reivindicação 1 em que a referida sequência é representada pela sequência mostrada na Figura 6.

3. Proteína purificada codificada pela referida sequência nucleotídica de acordo com as reivindicações 1 ou 2.

4. Método de produção de uma enzima alongase compreendendo os passos de:

- a) isolamento de uma sequência nucleotídica representada pela Figura 6;
- b) construção de um vector compreendendo: i) a referida sequência nucleotídica isolada operativamente ligada a ii) um promotor;
- c) introdução do referido vector numa célula hospedeira durante um tempo e sob condições suficientes para a expressão da referida enzima alongase.

5. Vector compreendendo: a) uma sequência nucleotídica tal como representada na Figura 6 operativamente ligada a b) um promotor.

6. Célula hospedeira compreendendo o referido vector de acordo com a reivindicação 5.

7. Célula vegetal, planta ou tecido vegetal compreendendo o referido vector de acordo com a reivindicação 5, em que a expressão da referida sequência nucleotídica do referido vector resulta na produção de pelo menos um ácido gordo seleccionado de entre o grupo que consiste num ácido gordo mono-insaturado e um ácido gordo poli-insaturado, pela referida célula vegetal, planta ou tecido vegetal.

8. Planta transgénica compreendendo o referido vector de acordo com a reivindicação 5, em que a expressão da referida sequência nucleotídica do referido vector resulta na produção de um ácido gordo poli-insaturado, em sementes da referida planta transgénica.

9. Mamífero não humano transgénico cujo genoma compreende uma sequência de ADN tal como definida na reivindicação 1 operativamente ligada a um promotor.

10. Mamífero não humano transgénico de acordo com a reivindicação 9 em que a referida sequência de ADN é representada pela Figura 6.

11. Método para produção de um ácido gordo poli-insaturado compreendendo os passos de:

- a) isolamento da referida sequência nucleotídica representada pela Figura 6;
- b) construção de um vector compreendendo a referida sequência nucleotídica isolada;
- c) introdução do referido vector numa célula hospedeira durante um tempo e sob condições suficientes para expressão de uma enzima alongase codificada pela referida sequência nucleotídica isolada; e
- d) exposição da referida enzima alongase expressa a um ácido gordo poli-insaturado substrato para converter o referido substrato num ácido gordo poli-insaturado produto.

12. Método de acordo com a reivindicação 11 compreendendo ainda o passo de exposição do referido ácido gordo poli-insaturado produto a pelo menos uma dessaturase para converter o referido ácido gordo poli-insaturado produto noutro ácido gordo poli-insaturado.

13. Método de acordo com a reivindicação 12 compreendendo ainda o passo de exposição do referido outro ácido gordo poli-insaturado a uma ou mais enzimas seleccionadas de entre o grupo que consiste em pelo menos uma alongase e pelo menos uma dessaturase adicional para converter o referido outro ácido gordo poli-insaturado num ácido gordo poli-insaturado final.

14. Sequência nucleotídica isolada codificando uma alongase funcionalmente activa possuindo identidade com pelo menos cerca de 35% da sequência nucleotídica mostrada na Figura 22, ou o seu complemento.

15. Sequência nucleotídica isolada de acordo com a reivindicação 14 em que a referida sequência é representada pela sequência mostrada na Figura 22.

16. Proteína purificada codificada pela referida sequência nucleotídica de acordo com as reivindicações 14 ou 15.

17. Método de produção de uma enzima alongase compreendendo os passos de:

- a) isolamento de uma sequência nucleotídica representada pela Figura 22;
- b) construção de um vector compreendendo: i) a referida sequência nucleotídica isolada operativamente ligada a ii) um promotor;
- c) introdução do referido vector numa célula hospedeira durante um tempo e sob condições suficientes para a expressão da referida enzima alongase.

18. Vector compreendendo: a) uma sequência nucleotídica tal como representada na Figura 22 operativamente ligada a b) um promotor.

19. Célula hospedeira compreendendo o referido vector de acordo com a reivindicação 18.

20. Célula vegetal, planta ou tecido vegetal compreendendo o referido vector de acordo com a reivindicação 18, em que a expressão da referida sequência nucleotídica do referido vector resulta na produção de um ácido gordo poli-insaturado pela referida célula vegetal, planta ou tecido vegetal.

21. Planta transgénica compreendendo o referido vector de acordo com a reivindicação 18, em que a expressão da

referida sequência nucleotídica do referido vector resulta na produção de um ácido gordo poli-insaturado em sementes da referida planta transgénica.

22. Mamífero não humano transgénico cujo genoma compreende uma sequência de ADN tal como definida na reivindicação 14, de *M. alpina*, codificando uma alongase operativamente ligada a um promotor.

23. Mamífero não humano transgénico de acordo com a reivindicação 22 em que a referida sequência de ADN é representada pela Figura 22.

24. Método para produção de um ácido gordo poli-insaturado compreendendo os passos de:

- a) isolamento da referida sequência nucleotídica representada pela Figura 22;
- b) construção de um vector compreendendo a referida sequência nucleotídica isolada;
- c) introdução do referido vector numa célula hospedeira durante um tempo e sob condições suficientes para expressão de uma enzima alongase codificada pela referida sequência nucleotídica isolada; e
- d) exposição da referida enzima alongase expressa a um ácido gordo poli-insaturado substrato para converter o referido substrato num ácido gordo poli-insaturado produto.

25. Método de acordo com a reivindicação 24 compreendendo ainda o passo de exposição do referido ácido gordo poli-insaturado produto a pelo menos uma dessaturase para converter o referido ácido gordo poli-insaturado produto noutro ácido gordo poli-insaturado.

26. Método de acordo com a reivindicação 25 compreendendo ainda o passo de exposição do referido outro ácido gordo poli-insaturado a uma ou mais enzimas seleccionadas de entre o grupo que consiste em pelo menos uma alongase e pelo menos uma dessaturase adicional para converter o referido outro ácido gordo poli-insaturado num ácido gordo poli-insaturado final.

27. Sequência nucleotídica isolada possuindo identidade ou complementaridade com a sequência mostrada na Figura 43, em que a referida sequência nucleotídica isolada codifica um polipéptido possuindo actividade de alongase.

28. Proteína purificada codificada pela referida sequência nucleotídica de acordo com a reivindicação 27.

29. Método de produção de uma enzima alongase compreendendo os passos de:

- a) isolamento de uma sequência nucleotídica representada pela Figura 43;
- b) construção de um vector compreendendo: i) a referida sequência nucleotídica isolada operativamente ligada a ii) um promotor;
- c) introdução do referido vector numa célula hospedeira durante um tempo e sob condições suficientes para expressão da referida enzima alongase.

30. Vector compreendendo: a) uma sequência nucleotídica tal como representada na Figura 43 operativamente ligada a b) um promotor.

31. Célula hospedeira compreendendo o referido vector de acordo com a reivindicação 30.

32. Célula vegetal, planta ou tecido vegetal compreendendo o referido vector de acordo com a reivindicação 30, em que a expressão da referida sequência nucleotídica do referido vector resulta na produção de pelo menos um ácido gordo seleccionado de entre o grupo que consiste num ácido gordo mono-insaturado e um ácido gordo poli-insaturado, pela célula vegetal, planta ou tecido vegetal.

33. Planta transgénica compreendendo o referido vector de acordo com a reivindicação 30, em que a expressão da referida sequência nucleotídica do referido vector resulta na produção de um ácido gordo poli-insaturado, em sementes da referida planta transgénica.

34. Mamífero não humano transgênico cujo genoma compreende uma sequência de ADN tal como definida na reivindicação 27 codificando uma alongase operativamente ligada a um promotor.

35. Mamífero não humano transgênico de acordo com a reivindicação 34 em que a referida sequência de ADN é representada pela Figura 43.

36. Método para produção de um ácido gordo poli-insaturado compreendendo os passos de:

- a) isolamento da referida sequência nucleotídica representada pela Figura 43;
- b) construção de um vector compreendendo a referida sequência nucleotídica isolada;
- c) introdução do referido vector numa célula hospedeira durante um tempo e sob condições suficientes para expressão de uma enzima alongase codificada pela referida sequência nucleotídica isolada; e
- d) exposição da referida enzima alongase expressa a um ácido gordo poli-insaturado substrato para converter o referido substrato num ácido gordo poli-insaturado produto.

37. Método de acordo com a reivindicação 36 compreendendo ainda o passo de exposição do referido ácido gordo poli-insaturado produto a pelo menos uma dessaturase para converter o ácido gordo poli-insaturado produto noutro ácido gordo poli-insaturado.

38. Método de acordo com a reivindicação 37 compreendendo ainda o passo de exposição do referido outro ácido gordo poli-insaturado a uma ou mais enzimas seleccionadas de entre o grupo que consiste em pelo menos uma alongase e pelo menos uma dessaturase adicional para converter o referido outro ácido gordo poli-insaturado num ácido gordo poli-insaturado final.

39. Sequência nucleotídica isolada possuindo identidade ou complementaridade com a sequência nucleotídica mostrada na Figura 46.

40. Sequência nucleotídica isolada de acordo com a reivindicação 39 em que a referida sequência é representada pela sequência mostrada na Figura 46.

41. Método de produção de uma enzima alongase compreendendo os passos de:

- a) isolamento de uma sequência nucleotídica representada pela Figura 46;
- b) construção de um vector compreendendo: i) a referida sequência nucleotídica isolada operativamente ligada a ii) um promotor;
- c) introdução do referido vector numa célula hospedeira durante um tempo e sob condições suficientes para expressão da referida enzima alongase.

42. Vector compreendendo: a) uma sequência nucleotídica tal como representada na Figura 46 operativamente ligada a b) um promotor.

43. Célula hospedeira compreendendo o referido vector de acordo com a reivindicação 42.

44. Célula vegetal, planta ou tecido vegetal compreendendo o referido vector de acordo com a reivindicação 42, em que a expressão da referida sequência nucleotídica do referido vector resulta na produção de um ácido gordo poli-insaturado pela célula vegetal, planta ou tecido vegetal.

45. Planta transgênica compreendendo o referido vector de acordo com a reivindicação 42, em que a expressão da referida sequência nucleotídica do referido vector resulta na produção de um ácido gordo poli-insaturado em sementes da referida planta transgênica.

46. Mamífero não humano transgênico cujo genoma compreende uma sequência de ADN tal como definida na reivindicação 40, de *C. elegans*, codificando uma alongase operativamente ligada a um promotor.

47. Método para produção de um ácido gordo poli-insaturado compreendendo os passos de:

- a) isolamento da referida sequência nucleotídica representada pela Figura 46;
- b) construção de um vector compreendendo a referida sequência nucleotídica isolada;
- c) introdução do referido vector numa célula hospedeira durante um tempo e sob condições suficientes para expressão de uma enzima alongase codificada pela referida sequência nucleotídica isolada; e
- d) exposição da referida enzima alongase expressa a um ácido gordo poli-insaturado substrato para converter o referido substrato num ácido gordo poli-insaturado produto.

48. Método de acordo com a reivindicação 47 compreendendo ainda o passo de exposição do referido ácido gordo poli-insaturado produto a pelo menos uma dessaturase para converter o referido ácido gordo poli-insaturado produto noutro ácido gordo poli-insaturado.

