



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)  
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0048945

(51)<sup>2020.01</sup> C12N 15/52; A23L 33/135; C12R 1/01; (13) B  
C12N 1/20; A23K 10/10; C07K 14/195

(21) 1-2022-01364

(22) 22/10/2020

(86) PCT/FI2020/050699 22/10/2020

(87) WO2021/084159 06/05/2021

(30) 19205786.7 29/10/2019 EP

(45) 25/07/2025 448

(43) 25/08/2022 413A

(73) Solar Foods Oyj (FI)

Harkalankki 11, 01730 Vantaa, Finland

(72) Sami HOLMSTRÖM (FI); Juha-Pekka PITKÄNEN (FI).

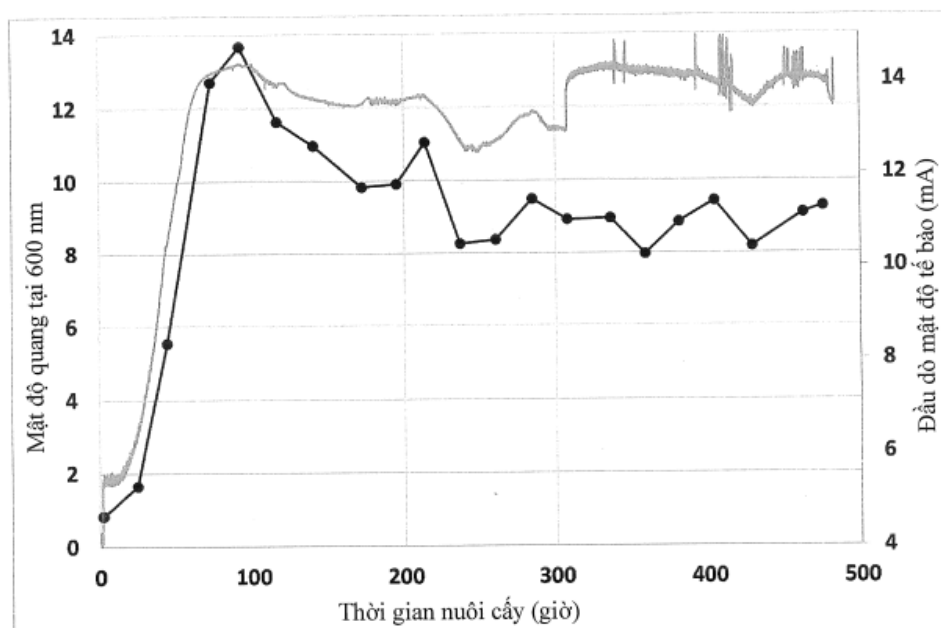
(74) Văn phòng Luật sư Ân Nam (ANNAM IP & LAW)

(54) CHŨNG VI KHUẨN ĐƯỢC PHÂN LẬP VTT-E-193585 VÀ MÔI TRƯỜNG  
NUÔI CÂY CHỨA CHŨNG NÀY

(21) 1-2022-01364

(57) Sáng chế đề cập đến chủng vi khuẩn thuộc chi *Xanthobacter* và quy trình nuôi cấy liên tục để sản xuất protein hoặc sinh khối bằng cách sử dụng vi khuẩn thuộc chi *Xanthobacter*, quy trình đã nêu bao gồm bước cung cấp khí và khoáng chất cho tế bào. Sáng chế cũng đề cập đến sản phẩm của các quy trình này và việc sử dụng các sản phẩm này ví dụ như trong lương thực hoặc thực phẩm.

**Fig. 1**



### **Lĩnh vực kỹ thuật được đề cập**

Sáng chế đề cập đến việc sản xuất protein và/hoặc đại phân tử khác sử dụng vi sinh vật. Cụ thể là, sáng chế đề cập đến chủng vi khuẩn mới và quy trình nuôi cấy liên tục để sản xuất protein hoặc sinh khối sử dụng vi khuẩn trong đó khí và khoáng chất được cung cấp cho tế bào. Sáng chế cũng đề cập đến sản phẩm của các quy trình này và việc sử dụng các sản phẩm này ví dụ như trong lương thực hoặc thực phẩm.

### **Tình trạng kỹ thuật của sáng chế**

Dân số thế giới phát triển, biến đổi khí hậu và sự thiếu hụt nước gây ra sự đe dọa đáng kể đến nông nghiệp truyền thống và do đó đe dọa việc cung cấp đủ lương thực và thực phẩm. Do đó, các nguồn thay thế của phân tử hữu cơ, chẳng hạn như protein, đang được nghiên cứu. Phương pháp thay thế tiềm năng là việc sản xuất tế bào đơn lẻ, tức là sản xuất protein và/hoặc đại phân tử khác sử dụng vi sinh vật.

Vi sinh vật hóa dưỡng được mô tả là có thể sinh trưởng trên môi trường khoáng chất tối thiểu với khí hydro làm nguồn năng lượng và cacbon dioxid làm nguồn cacbon duy nhất. Để xem xét các vi sinh vật này, xem ví dụ, Shively et al. (1998) *Annu Rev Microbiol* 52:191. Công bố sáng chế WO2018144965 mô tả các vi sinh vật và quy trình sinh học khác nhau để chuyển hóa chất khí thành sinh khối có hàm lượng protein cao. Andersen et al. (1979) *Biochim Biophys Acta* 585:1-11 mô tả chủng đột biến của *Alcaligenes eutrophus*, là vi khuẩn hydro sinh trưởng dễ dàng trong điều kiện dị dưỡng và tự dưỡng. Thể đột biến đặc trưng ở hoạt tính ribuloza-1,5-bisphosphat carboxylaza/oxygenaza (rubisco) bị thay đổi. Ohmiya et al. (2003) *J. Biosci. Bioeng.* 95:549-561 xem xét ứng dụng của gen vi sinh vật để chống lại việc sử dụng sinh khối. Yu Jian et al. (2013) *Int J hydro Ener* 38:8683-8690 mô tả việc cố định cacbon dioxid bằng thể phân lập vi khuẩn oxy hóa hydro. Hiệu quả năng lượng cao bằng 50% được đo ở nồng độ oxy trung bình (10 mol%).

Tuy nhiên, nhiều vi sinh vật hóa dưỡng khác nhau có đặc tính khác nhau về tốc độ sinh trưởng, hiệu suất, thành phần sinh khối cũng như các tính chất liên quan để được sử dụng làm thành phần thực phẩm chẳng hạn như độ an toàn cho việc tiêu thụ của con người, vị, mùi, cảm nhận trong miệng, đặc tính kỹ thuật và chức năng khi nấu, v.v.. Không phải mọi vi sinh vật hóa dưỡng đều có tốc độ sinh trưởng đủ và tạo ra hiệu suất

đủ và không phải mọi quy trình đều có thể được tăng quy mô trên thực tế thành quy trình có quy mô lớn khả thi về kinh tế. Để có đầu ra đủ của protein có chức năng, ví dụ như đối với ứng dụng lương thực hoặc thực phẩm, điều quan trọng là tìm ra sinh vật sản xuất thích hợp và quy trình thích hợp mà có thể được thực hiện ở quy mô lớn. Nhu cầu này được đáp ứng bởi sáng chế này.

### **Bản chất kỹ thuật của sáng chế**

Theo khía cạnh chính thứ nhất, sáng chế đề xuất chủng vi khuẩn được phân lập VTT-E-193585 hoặc dẫn xuất của chúng.

Theo các khía cạnh khác, sáng chế đề xuất môi trường nuôi cấy có chứa chủng vi khuẩn theo sáng chế hoặc dẫn xuất của chúng. Ngoài ra, sáng chế đề xuất quy trình sản xuất sinh khối và/hoặc protein, quy trình đã nêu bao gồm bước nuôi cấy chủng vi khuẩn theo sáng chế hoặc dẫn xuất của chúng.

Theo khía cạnh khác, sáng chế đề xuất quy trình sản xuất sinh khối và/hoặc protein, quy trình đã nêu bao gồm bước nuôi cấy chủng vi khuẩn thuộc chi *Xanthobacter* trong nuôi cấy liên tục với hydro làm nguồn năng lượng và nguồn cacbon vô cơ, trong đó nguồn cacbon vô cơ có chứa cacbon dioxit.

Theo các khía cạnh chính khác, sáng chế đề xuất protein khối, sinh khối hoặc thành phần tế bào không phải protein hoặc thành phần hóa học thu được hoặc có thể thu được bằng quy trình theo sáng chế, và đề cập đến sản phẩm lương thực hoặc thực phẩm thu được hoặc có thể thu được bằng quy trình theo sáng chế.

### **Mô tả vắn tắt các hình vẽ**

Fig. 1. Mật độ quang học được đo tại 600 nm (hình tròn màu đen) và kết quả đọc đầu dò mật độ quang học trong quá trình nuôi cấy 200-L hóa dưỡng của chủng vi khuẩn được phân lập được nộp lưu dưới dạng VTT-E-193585.

Fig. 2. Mật độ quang học được đo tại 600 nm trong quá trình nuôi cấy 200-mL hóa dưỡng song song của chủng vi khuẩn được phân lập được nộp lưu dưới dạng VTT-E-193585 trên các nguồn nitơ khác nhau.

### **Mô tả chi tiết sáng chế**

#### **Định nghĩa**

Khi được sử dụng trong bản mô tả này, thuật ngữ “được phân lập”, ví dụ, trong ngữ cảnh về chủng, có nghĩa là được phân lập từ môi trường tự nhiên của nó. Tốt hơn là, chủng được phân lập là tinh khiết, tức là không có các chủng khác.

Thuật ngữ “dẫn xuất”, khi được sử dụng trong bản mô tả này trong ngữ cảnh về chủng, dùng để chỉ chủng có nguồn gốc từ chủng tham chiếu, tức là được tạo ra bằng cách sử dụng chủng tham chiếu làm điểm bắt đầu. Ví dụ như chủng được thiết kế di truyền hoặc theo cách khác là bị đột biến hoặc được cải biến di truyền là một phương án của dẫn xuất này. Cải biến di truyền bao gồm đột biến điểm, cũng như là đột biến thêm hoặc mất, bao gồm thêm hoặc mất của toàn bộ locut hoặc mảnh của chúng. Dẫn xuất tốt hơn nếu có ít hơn 10 cải biến di truyền, ví dụ như ít hơn 5, chẳng hạn như 4, 3, 2 hoặc 1 (các) cải biến di truyền so với chủng tham chiếu.

Khi được sử dụng trong bản mô tả này, danh từ “môi trường nuôi cấy” dùng để chỉ huyền phù của các tế bào sống trong môi trường lỏng.

Thuật ngữ “sinh khối” có nghĩa thông thường của nó trong lĩnh vực lên men vi khuẩn và dùng để chỉ nguyên liệu tế bào.

Thuật ngữ “nuôi cấy liên tục”, khi được sử dụng trong bản mô tả này, dùng để chỉ quy trình nuôi cấy trong đó môi trường mới được bổ sung liên tục vào môi trường nuôi cấy và môi trường có môi trường nuôi cấy vi khuẩn được loại bỏ liên tục gần như ở cùng tốc độ.

Các khía cạnh và phương án của sáng chế

Theo khía cạnh chính thứ nhất, sáng chế đề xuất chủng vi khuẩn được phân lập VTT-E-193585 hoặc dẫn xuất của chúng.

Chủng VTT-E-193585 được phân lập từ bờ biển của biển Baltic ở Naantali, Phần Lan. Sinh vật này có thể sinh trưởng trong điều kiện bình phản ứng sinh học thích hợp với môi trường khoáng chất tối thiểu với hydro làm nguồn năng lượng và cacbon dioxide làm nguồn cacbon ở điều kiện oxy bị hạn chế. Giải trình tự 16S và giải trình tự metagenom Illumina cho thấy rằng chủng này có khả năng nhất là thành viên của chi *Xanthobacter*, chứ không phải là loài đã biết. Chủng vi khuẩn này rất thích hợp đối với ứng dụng lương thực và thực phẩm, vì bột tế bào được làm khô có hàm lượng protein cao và chứa tất cả các axit amin thiết yếu. Nó cũng chứa nhiều axit béo không bão hòa hơn là axit béo bão hòa và hàm lượng vitamin nhóm B cao. Hàm lượng peptidoglycan và lipopolysaccharit, mà có thể gây ra dị ứng hoặc gây độc, là thấp. Phân tích độc tố được thực hiện và không quan sát thấy sự gây độc gen hoặc gây độc tế bào nào đối với chủng này. Ngoài ra, chủng này thường nhạy đối với chất kháng sinh.

Chủng VTT-E-193585 (SoF1) được nộp lưu vào ngày 11 tháng 6 năm 2019 trong

Bộ Sưu Tập Nuôi Cây VTT tại Trung Tâm Nghiên Cứu Kỹ Thuật VTT ở Phần Lan, P.O. Box 1000, FI-02044 VTT, Finland, Cơ Quan Có Thảm Quyền Nộp Lưu Quốc Tế theo Hiệp Ước Budapest. Thông tin thêm về đặc điểm của chủng này và phương pháp nuôi cấy chủng này được đề xuất trong các Ví Dụ ở đây.

Theo một phương án được ưu tiên, nếu chủng này là dẫn xuất của chủng VTT-E-193585, dẫn xuất này giữ lại khả năng sinh trưởng bằng cách sử dụng khí hydro làm nguồn năng lượng và cacbon dioxid làm nguồn cacbon duy nhất.

Theo một phương án, nếu chủng này là dẫn xuất của chủng VTT-E-193585, dẫn xuất có chứa ARN ribosom 16S nêu trong SEQ ID NO:1 hoặc ARN ribosom 16S có đến 20 khác biệt nucleotit so với SEQ ID NO:1, ví dụ từ 1 đến 10, chẳng hạn từ 1 đến 5, ví dụ một, hai hoặc ba khác biệt nucleotit so với SEQ ID NO:1.

SEQ ID NO:1. Trình tự ARN ribosom 16S của chủng VTT-E-193585:

CTTGAGAGTTTGATCCTGGCTCAGAGCGAACGCTGGCGGCAGGCCTAACA  
CATGCAAGTCGAGCGCCCAGCAATGGGAGCGGCAGACGGGTGAGTAACGC  
GTGGGGATGTGCCCAATGGTACGGAATAACCCAGGGAACTTGGACTAAT  
ACCGTATGAGCCCTTCGGGGGAAAGATTTATCGCCATTGGATCAACCCGCG  
TCTGATTAGCTAGTTGGTGGGGTAACGGCCCAAGGCGACGATCAGTA  
GCTGGTCTGAGAGGATGATCAGCCACACTGGGACTGAGACACGGCCCAGA  
CTCCTACGGGAGGCAGCAGTGGGGAATATTGGACAATGGGCGCAAGCCTG  
ATCCAGCCATGCCGCGTGTGTGATGAAGGCCTTAGGGTTGTAAAGCACTTT  
CGCCGGTGAAGATAATGACGGTAACCGGAGAAGAAGCCCCGGCTAACTTC  
GTGCCAGCAGCCGCGGTAATACGAAGGGGGCTAGCGTTGCTCGGAATCAC  
TGGGCGTAAAGCGCACGTAGGCGGATCGTTAAGTCAGGGGTGAAATCCTG  
GAGCTCAACTCCAGAACTGCCCTTGATACTGGCGACCTTGAGTTCGAGAGA  
GGTTGGTGGAACTGCGAGTGTAGAGGTGAAATTCGTAGATATTCGCAAGA  
ACACCAGTGGCGAAGGCGGCCAACTGGCTCGATACTGACGCTGAGGTGCG  
AAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAA  
ACGATGGATGCTAGCCGTTGGGCAGCTTGCTGTTCAAGTGGCGCAGCTAACG  
CATTAAAGCATCCCCGCCTGGGGAGTACGGTCGCAAGATTAAACTCAAAGG  
AATTGACGGGGGGCCCGCACAAAGCGGTGGAGCATGTGGTTTAATTCGAAGC  
AACGCGCAGAACCTTACCAGCCTTTGACATGGCAGGACGATTTCCAGAGA  
TGGATCTCTTCCAGCAATGGACCTGCACACAGGTGCTGCATGGCTGTCGTC

AGCTCGTGTCGTGAGATGTTGGGTAAAGTCCCGCAACGAGCGCAACCCTCG  
 CCTCTAGTTGCCAGCATTGAGTTGGGCACTCTAGAGGGACTGCCGGTGATA  
 AGCCGAGAGGAAGGTGGGGATGACGTCAAGTCCTCATGGCCCTTACGGGC  
 TGGGCTACACACGTGCTACAATGGTGGTGACAGTGGGATGCGAAAGGGCG  
 ACCTCTAGCAAATCTCCAAAAGCCATCTCAGTTCGGATTGTACTCTGCAAC  
 TCGAGTGCATGAAGTTGGAATCGCTAGTAATCGTGGATCAGCATGCCACG  
 GTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCATGGGAGTT  
 GGCTTTACCCGAAGGCGCTGCGCTAACCCGCAAGGGAGGCAGGCGACCAC  
 GGTAGGGTCAGCGACTGGGGTGAAGTCGTAACAAGGTAGCCGTAGGGGAA  
 CCTGCGGCTGGATCACCTCCTTT

Theo khía cạnh tiếp theo, sáng chế đề xuất môi trường nuôi cấy có chứa chủng vi khuẩn theo sáng chế hoặc dẫn xuất của chúng. Theo phương án được ưu tiên, thể tích của môi trường nuôi cấy là 100 mL hoặc nhiều hơn, ví dụ như 1 L hoặc nhiều hơn, chẳng hạn như 10 L hoặc nhiều hơn, ví dụ như 1.000 L hoặc nhiều hơn, chẳng hạn như 10.000 L hoặc nhiều hơn, ví dụ như 50.000 L hoặc nhiều hơn, chẳng hạn như 100.000 l hoặc nhiều hơn, ví dụ như 200.000 L hoặc nhiều hơn.

Theo khía cạnh khác, sáng chế đề xuất quy trình sản xuất sinh khối và/hoặc protein, quy trình đã nêu bao gồm bước nuôi cấy chủng vi khuẩn theo sáng chế hoặc dẫn xuất của chúng. Theo một phương án, quy trình này là để sản xuất sinh khối. Theo phương án khác, quy trình này là để sản xuất protein. Theo một phương án, quy trình đã nêu bao gồm bước nuôi cấy chủng trong nuôi cấy liên tục với hydro làm nguồn năng lượng và nguồn cacbon vô cơ, trong đó nguồn cacbon vô cơ có chứa cacbon dioxit. Theo phương án khác, quy trình này là để sản xuất sinh khối và bao gồm bước nuôi cấy chủng trong môi trường nuôi cấy liên tục với hydro làm nguồn năng lượng và nguồn cacbon vô cơ, trong đó nguồn cacbon vô cơ có chứa cacbon dioxit. Các phương án khác nhau của quy trình được mô tả bên dưới.

Theo khía cạnh chính khác, sáng chế đề xuất quy trình sản xuất sinh khối và/hoặc protein, quy trình đã nêu bao gồm bước nuôi cấy chủng vi khuẩn của chi *Xanthobacter* trong nuôi cấy liên tục với hydro làm nguồn năng lượng và nguồn cacbon vô cơ, trong đó nguồn cacbon vô cơ có chứa cacbon dioxit. Theo một phương án, quy trình này là để sản xuất sinh khối. Theo phương án khác, quy trình này là để sản xuất protein. Các phương án khác nhau của quy trình được mô tả bên dưới.

Theo trình tự hệ gen, chúng được nộp lưu dưới mã số VTT-E-193585 rất có thể sử dụng chu trình Calvin-Benson-Bassham để cố định cacbon trong đó phân tử cacbon đioxit được nối với chuỗi 5 cacbon của ribuloza 1,5-bisphosphat tạo thành hai phân tử glyxerat 3-phosphat. Điều này cho phép chúng tổng hợp tất cả các phân tử hữu cơ khác mà chúng cần để sinh trưởng. Năng lượng từ hydro đi vào tế bào rất có thể thông qua hydrogenaza khử  $\text{NAD}^+$  và/hoặc NiFeSe-hydrogenaza. Thực chất là phản ứng oxy hóa khử trong đó hydro ( $\text{H}_2$ ) được oxy hóa thành  $\text{H}^+$  và  $\text{NAD}^+$  bị khử thành NADH. Ngoài ATP, NADH là một trong số các chất mang năng lượng chính bên trong sinh vật sống. Theo cách khác, một số dạng tương đương năng lượng khác bị khử bằng enzym hydrogenaza khác bằng cách sử dụng  $\text{H}_2$ . Chu trình Calvin-Benson-Bassham cần năng lượng ở dạng ATP và NADH/NADPH để cố định  $\text{CO}_2$ . Chúng rất có thể tạo ra ATP thông qua quá trình phosphoryl hóa oxy hóa, chỉ bao gồm bốn phức hợp protein tạo ra gradien proton qua màng. Gradien proton được tạo ra bằng cách sử dụng chủ yếu là năng lượng từ NADH. Gradien proton làm cho phức hợp ATP synthaza tạo ra ATP. Theo trình tự hệ gen, chúng này có ATP synthaza vi khuẩn kiểu F.

Cần hiểu rằng, khi xác định rằng quy trình bao gồm bước nuôi cấy chúng với nguồn cacbon vô cơ, thì nguồn cacbon vô cơ là nguồn cacbon *chính* trong môi trường nuôi cấy. Do đó, có thể có lượng nhỏ của nguồn cacbon hữu cơ tồn tại trong môi trường nuôi cấy, nhưng sự chuyển hóa và sự sinh trưởng chính của môi trường nuôi cấy dựa trên việc sử dụng nguồn cacbon vô cơ, tốt hơn là cacbon đioxit, làm nguồn cacbon. Tốt hơn nếu tỷ lệ của cacbon được cung cấp vào môi trường nuôi cấy mà là hữu cơ ít hơn 5%, chẳng hạn như ít hơn 1%, ví dụ như ít hơn 0,1% của tất cả cacbon được cung cấp vào môi trường nuôi cấy trong quy trình này. Tốt hơn là, không có nguồn cacbon hữu cơ nào được cung cấp vào quy trình.

Tương tự, cần hiểu rằng, khi xác định rằng quy trình bao gồm bước nuôi cấy chúng với hydro ( $\text{H}_2$ ) làm nguồn năng lượng, thì hydro là nguồn năng lượng chính trong môi trường nuôi cấy. Do đó, có thể có nguồn năng lượng thứ yếu khác tồn tại trong môi trường nuôi cấy chẳng hạn như amoniac, mà có thể được cung cấp làm nguồn nitơ, hoặc lượng nhỏ của hợp chất hữu cơ, nhưng sự chuyển hóa và sự phát triển chính của môi trường nuôi cấy dựa trên việc sử dụng hydro làm nguồn năng lượng. Trong toàn bộ quy trình hydro tốt hơn là được tạo ra bằng cách điện phân nước; tức là bằng cách tách nước bằng điện thành khí hydro và khí oxy. Do đó, khí hydro và khí oxy được cung cấp đến

bình phản ứng sinh học từ bình điện phân ở gần đó. Theo cách khác, điện cực có thể được đặt bên trong bình phản ứng sinh học để tạo ra hydro và oxy trong bình phản ứng sinh học chứ không phải trong bình điện phân riêng rẽ.