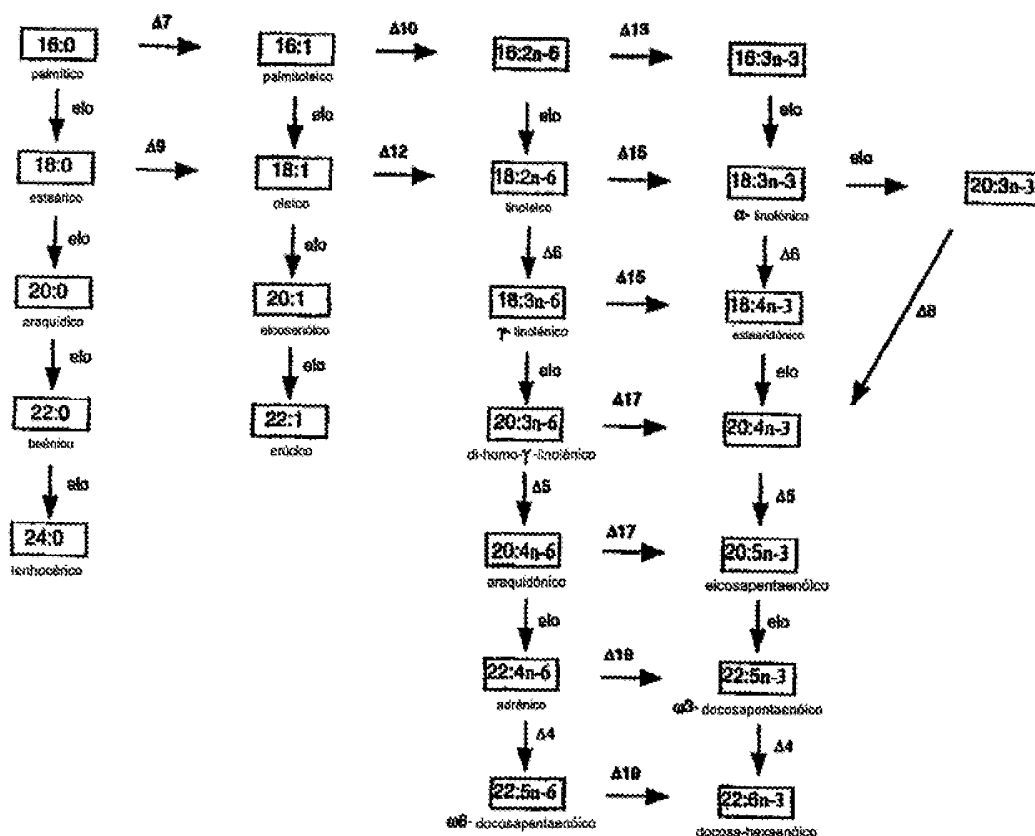
49. Método de acordo com a reivindicação 48 compreendendo ainda o passo de exposição do referido outro ácido gordo poli-insaturado a uma ou mais enzimas seleccionadas de entre o grupo que consiste em pelo menos uma alongase e pelo menos uma dessaturase adicional para converter o referido outro ácido gordo poli-insaturado num ácido gordo poli-insaturado final.

Lisboa, 2008-06-27

RESUMO

"Genes de alongases e suas utilizações"

Vias biossintéticas dos ácidos gordos



O presente invento refere-se à identificação de quatro genes envolvidos no alongamento de ácidos poli-insaturados (i.e., "alongases") e suas utilizações. Dois destes genes estão também envolvidos no alongamento de ácidos gordos mono-insaturados. Em particular, a alongase é utilizada na conversão de ácido gama-linolénico (GLA) em ácido di-homogama-linolénico (DGLA) e na conversão de DGLA ou 20:4n-3 em ácido eicosapentaenóico (EPA). O DGLA pode ser utilizado na produção de ácidos gordos poli-insaturados, tais como o ácido araquidónico (AA), ácido docosa-hexaenóico (DHA), EPA, ácido adrénico, ácido ω 6-docosapentaenóico ou ácido ω 3-docosapentaenóico que pode ser adicionado a composições farmacêuticas, composições nutricionais, rações alimentares, bem como outros produtos tais como cosméticos.

Figura 1
Vias biossintéticas dos ácidos gordos

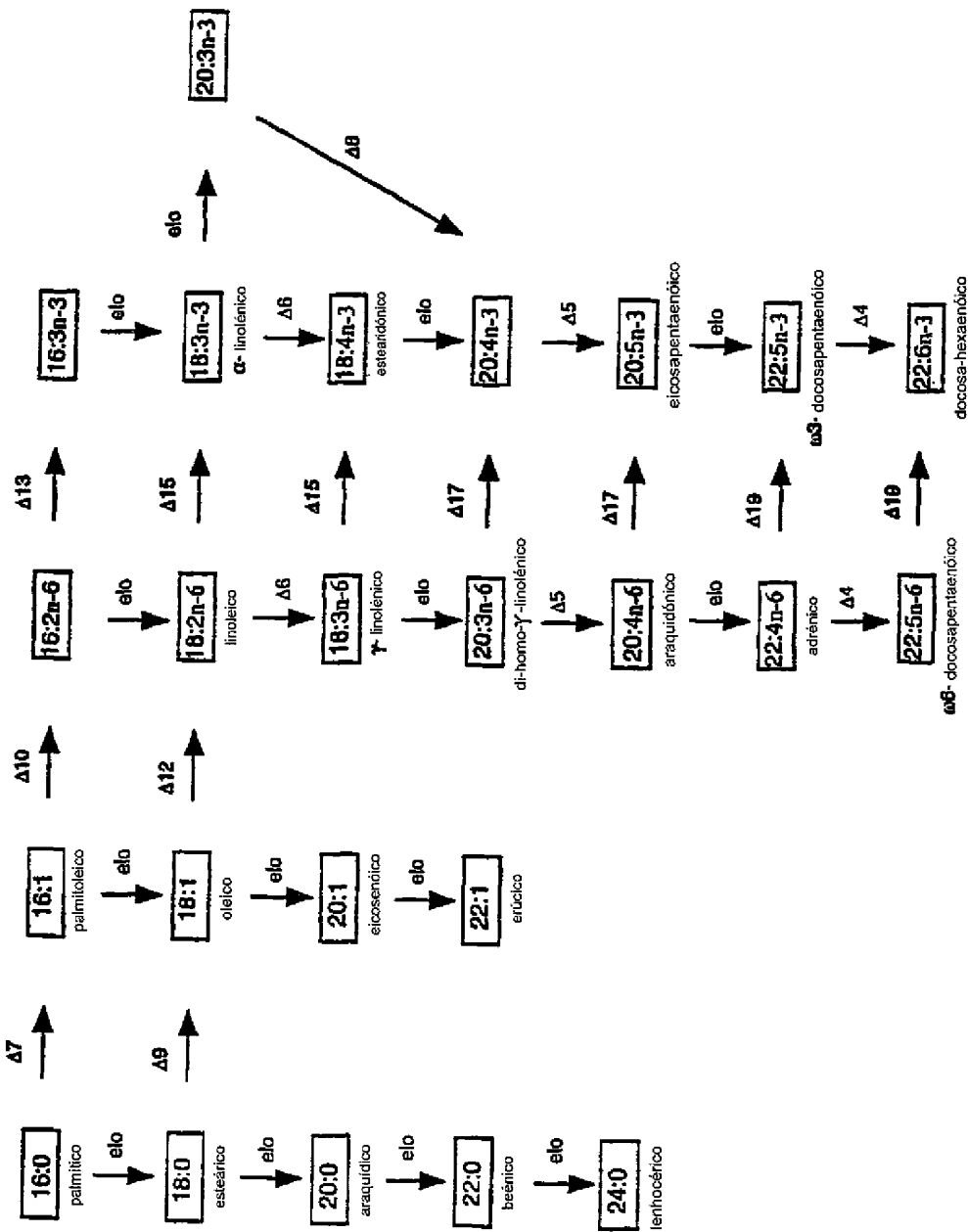


Figura 2

Peso do Intervalo:	6	Média de Coincidências:	2.912
Peso do Comprimento:	4	Média de Diferenças:	-2.003
Qualidade:	50	Comprimento:	84
Razão:	0.625	Intervalos:	4
Percentagem de Semelhança:	43.038	Percentagem de Identidade:	29.114

Limiares de apresentação de coincidências para o(s) alinhamento(s):

	=	IDENTIDADE
:	=	2
.	=	1

jojobakcs x ELO2 4 de Junho de 1998 08:23 ..

jojobakcs	24	ATLPNFKSSINLHHVKL.GYHYLISNALFLVFIPLLGLASAHLSFSFSAHD	72
		. . : : : : . ..	
ELO2	66	STLPPVLYAITAYYVIIFGGREFLLSKS..KPF.KLNGLFQLHNLVLTSL	112

jojobakcs	73	LSLLFDLLRRNLLPVVVCSEFLVLLATLHFLTRP	106
		. : .:: : : : .	
ELO2	113	LTLLL.LMVEQLVPITVQHGGLYFAICNIGAWTQP	145

Figura 3

ELO2 de *S. cerevisiae* (AA66-145) com codões preferenciais de *M. alpina*

S	T	L	P	P	V	L	Y	A	I	T	A	Y	Y	V	I	I	F	G	G	R	F	L
TCC	ACC	CTC	CCC	CCC	GTC	CTC	TAC	GCC	ATC	ACC	GCC	TAC	TAC	GTC	ATC	ATC	TTC	GTC	GGT	CGC	TTC	CTC
66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88

<--- RO339

L	S	K	S	K	P	F	K	L	N	G	L	F	Q	L	H	N	L	V	L	T	S	L
CTC	TCC	AAG	TCC	AAG	CCC	TTC	AAG	CTC	AAC	GGT	CTC	TTC	CAG	CTC	CAC	AAC	CTC	GTC	CTC	ACC	TCC	CTC
89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111

S	L	T	L	L	L	L	M	V	E	Q	L	V	P	I	I	V	Q	H	G	L	Y	F
TCC	CTC	ACC	CTC	CTC	CTC	CTC	ATG	GTC	GAG	CAG	CTC	GTC	CCC	ATC	ATC	GTC	CAG	CAC	GGT	CTC	TAC	TTC
112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134