Nguồn cacbon vô cơ có chứa cacbon dioxit có thể có chứa các nguồn cacbon vô cơ khác, chẳng hạn như ví dụ như cacbon monoxit. Theo một phương án, chỉ nguồn cacbon ở dạng khí được cung cấp cho môi trường nuôi cấy. Theo một phương án được ưu tiên, cacbon dioxit là nguồn cacbon vô cơ duy nhất, và thực sự là nguồn cacbon duy nhất, được cung cấp cho môi trường nuôi cấy. Theo một phương án, chỉ khí và khoáng chất được cung cấp cho môi trường nuôi cấy và hàm lượng cacbon dioxit trong khí được cung cấp nằm trong khoảng từ 10% đến 50%, ví dụ như trong khoảng từ 15% đến 45%, chẳng hạn như trong khoảng từ 20% đến 40%, ví dụ như trong khoảng từ 25% đến 35%, chẳng hạn như trong khoảng từ 26% đến 30%.

Theo phương án khác, khí và khoáng chất được cung cấp cho môi trường nuôi cấy và hàm lượng hydro ( $H_2$ ) trong khí được cung cấp nằm trong khoảng từ 30% đến 80%, ví dụ như trong khoảng từ 35% đến 75%, chẳng hạn như trong khoảng từ 40% đến 70%, ví dụ như trong khoảng từ 45% đến 65%, chẳng hạn như trong khoảng từ 50% đến 60%.

Theo phương án khác, khí và khoáng chất được cung cấp cho môi trường nuôi cấy và hàm lượng oxy ( $O_2$ ) trong khí được cung cấp nằm trong khoảng từ 10% đến 25%, ví dụ như trong khoảng từ 15% đến 20%, chẳng hạn như trong khoảng từ 16% đến 18%. Theo phương án khác, hàm lượng oxy được cung cấp sao cho hàm lượng oxy hòa tan trong môi trường nuôi cấy được duy trì ở trong khoảng từ 5% đến 10%.

Theo một phương án được ưu tiên, chỉ khí và khoáng chất được cung cấp cho môi trường nuôi cấy và khí được cung cấp bao gồm  $H_2$ ,  $CO_2$  và  $O_2$ , trong đó tỷ lệ phần trăm của  $H_2$  nằm trong khoảng từ 40% đến 70%, tỷ lệ phần trăm của  $CO_2$  nằm trong khoảng từ 18% đến 28% và tỷ lệ phần trăm của  $O_2$  nằm trong khoảng từ 12% đến 22%.

Thông thường, quy trình theo sáng chế bao gồm bước thêm nguồn nitơ. Ví dụ như nguồn nitơ có thể được cung cấp ở dạng amoni hydroxit, muối amoni, chẳng hạn như amoni sulphat hoặc amoni clorua, amoniac, ure hoặc nitrat, ví dụ như kali nitrat. Theo các phương án khác, khí nitơ ( $N_2$ ) được cung cấp làm nguồn nitơ. Theo một phương án được ưu tiên, nguồn nitơ là amoni hydroxit hoặc muối amoni, chẳng hạn như amoni sulphat.

Theo một phương án, nguồn nitơ được cung cấp là amoni hydroxit ở nồng độ nằm trong khoảng từ 100 mg/L đến 10 g/L, chẳng hạn như trong khoảng từ 250 mg/L đến 4 g/L, ví dụ như trong khoảng từ 0,5 g/L đến 2 g/L, chẳng hạn như trong khoảng từ 0,75 g/L đến 1,5 g/L.

Thông thường, quy trình theo sáng chế bao gồm bước thêm khoáng chất, chẳng hạn như khoáng chất chứa amoni, phosphat, kali, natri, vanadi, sắt, sulphat, magie, canxi, molybden, mangan, bo, kẽm, coban, selen, iot, đồng và/hoặc niken. Môi trường khoáng chất thích hợp đã được biết đến trong lĩnh vực, và ví dụ như được mô tả trong *Thermophilic Bacteria*, CRC Press, Boca Raton, FL, Jacob K. Kristjansson, ed., 1992, ví dụ như ở trang 87, Bảng 4.

Theo một phương án, khoáng chất được thêm vào bao gồm một hoặc nhiều chất sau đây: amoniac, amoni (ví dụ như, amoni clorua ( $\text{NH}_4\text{Cl}$ ), amoni sulphat ( $(\text{NH}_4)_2\text{SO}_4$ )), nitrat (ví dụ như, kali nitrat ( $\text{KNO}_3$ )), ure hoặc nguồn nitơ hữu cơ; phosphat (ví dụ như, dinatri phosphat ( $\text{Na}_2\text{HPO}_4$ ), kali phosphat ( $\text{KH}_2\text{PO}_4$ ), axit phosphoric ( $\text{H}_3\text{PO}_4$ ), kali dithiophosphat ( $\text{K}_3\text{PS}_2\text{O}_2$ ), kali orthophosphat ( $\text{K}_3\text{PO}_4$ ), dinatri phosphat ( $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ ) dikali phosphat ( $\text{K}_2\text{HPO}_4$ ) hoặc monokali phosphat ( $\text{KH}_2\text{PO}_4$ ); sulphat; dịch chiết nấm men; sắt được chelat hóa (ví dụ như được chelat hóa với EDTA hoặc axit xitric); kali (ví dụ như, kali phosphat ( $\text{KH}_2\text{PO}_4$ ), kali nitrat ( $\text{KNO}_3$ ), kali iodua ( $\text{KI}$ ), kali bromua ( $\text{KBr}$ )); và các muối vô cơ khác, khoáng chất, và chất dinh dưỡng vi lượng (ví dụ như, natri clorua ( $\text{NaCl}$ ), magie sulphat ( $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ ) hoặc magie clorua ( $\text{MgCl}_2$ ), canxi clorua ( $\text{CaCl}_2$ ), canxi sulphat ( $\text{CaSO}_4$ ) hoặc canxi cacbonat ( $\text{CaCO}_3$ ), mangan sulphat ( $\text{MnSO}_4 \cdot 7\text{H}_2\text{O}$ ) hoặc mangan clorua ( $\text{MnCl}_2$ ), sắt clorua ( $\text{FeCl}_2$ ), sắt sulphat ( $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ ) hoặc sắt clorua ( $\text{FeCl}_2 \cdot 4\text{H}_2\text{O}$ ), natri bicarbonat ( $\text{NaHCO}_3$ ) hoặc natri cacbonat ( $\text{Na}_2\text{CO}_3$ ), kẽm sulphat ( $\text{ZnSO}_4$ ) hoặc kẽm clorua ( $\text{ZnCl}_2$ ), amoni molybdat ( $\text{NH}_4\text{MoO}_4$ ) hoặc natri molybdat ( $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ ), đồng sulphat ( $\text{CuSO}_4$ ) hoặc đồng clorua ( $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ ), coban clorua ( $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ ) hoặc coban sulphat ( $\text{CoSO}_4$ ), nhôm clorua ( $\text{AlCl}_3 \cdot 6\text{H}_2\text{O}$ ), lithi clorua ( $\text{LiCl}$ ), axit boric ( $\text{H}_3\text{BO}_3$ ), niken clorua ( $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$ ) hoặc niken sulphat ( $\text{NiSO}_4$ ), thiếc clorua ( $\text{SnCl}_2 \cdot \text{H}_2\text{O}$ ), bari clorua ( $\text{BaCl}_2 \cdot 2\text{H}_2\text{O}$ ), đồng selenat ( $\text{CuSeO}_4 \cdot 5\text{H}_2\text{O}$ ), natri selenat ( $\text{Na}_2\text{SeO}_4$ ) hoặc natri selenit ( $\text{Na}_2\text{SeO}_3$ ), natri metavanadat ( $\text{NaVO}_3$ ), muối crom).

Theo một phương án được ưu tiên, quy trình theo sáng chế bao gồm bước thêm một, nhiều hoặc tất cả trong số:  $\text{NH}_4\text{OH}$ ,  $\text{KH}_2\text{PO}_4$ ,  $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ ,  $\text{NaVO}_3 \cdot \text{H}_2\text{O}$ ,

$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ ,  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ ,  $\text{CaSO}_4$ ,  $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ ,  $\text{MnSO}_4 \cdot 7\text{H}_2\text{O}$ ,  $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ ,  $\text{H}_3\text{BO}_3$ ,  $\text{CoSO}_4$ ,  $\text{CuSO}_4$ ,  $\text{NiSO}_4$ .

Theo một phương án, môi trường được cung cấp cho tế bào có chứa ít hơn 1 g/L muối clorua, chẳng hạn như ít hơn 0,25 g/L muối clorua, ví dụ như ít hơn 0,1 g/L muối clorua, chẳng hạn như ít hơn 0,025 g/L muối clorua, ví dụ như ít hơn 0,01 g/L clorua. Theo một phương án, không có muối clorua được cung cấp cho môi trường nuôi cấy.

Theo phương án khác, không có vitamin được cung cấp trong quy trình, tức là môi trường được cung cấp vào môi trường nuôi cấy không chứa vitamin.

Theo phương án khác, không có axit amin được cung cấp trong quy trình, tức là môi trường được cung cấp vào môi trường nuôi cấy không chứa axit amin.

Theo phương án khác, không có hợp chất hữu cơ được cung cấp trong quy trình, tức là môi trường được cung cấp vào môi trường nuôi cấy không chứa bất kỳ hợp chất hữu cơ nào.

Theo các phương án nhất định, pH của môi trường nuôi cấy vi khuẩn được kiểm soát ở mức nhất định. Theo các phương án nhất định, pH được kiểm soát trong khoảng tối ưu để duy trì và/hoặc phát triển và/hoặc sản xuất hợp chất hữu cơ của vi khuẩn. Theo một phương án, pH trong môi trường nuôi cấy được duy trì trong khoảng từ 5,5 đến 8,0, ví dụ như trong khoảng từ 6,5 đến 7,0, chẳng hạn như ở 6,8.

Theo các phương án nhất định, nhiệt độ của môi trường nuôi cấy vi khuẩn được kiểm soát. Theo các phương án nhất định, nhiệt độ được kiểm soát trong khoảng tối ưu để duy trì và/hoặc phát triển và/hoặc sản xuất hợp chất hữu cơ của vi khuẩn. Theo một phương án, môi trường nuôi cấy được sinh trưởng ở nhiệt độ nằm trong khoảng từ 25°C đến 40°C, ví dụ như trong khoảng từ 28°C đến 32°C, chẳng hạn như ở 30°C.

Thông thường, quy trình theo sáng chế được thực hiện trong bình phản ứng sinh học. Bình phản ứng sinh học được sử dụng để nuôi cấy tế bào, mà có thể được duy trì ở pha cụ thể trong đường cong sinh trưởng của chúng. Việc sử dụng bình phản ứng sinh học có lợi theo nhiều cách để nuôi cấy sinh trưởng hóa dưỡng. Nhìn chung, việc kiểm soát điều kiện sinh trưởng, bao gồm việc kiểm soát cacbon dioxit, oxy hòa tan, và các khí khác chẳng hạn như hydro, cũng như là các chất dinh dưỡng hòa tan khác, nguyên tố vết, nhiệt độ và pH, được tạo điều kiện thuận lợi trong bình phản ứng sinh học. Môi trường dinh dưỡng, cũng như là khí, có thể được thêm vào bình phản ứng sinh học dưới dạng bổ sung theo mẻ, hoặc định kì, hoặc tương ứng với sự cạn kiệt được phát hiện hoặc

điểm đặt được lập trình, hoặc liên tục trong quá trình môi trường nuôi cấy được sinh trưởng và/hoặc được duy trì. Trong quy trình nuôi cấy liên tục, môi trường dinh dưỡng, cũng như là khí, được thêm liên tục vào bình phản ứng sinh học. Hơn nữa, môi trường chứa vi khuẩn được lấy ra liên tục từ bình phản ứng sinh học.

Theo một phương án được ưu tiên, thể tích của môi trường nuôi cấy vi khuẩn là 100 mL hoặc nhiều hơn, chẳng hạn như 1 L hoặc nhiều hơn, ví dụ như 10 L hoặc nhiều hơn, chẳng hạn như 100 L hoặc nhiều hơn, ví dụ như 1.000 L hoặc nhiều hơn, chẳng hạn như 10.000 L hoặc nhiều hơn, ví dụ như 50.000 L hoặc nhiều hơn, chẳng hạn như 100.000 L hoặc nhiều hơn, ví dụ như 200.000 L hoặc nhiều hơn.

Theo một phương án, năng suất của môi trường nuôi cấy là hơn 0,1 g khối lượng tế bào khô trên mỗi lít trong mỗi giờ, chẳng hạn như hơn 0,2, ví dụ như lớn hơn 0,3, chẳng hạn như hơn 0,4, ví dụ như lớn hơn 0,5, chẳng hạn như hơn 0,6, ví dụ như lớn hơn 0,7, chẳng hạn như hơn 0,8, ví dụ như lớn hơn 0,9, chẳng hạn như hơn 1 g trên mỗi lít trong mỗi giờ.

Vi khuẩn có thể được cấy trực tiếp từ ngân hàng tế bào, hoặc thông qua môi trường nuôi cấy mầm ở quy mô nhỏ hơn. Tốt hơn là, việc cung cấp môi trường mới vào môi trường nuôi cấy và việc lấy môi trường đã được tận dụng hết với vi khuẩn ra diễn ra ở cùng tốc độ, sao cho thể tích trong bình phản ứng sinh học vẫn không đổi.

Theo một phương án, sau khi pha ban đầu đạt đến mật độ tế bào thích hợp, vi khuẩn sinh trưởng ở trạng thái ổn định hoặc trạng thái ổn định giả, vẫn liên tục trong pha log của chúng, ở OD600 trên 5, chẳng hạn như trên 10, ví dụ như trên 20, chẳng hạn như trong khoảng từ 50 đến 200, ví dụ như trong khoảng từ 50 đến 100.

Theo một phương án của quy trình theo sáng chế, chủng vi khuẩn có tốc độ sinh trưởng là 0,04 – 0,12 giờ<sup>-1</sup>.

Theo phương án khác của quy trình theo sáng chế, tốc độ cấp chất lỏng trong pha liên tục bằng 50-80% tốc độ sinh trưởng.

*Xanthobacter* là chi của vi khuẩn Gram-âm từ họ *Xanthobacteraceae*.

Theo một phương án, chủng *Xanthobacter* được sử dụng trong quy trình theo sáng chế là chủng sử dụng con đường Calvin Benson Bassham để chuyển hóa cacbon dioxid thành hợp chất hữu cơ, ví dụ như glucoza, thiết yếu đối với sinh vật sống.

Theo một phương án, chủng *Xanthobacter* được sử dụng trong quy trình theo sáng chế là chủng sử dụng NiFeSe-hydrogenaza để chuyển hóa hydro (H<sub>2</sub>) thành dạng

tương đương năng lượng tế bào.

Theo một phương án, chủng *Xanthobacter* được sử dụng trong quy trình theo sáng chế là chủng sử dụng hydrogenaza khử  $\text{NAD}^+$  để chuyển hóa hydro ( $\text{H}_2$ ) thành dạng tương đương năng lượng tế bào.

Theo một phương án, chủng *Xanthobacter* được sử dụng trong quy trình theo sáng chế có khả năng cố định nitơ.

Theo một phương án, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế được chọn từ nhóm bao gồm: *X. agilis*, *X. aminoxidans*, *X. autotrophicus*, *X. flavus*, *X. tagetidis*, *X. viscosus*, *Xanthobacter sp. 126*, *Xanthobacter sp. 91* và chủng VTT-E-193585.

Theo một phương án được ưu tiên, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế là VTT-E-193585 hoặc *X. tagetidis*. Tốt nhất là chủng được sử dụng trong quy trình theo sáng chế là VTT-E-193585.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế bao gồm ARN ribosom 16S nêu trong SEQ ID NO:1 hoặc ARN ribosom 16S có đến 20 khác biệt nucleotit so với SEQ ID NO:1, ví dụ như 1 đến 10, chẳng hạn như 1 đến 5, ví dụ như một, hai hoặc ba khác biệt nucleotit so với SEQ ID NO:1.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho chuỗi lớn ribuloza-1,5-bisphosphat carboxylaza/oxygenaza (rubisco) có trình tự nêu trong SEQ ID NO:3 hoặc trình tự tương đồng hơn 93%, ví dụ tương đồng hơn 95%, chẳng hạn như tương đồng hơn 96%, ví dụ tương đồng hơn 97%, chẳng hạn như tương đồng hơn 98%, ví dụ tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:3.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho chuỗi nhỏ ribuloza-1,5-bisphosphat carboxylaza/oxygenaza (rubisco) có trình tự nêu trong SEQ ID NO:5 hoặc trình tự tương đồng trình tự hơn 83%, ví dụ như hơn 86%, độ tương đồng chẳng hạn như tương đồng hơn 90%, ví dụ tương đồng hơn 95%, chẳng hạn như tương đồng hơn 96%, ví dụ tương đồng hơn 97%, chẳng hạn như tương đồng hơn 98%, ví dụ tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:5.

SEQ ID NO:2:

Trình tự nucleotit của chuỗi lớn ribuloza bisphosphat carboxylaza:

ATGGGTGCCGAAGCAACCGTCGGGCAGATCACGGACGCCAAGAAGA  
GATACGCCGCCGGCGTGCTGAAGTACGCCCAGATGGGCTACTGGAACGGC  
GACTACGTTCCCAAGGACACCGACCTCCTGGCGGTGTTCCGCATCACCCCC  
CAGGCGGGCGTGGACCCGGTGGAAAGCCGCCGCGGCGGTGCGCCGGCGAAA  
GCTCCACCGCTACCTGGACCGTGGTGTGGACCGACCGGCTCACCGCCGCCG  
ACGTCTACCGCGCCAAGGCCTACAAGGTGGAGCCGGTGCCGGGGCCAGGAA  
GGCCAGTATTTCTGCTACATCGCCTATGATCTCGATTTGTTTCGAGGAAGGC  
TCCATCGCCAACCTCACGGCGTCGATCATCGGCAACGTCTTCTCCTTCAAG  
CCGCTGAAGGCGGCGCGGCTGGAGGACATGCGGCTTCCCGTCGCCTATGT  
GAAGACCTTCCGCGGGCCCGCCACCGGCATCGTGGTCGAGCGCGAGCGCC  
TGGACAAGTTCGGCCGCCCCCTTCTGGGCGCCACCACCAAGCCGAAGCTTG  
GCCTCTCGGGCAAGAATTACGGCCGCGTGGTCTATGAGGCCCTCAAGGGC  
GGCCTCGACTTCGTGAAGGACGACGAGAACATCAACTCGCAGCCCTTCAT  
GCACTGGCGCGATCGCTTCCTCTATTGCATGGAGGCCGTCAACAAGGCCCA  
GGCCGAGACCGGCGAGGTGAAGGGGCACTATCTCAACATCACCGCCGGGA  
CCATGGAGGAGATGTACCGCCGCGCCGAGTTCGCCAAGGAACTGGGCTCC  
GTGGTGGTGATGGTGGATCTCATCATCGGCTGGACCGCCATCCAGTCCATG  
TCCAACTGGTGCCGCGAGAACGACATGATCCTGCACATGCACCGTGCGGG  
CCATGGCACCTACACGCGCCAGAAGAGCCACGGCGTCTCCTTCCGCGTCAT  
CGCCAAGTGGCTGCGGCTCGCCGGCGTCGACCACCTGCACACCGGCACCG  
CCGTGGGCAAGCTGGAAGGCGACCCCATGACCGTGCAGGGGCTTCTACAAT  
GTCTGCCGCGAGACGACGACGACGAGGACCTACCCGCGGCCTGTTCTTC  
GAGCAGGACTGGGGCGGCATCCGCAAGGTGATGCCGGTGGCCTCCGGCGG  
CATCCATGCGGGCCAGATGCACCAGCTCATCGACCTGTTTCGGCGAGGACG  
TGGTGCTCCAGTTCGGCGGCGGCACCATCGGCCACCCGGACGGCATCCAG  
GCCGGCGCCACCGCCAACCGCGTGGCGCTGGAAACCATGATCCTCGCCCG  
CAACGAGGGCCGCGACATCAGGAACGAGGGCCCGGAAATCCTGGTGGAA  
GCCGCCAAATGGTGCCGTCCGCTGCGCGCGGGCGCTCGATACCTGGGGCGA  
GGTGACCTTCAACTACGCCTCCACCGACACGTCCGATTACGTGCCACCGC  
GTCCGTCGCCTGA

SEQ ID NO: 3:

Trình tự axit amin của chuỗi lớn ribuloza bisphosphat carboxylaza

MGAEATVGQITDAKKRYAAGVLKYAQMGYWNGDYVPKDTDLLAVF  
 RITPQAGVDPVEAAA AVAGESSTATWTVVWTDRLTAADVYRAKAYKVEPVP  
 GQEGQYFCYIAYDLDLFEEGSIANLTASIIGNVFSFKPLKAARLEDMRLPVAYV  
 KTFRGPPTGIVVERERLDKFGRPLL GATT KPKLGLSGKNYGRVVYEALKGGL  
 DFKDDENINSQPFMHWRDRFLYCMEAVNKAQAETGEVKGHYLNITAGTME  
 EMYRRAEFAKELGSVVVMVDLIIGWTAIQSMSNWCRENDMILHMHRAHGHT  
 YTRQKSHGVSFRVIAKWLRLAGVDHLHTGTAVGKLEGDPMTVQGFYNVCRE  
 TTTQQDLTRGLFFEQDWGGIRKVMPPVASGGIHAGQMHLIDLFGEDVVLQFG  
 GGTIGHPDGIQAGATANRVALETMILARNEGRDIRNEGPEILVEAAKWCRPLR  
 AALDTWGEVTFNYASTDTSDYVPTASVA

SEQ ID NO: 4:

Trình tự nucleotit của chuỗi nhỏ ribuloza bisphosphat carboxylaza:

ATGCGCATCACCCAAGGCTCCTTCTCCTTCCTGCCGGACCTCACCGA  
 CACGCAGATCAAGGCCAGGTGCAATATTGCCTGGACCAGGGCTGGGCGG  
 TCTCGGTGGAGCACACCGACGATCCCCACCCGCGCAACACCTATTGGGAG  
 ATGTGGGGCCCGCCCATGTTCGATCTGCGCGACGCGGCCGGCGTCTTCGGC  
 GAGATCGAAGCCTGCCGGGCGCCAATCCCGAGCATTATGTGCGGGTGAA  
 CGCCTTCGATTCCAGCCGCGGATGGGAGACGATCCGCCTGTCCTTCATCGT  
 TCAGCGGCCACCGTGGAAGAGGGCTTCCGCCTCGACCGCACCGAAGGCA  
 AGGGCCGCAACCAGAGCTACGCCATGCGCTACCGGGCGCAGTTCGCGCCG  
 CGCTGA

SEQ ID NO: 5:

Trình tự axit amin của chuỗi nhỏ ribuloza bisphosphat carboxylaza:

MRITQGSFSFLPDLTDTQIKAQVQYCLDQGWAVSVEHTDDPHPRNTYW  
 EMWGPPMFDLRDAAGVFGEIEACRAANPEHYVRVNAFDSSRGWETIRLSFIV  
 QRPTVEEGFRLDRTEGKGRNQSYAMRYRAQFAPR

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị alpha của hydrogenaza khử  $\text{NAD}^+$  HoxS có trình tự nêu trong SEQ ID NO:7 hoặc trình tự tương đồng trình tự hơn 70%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID

NO:7.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị beta của hydrogenaza khử  $\text{NAD}^+$  HoxS có trình tự nêu trong SEQ ID NO:9 hoặc trình tự tương đồng trình tự hơn 77%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:9.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị gamma của hydrogenaza khử  $\text{NAD}^+$  HoxS có trình tự nêu trong SEQ ID NO:11 hoặc trình tự tương đồng trình tự hơn 70%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:11.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị delta của hydrogenaza khử  $\text{NAD}^+$  HoxS có trình tự nêu trong SEQ ID NO:13 hoặc trình tự tương đồng trình tự hơn 79%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:13.