A	I	C	N	I	G	A	W	T	Q	P
GCC	ATC	TGC	AAC	ATC	GGT	GCC	TGG	ACC	CAG	CCC
135	136	137	138	139	140	141	142	143	144	145

Figura 4

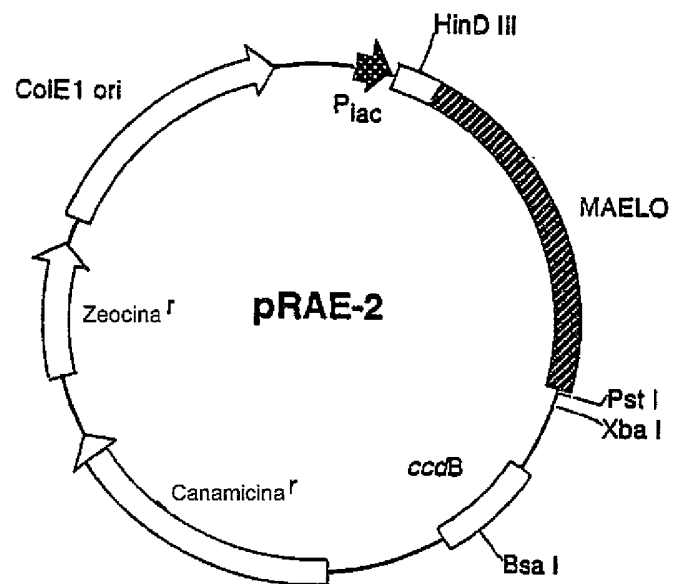
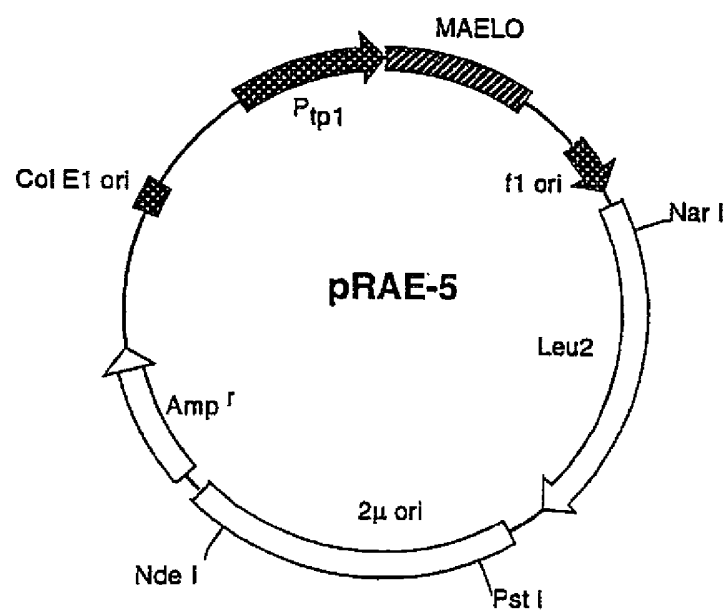
A**B**

Figura 5

pRAE-5	GAATTCAGG	*	*	*	*	*	*	*	*	*	*CATGGCCGCCCGCAATCTTGGACAA
pRAE-6	GAATTCAGGCATCTCATGGATCCGCCCATGGCCCGCCAAATCTTGGACAA	EcoRI	BamHI	NcoI							

Figura 6

```

1  ATGGCCGCCG CAATCTTGGA CAAGGTCAAC TTCGGCATTG ATCAGCCCTT
51  CGGAATCAAG CTCGACACCT ACTTTGCTCA GGCTATGAA CTCGTCACCG
101 GAAAGTCCAT CGACTCCTTC GTCTTCCAGG AGGGCGTCAC GCCTCTCTCG
151 ACCCAGAGAG AGGTCGCCAT GTGGACTATC ACTTACTTCG TCGTCATCTT
201 TGGTGGTCGC CAGATCATGA AGAGCCAGGA CGCCTTCAAG CTCAAGCCCC
251 TCTTCATCCT CCACAACCTC CTCCTGACGA TCGCGTCCGG ATCGCTGTTG
301 CTCCTGTTCA TCGAGAACCT GGTCCCCATC CTCGCCAGAA ACGGACTTTT
351 CTACGCCATC TCGGACGACG GTGCCTGGAC CCAGCGCCTC GAGCTCCTCT
401 ACTACCTCAA CTACCTGGTC AAGTACTGGG AGTTGGCCGA CACCGTCTTT
451 TTGGTCCTCA AGAAGAAGCC TCTTGAGTTC CTGCACTACT TCCACCACTC
501 GATGACCATG GTTCTCTGCT TTGTCCAGCT TGGAGGATAC ACTTCAGTGT
551 CCTGGGTCCC TATTACCCTC AACTTGACTG TCCACGTCTT CATGTACTAC
601 TACTACATGC GCTCCGCTGC CGGTGTTTCG ATCTGGTGGA AGCAGTACTT
651 GACCACTCTC CAGATCGTCC AGTTCGTTCT TGACCTCGGA TTCATCTACT
701 TCTGCGCCTA CACCTACTTC GCCTTCACCT ACTTCCCCTG GGCTCCCAAC
751 GTCGGCAAGT GCGCCGGTAC CGAGGGTGCT GCTCTCTTTG GCTGCGGACT
801 CCTCTCCAGC TATCTCTTGC TCTTTATCAA CTTCTACCGC ATTACCTACA
851 ATGCCAAGGC CAAGGCAGCC AAGGAGCGTG GAAGCAACTT TACCCCCAAG
901 ACTGTCAAGT CCGGCGGATC GCCCAAGAAG CCCTCCAAGA GCAAGCACAT
951 CTAA

```

Figura 7

1 MAAAILDKVN FGIDQPFGIK LDYFAQAYE LVTGKSIDSF VFQEGVTPLS
51 TQREVAMWTI TYFVVIFGGR QIMKSQDAFK LKPLFILHNF LLTIASGSLL
101 LLFIENLVPI LARNGLFYAI CDDGAWTQRL ELLYYLNYLV KYWELADTVF
151 LVLKKKPLEF LHYFHHSMTM VLCFVQLGGY TSVSWVPITL NLTVHVFMYY
201 YYMRSAAGVR IWWKQYLTTL QIVQFVLDLG FIYFCAYTYF AFTYFPWAPN
251 VGKCAGTEGA ALFGCGLLSS YLLLFINFYR ITYNAKAKAA KERGSNFTPK
301 TVKSGGSPKK PSKSKHI*

Figura 9

REGISTOS Inic1: 153 Inin: 199 Opt: 495

57,4% de identidade numa sobreposição de 549 pb

	150	160	170	180	190	200
MAELO	TCTCGACCCAGAGAGAGGTCGCCATGTGGAATATCACTTACTTCGTGTCATCTTTGGTG					
S78624	CATTAAAGCACTTTGCCCCCTGTGCTATACGCCATCACTGCCTATTACGTTATTATTTTG					
	5990	6000	6010	6020	6030	6040
	210	220	230	240	250	260
MAELO	GTGCGCCAGATCATGAAGAGCCAG--GACGCC--TTCAAGCTCAAC/CCCCTCTTCATCCTCC					
S78624	GTGGCAGGTTTTTGTAAAGTAAGTCGAACCATTTAAATTAATGGCCTTTTCCAATTGC					
	6050	6060	6070	6080	6090	6100
	270	280	290	300	310	320
MAELO	ACAACTTCCTCCTGACGATCGCGTCC--GGATCGCTGTGCTCCTGTTTCATCGAGAACCT					
S78624	ATAATTGGTTTTTAAC-TTCACTTTCATTGA-CGCTTTTATTGCTTATGGTTGAACAAT					
	6110	6120	6130	6140	6150	6160
	330	340	350	360	370	380
MAELO	GGTCCCCATCCTCGCCAGAAACGGAATTTCTACGCCATCTGCGACGACGGTGCCTGGAC					
S78624	AGTGCCAATTATTGTTTCAGCACGGGTTATACTTCGCTATCTGTAAATATTGGTGCCTGGAC					
	6170	6180	6190	6200	6210	6220
	390	400	410	420	430	440
MAELO	CCAGCGCCTCGAGCTCCTCTACTACCTCACTACCTGGTCAAGTACTGGGAGTTGGCCGA					
S78624	TCAACCGCTCGTTACATTATATTACATGAATTACATTGTCAAGTTATGAATTTATAGA					
	6230	6240	6250	6260	6270	6280
	450	460	470	480	490	500
MAELO	CACCGTCTTTTGGTCCTCAAGAAGAAGCCTCTTGAGTTCTGCACTACTTCCACCCTC					
S78624	CACCTTTTCTTGGTGCTAAAACATAAAAAATTGACATTTTGCA-TACTT--ATCA--C					
	6290	6300	6310	6320	6330	6340
	510	520	530	540	550	
MAELO	GATGACCATGGTTCTCTGCTTGT---CCAGCTTGGAGGATA-CACTTCAGTGTCTGG					
S78624	CATGGCGCTACTGCCTTATTATGTTACACCAATTGATGGGCACCACTCTATTCTTGG					
	6350	6360	6370	6380	6390	6400
	560	570	580	590	600	610
MAELO	GTCCCTATTACCTCAACTTGACTGTCCAGCTCTTCATGTACTACTACTACATGCCCTCC					
S78624	GTCCCTATTTCATTGAACCTTGGTGTTCAGTGGTTATGTATTGGTACTATT--CTTG					
	6410	6420	6430	6440	6450	
	620	630	640	650	660	670
MAELO	GCTGCC---GGTGTTCGCATCTGGTGAAGCAGTACTTGACCACTCTCCAGATCGTCCAG					
S78624	GCTGCCAGAGGCATCAGGGTCTGGTGAAGGAATGGGTACCAGATTTCAAATTATCCAA					
	6460	6470	6480	6490	6500	6510
	680	690	700	710	720	730
MAELO	TTGTTCTTGACCTCGGATTCATCTACTTCTGCGCCTACACCTACTTGCCTTCACCTAC					
S78624	TTTGTTTTGATATCGGTTTCATATATTTTGTCTGTCTACCAAAAAGCAGTTCACTTGTAT					
	6520	6530	6540	6550	6560	6570

Figura 10A

Hospedeiro (plasmídeo)	334(pCGN7875)	334(pYES2)	334(pYX242)	334(pRAE-5)	334(pRAE-6)	334(pYX242)	334(pRAE-5)
Substrato adicionado	25 µM OA	25 µM OA	25 µM GLA	25 µM GLA	25 µM GLA	sem substrato	sem substrato
Ácido gordo	Lípido (µg)	Lípido (µg)	Lípido (µg)	Lípido (µg)	Lípido (µg)	Lípido (µg)	Lípido (µg)
C16:0	11.948	23.601	35.123	92.011	85.160	16.294	25.34
C16:1	30.665	71.217	32.789	315.464	115.456	56.183	113.913
C18:0	6.185	9.704	10.515	22.628	18.879	5.535	11.092
C18:1n-9	35.340	57.429	33.989	154.386	106.881	28.388	51.538
C18:3n-6			48.856	58.084	12.434		
C20:0			0.474	0.710	0.244		
C20:1n-9	(0.375%)* 0.352	(0.309%)* 0.527		1.405	0.867		0.516
C20:3n-6	ND	ND	(0.092%)* 0.226	(0.324%)* 2.504	(0.269%)* 1.006	ND	ND
C22:0				0.460			
C22:1n-9				0.321	0.315		
C24:0					1.825		0.999
Lípidos Totais	93.760	170.490	245.090	771.690	374.420	112.99	256.52
ND = Não detectado							
*% de ácidos gordos totais							