SEQ ID NO:6:

Trình tự nucleotit của tiểu đơn vị alpha của hydrogenaza khử  $\text{NAD}^+$  HoxS:

ATGATGCCATCTGAGCCGCACGGCGCGGGCATGCCGCCCCACGGG  
AAGCGGCCGCGGTTCCCACCCCCAGGAGGTGAGCGCGGTGGTGGCCGAG  
GTGGTCGCGGATGCCGTGGCATCGGTGGGCGGCGCACGCACCCGGCTCAT  
GGACATCGTCCAGCTGGCCCAGCAGCGTCTCGGCCATCTCTCCGAAGAGA  
CCATGGCGGCCATTGCCGCGCGGCTCGCCATTCCGCCGGTGGAAGTGGCG  
GACATGGTGTCTTCTACGCCTTCCTGAACCGCGCGCCCAAGGGCCGCTAC  
CACATCCGCCTGTCGCGCAGCCCCATCTCGCTGATGAAGGGCGCCGAGGC  
GGTGGCTGCCGCCTTCTGCCAGATCCTCGGCATCGCCATGGGCGAGACCTC

GCAGGATGGCGACTTCACCCTGGAATGGACCAACGACATCGGCATGGCCG  
 ACCAGGAGCCGGCCGCCCTCGTCAACGGCACGGTGATGACGCAGCTCGCG  
 CCCGGCGATGCGGCCATCATCGTCGGCCGGCTGCGGGCCCATCACGCGCC  
 CAATGCCCTGCCGCTGTTCCCTGGAGCCGGCGTGGCCGGCTCCGGCCTGCC  
 CCATGCCCCGATCCGCCCCAGCCTGGTGATGCCGGGACAGCTTCTGTTCCG  
 CGAGGACCACACGACGCCGGGCGCCGGCATCAAGGCGGCACTCGCCCTCA  
 CCCC GGACGAAGTGGTGAGAAGGTCTCCGCCGCGCGCCTGCGCGGGCGG  
 GGTGGCGCCGGCTTTCCACCGGTCTCAAATGGAAGCTCTGCCGCCAGTCG  
 CCCGCCACCACCCGCCATGTGATCTGCAATGCGGACGAGGGCGAGCCCGG  
 CACCTTCAAGGATCGCGTGCTGCTCACGCAGGCGCCGCACCTCATGTTCTGA  
 CGGCATGACCATCGCCGGCTACGCCTTGGGGGCGCGGGAGGGCGTGTTCT  
 ATCTGCGCGGCGAGTACGCCTATCTGTGGGAGCCTCTGCATGCGGTCCTGC  
 GCGAGCGCTATGGGCTCGGGCTCGCCGGCGCGAACATCCTGGGACACGCG  
 GGCTTCGACTTCGACATCCGCATCCAGCTGGGCGCCGGCGCCTATATCTGC  
 GCGAGGAATCCGCGCTGGTGGAATCGCTGGAAGGCAAGCGCGGCTCGCC  
 CCGCGACCGCCCCCCTTCCCCACCGTGCGCGGCCATCTCCAGCAGCCCAC  
 CGCCGTGGACAATGTGGAGACCTTCGCCTGCGCCGCCCGCATCCTGGAGG  
 ATGGCGTGAGGCGTTTCGCGGGCATCGGCACGCCC GAATCCGCCGGCACG  
 AAGCTCCTCTCGGTGTCGGGCGATTGCCCGCGCCCCGGCGTGTATGAGGTG  
 CCCTTCGGCCTCACGGTGAACGCGCTGCTCGACCTTGTCGGCGCGCCGGAC  
 GCCGCCTTCGTGCAGATGGGTGGGCCGTCCGGCCAATGCGTGCGCGCCGAA  
 GGATTACGGCCGCCGCATCGCCTTCGAGGACCTGCCCACCGGCGGCTCGGT  
 GATGGTGTTTCGGCCCCGGGGCGCGACGTGCTCGCCATGGTGCGCGAGTTCG  
 CGGATTTCTTCGCCGGCGAATCCTGCGGCTGGTGACGCCCTGCCGGGTGG  
 GCACCACCTTGCTCAAGGAAGAGCTGGACAAGCTCCTCGCCAACCGCGCC  
 ACCCTCGCCGACATCCGCGCGCTGGAGACCCTGGCCACGACCGTCTCCCGC  
 ACCAGCCGCTGCGGCCTCGGCCAGACGGCGCCCAACCCCATCCTTTCCACC  
 ATGCGCAACCTGCCGGAAGCCTATGAGGCGAGGCTGAGGCCCCGAAGACTT  
 CCTGCCCTGGGCCTCGCTCGACGAGGCGCTGAAGCCCGCCATCGTCATCCA  
 GGGCCGCGCGCCCGTGCCGGAGGAAGAGGCATGA

SEQ ID NO: 7:

Trình tự axit amin của tiểu đơn vị alpha của hydrogenaza khử NAD<sup>+</sup> HoxS:

MMPSEPHGAGMPPPREAAA VPTPQEVSAVVAE VVADAVASVGGARTR  
 LMDIVQLAQQLGHLSEETMAAIAARLAIPPVEVADMVSFYAFLNRAPKGRY  
 HIRLSRSPISLMKGAEAVAAAFQCILGIAMGETSQDGDFTLEWTNDIGMADQE  
 PAALVNGTVMTQLAPGDAAIIVGRLRAHHAPNALPLFPGAGVAGSGLPHARIR  
 PSLVMPGQLLFREDHTTPGAGIKAALALTPDEVVQKVSAARLRGRGGAGFPT  
 GLKWKL CRQSPATTRHVICNADEGEPTFKDRVLLTQAPHL MFDGMTIAGYA  
 LGAREGVVYLRGEYAYLWEPLHAVLRERYGLGLAGANILGHAGFDFDIRIQL  
 GAGAYICGEESALVESLEGKRGSPDRPPFPTVRGHLQQPTAVDNVETFACAA  
 RILEDGVEAFAGIGTPESAGTKLLSVSGDCPRPGVYEVFGLTVNALDLVGAP  
 DAAFVQMGGPSGQCVAPKDYGRRIAFEDLPTGG SVMVFGPGRDVLAMVREF  
 ADFFAGESCGWCTPCRVTLLKEELDKLLANRATLADIRALET LATTVSRTS  
 RCGLGQTAPNPILSTMRNLPEAYEARLRPEDFLPWASLDEALKPAIVIQGRAPV  
 PEEEA

SEQ ID NO: 8:

Trình tự nucleotit của tiểu đơn vị beta của hydrogenaza khử NAD<sup>+</sup> HoxS:

ATGAGCCGGGGATCCCCCGATGCCGGGAAAGACCGCACCATGAGCG  
 CCACCGACGGCACCAACGCCCCCGCAAGATCGTCATCGATCCGGTGACC  
 CGCGTGGAGGGCCACGGCAAGGTCACCATCCGCCTGGATGAAGCCGGCGC  
 GGTGGAGGATGCGCGTTTCCACATCGTGGAGTTCCGCGGCTTCGAGCGGTT  
 CATCCAGGGCCGGATGTACTGGGAAGTGCCCCTTATCATCCAGCGGCTGTG  
 CGGCATCTGCCCCGGTGAGCCACCATCTGGCGGCGGCGAAAGCCATGGACC  
 AGGTGGCGGGCGTGACCGCGTACCGCCCACCGCCGAGAACTGCGCCGG  
 CTGATGCATTATGGGCAGGTGCTGCAATCCAACGCTTTGCACATCTTCCAC  
 CTCGCCTCGCCCGACCTCCTGTTTCGGCTTCGACGCGCCGGCCGAGCAGCGC  
 AACATCATCGCCGTGCTCCAGCGTTATCCGGAGATCGGCAAATGGGCGAT  
 CTTTCATCAGGAAGTTCGGCCAGGAGGTCATCAAGGCCACCGGCGGGCGCA  
 AGATCCATCCCACCGCGCCATTCCCGGGCGGGGTCAACCAGAACCTCGCC  
 GTGGAGGACCGCGACGCCCTGCGCGCCAAGGTGGGCGAGATCATCAGCTG  
 GTGCATGGCGGCGCTGGACCATACAAGGCCTATGTGGCGGAAAACCGGG  
 CGCTGCATGACAGCTTCGCCGCCTTCCCCTCCGCCTTCATGAGCCTCGTGG  
 GGCCGGATGGCGGCATGGACCTTTATGACGGCACCTGCGGGTGATCGAT  
 GCCGAGGGCGCCCCCTCATCGAAGGCGCGCCGCCCGCCCTCCTACCGCGA

CCACCTCATCGAGGAGGTGCGGCCCTGGAGCTATCTGAAATTCCCCCATCT  
 GCGCGCCTTCGGCCGCGACGATGGCTGGTATCGGGTCGGCCCCCTCGCCCA  
 GGTCAATTGCGCCGCGTCCATCGACACGCCCCGCGCCGAGGCGGCCCGGC  
 GGGACTTCATGGCCGAGGGCGGGCGGCAAGCCGGTGCATGCCACCCTCGCT  
 TATCACTGGGCGCGGCTCATCGTGCTGGTCCATTGCGCGGAGAAGATCGA  
 ACAGCTGCTGTTTCGACGACGACCTGCAAGGCTGCGATCTGCGTGCGGAGG  
 GCACCCGGCGCGGGGAAGGCGTCGCCTGGATCGAGGCGCCGCGCGGCACC  
 CTCATCCACCATTACGAGGTGGACGAGAACGACCAGGTGCGCCGCGCCAA  
 CCTCATCGTCTCCACCACCCACAATAACGAGGCCATGAACCGCGCCGTGCG  
 GCAGGTGGCGAAGACGGACCTTTCCGGTCGCGAGATCACCGAAGGGCTGC  
 TGAACCATATCGAGGTGGCCATCCGCGCCTTCGACCCCTGCCTGTCCTGCG  
 CCACCCATGCGCTGGGCCAGATGCCGCTGATCGTGACGCTTGAAGATGCCT  
 CCGGCGCAGAGATCGCCCGCGGAGTGAAGGAATGA

SEQ ID NO: 9:

Trình tự axit amin của tiểu đơn vị beta của hydrogenaza khử  $\text{NAD}^+$  HoxS:

MSRGSPDAGKDRTMSATDGTAPRKIVIDPVTRVEGHGKVTIRLDEAG  
 AVEDARFHIVEFRGFERFIQGRMYWEVPLIIQRLCGICPVSHHLAAAKAMDQV  
 AGVDRVPPTAEKLRRLMHYGQVLQSNALHIFHLASPDLLFGFDAPAEQRNIIA  
 VLQRYPEIGKWAIFIRKFGQEVIKATGGRKIHPTSAIPGGVNQNLAVEDRDALR  
 AKVGEIISWCMAALDHHKAYVAENRALHDSFAAFPSAFMSLVGPDGGMDLY  
 DGTLRVIDAEGAPLIEGAPPASYRDHLIEEVRPWSYLKFPHLRAFRDDGWYR  
 VGPLAQVNCAASIDTPRAEAARRDFMAEGGGKPVHATLAYHWARLIVLVHC  
 AEKIEQLLFDDDLQGCDLRAEGTRRGEGVAWIEAPRGTLIHHYEVDENDQVR  
 RANLIVSTTHNNEAMNRAVRQVAKTDLSGREITEGLLNHIEVAIRAFDPCLSC  
 ATHALGQMPLIVTLEDASGAELARGVKE

SEQ ID NO: 10:

Trình tự nucleotit của tiểu đơn vị gamma của hydrogenaza khử  $\text{NAD}^+$  HoxS:

ATGAGCGAGACCCCCTTCACCTTTACCGTGGACGGCATCGCGGTCCC  
 GGCCACCCCGGCCAGAGCGTCATCGAGGCGTGCGATGCGGCGGGCATCT  
 ATATCCCGCGCCTGTGCCACCACCCGGACCTGCCGCCGGCGGGGCCATTGCC  
 GGGTGTGCACCTGCATCATCGACGGGCGGCCGGCCAGCGCCTGCACCATG  
 CCCGCCGCCAGGGGCATGGTGGTGGAGAACGAGACGCCCCGCTTTGCTGGC

GGAGCGGCGCACGCTGATCGAGATGCTGTTCGCGGAAGGCAACCATTTCT  
 GCCAGTTCTGCGAGGCGAGCGGCGATTGCGAATTGCAGGCGCTGGGCTAC  
 CTGTTCGGCATGGTGGCCCCGCCCTTCCCCCATCTGTGGCCGAAGCGGCCG  
 GTGGATGCCAGCCATCCGGATATCTATATCGACCACAATCGCTGCATCCTG  
 TGCTCGCGCTGCGTGCGCGCCTCGCGCACCTGGACGGCAAGTCCGTGTTC  
 GGCTTCGAGGGGCGCGGCATCGAGATGCATCTGGCGGTGACCGGCGGGCA  
 CCTGGACGACAGCGCCATCGCCGCCGCCGACAGGGCGGTTGAGATGTGCC  
 CGGTGGGCTGCATCGTCCTCAAGCGCACCGGCTACCGCACGCCCTATGGCC  
 GGCGGCGCTACGACGCCGCGCCCATCGGCTCCGACATCACCGCCCCGGCGC  
 GGCGGCGCGAAGGACTGA

SEQ ID NO: 11:

Trình tự axit amin của tiểu đơn vị gamma của hydrogenaza khử  $\text{NAD}^+$  HoxS:

MSETPFTFTVDGIAVPATPGQSVIEACDAAGIYIPRLCHHPDLPPAGHCR  
 VCTCIIDGRPASACTMPAARGMVVENETPALLAERRTLIEMLF AEGNHFCQFC  
 EASGDCELQALGYLFGMVAPPPHLWPKRPVDASHPDIYIDHNRCILCSRCVR  
 ASRTL DGKSVFGFEGRGIEMHLAVTGGHLDDSAIAAADRAVEMCPVGCIVLK  
 RTGYRTPYGRRRYDAAPIGSDITARRGGA KD

SEQ ID NO: 12:

Trình tự nucleotit của tiểu đơn vị delta của hydrogenaza khử  $\text{NAD}^+$  HoxS:

ATGGCCAAGCCCAAACCTCGCCACCTGCGCGCTGGCCGGCTGCTTCGG  
 CTGCCACATGTCCTTCCTGGACATGGACGAGCGCATCGTCGAGCTCATCGA  
 CCTGGTGGACCTCGACGTCTCGCCCCTCGACGACAAGAAAACTTCACCG  
 GCATGGTGGAAATCGGCCTGGTGGAAAGGCGGCTGCGCCGACGAGCGCCAT  
 GTGAAGGTGCTGCGCGAGTTCCGCGAGAAATCCCGCATCCTGGTGGCGGT  
 GGGCGCCTGCGCCATCACCGGCGGCATCCCGGCATTGCGCAACCTCGCCG  
 GCCTCGACGAATGCCTGAGGGAAGCCTACCTCACCGGCCCCACGGTGGAA  
 GGCGGCGGGCTCATTCCCAACGACCCGGAGCTGCCGCTGCTGCTGGACAA  
 GGTCTATCCGGTGCAGGACTTCGTGAAGATCGACCATTTCTGCCCCGGCTG  
 CCCGCCCTCGGCCGACGCCATCTGGGCGGCTCTGAAGGCGCTGCTGACCG  
 GCACCGAGCCGCATCTGCCCTACCCGCTTTTCAAGTACGAATGA

SEQ ID NO: 13:

Trình tự axit amin của tiểu đơn vị delta của hydrogenaza khử  $\text{NAD}^+$  HoxS:

MAKPKLATCALAGCFGCHMSFLDMDERIVELIDLVDLDVSPLDDKKNF  
TGMVEIGLVEGGCADERHVKVLREFREKSRLVAVGACAITGGIPALRNLAGL  
DECLREAYLTGPTVEGGGLIPNDPELPLLLDKVYPVQDFVKIDHFLPGCPPSAD  
AIWAALKALLTGTEPHLPYPLFKYE

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị lớn của NiFeSe hydrogenaza có trình tự nêu trong SEQ ID NO:15 hoặc trình tự tương đồng trình tự hơn 84%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:15.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị nhỏ của NiFeSe hydrogenaza có trình tự nêu trong SEQ ID NO:17 hoặc trình tự tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:17.

SEQ ID NO: 14:

Trình tự nucleotit của tiểu đơn vị lớn của hydrogenaza [NiFeSe] tế bào chất:

TCCAGACCCGGGCAACATTGCTCCATGTGCTGGGCACCCTGGCCGGC  
CGCTGGCCCCATACCCTCGCGCTCCAGCCCGGCGGGGTGACCCGAAGCGC  
CGACCAGCACGACCGCATGCGCCTGCTCGCGACGCTGAAGGCGGTGCGGG  
CGGCGCTGGAAGAGACCTTGTTTCGGCGCGCCTTTGGAAGAGGTGGCGGCC  
CTGGACGGCGCCGCCGCGTGGAGGCCTGGCGCGCCAACGGCCCGGAAGG  
GGATTTCCGCCTGTTCTTGGAGATCGCCGCCGACCTGGAGCTGGACCGGCT  
CGGCCGCGCGCACGACCGCTTTCTCTCCTTCGGCGCCTACGCCAGGACGA  
GGGGCGCCTTTATGGCGCCGGCACCTTCGAGGCCGGGACGGCGGGAGGGC  
TCGATCCCAACGCCATCACCGAGGACCACGCCTTCGCCCCGCATGGAGGAC  
CGCGCGGCGCCCCATGCGCCCTTTGACGGCTCCACCTTCCCCGATGCCGAC  
GACACCGAGGGCTACACCTGGTGCAAGGCGCCGCGCCTTGCCGGCCTGCC  
CTTCGAGACCGGCGCCTTCGCCCCGGCAGGTGGTGGCGGGCCATCCGCTCGC  
CCGGGACCTCGTGACGCGGGAAGGCGGCACTGTGCGCAGCCGCGTGGTCG  
GCCGGCTGCTGGAACCGCGCGCACCTGATCGCCATGGAGGGCTGGGTG

AAGGAACTGCGGCCCGAAGGGCCCTGGTGCGCCCAGGGCCACCTGCCCCA  
 GGAAGGCCGCGCCTTCGGCCTCACCGAGGCGGCGCGCGGGGCGCTCGGCC  
 ACTGGATGGTGGTGGAGAAGGGCCGCATTGCCCCGCTACCAGATCATCGCC  
 CCCACCACCTGGAACCTTCTCCCCCGCGACGGCGCGGGCCTGCCCCGCCCCG  
 CTGGAGACGGCCCTGGTGGGCGCGCCCGTGCGGCAGGGAGAGACGACGCC  
 CGTGAGCGTGCAGCACATCGTGCGCTCCTTCGACCCGTGCATGGTCTGCAC  
 TGTGCATTGA

SEQ ID NO: 15:

Trình tự axit amin của tiểu đơn vị lớn của hydrogenaza [NiFeSe] tế bào chất:

MSAETRRLVVGPFNRVEGDLEVRDLVDGRVQQAFVSSPLFRGFERILE  
 GRDPRDALVIAPRICGICSVSQSHAAALALAGLQGIAPTHDGRIATNLIVAAEN  
 VADHLTHFHVFFMPDFARAVYEDRPWFAQAARRFKANQGVSVRRALQTRAT  
 LLHVLGTLAGRWPHTLALQPGGVTRSADQHDMRLLATLKAVRAALEETLF  
 GAPLEEVAALDGAAVEAWRANGPEGDFRLFLEIAADLELDRLGRAHDRFLS  
 FGAYAQDEGRLYGAGTFEAGTAGGLDPNAITEDHAFARMEDRAAPHAPFDGS  
 TFPDADDTEGYTWCKAPRLAGLPFETGAFARQVVAGHPLARDLVTREGGTVR  
 SRVVGRLLLETARTLIAMEGWVKELRPEGPWCAQGHLPQEGRAFGLTEAARGA  
 LGHWMVVEKGRIARYQIIAPTTWNFSPRDGAGLPGPLETALVGAPVRQGETTP  
 VSVQHIVRSFDPCMVCVTH

SEQ ID NO: 16:

Trình tự nucleotit của tiểu đơn vị nhỏ của hydrogenaza [NiFeSe] tế bào chất:

ACGGGGGAGGAAGCCCGCGCCATCTTCGACGCCATCCTTGCCGGCG  
 TTATCGTCCTCGACGCCCTGTGCGTGGAAGGCGCGCTGCTGCGCGGGCCGA  
 ACGGCACCGGGCGCTTCCATGTGCTGGCGGGCACGGACACCCCCACCATC  
 GACTGGGCGCGGCAGCTCGCCGGCATGGCGCGCCACGTGGTGGCGGTGGG  
 CACCTGCGCCGCCTATGGGGGCGTGACGGCGGCGGGCATCAACCCACCG  
 ATGCCTGCGGCCTCCAGTTCGACGGACGCCGGAAGGGTGGGGCGCTGGGG  
 GCGGACTTCCGCTCCCGCTCGGGGCTTCCGGTCATCAATGTGGCCGGCTGC  
 CCCACCCATCCCAACTGGGTGACGGAACCCTGATGCTGCTGCCTGCGGC  
 CTGCTGGGCGAGGCCGACCTCGACGTCTATGGCCGCCCGCGCTTCTATGCG  
 GACCTGCTGGTGCATCACGGCTGCCCGCGCAACGAATACTATGAATACAA  
 GGCGAGCGCCGAGAAGATGAGCGACCTCGGCTGCATGATGGAGCATCTGG

GCTGCCTCGGCACCCAGGCCACGCCGACTGCAACACGCGCCTTTGGAAT  
 GGCGAGGGCTCGTGCACCCGCGGGCTATGCCTGCATCAACTGCACGGC  
 GCCGGAATTCGAGGAGCCGGGCCACGCCTTCCTGGAGACGCCCAAGATCG  
 GCGGCATCCCCATCGGCCTGCCACCGACATGCCCAAGGCCTGGTTCATCG  
 CCTTGTCCTCCCTCGCCAAGGCGGCGACGCCGGAGCGGCTGCGCAAGAAC  
 GCGGTGTCCGACCATGTGGTCACGCCGCCCGCCGTCAAGGACATCAAGCG  
 GCGATGA

SEQ ID NO: 17:

Trình tự axit amin của tiểu đơn vị nhỏ của hydrogenaza [NiFeSe] tế bào chất:

MSTPFSVLWLQSGGCGGCTMSLLCAEAPDLATTLDAAGIGFLWHPALS  
 EETGEEARAIFDAILAGVIVLDALCVEGALLRGPNGTGRFHVLAGTDTPTIDW  
 ARQLAGMARHVVAVGTCAAYGGVTAAGINPTDACGLQFDGRRKGGALGAD  
 FRSRGLPVINVAGCPHPNWTETLMLLACGLLGEADLDVYGRPRFYADLL  
 VHHGCPRNEYEYKASAEKMSDLGCMMEHLGCLGTQAHADCNTRLWNGEG  
 SCTRGGYACINCTAPEFEPEGHAFLETPKIGGIPIGLPTDMPKAWFIALSSLAKA  
 ATPERLRKNAVSDHVVTTPPAVKDIKRR

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho chuỗi gamma của ATP synthaza atpG\_1 có trình tự nêu trong SEQ ID NO:19 hoặc trình tự tương đồng hơn 70%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:19.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị alpha của ATP synthaza atpA\_1 có trình tự nêu trong SEQ ID NO:21 hoặc trình tự tương đồng hơn 78%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:21.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị b của ATP synthaza atpF\_1 có trình tự nêu trong SEQ ID NO:23 hoặc trình tự tương đồng hơn 62%, ví dụ tương đồng hơn 70%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng

hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:23.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị c của ATP synthaza, đặc hiệu ion natri atpE\_1 có trình tự nêu trong SEQ ID NO:25 hoặc trình tự tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:25.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị a của ATP synthaza atpB\_1 có trình tự nêu trong SEQ ID NO:27 hoặc trình tự tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:27.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho chuỗi epsilon của ATP synthaza atpC\_1 có trình tự nêu trong SEQ ID NO:29 hoặc trình tự tương đồng hơn 71%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:29.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị beta của ATP synthaza atpD\_1 có trình tự nêu trong SEQ ID NO:31 hoặc trình tự tương đồng hơn 84%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:31.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị beta của ATP synthaza atpD\_2 có trình tự nêu trong SEQ ID NO:33 hoặc trình tự tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:33.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng

chế có chứa gen mã hóa cho chuỗi gamma của ATP synthaza atpG\_2 có trình tự nêu trong SEQ ID NO:35 hoặc trình tự tương đồng hơn 86%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:35.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị alpha của ATP synthaza atpA\_2 có trình tự nêu trong SEQ ID NO:37 hoặc trình tự tương đồng trình tự hơn 98%, chẳng hạn như lớn hơn 99% so với trình tự nêu trong SEQ ID NO:37.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị delta của ATP synthaza atpH có trình tự nêu trong SEQ ID NO:39 hoặc trình tự tương đồng hơn 85%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:39.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị b của ATP synthaza atpF\_2 có trình tự nêu trong SEQ ID NO:41 hoặc trình tự tương đồng hơn 87%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:41.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị b' của ATP synthaza atpG\_3 có trình tự nêu trong SEQ ID NO:43 hoặc trình tự tương đồng hơn 81%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:43.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho tiểu đơn vị c của ATP synthaza atpE\_2 có trình tự nêu trong SEQ ID NO:45 hoặc trình tự tương đồng trình tự hơn 98%, chẳng hạn như lớn hơn 99% so với trình tự nêu trong SEQ ID NO:45.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng

chế có chứa gen mã hóa cho tiểu đơn vị a của ATP synthaza atpB\_2 có trình tự nêu trong SEQ ID NO:47 hoặc trình tự tương đồng hơn 92%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:47.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho protein I của ATP synthaza atpI có trình tự nêu trong SEQ ID NO:49 hoặc trình tự tương đồng hơn 60%, ví dụ tương đồng hơn 70%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:49.

SEQ ID NO: 18:

Trình tự nucleotit của chuỗi gamma của ATP synthaza atpG\_1

GTGACCGAGCGCCTGTCCGACGTCAACGCCCGCATCGCCTCGGTGCG  
GCAGCTCTCATCGGTCATCACGGCCATGCGGGGCATTGCGGGCGGCGCGGG  
CGCGGGAGGCGCGGGGTCGGCTCGACGGCATCCGCGCCTATGCGCAGACC  
ATCGCCGAGGCCATCGGCCATGTGCTCGCCGTGCTGCCCCGAGGAGGCCCG  
CGCCCGGTCCTCCGGGCACCGGCATCGGGGCCATGCGGTCATCGCCCTGTG  
CGCGGAGCAGGGCTTTGCCGGCGTCTTCAACGAGCGGGTGCTGGACGAGG  
CCGCCCCGGCTGCTGACCGGCGGGGCGGGGCCGGCCGAGCTGCTGCTGGTG  
GGCGACCGGGGCCTGATGGTGGCCCGCGAGCGGGGGCTCGATGTCTCCTG  
GTCGGTGCCCATGGTGGCCCATGCGGGGCCAGGCCTCGGCGCTGGCGGACC  
GCATCAGCGAGGAGCTCTACCGGCGGATCGATGCGGGACGGGTGACGCGG  
GTGTCGGTGGTGCACGCCGAGCCCGCCGCGTCCGCCGCCATCGAGACGGT  
GGTGAAAGTGCTGGTGCCGTTCTGACTTCGCCCCGCTTCCCCCTGGCGCGGGT  
GGCATCCGCCCCGCTCATGACCATGCCGCCGCCGCGGCTGCTGGCCCAGCT  
GTCGGAGGAATATGTGTTCCGCCGAGCTGTGCGAGGCGCTCACCTTGTCCTT  
CGCGGCGGAGAACGAGGCCCGCATGCGGGGCCATGATCGCCGCCCGCGCCA  
ATGTGGCCGATAACCCTGGAGGGCCTCGTCGGCCGCGCCCGGCAGATGCGC  
CAGGAGGAGATCACCAACGAGATCATCGAGCTGGAAGGCGGCGCCGGCA  
GCGCCCGGCATGCGGATTGA

SEQ ID NO: 19:

Trình tự axit amin của chuỗi gamma của ATP synthaza atpG\_1

MTERLSDVNARIASVRQLSSVITAMRGIAAARAREARGRLDGIRAYAQT  
IAEAIGHVLAVLPPEARARSSGHRHRGHAVIALCAEQGFAGVFNERVLDEAAR  
LLTGGAGPAELLLVGDRGLMVARERGLDVSWSVPMVAHAGQASALADRISE  
ELYRRIDAGRVTRVSVVHAEPAAASAAIETVVKVLVPDFARFPLARVASAPLM  
TMPPPRLLAQLSEEYVFAELCEALTSFAAENEARMRAMIAARANVADTLEG  
LVGRARQMRQEEITNEIIELEGGAGSARHAD

SEQ ID NO: 20:

Trình tự nucleotit của tiểu đơn vị alpha của ATP synthaza atpA\_1

ATGAGCACGGGCGCGCAAGCGAGCGAGGATTGGCTCACCCGGAGCC  
GGGCGGCCCTGGCCGGGACGCGCCTTTCCCAGCAATCCCAATCGGTGGGC  
CGGGTGGAGGAGATGGCCGACGGCATCGCCCGCGTCTCCGGCCTGCCGGA  
TGTGCGGCTCGACGAGCTTCTCACCTTCGAGGGGCGGCCAGACCGGCTATGC  
CCTCACCTCGATCGCACCGAGATCGCCGTGGTGCTGCTGGATGACGCCTC  
CGGCGTGGAGGCGGGCGCCCGGGTGTTCGGCACCGGCGAGGTGGTGAAGG  
TGCCGGTGGGGCCGGGGCTGCTGGGGCCGCATCGTCGACCCCTCGGCCGG  
CCCATGGACCGCTCCGAGCCGGTGGTGGCGCAGGCGCACCATCCCATCGA  
GCGGCCGGCGCCGGCCATCATCGCCCGCGACCTGGTCTCGCAGCCGGTTCA  
GACCGGCACGCTGGTGGTGGATGCGCTGTTCTCCCTCGGCCGGGGCCAGC  
GCGAGCTCATCATCGGCGACCGGGCTACCGGCAAGACCGCCATCGCGGTG  
GACACCATCATCAGCCAGAAGCATTCGGACATCGTGTGCATCTACGTGGC  
GGTGGGCCAGCGCGCCCGCCCGGTGGAGCGGGTGGTGGAGGCGGTGCGCG  
CCCACGGGGCGATCGAGCGCTGCATCTTCGTGGTCGCCTCGGCCGCCGCCT  
CGCCAGGGCTGCAATGGATCGCGCCGTTTCGCCGGCATGACCATGGCGGAA  
TATTTCCGCGACAACGGCCAGCATGCGCTCATCATCATCGATGATCTCACC  
AAGCATGCGGCCACCCATCGCGAGCTGGCGCTGCTCACCCACGAGCCGCC  
GGGCCGCGAGGCCTATCCCGGCGACATCTTCTATGTGCACGCCCGCCTTCT  
GGAGCGGGCCGCCAAGCTCTCCGCCGAGCTGGGCGGTGGCTCGCTCACGG  
CCCTGCCCATCGCGGAGACGGACGCGGGAAACCTCTCCGCCTATATCCCCA  
CCAACCTCATCTCCATCACCGATGGGCAGATCGTGCTGGATTCGCGGGCTGT  
TCGCGGCCAACCAGCGCCCGGCGGTGGATGTGGGCCTCTCCGTGAGCCGG

GTGGGCGGCAAGGCGCAGCATCCCGCGCTTCGGGGCCGTGTCCGGGGCGCAT  
 CCGGCTCGATTATTCCCAGTTCCTGGAGCTGGAAATGTTACCCGCTTCGG  
 CGGCATCACCGATAACCCGCGTGAAGGCGCAGATCACCCGGGGCGAGCGCA  
 TCCGCGCGCTGCTCACCCAGCCGCGCTTTTCCACCCTGCGCCTTCAGGACG  
 AGGTGGCGCTGCTGGCCGCGCTGGCGGAGGGGGTGTTCGACACTTTGGCC  
 CCGGGGCTGATGGGCGCCGTGCGTGCCCGCATTCCGGGCCAGCTGGATGC  
 GCAGGTGAAGGACGTGGCCTCGGCCCTCGCCGAGGGCAAGGTGCTGGAGG  
 AGGGCTTGACGCCCCGTCTCGTGGCGGCCGTGCGGGCCGTGCGGGCGGAC  
 GTGGCCGCGACCGCGAAGGCCGGGCCGTGA

SEQ ID NO: 21:

Trình tự axit amin của tiểu đơn vị alpha của ATP synthaza atpA\_1

MSTGAQASEDWLTRSRAALAGTRLSSQSQSVGRVEEMADGIARVSGLP  
 DVRLDELLTFEGGQTGYALTLDRTEIAVVLLDDASGVEAGARVFGTGEVVKV  
 PVGPGLLGRIVDPLGRPMDRSEPVAQAHHPIERPAPAIARDLVSQPVQTGTL  
 VVDALFSLGRGQRELIIGDRATGKTAIAVDTIISQKHSDIVCIYVAVGQRAAAV  
 ERVVEAVRAHGAIERCIFVVASAAASPLQWIAPFAGMTMAEYFRDNGQHAL  
 IIIDDLTKHAATHRELALLTHEPPGREAYPGDIFYVHARLLERA AKLSAELGGG  
 SLTALPIAETDAGNLSAYIPTNLISITDGQIVLDSRLFAANQRPAVDVGLSVSRV  
 GGKAQHPALRAVSGRIRLDYSQFLELEMFTRFGGITDTRVKAQITRGERIRALL  
 TQPRFSTLRLQDEVALLAALAEGVFDLAPGLMGAVRARIPAQLDAQVKDVA  
 SALAEGKVLEEGLHARLVAAVRVAADVAATAKAGP

SEQ ID NO: 22:

Trình tự nucleotit của tiểu đơn vị b của ATP synthaza atpF\_1

ATGCAGATCGACTGGTGGACGCTGGGCCTGCAGACGGTCAACGTCC  
 TCGTTCTCATCTGGCTCCTGAGCCGCTTCCTGTTCAAGCCGGTGGCGCAGG  
 TCATCGCGCAGCGCCGTGCCGAGATCGAGAAGCTGGTGGAGGATGCGCGC  
 GCCGCCAAGGCCGCCGCCGAGGCCGAGCGGGACACGGCGAAGGCGGAGG  
 AGGCGCGCCTTGCCGCCGAGCGCGGGCGCCCGCATGGCGGCGGTGCGCAAG  
 GAGGCGGAGGCGCAGAAGGCGGCATTGCTGGCCGCCGCCAAGACCGAGG  
 CCGAGGCCCTGCACGCGGCCGCGGAAGCGGCCATCGTCCGGGGCGCGGGCG  
 AGCGAGGAGGAAGCCGCCGCCGACCGCGCCAGCCGCCTTGCCGTGGACAT  
 CGCCGCCAAGCTGCTGGACCGGCTGCCCCGACGACGCCCGGGTCGCGGGCT

TCATCGATGGCCTCGCCGAGGGGCTTGAAGCCCTGCCCGAGGCGAGCCGG  
 GCGGTGATCGGCGTCGACGGCGCGCCAGTGCGCGTGACGGCCGCGCGCGC  
 CCTTATGCCGGCGGAGGAGGAGGCCTGCCGCACGCGGCTCTCCCAGGCGC  
 TGGGCCGTCCGGTGACGCTGGCCGTGACCATCGACCCCGCCCTCATCGCCG  
 GCCTGGAGATGGAGACGCCCCACGCGGTGGTGCGCAATTCCTTCAAGGCC  
 GATCTCGACCGCGTCACCGCGGCGCTCACCCATCATGGGACCTGA

SEQ ID NO: 23:

Trình tự axit amin của tiểu đơn vị b của ATP synthaza atpF\_1

MQIDWWTLGLQTVNVLVLIWLLSRFLFKPVAQVIAQRRAEIEKLVEDA  
 RAAKAAAEAERDTAKAEEARLAAERGARMAAVAKEAEAQKAALLAAKTE  
 AEALHAAAEAAIVRARASEEEAAADRASRLAVDIAAKLLDRLPDDARVAGFI  
 DGLAEGLEALPEASRAVIGVDGAPVRVTAARALMPAEIEACRTRLSQALGRP  
 VTLAVTIDPALIAGLEMETPHAVVRNSFKADLDRVTAALTHHGT

SEQ ID NO: 24:

Trình tự nucleotit của tiểu đơn vị c của ATP synthaza, đặc hiệu ion natri atpE\_1

ATGACTGTCGAGATGGTCAGCATCTTCGCGGGCGGGCGCTCGCCGTCTC  
 CTTCGGCGCCATCGGGCCGGCCCTGGGCGAGGGCCGGGCGGTGGCCGCGG  
 CCATGGACGCCATCGCCCGCCAGCCGGAGGCGGCCGGAACCTTGTCGCGC  
 ACGCTCTTCGTCGGCCTCGCCATGATCGAGACCATGGCGATCTACTGCCTG  
 GTGATCGCGCTCCTGGTGCTCTTCGCCAATCCGTTCTGTGAAGTGA

SEQ ID NO: 25:

Trình tự axit amin của tiểu đơn vị c của ATP synthaza, đặc hiệu ion natri atpE\_1

MTVEMVSIFAAALAVSFGAIGPALGEGRAVAAAMDIAIRQPEAAGTLS  
 RTLFFVGLAMIEITMAIYCLVIALLVLFANPFVK

SEQ ID NO: 26:

Trình tự nucleotit của tiểu đơn vị a của ATP synthaza atpB\_1

ATGGGCTCGCCGCTGATCCTCGAACCCCTGTTCCATATCGGGCCCCGT  
 GCCCATCACCGCGCCGGTGGTGGTCACCTGGCTCATCATGGCCGCCTTCAT  
 TGGGCTGGCGCGGCTCATCACCCGGAAGCTTTCCACCGATCCCACCCGGAC  
 CCAGGCGGCGGTGGAAACGGTGCTGACCGCCATCGATTCCCAGATCGCCG  
 ACACCATGCAGGCCGATCCCGCGCCTTATCGCGCGCTCATCGGCACCATCT  
 TCCTTTATGTGCTGGTGGCCAACTGGTCCTCGCTCATCCCGGGCATCGAGC

CGCCACGGCGCATATCGAGACCGATGCGGGCGCTCGCTTTCATCGTGTTTCG  
 CCGCCACCATCGGGTTCGGGTGAAGACAAGGGGTGTGAAGGGCTATCTC  
 GCCACCTTCGCCGAACCCTCCTGGGTGATGATCCCGCTCAATGTGGTGGAG  
 CAGATCACCCGGACCTTCTCGCTCATCGTGCGCCTGTTTCGGCAACATCATG  
 AGCGGGGTGTTTCGTGGTCGGCATCATCCTGTCCCTCGCCGGGCTGCTGGTG  
 CCCATCCCCCTCATGGCGCTCGATCTCCTGACCGGCGCCGTGCAGGCCTAC  
 ATCTTCGCGGTGCTGGCCTGCGTGTTTCATCGGCGCGGCCATTGGCGAGGCG  
 CCGGCAAAGCCCCAATCGAAGGAGCCAGGGAAAACATCATGA

SEQ ID NO: 27:

Trình tự axit amin của tiểu đơn vị a của ATP synthaza atpB\_1

MGSPLILEPLFHIGPVPITAPVVVTWLIMAAFIGLARLITRKLSTDPTRTQ  
 AAVETVLTAIDSQIADTMQADPAPYRALIGTIFLYVLVANWSSLIPGIEPPTAHI  
 ETDAALAFIVFAATIGFGLKTRGVKGYLATFAEPSWVMIPLNVEQITRTFSLI  
 VRLFGNIMSGVFVVGIIISLAGLLVPIPLMALDLLTGAVQAYIFAVLACVFIGA  
 AIGEAPAKPQSKEPGKTS

SEQ ID NO: 28:

Trình tự nucleotit của chuỗi epsilon của ATP synthaza atpC\_1

GTGAGCGCGCCGCTGCACCTCACCATCACCACGCCGGCCGCGTTCT  
 GGTGGACCGTGCCGACATCGTGGCCCTGCGTGCCGAGGACGAGAGCGGCA  
 GCTTCGGCATCCTGCCCCGGCCATGCGGATTTCTGACCGTTCTGGAGGCCT  
 GCGTGGTGCGCTTCAAGGATGGGGCCGACGGCGTGCAATTATTGTGCTCTCA  
 GTGGTGGCGTGCTGTGCGTCGAGGAGGGCCGGCGCATCGCCATCGCCTGC  
 CGTCAGGGCACGGTGAGCGACGACCTGGTCGCCCTGGAAGGGGCGGTGGA  
 CGCCATGCGTTTCGGCGGAGAGCGATGCCGACAAGCGGGCCCCGGGTGGAGC  
 AGATGCGCCTTCATGCCACGCCGTGCGCCAGCTCCTGCACTATCTGCGGC  
 CCGGCCGGGCGGCGGCGTGCGCGCCGGCCGCGCGCCGGAGGAGGGGCC  
 GTCATGA

SEQ ID NO: 29:

Trình tự axit amin của chuỗi epsilon của ATP synthaza atpC\_1

MSAPLHLTITTPAAVLVDRADIVALRAEDESFSFILPGHADFLTVLEAC  
 VVRFKDGADGVHYCALSGGVLSVEEGRRIACRQGTVSDDLVALEGAVDA  
 MRSAESDADKRARVEQMRLHAHAVRQLLHYLRPGRAGGVAPAAAPEEGPS

SEQ ID NO: 30:

Trình tự nucleotit của tiểu đơn vị beta của ATP synthazaeta atpD\_1

ATGGCAGCGGCAGATGAGGAGGCGCAATCGGCCGCCGGCCCCGCCT  
CGGGCCGGGTGGTGGCCGTGCGCGGCGCGGTGATCGACATCGCCTTTGCC  
CAGCCTCCGCTGCCGCCGCTGGACGACGCCCTTCTCATCACCGACGGCCGG  
GGCGGCACGGTGCTGGTGGAGGTGCAGAGCCATATGGATCGGACACGGT  
GCGCGCCATCGCCCTTCAGGCCACCACCGGCCTCAGCCGGGGGCTGGAGG  
CGGCGCGGGTGGGCGGGCCGGTGAAGGTGCCGGTGGGAGACCATGTGCTC  
GGCCGCCTCCTGGATGTACCGGCGCCATCGGCGACAAGGGCGGGCCGCT  
GCCGGCCGACGTGCCACGCGGCCGATCCACCACGCGCCGCCATCCTTCGC  
CGCGCAGGGCGGCACGTCCGATCTGTTTCGCACCGGCATCAAGGTCATCG  
ACCTCCTGGCGCCCCTCGCCCAGGGCGGCAAGGCGGCCATGTTCGGCGGG  
GCCGGCGTGGGCAAGACCGTGCTGGTGGTGGAGCTGATCCACGCCATGGT  
GGCGAGCTACAAGGGCATCTCGGTGTTTGCCGGCGTGGGGGAGCGCTCCC  
GCGAGGGCCACGAGATGCTGCTGGACATGACCGATTCCGGCGTGCTCGAC  
CGCACCGTTCTGGTCTATGGCCAGATGAACGAGCCCCCGGGGCCCGCTG  
GCGGGTGCCCATGACGGCGCTGACCATCGCCGAATATTTCCGCGACGAGA  
AGCACCAGAACGTCCTGCTGCTGATGGACAACATCTTCCGCTTCGTCCAGG  
CGGGGGCGGAGGTCTCCGGCCTTTTGGGCGGTCCGCCCTCCCGGGTGGGAT  
ACCAGCCGACGCTGGCGAGCGAGGTGGCGGCGCTCCAGGAACGCATCACC  
TCCGTGGGCGAGGCCTCGGTGACCGCCATCGAGGCGGTCTACGTGCCGGC  
GGATGACTTCACCGATCCCGCCGTGACCACCATCGCCGCCACGTGGATTC  
CATGGTGGTGCTCTCCCGCGCCATGGCGGCGGAGGGCATGTATCCGGCGG  
TGGACCCCATCTCCTCCTCGTCGGTGCTGCTCGACCCGCTCATCGTGGGGG  
ACGAGCATGCGCGCGTCGCCAACGAGGTGCGCCGGACCATCGAGCATTAT  
CGCGAGCTTCAGGATGTGATCTCGCTGCTGGGCATGGAGGAATTGGGCAC  
CGAGGATCGCCGCATCGTGGAGCGGGCGCGCCGGCTCCAGCGCTTCCTCA  
CCCAGCCCTTCACGGTCACCGAGGCCTTCACCGGCGTGCCCGGCCGCTCGG  
TGGCCATCGCCGACACCATCGCCGGCTGCAGGATGATCCTGTCCGGCGCCT  
GCGACGACTGGCAGGAAAGCGCCCTCTACATGGTGGGCACCATCGACGAG  
GCCCCGCCAGAAGGAGGAGGCCGCTCGCGCCAAGGCGGGGCAGGGCGCCC  
CGGCCGGGACGGCAGCCGAGACGGCGGAGGCCGCCCGCCCGTGA

SEQ ID NO: 31:

Trình tự axit amin của tiểu đơn vị beta của ATP synthazaeta atpD\_1

MAAADEEAQSAAGPASGRVVAVRGAVIDIAFAQPPLPPLDDALLITDGR  
GGTVLVEVQSHMDRHTVRAIALQATTGLSRGLEAARVGGPVKVPVGDHVLG  
RLLDVTGAIGDKGGPLPADVPTRPIHHAPPSFAAQGGTSDLFRTGIKVIDLLAP  
LAQGGKAAMFGGAGVGKTVLVMELIHAMVASYKGISVFAGVGERSREGHE  
MLLDMTDSGVLDRTVLVYGQMNPPGARWRVPMTALTIAEYFRDEKHQNVL  
LLMDNIFRFVQAGAEVSGLLGRPPSRVGYQPTLASEVAALQERITSVGEASVT  
AIEAVYVPADDFTDPAVTTIAAHVDSMVVLSRAMAAEGMYPAPVDPISSSSVLL  
DPLIVGDEHARVANEVRRRTIEHYRELQDVISLLGMEELGTEDRRIVERARRLQ  
RFLTQPFTVTEAFTGVPGRSVAIADTIAGCRMILSGACDDWQESALYMGVTID  
EARQKEEAARAKAGQGAPAGTAAETAEAAP

SEQ ID NO: 32:

Trình tự nucleotit của tiểu đơn vị beta của ATP synthazaeta atpD\_2

ATGGCGAACAAGGTCGGACGCATCACCCAGATCATCGGCGCCGTCG  
TCGACGTGCAGTTCGACGGGCATCTGCCGGCGATTCTCAACGCGATCGAG  
ACCACCAACCAGGGCAACCGGCTGGTGCTCGAAGTGGCTCAGCATCTCGG  
CGAGAACACCGTGCGCTGCATCGCCATGGATGCCACTGAAGGCCTGGTGC  
GTGGCCAGGAGGTGGCCGACACCGATGCGCCCATCCAGGTGCCCCGTGGGC  
GCCGCCACCCTCGGCCGCATCATGAACGTGATCGGCGAGCCGGTGGACGA  
GCTGGGCCCCATCGAGGGCGAAGCGCTGCGCGGCATCCATCAGCCGGCCC  
CCTCCTATGCGGAGCAGGCCACGGAAGCTGAGATCCTCGTCACCGGCATC  
AAGGTGGTGGATCTGCTGGCGCCCTATTCCAAGGGCGGCAAGGTGGGCCT  
GTTTCGGCGGCGCCGGCGTGGGCAAGACCGTGCTCATCATGGAGCTGATCA  
ACAACGTGGCCAAGGCGCACGGCGGCTATTCCGTGTTTCGCCGGCGTGGGT  
GAGCGCACCCGCGAGGGCAACGACCTCTACCACGAGATGATCGAGTCCAA  
CGTGAACAAGGACCCGCACGAGAACAATGGCTCGGCGGCGCGGTTCCAAGT  
GCGCCCTGGTCTATGGCCAGATGAACGAGCCGCCCGGCGCCCGCGCCCGC  
GTGGCCCTCACCGGCCTCACCGTCGCCGAGCATTTCCGCGACCAGGGCCAG  
GACGTGCTGTTCTTCGTGGACAACATCTTCCGCTTCACCCAGGCGGGCTCC  
GAGGTGTCGGCGCTTCTCGGCCGCATCCCCCTCGGCGGTGGGCTACCAGCCG  
ACGCTGGCCACCGACATGGGCCAGCTGCAGGAGCGCATCACCAACCAC

CAAGGGCTCCATCACCTCGGTGCAGGCCATCTACGTGCCGGCGGACGATCT  
 GACCGATCCGGCGCCGGCCGCTCCTTCGCCCATCTGGACGCCACCACGGT  
 GCTGTCGCGCTCCATCGCGGAGAAGGGCATCTACCCGGCGGTGGATCCGC  
 TGGACTCCACCTCGCGCATGCTGTCTCCCGCCATCCTCGGCGACGAGCACT  
 ACAACACCGCGCGCCAGGTGCAGCAGACCCTGCAGCGCTACAAGGCGCTC  
 CAGGACATCATCGCCATCCTGGGCATGGACGAACTCTCCGAAGAGGACAA  
 GCTCACCGTGGCCCGCGCCCGCAAGATCGAGCGCTTCCTCTCCCAGCCCTT  
 CCACGTGGCCGAGGTGTTACCGGTTTCGCCCCGGCAAGCTGGTCGACCTCGC  
 CGACACCATCAAGGGCTTCAAGGGCCTGGTGGACGGCAAGTACGACTACC  
 TGCCCGAGCAGGCCTTCTACATGGTGGGCACCATCGAAGAAGCCATCGAG  
 AAGGGCAAGAAGCTGGCGGCCGAGGCGGCCTGA

SEQ ID NO: 33:

Trình tự axit amin của tiểu đơn vị beta của ATP synthaza atpD\_2

MANKVGRITQIIGAVVDVQFDGHLPAILNAIETTNQGNRLVLEVAQHLLG  
 ENTVRCIAMDATEGLVRGQEVADTDAPIQVPVGAATLGRIMNVIGEPVDELGP  
 IEGEALRGIHQPAPSYAEQATEAEILVTGIKVVDLLAPYSKGGKVGLFGGAGV  
 GKTVLIMELINNVAKAHGGYSVFAGVGERTREGNDLYHEMIESNVNKDPHEN  
 NGSAAAGSKCALVYGQMNEPPGARARVALTGLTVAEHFRDQGQDVLFFVDNI  
 FRFTQAGSEVSALLGRIPSAVG YQPTLATDMGQLQERITTTTKGSITSVQAIYV  
 PADDLTDPAASFAHLDATTVLSRSIAEKGIYPVDPLDSTSRMLSPAILGDE  
 HYNTARQVQQTLQRYKALQDIIAILGMDELSEEDKLTVARARKIERFLSQPFH  
 VAEVFTGSPGKLVDLADTIKGFKGLVDGKYDYLPQAFYMGVTIEEAIEKGK  
 KLAAEAA

SEQ ID NO: 34:

Trình tự nucleotit của chuỗi gamma của ATP synthaza atpG\_2

ATGGCGAGTCTGAAGGACCTGAGAAACCGCATTGCCTCGGTGAAGG  
 CGACGCAGAAGATCACCAAGGCGATGCAGATGGTCGCCGCGGCGAAGCTG  
 CGTCGCGCCCAGGCGGCGGCTGAAGCGGCCCCGTCCCTATGCGGAACGCAT  
 GGAGACGGTGCTCGGAAATCTTGCCTCCGGCATGGTGGTGGGCGCGCAGG  
 CGCCTGTTCTCATGACCGGGACGGGCAAGAGCGACACCCACCTGCTGCTG  
 GTGTGCACCGGCGAGCGCGGCCTGTGCGGCGCCTTCAACTCGTCCATCGTG  
 CGCTTCGCCCCGCGAGCGGGCGCAGCTGCTGCTGGCCGAGGGCAAGAAGGT

GAAAATCCTGTGCGTGGGCGCAAGGGCCACGAGCAGCTGCGCCGCATCT  
 ACCCGGACAACATCATCGACGTGGTGGACCTGCGCGCGGTGCGCAACATC  
 GGCTTCAAGGAGGCCGACGCCATCGCCCGCAAGGTGCTGGCCCTGCTCGA  
 TGAAGGCGCATTTCGACGTCTGCACGCTCTTCTACTCCCACTTCAGGAGCGT  
 GATCGCCCAGGTGCCGACGGCCCAGCAGCTCATTCGGGCCACCTTCGACG  
 AGCGGCCGGCCGTCGCCGATGCGCCGGTCTATGAATATGAGCCGGAGGAG  
 GAGGAGATCCTCGCCGAGCTGCTGCCGCGCAACGTGGCGGTGCAGATCTT  
 CAAGGCCCTCCTCGAGAACCAGGCTTCTTTCTATGGCTCCCAGATGAGCGC  
 CATGGACAACGCCACGCGCAATGCGGGCGAGATGATCAAGAAGCAGACG  
 CTCACCTACAACCGTACCCGCCAGGCCATGATCACGAAGGAACATCGA  
 GATCATCTCCGGCGCCGAGGCCGTCTGA

SEQ ID NO: 35:

Trình tự axit amin của chuỗi gamma của ATP synthaza atpG\_2

MASLKDLRNRIASVKATQKITKAMQMVA AAKLRRAQAAAEAA RPYAE  
 RMETVLGNLASGMVVG A QAPVLM TGTGKSDTHLLLVCTGERGLCGAFNSSI  
 VRFARERAQLLLAEGKKVKILCVGRKGHEQLRRIYPDNIIDVVDLRAVRNIGF  
 KEADAIARKVLALLDEGA F DVCTLFYSHFRSVIAQVPTAQQLIPATFDERPAV  
 ADAPVYEYEP EEEEEILAE LLPRNVAVQIFKALLENQASFYGSQMSAMD NATRN  
 AGEMIKKQTLTYNRTRQAMITKELIEIISGA EAV

SEQ ID NO: 36:

Trình tự nucleotit của tiểu đơn vị alpha của ATP synthaza atpA\_2

ATGGACATTCGAGCCGCTGAAATCTCTGCCATCCTGAAAGAGCAGA  
 TCCAGAATTTTCGGCCAGGAGGCGGAAGTCTCCGAGGTGGGTCAGGTTCTGT  
 CCGTGGGTGACGGCATCGCGCGCGTCTACGGCCTCGACAACGTCCAGGCG  
 GGCGAGATGGTCGAGTTCGAGAACGGCACGCGCGGCATGGCGCTGAACCT  
 CGAGCTCGACAATGTCGGCATCGTGATCTTCGGTTCCGACCGCGAGATCAA  
 GGAAGGCCAGACCGTCAAGCGGACCGGCGCCATCGTGGACGCCCCCGTCG  
 GCAAGGGCCTGCTCGGCCGCGTCGTGGACGCTCTCGGCAACCCGATCGAC  
 GGCAAGGGCCCGATCATGTTACCGAGCGTCGCCGGGTTCGACGTGAAGGC  
 GCCGGGCATCATCCCGCGCAAGTCGGTGCACGAGCCCATGCAGACCGGCC  
 TGAAGGCCATCGATGCGCTCATCCCCATCGGCCGCGGCCAGCGCGAGCTC  
 ATCATCGGCGACCGCCAGACCGGCAAGACCGCCGTGGCGCTCGACTCGAT

CCTGAACCAGAAGCCCATCAACCAGGGCGACGACGAGAAGGCCAAGCTCT  
 ACTGCGTCTATGTCGCGGTGGGCCAGAAGCGTTCCACTGTCGCGCAGTTTCG  
 TGAAGGTGCTCGAGGAGCACGGCGCGCTGGAATATTCCATCGTCGTCGCC  
 GCCACCGCCTCGGACGCGGCCCCCATGCAGTTCCTGGCGCCGTTACCGGC  
 ACCGCCATGGGCGAGTATTTCCGCGACAACGGCATGCACGCCCTCATCATC  
 CATGATGACCTGTCCAAGCAGGCCGTGGCCTACCGCCAGATGTCGCTGCTG  
 CTGCGCCGCCCGCCGGGCGCGAGGCCTATCCCGGCGATGTGTTCTACCTG  
 CACTCCCGCCTCTTGAGCGCGCCGCCAAGCTCAATGACGAGCACGGCGC  
 CGGCTCGCTGACCGCCCTGCCGGTGATCGAGACCCAGGCCAACGACGTGT  
 CGGCCTACATCCCGACCAACGTGATCTCCATCACCGACGGTCAGATCTTCC  
 TTGAATCCGATCTGTTCTACCAGGGCATCCGCCCCGGCGGTGAACGTGGGCC  
 TGTCGGTGTCGCGCGTGGGCTCTTCGGCCCAGATCAAGGCGATGAAGCAG  
 GTGGCCGGCAAGATCAAGGGCGAGCTCGCCCAGTATCGCGAGCTGGCGGC  
 CTTGCGCCAGTTCGGTTCGGACCTGGACGCGGCCACCCAGAAGCTGCTGAA  
 CCGCGGCGCCCGCCTACCGAGCTGCTGAAGCAGAGCCAGTTCTCGCCCCCT  
 CAAGGTGGAGGAGCAGGTGGCGGTGATCTATGCCGGCACCAATGGCTATC  
 TCGATCCGCTGCCGGTCTCCAAGGTGCGCGAGTTCGAGCAGGGTCTGCTCC  
 TGTCGCTGCGCTCGCAGCATCCGGAGATCCTGGACGCCATCCGCACGTCCA  
 AGGAGCTTTCCAAGGACACCGCCGAGAAGCTGACGAAGGCCATCGACGCC  
 TTCGCCAAGAGCTTCTCCTGA

SEQ ID NO: 37:

Trình tự axit amin của tiểu đơn vị alpha của ATP synthaza atpA\_2

MDIRAAEISAILKEQIQNFGQEAEVSEVGQVLSVGDGIARVYGLDNVQA  
 GEMVEFENGTRGMALNLELDNVGIVIFGSDREIKEGQTVKRTGAIVDAPVGKG  
 LLGRVVDALGNPIDGKGPIMFERRRVDVKAPGIIPRKSVHEPMQTGLKAIDA  
 LIPIGRGQRELIIGDRQTGKTAVALDSILNQKPINQGDDEKAKLYCVYVAVGQ  
 KRSTVAQFVKVLEEHLGALEYISIVVAATASDAAPMQFLAPFTGTAMGEYFRDN  
 GMHALIIHDDLKQAVAYRQMSLLLRPPGREAYPGDVFYLSRLLERAAKL  
 NDEHGAGSLTALPVIETQANDVSAYIPTNVISITDQGIFLESDFYQGIRPAVNV  
 GLSVSRVGSSAQIKAMKQVAGKIKGELAQYRELAFAAQFGSDLDAATQKLLN  
 RGARLTELLKQSQFSPLKVEEQVAVIYAGTNGYLDPLPVSKVREFEQGLLSL  
 RSQHPEILDAIRTSKELSKDTAEKLTKAIDAFKSF

SEQ ID NO: 38:

Trình tự nucleotit của tiểu đơn vị delta của ATP synthaza atpH

GTGGCGGAAACGATCGTGTTCAGGCATGGCGGGACGCTATGCGACCG  
CGCTGTTCGAGCTGGCGGACGAAGCCGGTGCCATCGATTCCGTCCAGGCG  
GATCTTGATCGCCTGTCCGGCCTTCTGGCCGAGAGCGCGGATCTGGCGCGG  
CTGGTCAAGAGCCCGGTCTTCACCGCCGAGCAGCAGCTCGGGCGCGATGGC  
GGCCATTCTCGATCAAGCAGGCATTTCCGGCCTTGCGGGCAAATTCGTGAA  
GCTGGTGGCGCAGAACCGCCGCCTGTTTCGCACTGCCGCGCATGATTGCCGA  
ATACGCCGTCCTGGTGGCCCGGAAGAAGGGCGAGACCTCGGGCGAGCGTGA  
CCGTTGCCACCCCCCTGAGCGATGAGCATCTGGCCACGCTCAAGGCGGCCC  
TGGCTGAAAAGACCGGCAAGGACGTGAAGCTCGACGTCACCGTCGATCCG  
TCCATCCTCGGTGGTCTCATCGTGAAGCTCGGCTCGCGCATGGTCGATGCT  
TCCCTGAAGACCAAACCTCAATTCTATCCGGCATGCGATGAAAGAGGTCCG  
CTGA

SEQ ID NO: 39:

Trình tự axit amin của tiểu đơn vị delta của ATP synthaza atpH

MAETIVSGMAGRYATALFELADEAGAIIDSVQADLDRLSGLLAESADLA  
RLVKSPVFTAEEQLGAMAILDQAGISGLAGKFKVLVAQNRRLFALPRMIAE  
YAVLVARKKGETSASVTVAATPLSDEHLATLKAALAEKTGKDVKLDVTVDPSI  
LGGLIVKLGSRMVDASLKTCLNSIRHAMKEVR

SEQ ID NO: 40:

Trình tự nucleotit của tiểu đơn vị b của ATP synthaza atpF<sub>2</sub>

ATGACCGAAATGGAAGTGGCTGAGCTCTGGGTCGCCATCGCCTTCCT  
GGTTTTTCGTAGGCCTCCTGATCTATGCGGGCGCCCACCGCGCCATCGTCTC  
CGCCCTGGATTCCCGCGGCTCGCGCATCGCCTCGGAAGTGGAGGAGGCCC  
GTCGGCTCAAGGAAGAGGCCAGAGCTGGTGGCCGAATTCAAGCGCAAG  
CAGCGCGAGGCCGAGGCCGAGGCCGAATCCATCGTCACCGGCGCCAAGGC  
CGAGGCCGAGCGCCTCGCCGCCGAGGCCAAGGCGAAGATCGAGGATTTTCG  
TCACCCGCCGCACCAAGATGGCCGAGGACAAGATCGCCCAGGCCGAGCAT  
CAGGCTCTGGCGGACGTGAAGTCCATCGCCGCCGAGGCGGCGGCCAAGGC  
GGCCGAGGTGATCCTCGGGCGCCCAGGCCACCGGCGCGGTGGCGGAGCGTC  
TGCTGTGCGGGCGCCATCTCCGAGGTCAAGACCAAGCTCAACTGA

SEQ ID NO: 41:

Trình tự axit amin của tiểu đơn vị b của ATP synthaza atpF\_2

MTEMELAE LWVAIAFLVFVGLLIYAGAHRAIVSALDSRGSRIASELEEA  
RRLKEEAQKLVAEFKRKQREAEAEAESIVTGAKAEAERLAAEAKAKIEDFVTR  
RTKMAEDKIAQAEHQALADVKSIAAEAAKAAEVILGAQATGAVAERLLSGA  
ISEVKTCLN

SEQ ID NO: 42:

Trình tự nucleotit của tiểu đơn vị b' của ATP synthaza atpG\_3

ATGATGATTGCATGGAAGCGGACCTTCGCAGTCGTGACCTTCGGGGC  
CGCCCTGATGGCCATGCCCCGTCGCGGGCGTGGTCGCAGCTGAGACTTCTCC  
CGCTCCGGCGGCAGTGGCGCAGGCCGATCATGCGGTGCCACCGAGGCGG  
CCGGCCAGGGCACCGCCGATGCGGCCCATGCCGCCGCGCCGGGCGAGGCC  
GCCCATGGTGGCGCGGCCAAGCACGAAACCCATTTCCCGCCCTTCGACGG  
CACCACCTTCGCTCCAGTTGCTGTGGCTCGCCGTCACCTTCGGCCTGCTT  
TACTACCTCATGAGCAAGGTCACGCTGCCGCGCATCGGCCGCATCCTGGAA  
GAGCGCCACGACCGCATCGCCGATGATCTGGAGGAAGCCTCCAAGCATCG  
CGCCGAGAGCGAGGCCGCCAGCGGGCCTATGAGAAGGCGCTGAGCGAG  
GCCCCGCGGAAGGCCATTCCATCGCCGCGGAAACCCGCGACCGCCTTGC  
CGCCACGCCGACACCAACCGCAAGGCGCTGGAGAGCGAGCTCACCGCCA  
AGCTGCAGGCGGCCGAGGAGCGCATCGCCACCACCAAGAGCGAAGCCCTC  
ACCCATGTGCGCGGCATCGCGGTGGACGCCACCCAATCCATCGTCTCCACC  
CTCATCGGTGTCGCGCCCCGCGGCGGCCGACGTGGAAAAAGCGGTGGACGG  
CGCCCTGTCCAGCACGGCCAGGCCTGA

SEQ ID NO: 43:

Trình tự axit amin của tiểu đơn vị b' của ATP synthaza atpG\_3

MMIAWKRTFAVVTFGAALMAMPVAGVVAAETSPAPAAVAQADHAVP  
TEAAGQGTADAAHAAAPGEAAHGGA AKHETHFPPFDGTTFASQLLWLA VTF  
GLLYYLM SKVTLPRI GRILEERHDRIADDLEEASKHRAESEAAQRAYEKAL SE  
ARAKAHSIAAETRDRLAAHADTNRKALESEL TAKLQAAEERIATTKSEAL THV  
RGIAVDATQSIVSTLIGVAPAAADVEKA VDGALSQH GQA

SEQ ID NO: 44:

Trình tự nucleotit của tiểu đơn vị c của ATP synthaza atpE\_2

ATGGAAGCGGAAGCTGGAAAGTTCATCGGTGCCGGCCTCGCCTGCC  
TCGGCATGGGTCTCGCTGGCGTCGGCGTCGGTAACATCTTCGGTAACTTCC  
TCTCCGGCGCCCTGCGCAACCCGTCCGCTGCCGACGGCCAGTTCGCCCCGCG  
CCTTCATCGGCGCCGCCCTCGCGGAAGGTCTCGGCATCTTCTCGCTGGTTCG  
TTGCGCTCGTCCTGCTGTTTCGTGGCCTGA

SEQ ID NO: 45:

Trình tự axit amin của tiểu đơn vị c của ATP synthaza atpE\_2

MEAEAGKFIGAGLACLGMGLAGVGVGNIFGNFLSGALRNPSAADGQF  
ARAFIGAALAEGLGIFSLVVALVLLFVA

SEQ ID NO: 46:

Trình tự nucleotit của tiểu đơn vị a của ATP synthaza atpB\_2

ATGACCGTCGATCCGATCCACCAGTTCGAGATCAAGCGCTACGTGG  
ATCTGCTGAACGTCGGCGGTGTCCAGTTCTCCTTACCAACGCAACGGTGT  
TCATGATTGGCATCGTCCTGGTGATTTTCTTCTTCCTGACTTTCGCGACACG  
CGGTCGCACCCTTGTGCCGGGCGGATGCAGTCGGCGGGCGGAGCTGAGCT  
ACGAGTTCATCGCCAAGATGGTGCGCGACGCGGCCGGCAGCGAGGGAATG  
GTGTTCTTTCCCTTCGTCTTCTCGCTCTTCATGTTTCGTGCTGGTGGCGAACG  
TATTGGGGCTCATCCCCTACACCTTCACGGTGACCGCCACCTCATCGTCA  
CCGCCGCCCTGGCGGGCGACGGTGATCCTCACCGTCATCATCTACGGCTTCG  
TGCGGCACGGCACCCACTTCCTGCACCTGTTCGTGCCGTCGGGCGTGCCGG  
GCTTCCTCCTGCCCTTCCTCGTGGTGATCGAGGTGGTGTCTTCCTGTCGCG  
GCCCATCAGCCTCTCGCTGCGTCTGTTCGCCAACATGCTGGCGGGGCCACAT  
CGCCCTCAAGGTGTTTCGCCTTCTTCGTGCTGGGACTGGCCTCGGCCGGCGC  
GATCGGCTGGTTCGGCGCCACCCTGCCCTTCTTCATGATCGTGGCGCTCAC  
CGCGCTGGAGCTGCTGGTGGCGGTGCTGCAGGCCTACGTGTTTCGCGGTGCT  
GACCTCGATCTACCTCAACGACGCCATCCATCCCGGCCACTGA

SEQ ID NO: 47:

Trình tự axit amin của tiểu đơn vị a của ATP synthaza atpB\_2

MTVDPIHQFEIKRYVDLLNVGGVQFSFTNATVFMIGIVLVIFFLTFATR  
GRTLVPGRMQSAAELSYEFIAKMVRDAAGSEGMVFFPFVFSLFMFVLVANVL  
GLIPYTFTVTAHLIVTAALAATVILTVIIYGFVRHGTHFLHLFVPSGVPGFLLPFL  
VVIEVVSFLSRPISLSLRLFANMLAGHIALKVFAFFVVGLASAGAIGWFGATLP

FFMIVALTALELLVAVLQAYVFAVLTSIYLNDAIHPGH

SEQ ID NO: 48:

Trình tự nucleotit của protein I của ATP synthaza atpI

ATGTCCGAGCCGAATGATCCATCCCGCAGGGACGGTGCGAAGGCGA  
AAGACGAGACGCAGGACTCCCGGCCCGGTGAGGCGGATCTTGCTCGGCGC  
CTCGATGCGCTCGGCACCTCCATCGGTCAAGTCCAGAAGCGGGGA  
GCCCCGCGGCGACGCCGCGCAAGGACACCTCCTCGGCCTCCGGCGCGGCCC  
TGGCGTTTCGGCTGGGCGCCGAGTTTGTTCAGGCGTGCTGGTGGGCTCGC  
TCATCGGCTACGGGTTGGATTATGCGTTTGCGATTTTCGCCCTGGGGGCTGA  
TCGCCTTCACGCTGATCGGCTTTGCCGCCGGCGTCCTGAACATGCTGCGCG  
TGGCGAACAGCGATGCCAAGCGCCACAGCGCGGACAGGTGA

SEQ ID NO: 49:

Trình tự axit amin của protein I của ATP synthaza atpI

MSEPNDPSRRDGAKAKDETQDSRPGEADLARRLDALGTSIGQVKSRSG  
EPAATPRKDTSSASGAALAFRLGAEFVSGVLVGSLIGYGLDYAFAISPWGLIAF  
TLIGFAAGVLNMLRVANSDAKRHSADR

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho chuỗi alpha của protein molybden-sắt của nitrogenaza nifD\_1 có trình tự nêu trong SEQ ID NO:51 hoặc trình tự tương đồng hơn 60%, ví dụ tương đồng hơn 70%, chẳng hạn như tương đồng hơn 92%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:51.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho chuỗi alpha của protein molybden-sắt của nitrogenaza nifD\_2 có trình tự nêu trong SEQ ID NO:53 hoặc trình tự tương đồng hơn 60%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:53.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho chuỗi beta của protein molybden-sắt của nitrogenaza nifK\_1 có trình tự nêu trong SEQ ID NO:55 hoặc trình tự tương đồng hơn 87%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng

hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:55.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho chuỗi beta của protein molybden-sắt của nitrogenaza nifK\_2 có trình tự nêu trong SEQ ID NO:57 hoặc trình tự tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:57.