Figura 10B

Hospedeiro (plasmídeo)	334(pYX242)	334(pYX242)	334(pRAE-5)	334(pRAE-5)	334(pRAE-6)
Substrato adicionado	25 µM GLA	25 µM GLA	25 µM GLA	25 µM GLA	25 µM GLA
Ácido gordo	Lípido (µg)	Lípido (µg)	Lípido (µg)	Lípido (µg)	Lípido (µg)
C16:0	60.683	61.487	100.998	96.193	66.761
C16:1	79.838	79.586	339.754	220.440	87.359
C18:0	9.784	10.106	15.317	15.165	16.744
C18:1n-9	38.536	39.936	108.472	89.637	71.631
C18:3n-6	17.974	17.833	82.866	56.596	17.766
C20:0			0.510	0.570	
C20:1n-9					
C20:3n-6	(0.136%)* 0.389	(0.130%)* 0.374	(0.336%)* 3.035	(0.401%)* 2.689	(0.353%)* 1.185
C22:0			0.414		
C22:1n-9				0.383	
C24:0			1.513	1.626	
Lípidos Totais	285.560	288.045	902.560	671.113	335.496
% de ácidos gordos totais					

Figura 11

Hospedeiro (plasmídeo)	334(pRAE-5/pCGR4)	334(pYX242/pYES2)	Hospedeiro (plasmídeo)	334(pRAE-5/pCGR4)	334(pYX242/pYES2)
Substrato adicionado	25 µM GLA	25 µM GLA	Substrato adicionado	25 µM GLA	25 µM GLA
Ácido gordo	Lípido (µg)	Lípido (µg)		Lípido (µg)	Lípido (µg)
C16:0	41.050	37.169	C16:0	96.986	32.221
C16:1	99.393	100.552	C16:1n-7	209.667	62.757
C18:0	34.432	27.852	C18:0	80.418	14.027
C18:1	110.631	92.786	C18:1n-9	207.104	28.701
C18:3n-6	15.004	7.924	C18:3n-6	25.264	10.543
C20:0	0.643	0.574	C20:0	2.038	
C20:1	1.996	1.684	C20:1n-9	3.591	
C20:3n-6	0.542	0.607	C20:3n-6	1.284	0.326
C20:4n-6	0.579		C20:4n-6	1.392	
C22:0	1.242	2.604	C22:0	1.124	
C24:0	4.754	4.563	C24:0	3.952	
Lípidos Totais	334	300	Lípidos Totais	756	197

Figura 12

Hospedeiro (plasmídeo)	334(pYX242)	334(pRAE-5)	334(pRELO-1)	334(pRELO-2)
Substrato adicionado	25 µM GLA	25 µM GLA	25 µM GLA	25 µM GLA
	25°C/48hrs	25°C/48hrs	25°C/48hrs	25°C/48hrs
Ácido gordo	Lípido (µg)	Lípido (µg)	Lípido (µg)	Lípido (µg)
C16:0	28.7	76.707	84.424	77.445
C16:1	0.729	2.513	1.532	1.056
C18:0	7.432	15.761	27.17	21.32
C18:1n-9	28.9	77.323	109.419	82.844
C18:3n-6	9.729	29.236	19.085	18.804
C20:0		0.643	0.522	0.537
C20:1n-9		0.77	0.426	0.299
C20:3n-6	(0.185%)* 0.374	(0.279%)* 1.472	(0.153%)* 0.748	(0.200%)* 0.832
C22:0		0.451		
C22:1n-9			0.224	
C24:0		0.918		
Lípidos Totais	202	527	490	416
*% de ácidos gordos totais				

Figura 14

REGISTOS Inic1: 178 Inin: 178 Opt: 318
 .Registo de Smith-Waterman: 318; 33,0% de identidade numa sobreposição de 188 aa

	50	60	70	80	90	100
Z68749	SLLTNQDEVFPHIRARRFIQEHFGLFVQMAIAYVILVFSIKRFMRDREPFQLTALRLWN					
	: ::: : ::: ::: : :					
MAELO	ELVTGKSIDSFVFQEGVTPLSTOREVAMWTITYFVVIFGGRQIMKSQDAFKLKPLFILHN					
	30	40	50	60	70	80
	110	120	130	140	150	160
Z68749	FFLSVFSIYGSWTMFPP--MVQIRLYGLYGCGCEALSNLPSQAEYNLFLTILSKAVEFV					
	: :: ::: : : : : : :: : : ::					
MAELO	FLLTIAS--GSLLLLFIEINLVPILARNGLFYAICDD-GAWTORLELLYYLNYLVKYWELA					
	90	100	110	120	130	140
	170	180	190	200	210	220
Z68749	DTFFLVLRKKPLIFLHWYHHMATFVFFCSNYPTPSSQSRVGVIVNLFVHAFMYPYYFTRS					
	: :: : : : : : : : : :					
MAELO	DTVFLVLKKKPLEFLHYFHHSMIMVLCFVQLGGYTSVSWVPITLNLTVHVFMYYYMRS					
	150	160	170	180	190	200
	230	240	250	260	270	
Z68749	MNIKVPAKISMAVTVLQLTQF---MCFIYGCTLMYYSLATNQARYPSNTPATLQCLSYTL					
	::: : : : :: : : :::					
MAELO	AGVRIWKK--QYLTTLQIVQFVLDLGFIFYFCAYTYFAFTYFPWAPNVGKCAGTEGAALFG					
	210	220	230	240	250	260
Z68749	280 HLL					
MAELO	CGLLSSYLLLFINFYRITYNAKAKAAKERGSNFTPKTVKSGGSPKKPSKSKHIX					
	270	280	290	300	310	

Figura 15

REGISTOS Inic1: 30 Inin: 30 Opt: 40
 Registo de Smith-Waterman: 49; 22,1% de identidade numa sobreposição de 86 aa

		10	20	30	40
AF003134		MLYSITRRCYTFVFVTSLSHFYQLYVTECLENVIFNVLVNGQSINSRWKD			
		: : : : : : :: : : : : :: :			
MAELO	MAAAILDKVNFGIDQPFQIKLDTYFAQA---	YELVTGKSIDSFVFQEGVT--PLSTQREV			
		10	20	30	40
		50	60	70	80
AF003134	AEKTITSFPPHF-----PQTFEQPHILTLHFLFFVFSVTLVTVFKPKCEFPHSLA	90	100		
	: :: ::::: : : :				
MAELO	AMWTITYFVVIFGGRQIMKSQDAFKLPLFILHNFLLTIASGSLLLFIENLVPILARNG				
		60	70	80	90
				100	110

Figura 16

Ratinho

REGISTOS Inic1: 161 Inin: 191 Opt: 325

Registo de Smith-Waterman: 325; 28,8% de identidade numa sobreposição de 285 aa

	10	20	30	39	40
U97107	MDTSMNFSRGLKMD--LMQPYDFETFDLRPFLEEWVSSF-----LIVV				
	: : : : : : : : : : : :				
MAELO	MAAAILDKNVFGIDQPFGIKLDITYFAQAYELVTGKSIDSFVFOEGVTPLSQREVAMWTI				
	10	20	30	40	50 60

	50	60	70	80	90	100
U97107	VYLLIIVVGQTYMRTRKSFSLQREPLILWSFFLAIFSLGTLRMWKMATVMFTVGLKQTV					
	: :: : : : : : : : : : : : : : : : :					
MAELO	TYFVVFEGGRQIMKSQDAFKLPLFLLHNFLLTIAS--GSL-LLLFIENLV-PILARNGL					
	70	80	90	100	110	

	110	120	130	140	150
U97107	CEAIYTDDAVVRWFSEFLLSKV---ELGDTAFIILRKRPLIEVHWYHST--VLLPTS				
	: : : : : : : : : : : : :				
MAELO	FYAICDDGAWTORLELLYYLNYLVKYWELADTVFLVLKKKPLEFLHYFHHSMTMVLCEVQ				
	120	130	140	150	160 170

	160	170	180	190	200	210
U97107	FGYKNKVPSSGGWF-MTMNEGVHSVMYTYTMTKAAKLKHPNLLPMVITSIQIQLQVLG---					
	: : : : : : : : : : : : : :					
MAELO	LGGYTSV---SWVPITLNLTVHVFMYYYMRSAGVR--IWWKQYLTTLQIVQFVLDLGF					
	180	190	200	210	220	230

	220	230	240	250	260
U97107	-----TIFGIILNYIWRQEKG-CHTTTEHFFWSFMYLGYTIFLFAHFFHRAYLAPKGVKVA				
	: : : : : : : : : : : : :				
MAELO	IYFCAYTYFAFTYFFWAPNVGKCACTEGAALFGCGLLSSYLLLFNFYRITY-NAKAKAA				
	240	250	260	270	280 290

	270
U97107	SKSQX
	::
MAELO	KERGSNFTPKTVKSGGSPKKPSKSHIX
	300 310

Humano

REGISTOS Inic1: 147 Inin: 147 Opt: 211

Registo de Smith-Waterman: 211; 28,7% de identidade numa sobreposição de 150 aa

	110	120	130	140	150	160
MAELO	NLVPILARNGLFYAICDDGAWTORLELLYYLNYLVKYWELADTVFLVLKKKPLEFLHYFH					
	: : : : : : :					
AC004050	SLLVVKDLTYLLPLCLPGDTIFITLRKQKLIIFLHWYH					
	10	20	30			

	170	180	190	200	210	220
MAELO	HSMTMVLCEVQLGGYTSVSWVPITLNLTVHVFMYYYMRSAGVR IWWK--QYLTTLQIV					
	: : : : : : :					
AC004050	HITVLLYSWYSYKDMVAGGGWFMTMNYGVHVMYSYALRAAGFRVSRKFAMFTITLSQIT					
	40	50	60	70	80	90

	230	240	250	260	270	280
MAELO	QFVLDLGFYFCAYTYFAFTYFFWAPNVGKCACTEGAALFGCGLLSSYLLLFNFYRITY					
	: : : : : : : : : : : : :					
AC004050	QMLMG-----CVVNYLVFC---WMQH-DQCHSHFQNI FWSLSMYLSYLVLCRFFFEAY					
	100	110	120	130	140	

Figura 18

1 MGTDQKTFE WEELAAHNTK DDLLLAIRGR VYDVKFLSR HPGGVDTLLE
51 GAGRDTVTFV EMYHAFGAAD AIMKKYYVGT LVSNELPIFP EPTVFHKTIK
101 TRVEGYFTDR NIDPKNRPEI WGRYALIFGS LIASYYAQLF VPFVVERTWL
151 QVVFAIIMGF ACAQVGLNPL HDASHFSVTH NPTVWKILGA THDFENGASY
201 LVWVYQHMLG HHPYTNIAGA DPDVSTSEPD VVRIKPNQKW FVNHINQHMF
251 VPFLYGLLAF KVRIQDINIL YFVKTNDAIR VNPISWHTV MFWGGKAFFV
301 WYRLIVPLQY LPLGKVLLLF TVADMVSSYW LALTFQANHV VEEVQWPLPD
351 ENGLIQKWA AMQVETTQDY AHDShLWTSI TGSLNYQAVH HLFPNVSQHH
401 YPDILAIKN TCSEYKVPYL VKDTFWQAFA SHLEHLRVLG LRPKEE*

Figura 19

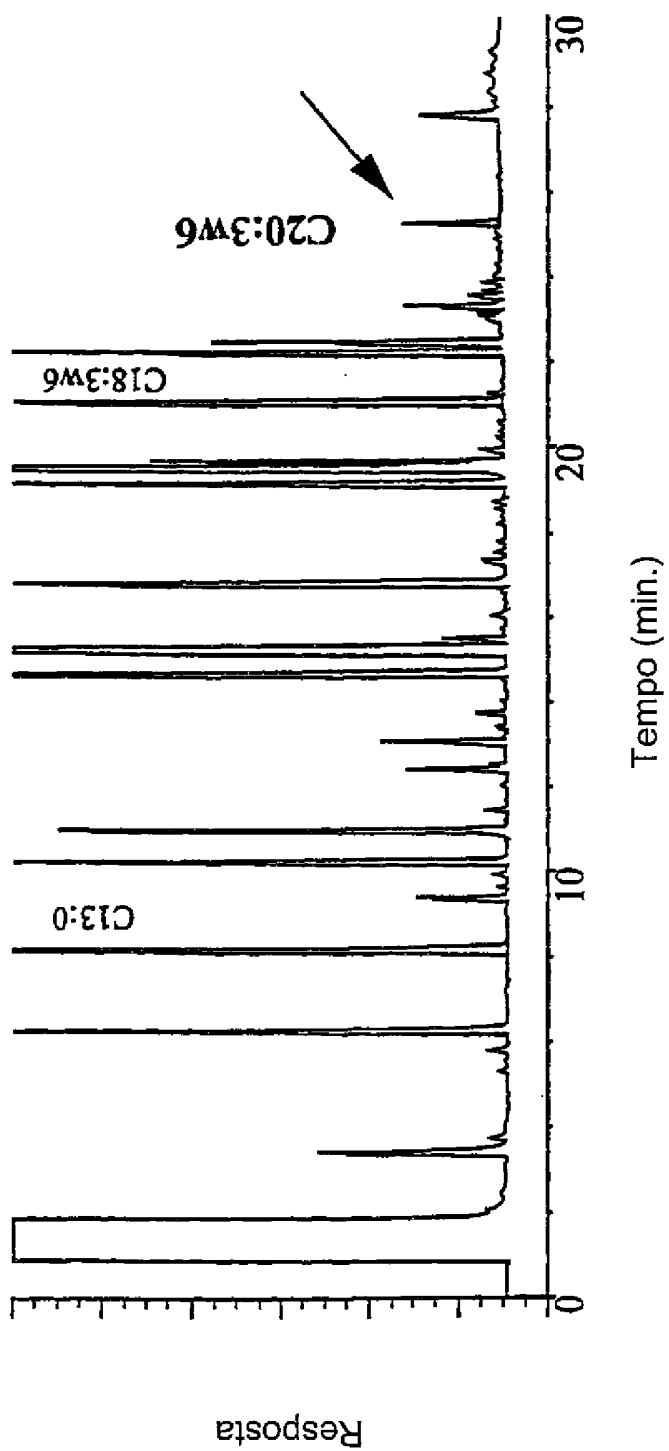


Figura 21

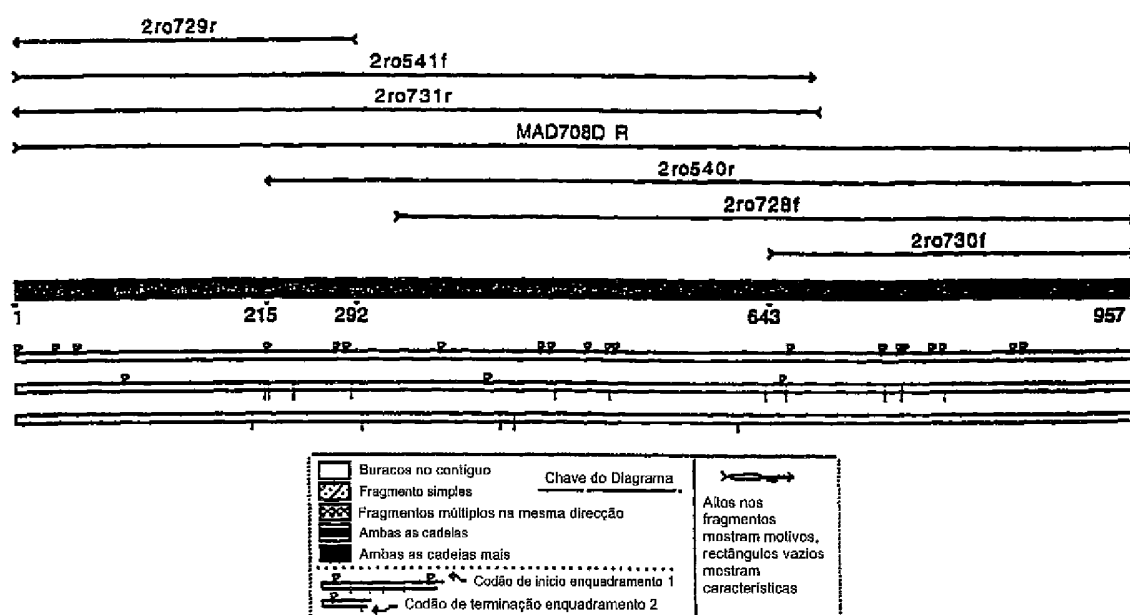


Figura 22

```

1  ATGGAGTCGA TTGCGCCATT CCTCCCATCA AAGATGCCGC AAGATCTGTT
51 TATGGACCTT GCCACCGCTA TCGGTGTCCG GGCCGCGCCC TATGTGATC
101 CTCTCGAGGC CGCGCTGGTG GCCCAGGCCG AGAAGTACAT CCCACGATT
151 GTCCATCACA CGCGTGGGTT CCTGGTCGCG GTGGAGTCGC CTTTGGCCCCG
201 TGAGCTGCCG TTGATGAACC CGTTCCACGT GCTGTTGATC GTGCTCGCTT
251 ATTTGGTCAC GGTCTTTGTG GGCATGCAGA TCATGAAGAA CTTTGAGCGG
301 TTCGAGGTCA AGACGTTTTT GCTCCTGCAC AACTTTTGTG TGGTCTCGAT
351 CAGCGCCTAC ATGTGCGGTG GGATCCTGTA CGAGGCTTAT CAGGCCAACT
401 ATGGACTGTT TGAGAACGCT GCTGATCATA CCTTCAAGGG TCTTCCTATG
451 GCCAAGATGA TCTGGCTCTT CTACTTCTCC AAGATCATGG AGTTTGTCTGA
501 CACCATGATC ATGGTCCTCA AGAAGAACAA CCGCCAGATC TCCTTCTTGC
551 ACGTTTACCA CCACAGCTCC ATCTTCACCA TCTGGTGGTT GGTACCTTT
601 GTTGCACCCA ACGGTGAAGC CTACTTCTCT GCTGCGTTGA ACTCGTTCAT
651 CCATGTGATC ATGTACGGCT ACTACTTCTT GTCGGCCTTG GGCTTCAAGC
701 AGGTGTCGTT CATCAAGTTC TACATCACGC GCTCGCAGAT GACACAGTTC
751 TGCATGATGT CGGTCCAGTC TTCCTGGGAC ATGTACGCCA TGAAGGTCCT
801 TGGCCGCCCC GGATACCCCT TCTTCATCAC GGCTCTGCTT TGGTTCTACA
851 TGTGGACCAT GCTCGGTCTC TTCTACAACT TTTACAGAAA GAACGCCAAG
901 TTGGCCAAGC AGGCCAAGGC CGACGCTGCC AAGGAGAAGG CAAGGAAGTT
951 GCAGTAA

```

Figura 23

1 MESIAPFLPS KMFQDLFMDL ATAIGVRAAP YVDPLEAALV AQAEKYIPTI
51 VHHTRGFLVA VESPLARELP LMNPFHVLLI VLAYLVTVFV GMQIMKNFER
101 FEVKTFSLLE NFCLVSISAY MCGGILYEAY QANYGLFENA ADHTFKGLPM
151 AKMIWLFYFS KIMEFVDIMI MVLKKNRQI SFLHVEHSS IFTIWWLVTF
201 VAPNGEAYFS AALNSFIEVI MYGYFELSAL GFKQVSFIKF YITRSQMTQF
251 CMMSVQSSWD MYAMKVLGRP GYPFFITALL WFYMWTLGL FYNFYRKNK
301 LAKQAKADAA KEKARKLQ*

Figura 24

Hospedeiro (plasmídeo)	334(pRPB2)	334 (pYES2)
Substrato adicionado	25µM GLA	25µM GLA
	(n=4)	
Ácido gordo	% de lípidos totais	
C16:0	15.65	15.23
C16:1	35.2	38.59
C18:0	5.68	5.55
C18:1n-9	25.55	25.27
C18:3n-6	3.1	6.75
C20:0	0.36	0.14
C20:3n-6	(62.0%) 5.06	(2.6%) 0.18
Lípidos Totais (µg)	314	247
(% de conversão) = produto/(substrato+produto)		

Figura 25

Hospedeiro (plasmídeo)	334(pRPB2)	334(pRPB2)	334(pRPB2)	334(pRPB2)	334(pRPB2)	334(pRPB2)
Substrato adicionado	25µM SA C18:0	25µM OA C18:1n-9	25µM LA C18:2n-6	25µM DLGA C20:3n-6	25µM AA C20:4n-6	25µM Adrenico C22:4n-6
Ácido gordo	% de lípidos totais					
C16:0	15.07	14.52	15.74	15.69	16.06	15.15
C16:1	33.7	32.37	32.23	25.65	33.65	33.39
C18:0	*9.78	5.83	5.61	8.33	4.52	5.35
C18:1n-9	31.2	*37.25	26.05	20.15	24.54	28.54
C18:2n-6			*10.4			
C18:3n-6						
C20:2n-6			0.29			
C20:3n-6				*16.5		
C20:4n-6				0.27	*11.7	
C22:4n-6						*7.46
Lípidos Totais (µg)	132	130	171	55	225	163

Hospedeiro (plasmídeo)	334(pRPB2)	334(pRPB2)	334(pRPB2)
Substrato adicionado	25µM ALA C18:3n-3	25µM STA C18:4n-3	25µM EPA C20:5n-3
Ácido gordo	% de lípidos totais		
C16:0	17.32	16.01	20.67
C16:1	27.68	34.31	50.7
C18:0	6.75	5.39	6.14
C18:1n-9	28.4	28.54	
C18:3n-3	*8.39		
C18:4n-3		*1.95	
C20:4n-3		(73.2%) 5.35	
C20:5n-3			*10.33
C22:5n-3			0.25
Lípidos Totais (µg)	114	199	201

* Indica substrato adicionado

(% de conversão) = produto/(substrato+produto)

Figura 26

Hospedeiro (plasmídeo)	334(pRPB2+PRPE31)	334(pYES2+pYX242)
Substrato adicionado	25µM GLA	25µM GLA
Ácido gordo	% de lípidos totais	
C16:0	15.54	18.26
C16:1	30.16	33.51
C18:0	8.76	5.58
C18:1n-9	27	27.37
C18:3n-6	*2.6	*5.6
C20:0	0.4	0.32
C20:3n-6	(57.4%) 3.55	(2.9%) 0.17
C20:4n-6	(27.6%) 1.32	ND
Lípidos Totais (µg)	254	258

A.