Theo phương án khác, chủng vi khuẩn được sử dụng trong quy trình theo sáng chế có chứa gen mã hóa cho protein sắt của nitrogenaza nifH có trình tự nêu trong SEQ ID NO:59 hoặc trình tự tương đồng trình tự hơn 98,5% so với trình tự nêu trong SEQ ID NO:59.

SEQ ID NO: 50:

Trình tự nucleotit của chuỗi alpha của protein molybden-sắt của nitrogenaza nifD\_1

```

ATGAGTTCGCTCTCCGCCACTATTCAACAGGTCTTCAACGAGCCGGG
CTGCGCGAAGAACCAGAATAAGTCCGAGGCGGAGAAGAAGAAGGGCTGC
ACCAAGCAGCTGCAACCCGGCGGAGCGGCCGGCGGCTGCGCGTTTCGACGG
CGCGAAGATCGCGCTCCAGCCCTTGACCGACGTCGCCCACCTGGTGCACG
GCCCCATCGCCTGCGAAGGCAATTCCTGGGACAATCGTGGCGCCAAGTCCT
CCGGCTCGAACATCTGGCGCACCGGCTTCACCACGGACATCAACGAAACC
GACGTGGTGTTCGGCGGCGAGAAGCGTCTGTTCAAGTCCATCAAGGAAAT
CATCGAGAAGTACGACCCGCCGGCGTCTTCGTCTATCAGACCTGCGTCCC
CGCCATGATCGGCGACGACATCGACGCGGTGTGCAAGGCGGCCAGGGAGA
AGTTCGGAAAGCCGGTGATCCCGATCAATCCCCCGGCTTCGTGGGGCCGA
AGAATCTCGGCAACAAGCTCGCCGGCGAGGCGCTCCTCGACCATGTGATC
GGCACCGAGGAGCCCGATTACACGACGGCCTACGACATCAACATCATCGG
CGAATACAATCTCTCCGGCGAGTTGTGGCAGGTGAAGCCGCTGCTGGACG
AGCTGGGCATCCGCATCCTCGCCTGCATCTCCGGCGACGGGAAGTACAAG
GATGTGGCGTCCTCCCACCGCGCCAAGGCGGCGATGATGGTGTGCTCCAA
GGCCATGATCAACGTGGCCCGCAAGATGGAGGAGCGCTACGACATCCCCT
TCTTCGAAGGCTCCTTCTACGGCATCGAGGATAGCTCCGATTCCCTGCGCG
AGATTGCGCGCATGCTCATCGAGAAGGGCGCCGATCCGGAGCTGATGGAC

```

CGCACCGAGGCGCTGATTGAGCGGGAAGAGAAGAAGGCGTGGGACGCCA  
 TCGCCGCCTACAAGCCCCGCTTCAAGGACAAGAAGGTGCTGCTCATCACC  
 GGCGGCGTGAAATCCTGGTCGGTGGTGGCAGCGCTCCAGGAAGCCGGCCT  
 CGAACTGGTGGGCACCTCGGTGAAGAAGTCCACCAAGGAGGACAAGGAG  
 CGCATCAAGGAACTGATGGGCCAGGACGCCCACATGATCGACGACATGAC  
 GCCCCGCGAAATGTACAAGATGCTGAAGGACGCCAAGGCGGACATCATGC  
 TCTCGGGCGGGCGCTCGCAATTCATCGCGCTCAAGGCCGCCATGCCCTGGC  
 TCGACATCAACCAGGAGCGCCACCACGCCTATATGGGCTATGTGGGCATG  
 GTGAAGCTGGTCGAGGAGATCGACAAGGCGCTCTACAATCCCGTGTGGGA  
 ACAGGTGCGCAAGCCCGCCCCGTGGGAAAATCCGGAAGACACCTGGCAGG  
 CCCGTGCGCTCGCCGAAATGGAGGCGGAGGCCGCCGCGCTCGCCGCCGAT  
 CCGGTGCGCGCGGAAGAGGTGCGCCGGTCCAAGAAGATCTGCAATTGCAA  
 GAGCGTCGACCTCGGAACCATTTAGGACGCCATCAAGGCTCACGCGCTGA  
 CCACCGTGGAGGGTGTGCGAGAGCACACCAATGCCTCGGGAGGCTGCGGA  
 GCCTGCAGCGGGCGGATCGAGGAGATCTTCGAGGCCGTGGGGCGTTGTGCG  
 CGCCCCGCCTCCCGCGGAGGCCGCCCGTCTCCGCAGGAGATCGCGCCCCG  
 ATCCGCTCGCTGCGGAGGAAAAGCGCCGCGCCAAGAAGGCCTGCGGCTGC  
 AAGGAGGTAGCGGTTCGGCACCATTGAGGATGCCATCCGCGCCAAGGGTCT  
 GCGAAACATCGCGGAGGTGCGTGCGGCCACCGATGCCAACACCGGCTGCG  
 GCAATTGCCAGGAGCGGGTGGAGGGCATCCTCGACCGGGTTCTCGCCGAG  
 GCGGCCTCAGAACTCCAGGCGGCGGAATAG

SEQ ID NO: 51:

Trình tự axit amin của Chuỗi alpha của protein molybden-sắt của nitrogenaza  
 nifD\_1

MSSLSATIQVFNEPGCAKNQNKSEAEKKKGCTKQLQPGBAAGGCAF  
 DGAKIALQPLTDVAHLVHGPIACEGNSWDNRGAKSSGSNIWRTGFTTDINETD  
 VVFGGEKRLFKSIKEIIEKYDPPAVFVYQTCVPAMIGDDIDAVCKAAREKFGKP  
 VIPINSPGFVGPKNLGNKLAGEALLDHVIGTEEPDYTTAYDINIIGEYNLSGELW  
 QVKPLLDELGIRILACISGDGKYKDVASSHRAKAAMMVCSKAMINVARKMEE  
 RYDIPFFEGSFYGIEDSSDSLREIARMLIEKGADPELMDRTEALIEREEKKAWD  
 AIAAYKPRFKDKKVLLITGGVKSWSVVAALQEAGLELVGTSVKKSTKEDKERI  
 KELMGQDAHMIDDMTPREMYKMLKDAKADIMLSGGRSQFIALKAAMPWLDI

NQERHHAYMGYVGMVKLVEEIDKALYNPVWEQVRKPAPWENPEDTWQARA  
 LAEMEAEEAALAADPVRAEEVRRSKKICNCKSVDLGTIEDAIKAHALTTVEG  
 VREHTNASGGCGACSGRIEEIFEAVGVVAAPPPAEAAPSPQEIAPDPLAAEEKR  
 RAKKACGCKEVAVGTTIEDAIRAKGLRNIAEVRAATDANTGCGNCQERVEGIL  
 DRVLAEAASELQAAE

SEQ ID NO: 52:

Trình tự nucleotit của chuỗi alpha của protein molybden-sắt của nitrogenaza  
 nifD\_2

ATGAGTGTCGCACAGTCCCAGAGCGTCGCCGAGATCAAGGCGCGCA  
 ACAAGGAACTCATCGAAGAGGTCTCAAGGTCTATCCCGAGAAGACCGCC  
 AAGCGCCGCGCCAAGCACCTGAACGTCCACGAAGCCGGCAAGTCCGACTG  
 CGGCGTGAAGTCCAACATCAAGTCCATCCCGGGCGTGATGACCATCCGCG  
 GTTGCGCTTATGCCGGCTCCAAGGGTGTGGTGTGGGGTCCCATCAAGGACA  
 TGATCCACATCTCCACGGCCCGGTGGGCTGCGGCCAGTATAGCTGGGCCG  
 CCCGCCGCAACTACTATATCGGCACGACCGGCATCGACACCTTCGTGACGA  
 TGCAGTTCACCTCCGACTTCCAGGAGAAGGACATCGTCTTCGGCGGCGACA  
 AGAAGCTCGCCAAGATCATGGACGAGATCCAGGAGCTGTTCCCGCTGAAC  
 AACGGCATCACCGTTCAGTCCGAGTGCCCCATCGGCCTCATCGGCGACGAC  
 ATCGAGGCCGTCTCCAAGCAGAAGTCCAAGGAGTATGAGGGCAAGACCAT  
 CGTGCCGGTGCGCTGCGAGGGCTTCCGCGGCGTGTCCCAGTCCCTGGGCCA  
 CCACATCGCCAACGACGCCATCCGCGATTGGGTGTTTCGACAAGATCGCGC  
 CCGACGCCGAGCCGCGCTTTGAGCCGACCCCGTACGACGTCGCCATCATCG  
 GCGACTACAATATCGGTGGTGACGCCTGGTCGTCCCGTATCCTCCTGGAGG  
 AGATGGGCCTGCGCGTGATCGCCCAGTGGTCCGGCGACGGTTCGCTCGCTG  
 AGCTGGAGGCCACCCCGAAGGCCAAGCTCAACGTGCTGCACTGCTACCGC  
 TCCATGAACTACATCTCGCGCCACATGGAAGAGAAGTACGGTATCCCGTG  
 GTGCGAGTACAACCTTCTTCGGTCCTTCCAAGATCGCCGAGTCCCTGCGCAA  
 GATCGCCAGCTACTTCGACGACAAGATCAAGGAAGGCGCGGAGCGCGTCA  
 TCGCCAAGTATCAGCCGCTCATGGATGCGGTGATCGCGAAGTATCGTCCCC  
 GCCTCGAGGGCAAGACCGTGATGCTGTACGTGGGCGGCCTGCGTCCCCGT  
 CACGTCATCGGCGCCTACGAGGACCTGGGCATGGAAGTGGTCGGCACGGG  
 CTACGAGTTCGCCCATACGACGACTACCAGCGCACCGCCAGCACTACG

TCAAGGATGGCACCATCATCTATGACGACGTGACCGGCTACGAGTTCGAG  
 AAGTTCGTCGAGAAGATCCAGCCGGACCTGGTTCGGTTCGGGCATCAAGGA  
 AAAGTACGTCTTCCAGAAGATGGGCGTGCCGTTCCGCCAGATGCACTCCTG  
 GGACTACTCGGGCCCGTACCACGGCTATGACGGCTTCGCGATCTTCGCGCG  
 CGACATGGACATGGCCATCAACAGCCCCGTGTGGAAGATGACCCAGGCTC  
 CGTGGAAGAGCGTCCCCAAGCCGACGATGCTCGCGGCTGAATGA

SEQ ID NO: 53:

Trình tự axit amin của chuỗi alpha của protein molybden-sắt của nitrogenaza  
 nifD\_2

MSVAQSQSVAEIKARNKELIEVLKVYPEKTAKRRRAKHLNVHEAGKSD  
 CGVKSNIKSIPGVM TIRGCAYAGSKGVVWGPIKDMIHISHGPVGCQYSWAA  
 RRNYIIGTTGIDTFVTMQFTSDFQEKDIVFGGDKKLAKIMDEIQELFPLNNGIT  
 VQSECPIGLIGDDIEAVSKQKSKEYEGKTIVPVRCEGFRGVSQSLGHHIANDAI  
 RDWVFDKIAPDAEPRFEPTPYDVAIIGDYNIGGDAWSSRILLEEMGLRVIAQW  
 SGDGLAELEATPKAKLNLHCYRSMNYISRHMEEKY GIPWCEYNFFGPSKIA  
 ESLRKIASYFDDKIKEGAERVIAKYQPLMDAVIAKYRPRLEGKTVMLYVGGLR  
 PRHVIGAYEDLGMEVVGTGYEFAHNDDYQRTAQHYVKDGTIIYDDVTGYEFE  
 KFVEKIQPD LVSGIKEKYVFQKMGVPFRQMHSWDYSGPYHGYDGFAIFARD  
 MDMAINS PVWKMTQAPWKSVPKPTMLAAE

SEQ ID NO: 54:

Trình tự nucleotit của chuỗi beta của protein molybden-sắt của nitrogenaza  
 nifK\_1

ATGGCCACCGTTTCCGTCTCCAAGAAGGCCTGCGCGGTCAACCCCCT  
 CAAGATGAGCCAGCCGGTGGGCGGCGCGCTCGCCTTCATGGGCGTGCGCA  
 AGGCCATGCCGCTGCTGCACGGCTCGCAGGGCTGCACCTCCTTCGGCCTGG  
 TGCTGTTCGTGCGCCACTTCAAGGAAGCCATCCCCATGCAGACCACCGCCA  
 TGAGCGAGGTGGCGACGGTTCTGGGCGGCCTTGAGAATGTGGAGCAGGCC  
 ATTCTCAACATCTACAATCGCACCAAGCCGGAGATCATCGGCATCTGCTCC  
 ACCGGCGTCACCGAGACCAAGGGCGATGATGTGACGGCTACATCAAGCT  
 GATCCGGGACAAGTATCCCCAGCTGGCCGACTTCCCGCTGGTCTATGTCTC  
 CACCCCCGATTTCAAGGACGCCTTCCAGGACGGTTGGGAGAAGACCGTGG  
 CGAAGATGGTGGAGGCGCTGGTGAAGCCCGCCGCGACAAGCAGAAGGA

CAAGACCCGCGTCAACGTCCTGCCCCGGCTGCCACCTCACGCCCCGGCGATCT  
 GGATGAGATGCGGACCATCTTCGAGGATTTCGGGGCTCACACCCTATTTCT  
 GCCGGATCTGGCCGGCTCGCTGGATGGGCATATCCCCGAGGACTTCTCGCC  
 CACCACCATCGGCGGCATCGGCATCGATGAGATCGCCACCATGGGCGAGG  
 CGGCCCACACCATCTGCATCGGCGCGCAGATGCGCCGGGCGGGCGAGGCC  
 ATGGAGAAGAAGACCGGCATTCCCTTCAAGCTGTTCGAGCGCCTGTGCGG  
 CCTGGAGGGCGAACGACGCCTTCATCATGCACCTGTGCGAGATCTCCGGCCG  
 GCCGGTGCCGGTGAAGTATCGCCGGCAGCGGGGCCAGCTGGTGGATGCCA  
 TGCTGGACGGCCACTTCCATCTGGGCGGTTCGCAAGGTGGCCATGGGGGCG  
 GAGCCGGACCTGCTCTACGACGTGGGCTCCTTCCTGCACGAGATGGGCGCC  
 CACATCCTTTCCGCGGTCACCACCACCCAGTCGCCGGTGCTGGCGCGCCTG  
 CCTGCCGAGGAGGTGCTTATCGGCGACCTGGAGGATCTGGAGACCCAGGC  
 GAAGGCGCGCGGATGCGATCTCCTGCTCACCCATTCCCATGGGCGCCAGG  
 CGGCGGAGCGCCTCCACATCCCCTTCTACCGGATCGGCATTCCCATGTTTG  
 ACCGGCTGGGGGCGGGGCATCTGTTGTGCGGTGGGCTATCGCGGCACCCGC  
 GACCTCATCTTCCATCTCGCCAACCTTGTGATCGCCGACCACGAGGAAAAT  
 CACGAGCCGACGCCCCGACACCTGGGGCCACCGGCCATGGCGAGCATGCCGC  
 CGCCCCCACTTCCCATTGA

SEQ ID NO: 55:

Trình tự axit amin của chuỗi beta của protein molybden-sắt của nitrogenaza  
 nifK\_1

MATVSVSKKACAVNPLKMSQPVGGALAFMGVRKAMPLHGSQGCTSF  
 GLVLFVRHFKEAIPMQTTAMSEVATVLGGLENVEQAILNIYNRTKPEIIGICSTG  
 VTETKGDDVDGYIKLIRDKYPQLADFPVYVSTPDFKDAFQDGWEKTVAKM  
 VEALVKPAADKQKDKTRVNVLPGCHLTPGDLEDMRTIFEDFGLTPYFLPDLA  
 GSLDGHIPEDFSPTTIGGIGIDEIATMGEEAHTICIGAQMRRAGEAMEKKTGIPF  
 KLFERLCGLEANDAFIMHLSQISGRPVVPKYRRQRGQLVDAMLDGHFHLGGR  
 KVAMGAEPDLLYDVGSFLHEMGAHILSAVTTTQSPVLARLPAEEVLIGDLEDL  
 ETQAKARGCDLLLTHSHGRQAAERLHIPFYRIGIPMFDRLGAGHLLSVGYRGT  
 RDLIFHLANLVIADHEENHEPTPDTWATGHGEHAAAPTSH

SEQ ID NO: 56:

Trình tự nucleotit của chuỗi beta của protein molybden-sắt của nitrogenaza

nifK\_2

ATGCCACAAAATGCTGACAATGTGCTCGATCACTTCGAGCTCTTCCG  
TGGTCCCGAATACCAGCAGATGCTGGCCAATAAGAAAAAGATGTTTCGAGA  
ACCCCCGCGATCCGGCCGAAGTCGAGCGCGTGCGGGAATGGGCGAAGACT  
CCTGAATACAAGGAGCTGAACTTCGCCCCGCGAGGCGCTCACCGTGAATCC  
GGCCAAGGCTTGTCAGCCGCTGGGCGCGGTGTTTCGTGCGCGTCGGCTTCGA  
GAGCACGATCCCCCTTCGTGCACGGCTCGCAGGGTTGCGTCGCGTATTACCG  
CTCGCACCTCTCCCGCCACTTCAAGGAGCCGTCTCCTGCGTCTCCTCGTCC  
ATGACCGAGGATGCGGCGGTGTTTCGGCGGCCTCAACAACATGATTGACGG  
CCTCGCCAACACCTACAACATGTACAAGCCGAAGATGATCGCCGTCTCCAC  
CACCTGCATGGCGGAAGTCATCGGCGACGATCTGAACGCCTTCATCAAGA  
CCGCGAAGGAAAAGGGCTCGGTTCCGGCCGAATACGACGTGCCCTTCGCC  
CACACCCCGGCGTTCGTGCGCAGCCATGTCACCGGCTACGACAATGCGCTC  
AAGGGCATCCTCGAGCACTTCTGGGACGGCAAGGCCGGCACCGCGCCGAA  
GCTGGAGCGCGTTCCCAACGAGAAGATCAACTTCATCGGCGGCTTCGACG  
GCTACACCGTCGGCAACACTCGCGAAGTGAAGCGCATCTTCGAGGCGTTC  
GGCGCCGATTACACCATCCTCGCCGACAATTCCGAAGTGTTTCGACACCCCG  
ACCGACGGCGAGTTCCGCATGTATGACGGCGGCACGACCCTGGAGGACGC  
GGCGAACGCGGTGCACGCCAAGGCCACCATCTCCATGCAGGAATACTGCA  
CGGAGAAGACCCTGCCCATGATCGCCGGTCATGGCCAGGACGTGGTTCGCC  
CTCAACCACCCCGTGGGCGTGGGCGGCACCGACAAGTTCCTCATGGAGAT  
CGCCCGCCTCACCGGCAAGGAGATCCCCGAGGAGCTGACCCGCGAGCGCG  
GCCGTCTCGTGGACGCTATCGCGGACTCTTCCGCGCACATCCACGGCAAGA  
AGTTCGCCATCTACGGCGATCCGGATCTGTGCCTGGGCCTCGCCGCGTTCC  
TGCTGGAGCTGGGCGCCGAGCCGACCCATGTGCTGGCCACCAACGGCACC  
AAGAAGTGGGCCGAGAAGGTTTCAGGAAGTTCGACTCTTCGCCGTTCGG  
CGCCAAGTGAAGGTCTATCCCGGCAAGGACCTGTGGCACATGCGCTCGCT  
CCTGTTTCGTGGAGCCGGTGGATTTCATCATCGGCAACACCTACGGCAAGTA  
TCTCGAGCGCGACACGGGCACCCCGCTGATCCGTATCGGCTTCCCGGTGTT  
CGACCGTCACCACCACCACCGCCGTCCGGTGTGGGGCTATCAGGGCGGCA  
TGAACGTCCTGATCACGATCCTCGACAAGATCTTTGACGAGATCGACCGCA  
ACACCAACGTGCCGGCCAAGACCGACTACTCGTTCGACATCATTCGTTGA

SEQ ID NO: 57:

Trình tự axit amin của chuỗi beta của protein molybden-sắt của nitrogenaza nifK\_2

MPQNADNVLDHFELFRGPEYQQMLANKKKMFENPRDPAEVERVREW  
 AKTPEYKELNFAREALTVNPAKACQPLGAVFVAVGFESTIPFVHGSQGCVA  
 YRSHLSRHFKEPSSCVSSSMTEDAAVFGGLNNMIDGLANTYNMYKPKMIAVS  
 TTCMAEVIGDDLNAFIKTAKEKGSVPAEYDVPFAHTPAFVGSHVTGYDNALK  
 GILEHFWDGKAGTAPKLERVPNEKINFIGGFDGYTVGNTREVKRIFEAFGADY  
 TILADNSEVFDTPDGEFRMYDGGTTLEDAANAVHAKATISMQEYCTEKTLP  
 MIAGHGQDVVALNHPVGVGGTDKFLMEIARLTGKEIPEELTRERGRLVDAIAD  
 SSAHIHGKKFAIYGDPDLCLGLAAFLLELGAEPHVLAATNGTKKWAEEKVQELF  
 DSSPFGANCKVYPGKDLWHMRSLLFVEPVDFIIGNTYGKYLERDTGTPLIRIGF  
 PVFDRHHHHRRPVWGYQGGMNVLITLDKIFDEIDRNTNVPKTDYSFDIIR

SEQ ID NO: 58:

Trình tự nucleotit của protein sắt của nitrogenaza nifH

GTGGAGTCCGGTGGTCCTGAGCCGGGCGTGGGCTGCGCCGGCCGCG  
 GCGTGATCACCTCCATCAACTTCCTGGAGGAGAACGGCGCCTACGAGGAC  
 ATCGACTATGTGTCCTACGACGTGCTGGGCGACGTGGTGTGCGGCGGCTTC  
 GCCATGCCCATCCGCGAGAACAAAGGCGCAGGAAATCTACATCGTGATGTC  
 CGGCGAGATGATGGCCATGTATGCGGCCAACAAACATCTCCAAGGGCATCC  
 TGAAGTATGCCAATTCCGGCGGCGTGCCTGGGCGGGCTGGTCTGCAAC  
 GAGCGCCAGACCGACAAGGAGCTGGAGCTGGCGGAGGCTCTGGCGAAGA  
 AGCTCGGCACCGAGCTGATCTACTTCGTGCCGCGCGACAACATCGTGCAGC  
 ATGCCGAGCTGCGCCGCATGACAGTGATCGAGTATGCGCCCGATTCCGCCC  
 AGGCCAGCACTACCGGAACCTGGCCGAGAAGGTGCACGCCAACAAAGGGC  
 AACGGCATCATCCCGACCCCGATCACCATGGACGAGCTGGAAGACATGCT  
 CATGGAGCACGGCATCATGAAGGCCGTGGACGAGAGCCAGATCGGCAAG  
 ACCGCCGCGAGCTCGCCGTCTGA

SEQ ID NO: 59:

Trình tự axit amin của protein sắt của nitrogenaza nifH

MESGGPEPGVGCAGRGVITSINFLEENGAYEDIDYVSVDVLGDVVC GG  
 FAMPIRENKAQEIIYVMSGEMMAMYAANNISKILKYANSGGVRLGGLVCNE

RQTDKELELAEALAKKLGTELIYFVPRDNIVQHAE LRRMTVIEYAPDSAQAQH  
YRNLAEKVHANKGNIGIIPITMDELEDMLMEHGIMKAVDESQIGKTAAELA  
V

*Việc xử lý xuôi dòng*

Theo một phương án, quy trình theo sáng chế bao gồm bước tiếp theo để thu hoạch sinh khối được tạo ra trong môi trường nuôi cấy. Ví dụ sinh khối có thể được thu hoạch bằng cách lắng cặn (lắng trên cơ sở trọng lực), lọc, ly tâm hoặc keo tụ. Keo tụ có thể đòi hỏi phải bổ sung chất keo tụ. Ly tâm ví dụ có thể được thực hiện sử dụng máy ly tâm dòng liên tục.