Hospedeiro (plasmídeo)	334(pRPB2+PRPE31)	334(pYES2+pYX242)
Substrato adicionado	25µM STA	25µM STA
Ácido gordo	% de lípidos totais	
C16:0	18	16.4
C16:1	28.37	34.78
C18:0	7.42	5.71
C18:1n-9	26.44	30.15
C18:4n-3	*2.93	*4.57
C20:0	0.25	0.17
C20:4n-3	4.13	0.32
C20:5n-3	(39%) 1.87	(2.1%) .10
Lípidos Totais (µg)	257	304

B.

* indica substrato adicionado
 (% de conversão) = produto/(substrato+produto)

Figura 27

REGISTOS Inic1: 114 Inin: 278 Opt: 278
 Registo de Smith-Waterman: 308; 30,9% de identidade numa sobreposição de 259 aa

	40	50	60	70	80	90	99
GLELO	VAQAERYIPTIVHETRGFLVAVESPLARELPLMPPHVLIVLAYLVTTFVGMQIMKNFE						
					: : : : : : : : :		
MAELO	GIRLDTYFAQAYELVTGKSIDSFVFQEGVTPFLSTQREVAMWTITYFVVIFGGRQIMKSQD						
	20	30	40	50	60	70	

	100	110	120	130	140	150
GLELO	RFEVKTFSLLHNFCLVSI SAYMCGGILYE--AYQANYGLFENAADETFKGLPMAXMIWLF					
	: : : : : : : : : : :					
MAELO	AFKLPFLFLENFLITIASGSLLLLFIEENLVPIARNGLFYAI CDDGANTQRLLELYILN					
	80	90	100	110	120	130

	160	170	180	190	200	210
GLELO	YFSKIMEFVDTMIMVLKKNNRQISFLBVIHSSIFTIWWLVTFVAPNGEAYFSAALNSFI					
	: : : : : : : : : : : : : :					
MAELO	YLVKIWELADTVFLVLKK--KPLEFLHYFHS-MTMVLCFVQLGGYTSVSWVPITIALTV					
	140	150	160	170	180	190

	220	230	240	250	260
GLELO	HVIMYGYFFLSALGFKQVSFIKFFYITRSQMTQF-----CMMSVQS----SWDMYAM				
	: : : : : : : : : : : :				
MAELO	HVFMYYYMRSAAAGVRI--WWKQYLTTLQIVQFVLDLGFIFYFCAYTYFAFTYFPWAPNVG				
	200	210	220	230	240

	270	280	290	300	310
GLELO	KVLGRPGYPFFITALLWFYMTMLGLFYNFYRKNAKLAQAKADAAKEKARKLQ				
	: : : : : : : :				
MAELO	KCAGTEGAALFGCGLLSYLL----LFINFYR----ITYNAKAKAAKERGSNFTPKTVKS				
	260	270	280	290	300

MAELO	GGSPKKPSKSKHIX
	310

Figura 29

REGISTOS Inic1: 83 Inin: 186 Opt: 271

Registo de Smith-Waterman: 297; 28,5% de identidade numa sobreposição de 242 aa

	30	40	50	60	70	80
MAELO	YELVTGKSIDSFVFOEGVTPPLSTQREVAMWTITYFVVIFGGRQIMKSQDAFKLKPLFILH					
					:: - :: :: : :: :: : : :: ::	
HS1	STYFKALLGPRDTRVKGWFLLDNYIPTFICSVIYLLIVWLGPKYMRNKQPFSCRGLVVY					
	10	20	30	40	50	60
	90	100	110	120	130	140
MAELO	NFLTITASGSLLLLFIEENLVPIARNGLFYAICDDGAWTORLELLYYL--NYLVKYWELA					
	: : : : : : : : : :: :: : :					
HS1	NLGLTLLS---LYMFCELVTVGWEGKYNFFCQGTRTAGESDMKLIIRVLWVWYFYSKLIEFM					
	70	80	90	100	110	120
	150	160	170	180	190	200
MAELO	DTVFLVLKK--KPLEFLHYFHH-SMTMVLCFVQLGGYTSVSWVPITLNLTVHVFMYYYY-					
	: : : : : : : : : : : : : :					
HS1	DTFFFILRKNNHQITVLHVYHHASMLNIWWFVMNWVPCGHSYFGATLNSFIHVLMYSSYG					
	130	140	150	160	170	180
	210	220	230	240	250	260
MAELO	MRSAAGVR--IWWKQYLTTLQIVQFVLDLGFIFYFCAYTYFAFTYFPWAPNVGKCAGTEGA					
	: : : : : : : : : : : : : : : :					
HS1	LSSVPSMRPYLWWKKYITQGQLLQFVLTII-IQTSCGVI-----W-P-----CTFPLGW					
	190	200	210	220		230
	270	280	290	300	310	
MAELO	ALFGCGLSSYLLLFINFYRITYNAKAKAAKERGSNFTPKTVKSGGSPKKPSKSKHI					
	: : : : : :					
HS1	LYFQIGYMISLIALFTNFYIQTYNKKGASRRKDHLKDHQNGSMAAVNGHTNSFSPLENNV					
	240	250	260	270	280	290
HS1	KPRKLRKDX					
	300					

Figura 30

REGISTOS Inic1: 88 Inin: 208 Opt: 272
 Registo de Smith-Waterman: 279; 28,2% de identidade numa sobreposição de 266 aa

	30	40	50	60	70	80
MAELO	QAYELVTGKSIDSFVFOEGVTPLSTQREVAMWTITYFVVIFGGROIMKSQDAFKLKPLFI					
	: : : :: : : :					
HS2	VNLVQEVMMKHADPRIQGYPLMGSPLLMTSILLTYVYFVLSLGR-IMANRKPFQLRGFMI					
	10	20	30	40	50	60
	90	100	110	120	130	140
MAELO	LHNEFLTITASGSLLLFTEN--LVPILAR-NGLFYAICDDGAWTORLELLYYLNYLVKXW					
	: : : : : : : : : : : : : :					
HS2	VYNFSLVALSLYIVYEFILMSGWLSTYIWRCDPVDYSNSPEALRMVRVANLFLFS---KFI					
	70	80	90	100	110	120
	150	160	170	180	190	
MAELO	ELADTVFLVLKKK--PLEFLHYFHHSMT----MVLCFVQLGGYTSVSWVPITLALTVHVF					
	: : : : : : : : : : :					
HS2	ELMDIVIFILRKKGQVITFLHVFHHSVLPWSWWGVKIAPGGMGSGFHAM---INSSVHVI					
	130	140	150	160	170	
	200	210	220	230	240	
MAELO	MYYYVMRSAAGV----RIWMKQYLTTLOIVQFVL---DLGFTYF---CAYTYFAFTYFPW					
	: : : : : : : : : : : : : :					
HS2	MYLYYGLSAFGPVAQPYLWKKHMTAIQLIQFVLVSLHISQVYFMSSCNQYFVLIHLIW					
	180	190	200	210	220	230
	250	260	270	280	290	300
MAELO	APNVGKCACTEGAALFGCGLLSYLLLFIFYRITYNAKAKAAKERGSNFTPKTVKSGGS					
	: : : : : : : : : : : : : :					
HS2	-----MYG-----TIFFMLFSNFWYHSYTKGKRLPRALQONGAPGLAKVKAN					
	240	250	260	270		
	310					
MAELO	PKKPSKSKHI					
HS2	X					
	280					

Figura 31

REGISTOS : Inic1: 88 Inin: 207 Opt: 223

Registo de Smith-Waterman: 236; 30,4% de identidade numa sobreposição de 191 aa

	100	110	120	130	140	150
MAELO	LLLLFIENLVPTILARNGLFYAICDDGAWTQRLLELLYYLVNVLVKYWEADTVFLVLKKKP-					
				:	:	:
MM2	IVYEFLMSGWLSSTYTWRCDPIDFSNSPEALRMVRVAWLFLSKVIELMDTVIFILRKKDG					
	20	30	40	50	60	70
	160	170	180	190	200	209
MAELO	-LEFLHYFHHSMIMVLCF----VOLGGYTSVSWVPITLNLTVHVFMYVYMRSAAGV----					
	:		:	:	:	:
MM2	QVTFLHVFHHSVLPWSWWGKIAPGGMGSFHAM---INSSVHVVMYLYYGLSALGPVAQ					
	80	90	100	110	120	130
	210	220	230	240	250	260
MAELO	-RIWWKQYLTTLQIVQFVL---DLGFIYF---CAYTYFAFTYFFWAPNVGKCAGTEGAAL					
	:		:	:	:	:
MM2	PYLWKKHMTAIQLIQFVLVSLHISQYFMPSCNYQYFVLIHLIW-----M					
	140	150	160	170		
	270	280	290	300	310	
MAELO	FGCGLLSYLLLFINFYRITYNAKAKAAKERGSNFTPKTVKSGGSPKKPSKSKI					
	:		:	:	:	:
MM2	YG---TIFFILFSNFWYHSYTKGKRLPRAVQONGAPATTKVKAN					
	180	190	200	210		

Figura 32

REGISTOS Inic1: 51 Inin: 115 Opt: 168
 Registo de Smith-Waterman: 168; 30,4% de identidade numa sobreposição de 115 aa

	30	40	50	60	70	80
MAELO	YELVTGKSIDSFVFOEGVTPLSTQREVAMWTITYFVVIFGGRQIMKSQDAFKLKPLFILH					
AI225632	NAFLDNMFGPRDSRVRGWFLLD SYLPTFILTIYLLSIWLGNKYMKNRPALSLRGITILY					
	20	30	40	50	60	70
	90	100	110	120	130	140
MAELO	NFLLTIASGSLLLLFIEENLVPIARNGLFYAICDD----GAWTORLELLYYLNYLVKYWE					
AI225632	NLAITLLSAYMLVELI-----LSSWEGGYNLQCONLDSAGEGDVRVAKVLVWYVFSKLVE					
	80	90	100	110	120	
	150	160	170	180	190	200
MAELO	LADTVFLVLKKK--PLEFLHYFHHSMVMVLCFVQLGGYTSVSWVPITLNLTVHVFMYYYY					
AI225632	FLDTIFFVLRKKANQITFLHVYHHASMFNI					
	130	140	150			

Figura 33

REGISTOS Quadro: (3) Inic: 332 Inin: 332 Opt: 384

40,3% de identidade numa sobreposição de 144 aa

	80	90	100	110	120	130
GLELO	LIVLAYLVTVFVGMQIMKNFERFEVKTFSLLENFCLVSI SAYMCGGILYEAYQANYGL-F					
AI815960	: : : : : : : : : :					
					10	20
						30
	140	150	160	170	180	190
GLELO	ENAADHTFKGLPMAKMIWLFYFSKIMEFVDTMIMVLKKNRQISFLHVVYHHSSIFTIWWL					
AI815960	: : : : : : : : : : : : : : : : :					
	40	50	60	70	80	90
	200	210	220	230	240	250
GLELO	VTFVAPNGEAYFSAALNSFIHVIMYGYFLSAL-GFKQVSFIKFIYITRSQMTQFCMMSVQ					
AI815960	: : : : : : : : : : : :					
	100	110	120	130	140	
	260	270	280	290	300	310
GLELO	SSWDMYAMKVLGRPGYPFFITALLWFYMWMTMLGLFYNFYRKNKAKLAKQAKADAAKEKARK					

Figura 34

.REGISTOS Inic1: 316 Inin: 384 Opt: 477

.Registo de Smith-Waterman: 477; 34,2% de identidade numa sobreposição de 240 aa

	50	60	70	80	90	100
GLELO	AQAEKYIPTIVHHTRGFLVAVESPLARELPLMNPFFHVLLIVLAYLVTVFVGMQIMKNFER					
	: : : : : : : : : :					
HS1	MEHFDASLSTYFKALLGPRDTRVKGWFLLDNYIPTFICSVIYLLIVWLGPKYMRNKQP					
	10	20	30	40	50	
	110	120	130	140	150	159
GLELO	FEVKTFSLHNFCLVSISAYMCGGILYEAYQANYGLF-ENADHTFKGLPMAKMIWLFYF					
	: : : : : : : : : : : : : : : : : : : : :					
HS1	FSCRGILVVYNLGLTLLSLYMFCELVTVGWEGKYNFFCQGTRTAGESDMKIIIRVLWWYYF					
	60	70	80	90	100	110
	160	170	180	190	200	210
GLELO	SKIMEFVDTMIMVLKKNRQISFLHVVYHHSSIIFTIWWLVTFFVAPNGEAYFSAALNSFIHV					
	: : : : : : : : : : : : : : : : : :					
HS1	SKLIEFMDTFFFILRKNNHQITVLHVVYHHASMLNIWWFVMNWVPCGHSYFGATLNSFIHV					
	120	130	140	150	160	170
	220	230	240	250	260	270
GLELO	IMYGYFFLSAL-GFKQVSFIKFYITRSQMTQFCMMSVQSSWDMYAMKVLGRPGYPFFITA					
	: : : : : : : : : : : : : : : : : : : :					
HS1	LMYSYGLSSVPSMRPYLWKKYITQGQLLQFVLTIIQTS-----CGVIWPCTFPLGWLY					
	180	190	200	210	220	230
	280	290	300	310		
GLELO	LLWFYMWMTMLGLFYNFYRK--NAKLAKQAKADAAKEKARKLQ					
	: : : : : : : : : :					
HS1	FQIGYMISLIALFTNFYIQTYNKKGASRRKDHLKDHQNGSMAAVNGHTNSFSPLENNVKP					
	240	250	260	270	280	290

Figura 35

REGISTOS Inic1: 80 Inin: 114 Opt: 178
 Registo de Smith-Waterman: 178; 28,8% de identidade numa sobreposição de 146 aa

	140	150	160	170	180	190
GLELO	FENAADHTFKGLPMAKMIWLFYFSKIMEFVDTMIMVLKKNNRQISFLHVYHHSSIFTIWW					
AC004050	DTIFIILRK--QKLIFLHWYHHITVLLYSW					
				10	20	

	200	210	220	230	240	250
GLELO	LVTFFVAPNGEAYFSAALNSFIHVIMYGYFELSALGFKQVSFIKFYITRSQMTQFCMMSVQ					
AC004050	YSYKDMVAGGGWF-MTMNYGVHAVMYSYVALRAAGFRVSRKFAMFITLSQITQMLMGCVV					
	30	40	50	60	70	80

	260	270	280	290	300	310
GLELO	SSWDMYAMKVLGRPGYFFFITALLW--FYMWTMLGLFYNFYRKN--AKLAKQAKADAAKE					
AC004050	NYLVFCWMQ--HDQCHSHF-QNIFWSSLMYLSYLVLFCHFFFEAYIGMRKTTKAEX					
	90	100	110	120	130	140

GLELO	KARKLQ
-------	--------

Figura 36

REGISTOS Inic1: 288 Inin: 288 Opt: 399

Registo de Smith-Waterman: 399; 34,6% de identidade numa sobreposição de 211 aa

	80	90	100	110	120	130
GLELO	LLIVLAYLVTVFVGMQIMKNFERFEVKTFSLLHNFCLVSI SAYMCGGILYEAYQANYGLF					
	:: : : :: ::					
MM2	IVYNFSLVILSLYIVYEFLMSGWLSTYTW					
			10		20	30
	140	150	160	170	180	190
GLELO	ENAAD--HTFKGLPMAMWLFYFSKIMEFVDTMIMVLKKNNRQLSFLHVYHHSSIFTIW					
	: :: :: :: :: :: :: :: :: :: :: :: :: :: ::					
MM2	CDPIDFSNSPEALRMVRVAWLEMLSKVIELMDTVIFILRKKGQVTFHVFHHSVLPWSW					
	40	50	60	70	80	90
	200	210	220	230	240	250
GLELO	WLVTVPAPNGEAYFSAALNSFIHVIMYGYFLSALGFKQVSFI--KFYITRSQMTQFCMM					
	:: : : : : : :: : : ::					
MM2	WWGIKIAPGGMGSFHAMINSSVHVVMYLYGLSALGPVAQPYLWKKHMTAIQLIQFVLV					
	100	110	120	130	140	150
	260	270	280	290	300	309
GLELO	SVQSSWDMYAMKVLGRPGYPFFITALLWFYMWTLGLFYNF----YRKNAKLAKQAKADA					
	:: :: : : : : :: : : :: ::					
MM2	SLHIS-QYYFMPSCNYQ-YPVITH-LIWMYGTIFFILFSNFWYHSYTKGKRLPRAVQQNG					
	160	170	180	190	200	
	310					
GLELO	AKEKARKLQ					
MM2	APATTKVKAN					
	210					

Figura 38

,REGISTOS Inic1: 64 Inin: 129 Opt: 233

, Registo de Smith-Waterman: 239; 23,7% de identidade numa sobreposição de 279 aa

	20	30	40	50	60	70
GLELO	FMDLATAIGVRAAPYVDPLEAALVAQAEKYIPTIVHHTRGFLVAVESPLAREL-----PL					
U97107	MDTSMNFSRGLKMDLMQPYDFETFDLRF					
			10	20	30	
	80	90	100	110	120	129
GLELO	MNPFFHV--LLIVLAYLVTVFVGMQIMKNFERFEVKTFSLLHNFCLVSISAYMCGGILYEA					
U97107	LEEYWVSSFLIVVVYLLLVVGQTYMTRKSFSLQRPILWSSFPLAIFS--ILGTLRMWK					
	40	50	60	70	80	
	130	140	150	160	170	180
GLELO	YQAN----YGLFENAADHTFKGLPMAMKIW--LFYFSKIMEFVDTMIMVLKQNNRQISFL					
U97107	FMAIVMFTVGLKQTVCFAYTDDAVVRF--WSFLFLLSKVVELGDTAFIILRK--RPLIFV					
	90	100	110	120	130	140
	190	200	210	220	230	240
GLELO	HVYHHSSI--FTIWWLVTFVAPNGEAYFSAALNSFIHVIMYGYFSLALGFRQVSVFIKPY					
U97107	HWYHHSTVLLFTSFGYKNKV-PSGGWFMT--MNEGVHSMYTYTMTKAAKLKHPNLLPMV					
	150	160	170	180	190	200
	250	260	270	280	290	
GLELO	ITRSQMTQFCMMSVQSSWDMYAMKVLG--RPGYPFFITALLWFYMWMTMLGLFYN--FYRK					
U97107	ITSLQILQMVLTIFGILNYIWRQEKGCHTTTEHFFWSFMLYGTIFILFAHFFHRAYLRLP					
	210	220	230	240	250	260
	300	310				
GLELO	NAKLAKQAKADAAKEKARKLQ					
U97107	KGKVASKSQ					
	270					

Figura 39

REGISTOS Inic1: 100 Inin: 205 Opt: 271
 Registo de Smith-Waterman: 271; 30,7% de identidade numa sobreposição de 218 aa

	60	70	80	90	100	110
GLELO	TRGFLVAVESPLARELPLMNPFFHVLLIVLAYLVTVFVGMQIMKNFERFEVKTFSLLHNFC					
				: :	: :	: :
U68749	ATHGPKNFPDAEGRKFFADHFDVTIQASILYMVVVFGTKWFMNRNRPQFQLTIPLNINWFI					
(F56H11.4)	30	40	50	60	70	80
	120	130	140	150	160	
GLELO	LVSISAYMCGGILYEAYQ--ANYGL---FENAADHTFKGLPMAKMTWLFYFSKIMEFVDT					
	::	:	:	:	: :	: :
U68749	LAAFSIAGAVKMTPEFFGTIANKGIVASYCKVDFDT-KG-ENGYVWVWLFMASKLFELVDT					
(F56H11.4)	90	100	110	120	130	140
	170	180	190	200	210	220
GLELO	MIMVLKKNRQISFLHVYHHSSIFTIWWLVTFVAPNGEAYFSAALNSFIHVIMYGYFVLS					
	::: :	:			:	:
U68749	IFLVLRK--RPLMFLHWYHHILTMIYAWYSHPLTP-GFNRYGIYLNFEVVFHAFMYSYYFLR					
(F56H11.4)	150	160	170	180	190	
	230	240	250	260	270	280
GLELO	ALGFKQVSFIKFYITRSQMTQFCMMSVQSSWDMYAMKVLGRP-GYPFFITALLWFYMWTM					
	:::	: :		: :	:	:
U68749	SMKIRVPGFIAQAITSLOIVQFIISCAVLAHLGYLMHFTNANCDFEPSVFKLAVFMDTTY					
(F56H11.4)	200	210	220	230	240	250
	290	300	310			
GLELO	LGLFYNFYRKNAKLAKQAKADAAKEKARKLQ					
	:	:				
U68749	LALFVNFFLQSYVLRGGGDKYKAVPKKKNN					
(F56H11.4)	260	270	280			

Figura 41

REGISTOS Inic1: 77 Inin: 155 Opt: 264
 Registo de Smith-Waterman: 264; 27,2% de identidade numa sobreposição de 206 aa

	40	50	60	70	80	90
GLELO	AALVAQAQEKYIPTIVHHTRGFLVAVESPLARELPLMNPFHVLLIVLAYLVTVFVGMQIMK					
				: :	::: ::: :	
DM1	PTKMINMDISVTPNYSYIFDFENDFIHQTRKWMLNWTWVFYCYGIMLVIFGGQHF MQ					
	10	20	30	40	50	60
	100	110	120	130	140	150
GLELO	NFERFEVKTFSLHNFCLVSISAYMCGGILYEAYQA--NYGLFENAADHTF--KGLPMAK					
	::: :: : :	: :	: :	: :	: :	: :
DM1	NRPRFQLRGPLIIWNTLLAMFSIMGAARTAPELIHVLRHYGLFHSVCVPSYIEQDRVCGF					
	70	80	90	100	110	120
	160	170	180	190	200	210
GLELO	MIWLFYFSKIMEFVDTMIMVLKKNNRQISFLHVYHHSSIFTIWWLVTFVAPNGEAYFSAA					
	: : :	: ::: :	: :	: :	: :	: :
DM1	WTWLFVLSKLPGLGDTIFIVLRK--QPLIFLHWYHHITVLIYSWF-SYTEYTSSARWFIV					
	130	140	150	160	170	180
	220	230	240	250	260	270
GLELO	LNSFIHVIMYGYFFLSALGFKQVSFIKFYITRSQMTQFCMMSVQSSWDMYAMKVLGRPGY					
	: : : : :	: :	: :	: :	: :	: :
DM1	MNYCVHSVMYSYYALKAAARNPPRFISMIITSLQLAQMIIGCAINWANGFLKTHGTXSC					
	190	200	210	220	230	240
	280	290	300	310		
GLELO	PFFITALLWFYMWMTMLGLFYNFYRKNAKLAKQAKADAAKEKARKLQ					
DM1	HISQRNINLSIAMYSSYFVLFARFFYKAYLAPGGHKSRMA					
	250	260	270	280		