Theo một phương án, sinh khối đã thu hoạch sau đó được sấy khô. Sấy khô ví dụ có thể được thực hiện sử dụng phương pháp đã được biết rõ, bao gồm ly tâm, sấy khô bằng trống, bay hơi, đông khô lạnh, gia nhiệt, sấy phun sương, sấy chân không và/hoặc lọc chân không. Tiếp đó sinh khối đã sấy khô có thể được sử dụng trong sản phẩm, ví dụ như sản phẩm lương thực hoặc thực phẩm hoặc thành phần lương thực hoặc thực phẩm.

Theo phương án khác, tế bào của sinh khối đã được thu hoạch được phân giải. Theo một số phương án dịch sản phẩm phân giải có thể được tách thành các phân đoạn không tan và tan, một trong hai hoặc cả hai sau đó có thể được cô hoặc/và được sấy khô, và sau đó được sử dụng trong sản phẩm, ví dụ như sản phẩm lương thực hoặc thực phẩm.

Theo một phương án, sinh khối được thu hoạch và protein được phân lập từ sinh khối này, dẫn đến phân đoạn protein và phân đoạn chứa thành phần không phải protein. Do đó, theo một phương án, quy trình này là để sản xuất protein và bao gồm bước nuôi cấy chủng VTT-E-193585 hoặc dẫn xuất của chúng, sau đó là bước thu hoạch sinh khối và bước tiếp theo là phân lập protein từ sinh khối này. Theo phương án khác, quy trình này là để sản xuất protein và bao gồm bước nuôi cấy chủng vi khuẩn của chi *Xanthobacter* trong nuôi cấy liên tục với hydro làm nguồn năng lượng và nguồn cacbon vô cơ, trong đó nguồn cacbon vô cơ có chứa cacbon dioxit, sau đó là bước thu hoạch sinh khối và bước tiếp theo là phân lập protein từ sinh khối này. Tùy thuộc vào phương pháp để phân lập protein, các phân đoạn thu được có thể tinh khiết hơn hoặc ít tinh khiết hơn. Do đó, thuật ngữ “phân đoạn protein” có nghĩa là phân đoạn được làm giàu trong protein. Phân đoạn protein có thể vẫn có chứa lượng đáng kể của các thành phần khác và lượng đáng kể của protein cũng có thể dừng lại ở trong “phân đoạn chứa thành phần

không phải protein”.

Việc phân lập protein có thể được thực hiện bằng cách sử dụng phương pháp thích hợp bất kỳ. Ví dụ như, theo một phương án, protein được phân lập bằng cách phá vỡ tế bào một cách cơ học và tách protein ra khỏi mảnh vỡ tế bào thông qua một hoặc nhiều bước lọc, ví dụ như lọc liên tiếp thông qua nhiều bộ lọc với kích thước lỗ giảm dần. Việc phá vỡ cơ học có thể được thực hiện bằng cách sử dụng phương pháp thích hợp bất kỳ, ví dụ như nghiền bi, nghiền bằng siêu âm, đồng nhất hóa, đồng nhất hóa áp suất cao, trượt cơ học, v.v.. Phân đoạn protein được lọc thu được sẽ được làm giàu trong protein, nhưng cũng vẫn chứa các thành phần nhỏ hơn khác. Protein có thể tùy ý tinh sạch tiếp từ phân đoạn này bằng cách sử dụng phương pháp thích hợp bất kỳ.

Theo phương án khác, phân đoạn protein được phân lập bằng cách thực hiện việc chiết etanol sau đó là một hoặc nhiều bước lọc. Ví dụ như các phương pháp này được biết từ việc điều chế protein đậu tương (xem ví dụ như Chương 5 “Soybean Protein Concentrates” trong “Technology of production of edible flours and protein products from soybeans” của Berk FAO Agricultural Services Bulletin Số 97 (1992)). Phân đoạn protein thu được sẽ được làm giàu trong protein, nhưng cũng vẫn chứa các thành phần khác. Protein có thể còn được tinh sạch tùy ý từ phân đoạn này bằng cách sử dụng phương pháp thích hợp bất kỳ.

Theo một phương án, quy trình theo sáng chế bao gồm bước tiếp theo để thủy phân phân đoạn protein thu được từ quy trình theo sáng chế để thu được axit amin và peptit nhỏ.

Theo một phương án của quy trình theo sáng chế, quy trình bao gồm bước tiếp theo để sản xuất sản phẩm lương thực hoặc thực phẩm từ sinh khối này, từ phân đoạn protein đã nêu hoặc từ phân đoạn đã nêu chứa thành phần không phải protein. Bước tiếp theo đã nêu có thể đơn giản bao gồm kết hợp sinh khối đã nêu, phân đoạn protein hoặc phân đoạn chứa thành phần không phải protein trong sản phẩm lương thực hoặc thực phẩm, bằng cách bổ sung nó trong suốt quá trình sản xuất sản phẩm lương thực hoặc thực phẩm. Theo các phương án khác, việc tinh sạch hoặc cải biến tiếp theo của sinh khối hoặc phân đoạn của chúng được thực hiện trong khoảng thời gian kết hợp của nó vào sản phẩm lương thực hoặc thực phẩm.

Theo khía cạnh khác, sáng chế đề cập đến sản phẩm, chẳng hạn như sinh khối, protein, hoặc thành phần không phải protein, thu được hoặc có thể thu được bằng quy

trình theo sáng chế.

Theo một phương án, sản phẩm được thu từ quy trình theo sáng chế chứa hơn 40% protein, chẳng hạn như trong khoảng từ 40% đến 99% protein, ví dụ trong khoảng từ 40% đến 90% protein, chẳng hạn như trong khoảng từ 40% đến 60% protein. Theo phương án cụ thể, sản phẩm có chứa từ 25% đến 75% protein, từ 0% đến 20% lipit và từ 5% đến 40% carbohydrat. Theo phương án khác, sản phẩm có chứa từ 40% đến 60% protein, từ 0% đến 15% lipit và từ 10% đến 25% carbohydrat. Theo phương án khác nữa, sản phẩm được thu từ quy trình theo sáng chế có chứa từ 45% đến 55% protein, từ 5% đến 10% lipit và từ 10% đến 20% carbohydrat.

Như được mô tả ở trên, sáng chế theo khía cạnh tiếp theo đề cập đến sản phẩm lương thực hoặc thực phẩm thu được hoặc có thể thu được bằng quy trình theo sáng chế. Khi được sử dụng trong bản mô tả này, các thuật ngữ “lương thực” và “thực phẩm” được dự định là bao gồm không chỉ sản phẩm lương thực và thực phẩm thông thường, chẳng hạn như lương thực đã xử lý, mà còn bao gồm sản phẩm liên quan, chẳng hạn như lương thực và thực phẩm bổ sung, ví dụ như thanh protein, bột hoặc đồ lắc, sản phẩm thay thế bữa ăn, thành phần lương thực, lợi khuẩn, thức ăn cho lợi khuẩn, dược thực phẩm và dạng tương tự. Theo các phương án nhất định, sinh khối, phân đoạn protein hoặc phân đoạn chứa thành phần không phải protein được sử dụng trong việc sản xuất sản phẩm lương thực không thịt hoặc thuần chay.

Sáng chế còn được minh họa bằng các ví dụ, không làm giới hạn, sau đây:

### **Ví dụ thực hiện sáng chế**

Ví dụ 1. Phân lập chủng vi khuẩn có khả năng sinh trưởng hóa dưỡng

Mẫu 50 mL chứa đất và nước biển được thu thập trong ống falcon vô trùng từ bờ biển của biển Bantic ở Naantali tại Phần Lan. Một phần của mẫu đất được trộn với 10 mL môi trường khoáng chất trong bình thót cổ Erlenmeyer vô trùng. Môi trường bao gồm 1 g/L  $\text{NH}_4\text{OH}$ , 0,23 g/L  $\text{KH}_2\text{PO}_4$ , 0,29 g/L  $\text{Na}_2\text{HPO}_4 \cdot 2 \text{H}_2\text{O}$ , 0,005 g/L  $\text{NaVO}_3 \cdot \text{H}_2\text{O}$ , 0,2 g/L  $\text{FeSO}_4 \cdot 7 \text{H}_2\text{O}$ , 0,5 g/L  $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$ , 0,01 g/L  $\text{CaSO}_4$ , 0,00015 g/L  $\text{Na}_2\text{MoO}_4 \cdot 2 \text{H}_2\text{O}$ , 0,005 g/L  $\text{MnSO}_4$ , 0,0005 g/L  $\text{ZnSO}_4 \cdot 7 \text{H}_2\text{O}$ , 0,0015 g/L  $\text{H}_3\text{BO}_3$ , 0,001 g/L  $\text{CoSO}_4$ , 0,00005 g/L  $\text{CuSO}_4$  và 0,0001 g/L  $\text{NiSO}_4$  được điều chế trong nước máy. Huyền phù của đất và môi trường được ủ trong thiết bị ủ lắc ở nhiệt độ +30 °C trong hộp thép bịt kín mà được phun rửa liên tục bằng hỗn hợp khí: 150 mL/phút  $\text{N}_2$ , 18 mL/phút  $\text{H}_2$ , 3 mL/phút  $\text{O}_2$  và 6 mL/phút  $\text{CO}_2$ . Nuôi cấy được làm mới ở các khoảng

thời gian bảy ngày bằng cách lấy 1 mL huyền phù, mà được thêm trong điều kiện vô trùng vào 9 mL môi trường trong bình thót cổ Erlenmeyer, và sau đó đặt lại vào tủ ủ. Sau lần pha loãng thứ tư, không có đất có thể nhìn thấy được còn sót lại trong huyền phù. Thể tích của huyền phù tế bào được tăng lên thành 100 mL để sinh trưởng sinh khối để nuôi cấy bằng bình phản ứng sinh học. Mật độ quang học ( $OD_{600}$ ) của huyền phù là 1,53 khi nó được cấy vào 190 mL môi trường khoáng chất trong hệ thống bình phản ứng sinh học song song 15 bình 200 mL (Medicel Explorer, Medicel Oy, Phần Lan). Điều kiện nuôi cấy là khuấy 800 vòng/phút, nhiệt độ  $+30^{\circ}\text{C}$  và độ pH được thiết lập thành 6,8, kiểm soát nó bằng NaOH 1 M. Khí được cấp thông qua máy sục với hỗn hợp khí bao gồm 14 mL/phút  $\text{H}_2$ , 3 mL/phút  $\text{O}_2$  và 6 mL/phút  $\text{CO}_2$ . Khoảng trống bên trên của bình phản ứng được phun rửa bằng 300 mL/phút không khí. Việc nuôi cấy liên tục được cấp môi trường khoáng chất 6 ml/giờ và huyền phù tế bào được lấy ra từ bình phản ứng thông qua mao dẫn giữ thể tích không đổi ở 200 ml. Huyền phù tế bào được lấy ra từ bình phản ứng được bảo quản ở  $+4^{\circ}\text{C}$ . Mẫu được đưa từ bình phản ứng sinh học tự động mỗi ngày, và độ hấp thụ tại 600 nm được đo để theo dõi sự sinh trưởng. Sau 498 giờ nuôi cấy bằng bình phản ứng sinh học, các mẫu được lấy ra một cách vô trùng và huyền phù được pha loãng và được đặt vào các đĩa môi trường khoáng chất aga chứa các khoáng chất ở trên và 2% aga vi khuẩn. Các đĩa được ủ trong cùng điều kiện như được mô tả ở trên đối với các bình thót cổ Erlenmeyer. Sau đó các khuẩn lạc được nhặt từ các đĩa aga và được vạch vào các đĩa aga mới để phân lập một sinh vật trong một khuẩn lạc. Việc này được lặp lại hai lần. Các khuẩn lạc đơn lẻ được nhặt và được tạo huyền phù vào 200  $\mu\text{l}$  môi trường trong đĩa vi chuẩn độ 96 giếng. Huyền phù được ủ ở nhiệt độ  $+30^{\circ}\text{C}$  và được lắc 625 vòng/phút trong hộp kín khí EnzyScreen mà được phun rửa liên tục bằng 150 mL/phút  $\text{N}_2$ , 18 mL/phút  $\text{H}_2$ , 3 mL/phút  $\text{O}_2$  và 6 mL/phút  $\text{CO}_2$ . Huyền phù từ một giếng được chuyển vào bình thót cổ Erlenmeyer và bổ sung với môi trường mới. Thể tích tăng lên đến khi có đủ sinh khối để tiến hành nuôi cấy bằng bình phản ứng sinh học. Sinh vật được nộp lưu trong bộ sưu tập nuôi cấy VTT dưới dạng VTT-E-193585.

Giải trình tự rARN 16S của mẫu chứng minh rằng mẫu chỉ chứa một sinh vật. Cùng mẫu này được sử dụng để giải trình tự Illumina NextSeq tạo ra trình tự shotgun metagenom 1x150 bp. Bằng cách sử dụng Unicycler (Wick et al, 2017 PLoS computational biology 13:e1005595), sự lắp ráp de novo được tạo ra đối với trình tự

metagenom gồm có 101 contig. Tổng chiều dài hệ gen là 4.846.739 bp và hàm lượng GC là 67,9 %. Việc dự đoán gen và chú thích chức năng được thực hiện bằng cách sử dụng Prokka (Seemann, 2014 Bioinformatics 30:2068). Chú thích hệ gen tạo ra 4.429 gen. Sắp thẳng hàng pan-genomics (Page et al, 2015 Bioinformatics 31:3691) nhóm lại VTT-E-193585 trong số các loài *Xanthobacter*. Do đó chủng được xác định là *Xanthobacter* sp., hệ gen gần nhất là *Xanthobacter tagetidis*. Việc tính toán độ tương đồng nucleotit trung bình dựa trên việc sắp thẳng hàng mà chỉ tính đến các mảnh trực giao (OrthoANI) (Lee et al, 2016 Int J Syst Evol Microbiol 66:1100) tạo ra độ ăn khớp tốt nhất bằng 80,4% với *Xanthobacter tagetidis* (ATCC 700314; GCF\_003667445.1), trong khi đó ngưỡng đường biên loài được đề xuất là 95-96% (xem ví dụ như, Chun et al., 2018 Int J Syst Evol Microbiol, 68: 461-466). *Xanthobacter autotrophicus* Py2 có độ ăn khớp bằng 79,6%, trong khi độ ăn khớp đối với *Xanthobacter* sp. 91 bằng 79,0%. Do đó có thể kết luận rằng chủng vi khuẩn được phân lập được nộp lưu dưới dạng VTT-E-193585 thuộc Ngành: Proteobacteria; Lớp: Alpha Proteobacteria; và Bộ: Rhizobiales. Họ có khả năng nhất là Xanthobacteraceae, và chi *Xanthobacter*. Chủng vi khuẩn VTT-E-193585 không thể được gán rõ ràng vào loài đã biết bất kỳ.

Nghiên cứu đối với gen kháng kháng khuẩn giả định được thực hiện. Công cụ ABRicate (<https://github.com/tseemann/abricate>) được sử dụng để tra cứu hệ gen trên các cơ sở dữ liệu Arg-Annot, NCBI, ResFinder, ecOH, Megares và VFDB bằng cách sử dụng blastn hoặc blastp. Ngưỡng bằng 50 % được thiết lập cho cả độ tương đồng và độ bao phủ, cả ở cấp độ nucleotit và protein. Chỉ hai gen kháng kháng khuẩn giả định được xác định. Hai gen này không chứa sự thay đổi axit amin được liên kết với tính kháng kháng sinh và do đó kiểu hình kháng không được mong đợi.

Ví dụ 2. Việc nuôi cấy và phân tích thí điểm của chủng vi khuẩn được phân lập

Chủng vi khuẩn được phân lập được nộp lưu dưới dạng VTT-E-193585 được nuôi cấy trong bình phản ứng sinh học dạng bể có khuấy 200 lít thông thường (MPF-U, Marubishi Ltd, Nhật Bản). Việc trộn được thực hiện với chong chóng kiểu Rushton ở tốc độ 400 vòng/phút. Nhiệt độ trong việc nuôi cấy được duy trì ở +30°C. Độ pH được duy trì ở 6,8±0,2 bằng cách bổ sung NaOH 8 M hoặc H<sub>3</sub>PO<sub>4</sub> 3,6 M bằng sự điều khiển của phần mềm. Môi trường nuôi cấy chứa 1 g/L NH<sub>4</sub>OH, 0,23 g/L KH<sub>2</sub>PO<sub>4</sub>, 0,29 g/L Na<sub>2</sub>HPO<sub>4</sub> · 2 H<sub>2</sub>O, 0,005 g/L NaVO<sub>3</sub> · H<sub>2</sub>O, 0,2 g/L FeSO<sub>4</sub> · 7 H<sub>2</sub>O, 0,5 g/L MgSO<sub>4</sub> · 7 H<sub>2</sub>O, 0,01 g/L CaSO<sub>4</sub>, 0,00015 g/L Na<sub>2</sub>MoO<sub>4</sub> · 2 H<sub>2</sub>O, 0,005 g/L MnSO<sub>4</sub>, 0,0005 g/L

$\text{ZnSO}_4 \cdot 7 \text{H}_2\text{O}$ , 0,0015 g/L  $\text{H}_3\text{BO}_3$ , 0,001 g/L  $\text{CoSO}_4$ , 0,00005 g/L  $\text{CuSO}_4$  và 0,0001 g/L  $\text{NiSO}_4$  được điều chế trong nước máy. Hỗn hợp chứa 1,8-10,5 l/phút khí hydro, 0,6-2,5 l/phút khí oxy và 1,8-5 l/phút khí cacbon dioxid được cung cấp liên tục dưới dạng nguồn năng lượng và cacbon chính. Hàm lượng oxy hòa tan được duy trì ở  $7,2 \pm 0,5$  % bằng cách điều chỉnh hợp phần hỗn hợp khí. Chất cấy để nuôi cấy được điều chế như được mô tả trong Ví Dụ 1. Sinh trưởng được theo dõi bằng cách lấy mẫu thủ công và phân tích mật độ tế bào dưới dạng mật độ quang học bằng cách đo độ hấp thụ ở mức 600 nm (quang phổ kế Ultrospec 2100 pro Uv/ánh sáng nhìn thấy, Biochrom Ltd., Anh) và bằng cách đo khối lượng khô của tế bào (CDW) bằng cách sấy khô trong lò qua đêm ở nhiệt độ  $105^\circ\text{C}$ . Mật độ quang học cũng được theo dõi bằng cách sử dụng đầu dò độ hấp thụ in situ (Trucell 2, Finesse Ltd, Mỹ). Đường cong sinh trưởng của sự nuôi cấy được thực hiện trên Fig. 1. Tốc độ sinh trưởng lớn nhất trong pha theo mẻ là  $0,06 \text{ giờ}^{-1}$ . Mật độ tế bào tối đa là  $4,5 \text{ g\_CDW/l}$  ở 92 giờ. Sau 92 giờ nuôi cấy, việc cấp môi trường nuôi cấy mới như được mô tả ở trên được bắt đầu ở tỉ lệ pha loãng là  $0,01 \text{ giờ}^{-1}$ . Trong quá trình cấp liên tục, mật độ tế bào trung bình là  $2,9 \text{ g\_CDW/l}$ . Dịch lỏng nuôi cấy được thu thập liên tục vào bể đã được làm mát ( $+10^\circ\text{C}$ ) mà nó được cấp từ đó trong các mẻ 300 lít đến máy tách ly tâm liên tục (BTPX-205, Alfa-Laval AB, Thụy Điển). Bùn sệt chứa tế bào được thu từ máy phân tách được nạp vào máy sấy kiểu trống kép khí quyển (Buflovak 6x8 ADDD, Hebel process solutions Llc., Mỹ), được gia nhiệt bằng hơi 4 bar và trống quay ở tốc độ 3,5 vòng/phút. Điều này thu được bột tế bào được làm khô với hàm lượng vật chất khô xấp xỉ 96%. Kết quả phân tích của bột tế bào được làm khô được trình bày trong Bảng 1 đối với thành phần quan trọng, trong Bảng 2 đối với thành phần axit amin, trong Bảng 3 đối với thành phần axit béo, và trong Bảng 4 đối với hàm lượng vitamin. Phân tích chứng minh rằng bột tế bào được làm khô có hàm lượng protein cao với tất cả các axit amin thiết yếu. Nó cũng chứa nhiều axit béo không bão hòa hơn axit béo bão hòa và nhiều vitamin nhóm B. Hàm lượng peptidoglycan chỉ bằng  $0,002 \text{ mg/g\_CDW}$  và hàm lượng lipopolysacarit bằng  $0,01 \text{ mg/g\_CDW}$ . Sẽ có lợi nếu các nồng độ này nhỏ nhất có thể. Để so sánh, trong việc điều chế vi khuẩn axit lactic trên thị trường được phân tích cùng lúc, hàm lượng peptidoglycan bằng  $0,244 \text{ mg/g\_DW}$  và hàm lượng lipopolysacarit bằng  $0,015 \text{ mg/g\_DW}$ . Thử nghiệm gây độc tế bào và gây độc gen được thực hiện bằng cách sử dụng mẫu dịch nổi nuôi cấy. Không quan sát thấy có sự gây độc tế bào nào chống lại dòng tế bào người HepG2 hoặc HeLa229. Không

quan sát thấy có sự gây độc gen nào chống lại chủng *Escherichia coli* WP2 *trp*- hoặc CM871 *uvrA recA lexA*.