Figura 43

```

1   ATGGAACATT TTGATGCATC ACTTAGTACC TATTTCAAGG CATTGCTAGG
51  CCCTCGAGAT ACTAGAGTAA AAGGATGGTT TCTTCTGGAC AATTATATAC
101 CCACATTTAT CTGCTCTGTC ATATATTTAC TAATTGTATG GCTGGGACCA
151 AAATACATGA GGAATAAACA GCCATTCTCT TGCCGGGGGA TTTTAGTGGT
201 GTATAACCTT GGA CTCACAC TGCTGTCTCT GTATATGTTT TGTGAGTTAG
251 TAACAGGAGT ATG GGAAGGC AAATACA ACT TCTTCTGTCA GGGCACACGC
301 ACCGCAGGAG AATCAGATAT GAAGATTATC CGTGTCTCT GGTGGTACTA
351 CTTCTCCAAA CTCATAGAAT TTATGGACAC TTTCTTCTTC ATCCTGCGCA
401 AGAACAACCA CCAGATCACG GTCCTGCACG TCTACCACCA TGCCCTCGATG
451 CTGAACATCT GGTGGTTTGT GATGAACTGG GTCCCTGCG GCCACTCTTA
501 TTTTGGTGCC AACTTAATA GCTTCATCCA CGTCTCATG TACTCTTACT
551 ATGGTTTGTC GTCAGTCCCT TCCATGCGTC CATACTCTG GTGGAAGAAG
601 TACATCACTC AGGGGCAGCT GCTTCAGTTT GTGCTGACAA TCATCCAGAC
651 CAGCTGCGGG GTCATCTGGC CGTGCACATT CCCTCTTGGT TGGTTGTATT
701 TCCAGATTGG ATACATGATT TCCCTGATTG CTCTCTTCAC AAACCTCTAC
751 ATTCAGACCT ACAACAAGAA AGGGGCCTCC CGAAGGAAAG ACCACCTGAA
801 GGACCACCAG AATGGGTCCA TGGCTGCTGT GAATGGACAC ACCAACAGCT
851 TTTCACCCCT GGAAAACAAT GTGAAGCCAA GGAAGCTGCG GAAGGATTGA
901 AGTCAAAGAA TTGA

```

Figura 44

```
1  MEHFDASLST YFKALLGPRD TRVKGWFLLD NYIPTFICSV IYLLIVWLGP
51  KYMRNKQPF S CRGILVVYNL GLTLLSLYMF CELVTGWEG KYNFFCQGR
101  TAGESDMKII RVLWWYYFSK LIEFMDTFFF ILRKNNHQIT VLVVYHHASM
151  LNIWVFMNW VPCGHSYFGA TLNSFIHVLN YSYYGLSSVP SMRPYLWKK
201  YITQGQLLQF VLTIIQTSCG VIWPCTFPLG WLYFQIGYMI SLIALFTNFY
251  IQTYNKKGAS RRKDHLKDQ NGSMMAVNGH TNSFSPLENN VKPRKLRKD*
```

Figura 45

Hospedeiro (plasmídeo)	334(pYX242)	334(pRAE-58-A1)	334(pYX242)	334(pRAE-58-A1)
Substrato adicionado	25 µM GLA	25 µM GLA	25 µM AA	25 µM AA
Ácido gordo	% de ácido gordo total	% de ácido gordo total	% de ácido gordo total	% de ácido gordo total
C18:3n-6	4.40	2.71	0.03	0.04
C20:3n-6	0.09	(50.34%)* 2.75	0.02	0.02
C20:4n-6			7.84	3.97
C22:4n-6			ND	(23.37%)* 1.21
C16:1n-7	41.11	34.72	41.49	35.07
C18:1n-7	1.85	11.33	2.01	11.57
C20:1n-7	0.04	1.48	0.04	1.62
C18:1n-9	15.60	15.66	15.16	14.57
C20:1n-9	0.06	0.22	0.06	0.23
C18:1n-5	0.11	0.62	0.12	0.58
Lípidos Totais	370	969	359	514
*(% de conversão) = produto/(substrato+produto)				

Figura 46

```

1  ATGGCTCAGC ATCCGCTCGT TCAACGGCTT CTCGATGTCA AATTCGACAC
51  GAAACGATTT GTGGCTATTG CTACTCATGG GCCAAAGAAT TTCCCTGACG
101 CAGAAGGTCG CAAGTTCCTT GCTGATCACT TTGATGTTAC TATTCAGGCT
151 TCAATCCTGT ACATGGTCGT TGTGTTCCGA ACAAATGGT TCATGCGTAA
201 TCGTCAACCA TTCCAATTGA CTATTCCACT CAACATCTGG AATTTTCATCC
251 TCGCCGCATT TTCCATCGCA GGAGCTGTCA AAATGACCCC AGAGTTCCTT
301 GGAACCATTG CCAACAAAGG AATTGTGCGA TCCTACTGCA AAGTGTGTTGA
351 TTTCACGAAA GGAGAGAATG GATACTGGGT GTGGCTCTTC ATGGCTTCCA
401 AACTTTTCGA ACTTGTTGAC ACCATCTTCT TGGTTCTCCG TAAACGTCCA
451 CTCATGTTCC TTCCTGGTA TCACCATATT CTCACCATGA TCTACGCCTG
501 GTACTCTCAT CCATTGACCC CAGGATTCAA CAGATACGGA ATTTATCTTA
551 ACTTTGTCGT CCACGCCTTC ATGTACTCTT ACTACTTCCT TCGCTCGATG
601 AAGATTCGCG TGCCAGGATT CATCGCCCAA GCTATCACAT CTCCTCAAAT
651 CGTTCAATTC ATCATCTCTT GCGCCGTTCT TGCTCATCTT GGTTATCTCA
701 TGCACCTCAC CAATGCCAAC TGTGATTTG AGCCATCAGT ATTCAAGCTC
751 GCAGTTTTCA TGGACACAAC ATACTTGGCT CTTTTCGTCA ACTTCTTCCT
801 CCAATCATAT GTTCTCCGCG GAGGAAAAGA CAAGTACAAG GCAGTGCCAA
851 AGAAGAAGAA CAACTAA

```

Figura 47

1 MAQHPLVQRL LDVKFDTKRF VAIATHGPKN FPDAEGRKFF ADHFDVTIQA
51 SILYVVVFG TKWFMRNRQP FQLTIPLNIW NFILAAFSIA GAVKMTPEFF
101 GTIANKGIVA SYCKVFDFTK GENGYVWVLF MASKLFELVD TIFLVLKRP
151 LMFLHWYHHI LTMIAWYSH PLTPGFNRYG IYLNFFVHAF MYSYYFLRSM
201 KIRVPGFIAQ AITSLQIVQF IISCAVLAHL GYLMHFTNAN CDFEPSVFKL
251 AVFMDTTYLA LFVNFFLQSY VLRGGKDKYK AVPKKKNN

Figura 48

Hospedeiro (plasmídeo)	334(pYX242)	334(pRET-21)	334(pRET-22)
Substrato adicionado	50 μ M GLA +	50 μ M GLA +	50 μ M GLA +
	50 μ M AA	50 μ M AA	50 μ M AA
Ácido gordo	% de ácido gordo total	% de ácido gordo total	% de ácido gordo total
C16:0	9.22	12.46	9.9
C16:1	0.09	0.18	0.13
C18:0	1.46	2.41	1.49
C18:1n-9	4.03	4.92	3.91
C18:3n-6	10.02	11.89	8.69
C20:3n-6	(1.28%)* 0.13	(11.1%)* 1.48	(19.4%)* 2.09
C20:4n-6	46.98	28.87	35.25
C22:4n-6	0	0	0
* Lípidos Totais (mg)	212	174	187
*(% de conversão) = produto/(substrato+produto)			

Figura 49

```
1  ATGAACATGT CAGTGTTGAC TTTACAAGAA TATGAATTCG AAAAGCAGTT
51 CAACGAGAAT GAAGCCATCC AATGGATGCA GGAAAAC TGG AAGAAATCTT
101 TCCTGTTTTT TGCTCTGTAT GCTGCCTTTA TATTCGGTGG TCGGCACCTA
151 ATGAATAAAC GAGCAAAGTT TGAAGTGAAG AAGCCATTAG TGCTCTGGTC
201 TCTGACCCTT GCAGTCTTCA GTATATTCGG TGCTCTTCGA ACTGGTGCTT
251 ATATGGTGTA CATTTTGATG ACCAAAGGCC TGAAGCAGTC AGTTTGTGAC
301 CAGGGTTTTT ACAATGGACC TGTCAGCAA TTCTGGGCTT ATGCATTGT
351 GCTAAGCAA GCACCCGAAC TAGGAGATAC AATATTCATT ATTCTGAGGA
401 AGCAGAAGCT GATCTTCCTG CACTGGTATC ACCACATCAC TGTGCTCCTG
451 TACTCTTGGT ACTCCTACAA AGACATGGT GCCGGGGGAG GTTGGTTCAT
501 GACTATGAAC TATGGCGTGC ACGCCGTGAT GTACTCTTAC TATGCCTTGC
551 GGGCGGCAGG TTTCCGAGTC TCCCGGAAGT TTGCCATGTT CATCACCTTG
601 TCCCAGATCA CTCAGATGCT GATGGGCTGT GTGGTTAACT ACCTGGTCTT
651 CTGCTGGATG CAGCATGACC AGTGTCAC TC ACTTTCAG AACATCTTCT
701 GGTCCCTACT CATGTACCTC AGCTACCTTG TGCTCTTCTG CCATTTCTTC
751 TTTGAGGCCT ACATCGGCAA AATGAGGAAA ACAACGAAAG CTGAATAG
```

Figura 50

1 MNMSVLTQLQ YEFKQFNEN EAIQWMQENW KKSFLFSALY AAFIFGGRHL
51 MNKRAKFEELR KPLVLWSLTL AVFSIFGALR TGAYMVYILM TKGLKQSVCD
101 QGFYNGFVSK FWAYAFVLSK APELGDTIFI ILRRQKLIFL HWYHHITVLL
151 YSWYSYKDMV AGGGWFMTMN YGVHAVMYSY YALRAAGFRV SRKFAMFITL
201 SQITQMLMGC VVNYLVFCWM QHDQCHSHFQ NIFWSSLMYL SYLVLFCHFF
251 FEAYIGKMRK TTKAE*