Bảng 1. Kết quả phân tích bột tế bào được làm khô của chủng vi khuẩn được phân lập được nộp lưu dưới dạng VTT-E-193585.

| Thông số                       | Phương pháp               | Đơn vị    | Giá trị |
|--------------------------------|---------------------------|-----------|---------|
| Độ ẩm                          | Sấy khô ở 103 °C          | g/100g    | 3,3     |
| Protein                        | Kjeldahl (N x 6,25)       | g/100g    | 72,2    |
| Chất béo                       | Weibull-Stoldt            | g/100g    | 6,0     |
| Axit béo no                    | Được tính dựa trên Bảng 3 | g/100g    | 1,8     |
| Axit béo đơn không bão hòa đơn | Được tính dựa trên Bảng 3 | g/100g    | 3,8     |
| Axit béo không bão hòa đa      | Được tính dựa trên Bảng 3 | g/100g    | 0,4     |
| Axit béo omega 3               | Được tính dựa trên Bảng 3 | g/100g    | < 0,01  |
| Axit béo omega 6               | Được tính dựa trên Bảng 3 | g/100g    | 0,4     |
| Chất xơ thực phẩm              | Phân tích trọng lượng     | g/100g    | 10,5    |
| Tro                            | Tạo tro ở nhiệt độ 550 °C | g/100g    | 5,8     |
| Glucosa                        | HPLC-ELSD                 | g/100g    | < 0,15  |
| Fructoza                       | HPLC-ELSD                 | g/100g    | < 0,1   |
| Sucroza                        | HPLC-ELSD                 | g/100g    | < 0,1   |
| Lactoza                        | HPLC-ELSD                 | g/100g    | < 0,25  |
| Maltoza                        | HPLC-ELSD                 | g/100g    | < 0,2   |
| Tổng lượng đường               | Tính                      | g/100g    | < 0,8   |
| Carbohydrat                    | Tính                      | g/100g    | 2,2     |
| Năng lượng                     | Tính                      | kJ/100g   | 1572    |
| Năng lượng                     | Tính                      | kcal/100g | 373     |

Bảng 2. Thành phần axit amin của bột tế bào được làm khô của chủng vi khuẩn được phân lập được nộp lưu dưới dạng VTT-E-193585.

| Thông số  | Phương pháp | Đơn vị | Giá trị |
|-----------|-------------|--------|---------|
| Lysin     | Sắc ký ion  | %      | 3,95    |
| Methionin | Sắc ký ion  | %      | 1,60    |

|              |            |   |      |
|--------------|------------|---|------|
| Xystin       | Sắc ký ion | % | 0,39 |
| Aspartic     | Sắc ký ion | % | 6,82 |
| Threonin     | Sắc ký ion | % | 3,47 |
| Serin        | Sắc ký ion | % | 2,75 |
| Glutamic     | Sắc ký ion | % | 8,84 |
| Prolin       | Sắc ký ion | % | 3,14 |
| Glyxin       | Sắc ký ion | % | 4,40 |
| Alanin       | Sắc ký ion | % | 6,94 |
| Valin        | Sắc ký ion | % | 4,96 |
| Isoloxin     | Sắc ký ion | % | 3,34 |
| Loxin        | Sắc ký ion | % | 6,08 |
| Tyrosin      | Sắc ký ion | % | 2,99 |
| Phenylalanin | Sắc ký ion | % | 4,58 |
| Histidin     | Sắc ký ion | % | 1,66 |
| Arginin      | Sắc ký ion | % | 4,96 |
| Tryptophan   | HPLC       | % | 1,34 |

Bảng 3. Thành phần axit béo của bột tế bào được làm khô của chủng vi khuẩn được phân lập được nộp lưu dưới dạng VTT-E-193585.

| Thông số                       | Phương pháp | Đơn vị | Giá trị |
|--------------------------------|-------------|--------|---------|
| C16:0 (Axit Palmitic)          | GC-MS       | %      | 24,8    |
| C16:1 (Axit Palmitoleic)       | GC-MS       | %      | 3,0     |
| C18:0 (Axit Stearic)           | GC-MS       | %      | 4,4     |
| C18:1n9 (Axit Oleic)           | GC-MS       | %      | 59,9    |
| C18:2n6 (Axit Linoleic)        | GC-MS       | %      | 6,1     |
| C18:3n3 (Axit alpha-Linolenic) | GC-MS       | %      | 0,4     |

Bảng 4. Hàm lượng vitamin của bột tế bào được làm khô của chủng vi khuẩn được phân lập được nộp lưu dưới dạng VTT-E-193585.

| Thông số       | Phương pháp | Đơn vị   | Giá trị |
|----------------|-------------|----------|---------|
| Vitamin A (RE) | HPLC        | IU/100 g | <100    |

|                            |                 |          |       |
|----------------------------|-----------------|----------|-------|
| Vitamin E (TE)             | HPLC            | mg/100 g | 0,33  |
| Vitamin D3                 | HPLC            | IU/100 g | <10   |
| Vitamin D2                 | HPLC            | IU/100 g | 21,6  |
| Vitamin C                  | HPLC            | mg/100 g | <1    |
| Thiamin clorua Hydroclorua | LC-MS/MS        | mg/100 g | 0,9   |
| Vitamin B1 (Thiamin)       | Tính            | mg/100 g | 0,708 |
| Vitamin B2 (Riboflavin)    | HPLC            | mg/100 g | 6,27  |
| Pyridoxin hydroclorua      | HPLC            | mg/100 g | 3,39  |
| Vitamin B6 (Pyridoxin)     | Tính            | mg/100 g | 2,79  |
| Vitamin B12                | LC-MS/MS        | µg/100 g | 224   |
| Cholin clorua              | LC-MS/MS        | mg/100 g | 14,3  |
| Biotin                     | LC-MS/MS        | µg/100 g | 15,6  |
| Axit folic                 | Vi sinh vật học | µg/100 g | 1270  |
| Niixin (Vitamin B3)        | Vi sinh vật học | mg/100 g | 23,2  |
| Axit Pantothenic           | Vi sinh vật học | mg/100 g | 6,53  |

Ví dụ 3. Nuôi cấy chủng vi khuẩn được phân lập trên các nguồn nitơ khác nhau.

Chủng vi khuẩn được phân lập được nộp lưu dưới dạng VTT-E-193585 được nuôi cấy trong hệ thống bình phản ứng sinh học song song 15 bình ở thể tích 200 mL (Medicel Explorer, Medicel Oy, Phần Lan). Việc trộn được thực hiện với chong chóng kiểu Rushton ở tốc độ 800 vòng/phút. Nhiệt độ trong việc nuôi cấy được duy trì ở +30°C. Độ pH được duy trì ở 6,8 bằng cách bổ sung NaOH 1 M. Môi trường nuôi cấy chứa 0,23 g/L  $\text{KH}_2\text{PO}_4$ , 0,29 g/L  $\text{Na}_2\text{HPO}_4 \cdot 2 \text{H}_2\text{O}$ , 0,005 g/L  $\text{NaVO}_3 \cdot \text{H}_2\text{O}$ , 0,2 g/L  $\text{FeSO}_4 \cdot 7 \text{H}_2\text{O}$ , 0,5 g/L  $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$ , 0,01 g/L  $\text{CaSO}_4$ , 0,00015 g/L  $\text{Na}_2\text{MoO}_4 \cdot 2 \text{H}_2\text{O}$ , 0,005 g/L  $\text{MnSO}_4$ , 0,0005 g/L  $\text{ZnSO}_4 \cdot 7 \text{H}_2\text{O}$ , 0,0015 g/L  $\text{H}_3\text{BO}_3$ , 0,001 g/L  $\text{CoSO}_4$ , 0,00005 g/L  $\text{CuSO}_4$  và 0,0001 g/L  $\text{NiSO}_4$  được điều chế trong nước máy. Hơn nữa, nguồn nitơ được thay đổi khi nuôi cấy sao cho bốn nuôi cấy chứa  $\text{NH}_4\text{OH}$  18,7 mM, bốn nuôi cấy chứa ure 9,34 mM ( $\text{OC}(\text{NH}_2)_2$ ), bốn nuôi cấy chứa kali nitrat 18,7 mM ( $\text{KNO}_3$ ), và ba nuôi cấy được để mà không có nguồn nitơ trong môi trường. Hỗn hợp chứa 22 mL/phút khí hydro, 3,2 mL/phút không khí và 6,4 mL/phút khí cacbon dioxid được cung cấp liên tục làm nguồn năng lượng và cacbon chính. Do đó, với không khí, tất cả việc nuôi cấy

cũng được cung cấp với khí nitơ. Sự sinh trưởng được theo dõi bằng cách lấy mẫu tự động và phân tích mật độ tế bào dưới dạng mật độ quang học bằng cách đo độ hấp thụ ở mức 600 nm (quang phổ kế Ultrospec 2100 pro Uv/ánh sáng nhìn thấy được, Biochrom Ltd., Anh). Đường cong sinh trưởng của việc nuôi cấy được thực hiện trên Fig. 2. Sự sinh trưởng trên amoniac và ure có thể so sánh được. Sự sinh trưởng trên nitrat hoặc khí nitơ thấp hơn rõ ràng so với trên amoniac hoặc ure. Đến khi kết thúc nuôi cấy, sinh trưởng trên nitrat tốt hơn sinh trưởng trên khí nitơ làm nguồn nitơ duy nhất. Tuy nhiên cũng có sinh trưởng trong nuôi cấy mà tại đó khí nitơ là nguồn nitơ duy nhất chứng minh rằng chủng vi khuẩn được phân lập được nộp lưu dưới dạng VTT-E-193585 có khả năng cố định nitơ.

#### Ví dụ 4. Xác định đặc điểm của tính miễn cảm kháng sinh

Tính miễn cảm kháng sinh của gentamixin, kanamixin, streptomycin, tetracyclin, ampicillin, ciprofloxacin, colistin và fosfomycin đối với chủng vi khuẩn được phân lập được nộp lưu dưới dạng VTT-E-193585 được phân tích theo tiêu chuẩn CLSI M07-A111 (Clinical and laboratory standards institute. Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically, 11th ed. CLSI standard M07, 2018) với đĩa vi pha loãng làm bằng tay đối với ampicillin, ciprofloxacin và colistin, với đĩa VetMIC Lact-1 (SVA National Veterinary Institute, Uppsala, Thụy Điển) đối với gentamixin, kanamixin, streptomycin và tetracyclin bằng cách sử dụng cả phương pháp vi pha loãng và đối với fosfomycin bằng cách sử dụng phương pháp pha loãng aga trong điều kiện ươm khí ở nhiệt độ  $+35 \pm 2^\circ\text{C}$  trong  $48 \pm 1$  giờ bằng cách sử dụng môi trường Mueller Hinton Broth được điều chỉnh cation (LabM, LAB114, cation  $\text{Mg}^{2+}$  và  $\text{Ca}^{2+}$  được bổ sung riêng rẽ). *Escherichia coli* ATCC 25922 được sử dụng làm chủng đối chứng chất lượng và nó được ủ trong điều kiện ươm khí, ở nhiệt độ  $+35 \pm 2^\circ\text{C}$  trong  $18 \pm 2$  giờ. Kết quả của tính miễn cảm kháng sinh của chủng được thể hiện trên Bảng 5. Chủng vi khuẩn phân lập được tìm thấy là thường nhạy cảm đối với chất kháng sinh. Đối với giá trị nồng độ ức chế tối thiểu (MIC) gentamixin, kanamixin, streptomycin và tetracyclin đối với VTT-E-193585 thấp hơn hoặc có thể so sánh được với *E. coli* ATCC 25922, trong khi đối với ampicillin, ciprofloxacin, colistin và fosfomycin giá trị MIC cao hơn ở VTT-E-193585.

Bảng 5. Giá trị Nồng Độ Ức Chế Tối Thiểu (MIC,  $\mu\text{g/ml}$ ) của chất kháng sinh đối với chủng VTT-E-193585 và *Escherichia coli* ATCC 25922

|               |                                    |                                                 |
|---------------|------------------------------------|-------------------------------------------------|
|               | VTT-E-193585<br>48 giờ $\pm$ 1 giờ | <i>E. coli</i> ATCC 25922<br>18 giờ $\pm$ 2 giờ |
| Gentamixin    | 0,5                                | 0,5                                             |
| Kanamycin     | 2                                  | 4                                               |
| Streptomycin  | 0,5                                | 4                                               |
| Tetracyclin   | 0,25                               | 1                                               |
| Ampicillin    | 16                                 | 8                                               |
| Ciprofloxacin | 0,06                               | 0,008                                           |
| Colistin      | 4                                  | 2                                               |
| Fosfomycin    | 32                                 | 0,5                                             |

## **YÊU CẦU BẢO HỘ**

1. Chủng vi khuẩn được phân lập VTT-E-193585 hoặc chủng được cải biến di truyền hoặc được gây đột biến được tạo ra sử dụng chủng vi khuẩn VTT-E-193585, trong đó chủng được cải biến di truyền hoặc được gây đột biến đã nêu giữ lại khả năng để sinh trưởng sử dụng khí hydro làm nguồn năng lượng và cacbon dioxit làm nguồn cacbon duy nhất.

2. Môi trường nuôi cấy có chứa chủng vi khuẩn theo điểm 1.

3. Quy trình sản xuất sinh khối, quy trình đã nêu bao gồm bước nuôi cấy chủng vi khuẩn theo điểm 1.

4. Quy trình theo điểm 3, bao gồm bước nuôi cấy chủng vi khuẩn trong nuôi cấy liên tục với hydro làm nguồn năng lượng và nguồn cacbon vô cơ, trong đó nguồn cacbon vô cơ có chứa cacbon dioxit.

5. Quy trình theo điểm 4, trong đó oxy hòa tan trong môi trường nuôi cấy được duy trì trong khoảng từ 5% đến 10%.

6. Quy trình theo điểm bất kỳ trong số các điểm 4 hoặc 5, trong đó amoni, urê, nitrat và/hoặc khí nitơ được sử dụng làm nguồn nitơ.

7. Quy trình theo điểm bất kỳ trong số các điểm 4 đến 6, trong đó môi trường nuôi cấy có chứa khoáng chất trong đó khoáng chất đã nêu chứa ít hơn 1 g/L muối clorua, chẳng hạn như ít hơn 0,25 g/L muối clorua, ví dụ ít hơn 0,1 g/L muối clorua, tốt hơn nếu không chứa muối clorua.

8. Quy trình theo điểm bất kỳ trong số các điểm từ 4 đến 7, trong đó môi trường nuôi cấy không chứa vitamin.

9. Quy trình theo điểm bất kỳ trong số các điểm từ 4 đến 8, trong đó pH trong môi trường nuôi cấy được duy trì trong khoảng từ 5,5 đến 8,0, ví dụ trong khoảng từ 6,5 đến 7,0, chẳng hạn tại 6,8.

10. Quy trình theo điểm bất kỳ trong số các điểm từ 4 đến 9, trong đó môi trường nuôi cấy được cho sinh trưởng ở nhiệt độ trong khoảng từ 25°C và 40°C, ví dụ trong khoảng từ 28°C và 32°C, chẳng hạn tại 30°C.

11. Quy trình theo điểm bất kỳ trong số các điểm từ 4 đến 10, trong đó chủng vi khuẩn có tốc độ sinh trưởng là 0,04-0,12 giờ<sup>-1</sup>.

12. Quy trình theo điểm bất kỳ trong số các điểm từ 4 đến 11, trong đó chủng vi khuẩn có chứa ARN ribosom 16S nêu trong SEQ ID NO:1 hoặc ARN ribosom 16S có

đến 20 khác biệt nucleotit so với SEQ ID NO:1, ví dụ từ 1 đến 10, chẳng hạn từ 1 đến 5, ví dụ một, hai hoặc ba khác biệt nucleotit so với SEQ ID NO:1.

13. Quy trình theo điểm bất kỳ trong số các điểm từ 4 đến 12, trong đó chủng vi khuẩn có chứa:

i) gen mã hóa cho chuỗi lớn ribuloza-1,5-bisphosphat carboxylaza/oxygenaza (rubisco) có trình tự nêu trong SEQ ID NO:3 hoặc trình tự tương đồng hơn 93%, ví dụ tương đồng hơn 95%, chẳng hạn như tương đồng hơn 96%, ví dụ tương đồng hơn 97%, chẳng hạn như tương đồng hơn 98%, ví dụ tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:3, và/hoặc

ii) gen mã hóa cho chuỗi nhỏ ribuloza-1,5-bisphosphat carboxylaza/oxygenaza (rubisco) có trình tự nêu trong SEQ ID NO:5 hoặc trình tự tương đồng trình tự hơn 83%, ví dụ lớn hơn 86%, tương đồng chẳng hạn như tương đồng hơn 90%, ví dụ tương đồng hơn 95%, chẳng hạn như tương đồng hơn 96%, ví dụ tương đồng hơn 97%, chẳng hạn như tương đồng hơn 98%, ví dụ tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:5, và/hoặc

iii) gen mã hóa cho tiểu đơn vị alpha của HoxS hydrogenaza khử  $\text{NAD}^+$  có trình tự nêu trong SEQ ID NO:7 hoặc trình tự tương đồng trình tự hơn 70%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:7, và/hoặc

iv) gen mã hóa cho tiểu đơn vị beta của HoxS hydrogenaza khử  $\text{NAD}^+$  có trình tự nêu trong SEQ ID NO:9 hoặc trình tự tương đồng trình tự hơn 77%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:9, và/hoặc

v) gen mã hóa cho tiểu đơn vị gamma của HoxS hydrogenaza khử  $\text{NAD}^+$  có trình tự nêu trong SEQ ID NO:11 hoặc trình tự tương đồng trình tự hơn 70%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID

NO:11, và/hoặc

vi) gen mã hóa cho tiểu đơn vị delta của HoxS hydrogenaza khử  $\text{NAD}^+$  có trình tự nêu trong SEQ ID NO:13 hoặc trình tự tương đồng trình tự hơn 79%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:13, và/hoặc

vii) gen mã hóa cho tiểu đơn vị lớn của NiFeSe hydrogenaza có trình tự nêu trong SEQ ID NO:15 hoặc trình tự tương đồng trình tự hơn 84%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:15, và/hoặc

viii) gen mã hóa cho tiểu đơn vị nhỏ của NiFeSe hydrogenaza có trình tự nêu trong SEQ ID NO:17 hoặc trình tự tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:17, và/hoặc

ix) gen mã hóa cho chuỗi gamma của ATP synthaza atpG\_1 có trình tự nêu trong SEQ ID NO:19 hoặc trình tự tương đồng hơn 70%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:19, và/hoặc

x) gen mã hóa cho tiểu đơn vị alpha của ATP synthaza atpA\_1 có trình tự nêu trong SEQ ID NO:21 hoặc trình tự tương đồng hơn 78%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:21, và/hoặc

xi) gen mã hóa cho tiểu đơn vị b của ATP synthaza atpF\_1 có trình tự nêu trong SEQ ID NO:23 hoặc trình tự tương đồng hơn 62%, ví dụ tương đồng hơn 70%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ

ID NO:23, và/hoặc

xii) gen mã hóa cho tiểu đơn vị c của ATP synthaza, đặc hiệu ion natri atpE\_1 có trình tự nêu trong SEQ ID NO:25 hoặc trình tự tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:25, và/hoặc

xiii) gen mã hóa cho tiểu đơn vị a của ATP synthaza atpB\_1 có trình tự nêu trong SEQ ID NO:27 hoặc trình tự tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:27, và/hoặc

xiv) gen mã hóa cho chuỗi epsilon của ATP synthaza atpC\_1 có trình tự nêu trong SEQ ID NO:29 hoặc trình tự tương đồng hơn 71%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:29, và/hoặc

xv) gen mã hóa cho tiểu đơn vị beta của ATP synthazaeta atpD\_1 có trình tự nêu trong SEQ ID NO:31 hoặc trình tự tương đồng hơn 84%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:31, và/hoặc

xvi) gen mã hóa cho tiểu đơn vị beta của ATP synthazaeta atpD\_2 có trình tự nêu trong SEQ ID NO:33 hoặc trình tự tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:33, và/hoặc

xvii) gen mã hóa cho chuỗi gamma của ATP synthaza atpG\_2 có trình tự nêu trong SEQ ID NO:35 hoặc trình tự tương đồng hơn 86%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:35, và/hoặc

xviii) gen mã hóa cho tiểu đơn vị alpha của ATP synthaza atpA\_2 có trình tự nêu trong SEQ ID NO:37 hoặc trình tự tương đồng trình tự hơn 98%, chẳng hạn như lớn hơn 99% so với trình tự nêu trong SEQ ID NO:37, và/hoặc

xix) gen mã hóa cho tiểu đơn vị delta của ATP synthaza atpH có trình tự nêu trong SEQ ID NO:39 hoặc trình tự tương đồng hơn 85%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:39, và/hoặc

xx) gen mã hóa cho tiểu đơn vị b của ATP synthaza atpF\_2 có trình tự nêu trong SEQ ID NO:41 hoặc trình tự tương đồng hơn 87%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:41, và/hoặc

xxi) gen mã hóa cho tiểu đơn vị b' của ATP synthaza atpG\_3 có trình tự nêu trong SEQ ID NO:43 hoặc trình tự tương đồng hơn 81%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:43, và/hoặc

xxii) gen mã hóa cho tiểu đơn vị c của ATP synthaza atpE\_2 có trình tự nêu trong SEQ ID NO:45 hoặc trình tự tương đồng trình tự hơn 98%, chẳng hạn tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:45, và/hoặc

xxiii) gen mã hóa cho tiểu đơn vị a của ATP synthaza atpB\_2 có trình tự nêu trong SEQ ID NO:47 hoặc trình tự tương đồng hơn 92%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:47, và/hoặc

xiv) gen mã hóa cho protein I của ATP synthaza atpI có trình tự nêu trong SEQ ID NO:49 hoặc trình tự tương đồng hơn 60%, ví dụ tương đồng hơn 70%, chẳng hạn như tương đồng hơn 80%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:49, và/hoặc

xv) gen mã hóa cho chuỗi alpha của protein molybden-sắt của nitrogenaza nifD\_1 có trình tự nêu trong SEQ ID NO:51 hoặc trình tự tương đồng hơn 60%, ví dụ tương đồng hơn 70%, chẳng hạn như tương đồng hơn 92%, chẳng hạn như tương đồng hơn

95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:51, và/hoặc

xvi) gen mã hóa cho chuỗi alpha của protein molybden-sắt của nitrogenaza nifD\_2 có trình tự nêu trong SEQ ID NO:53 hoặc trình tự tương đồng hơn 60%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:53, và/hoặc

xvii) gen mã hóa cho chuỗi beta của protein molybden-sắt của nitrogenaza nifK\_1 có trình tự nêu trong SEQ ID NO:55 hoặc trình tự tương đồng hơn 87%, ví dụ tương đồng hơn 90%, chẳng hạn như tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:55, và/hoặc

xviii) gen mã hóa cho chuỗi beta của protein molybden-sắt của nitrogenaza nifK\_2 có trình tự nêu trong SEQ ID NO:57 hoặc trình tự tương đồng hơn 95%, ví dụ tương đồng hơn 96%, chẳng hạn như tương đồng hơn 97%, ví dụ như lớn hơn 98%, chẳng hạn như tương đồng trình tự hơn 99% so với trình tự nêu trong SEQ ID NO:57, và/hoặc

xxix) gen mã hóa cho protein sắt của nitrogenaza nifH có trình tự nêu trong SEQ ID NO:59 hoặc trình tự tương đồng trình tự hơn 98,5% so với trình tự nêu trong SEQ ID NO:59, trong đó chủng tốt hơn nếu bao gồm:

- gen được mô tả trong iii), iv), v) và vi),
- gen được mô tả trong vii) và viii), hoặc
- gen được mô tả trong iii), iv), v), vi), vii) và viii).

14. Quy trình theo điểm bất kỳ trong số các điểm từ 4 đến 13, bao gồm bước tiếp theo là thu hoạch sinh khối được sản xuất trong suốt quá trình nuôi cấy, tùy ý bao gồm bước tiếp theo là sấy khô sinh khối.

15. Quy trình sản xuất protein, bao gồm bước thực hiện quy trình theo điểm 14 và bước tiếp theo là phân lập protein từ sinh khối đã nêu, trong đó quy trình dẫn đến thu được phân đoạn protein và phân đoạn chứa thành phần không phải protein.

16. Quy trình theo điểm 14 hoặc 15, bao gồm bước tiếp theo là sản xuất sản phẩm lương thực hoặc thực phẩm từ sinh khối đã nêu, từ phân đoạn protein đã nêu hoặc từ

phân đoạn chứa thành phần không phải protein đã nêu.

17. Sản phẩm sinh khối chứa vật liệu tế bào của chủng theo điểm 1, trong đó sinh khối đã nêu có thể thu được bằng quy trình theo điểm bất kỳ trong số các điểm từ 4 đến 16.

Fig. 1

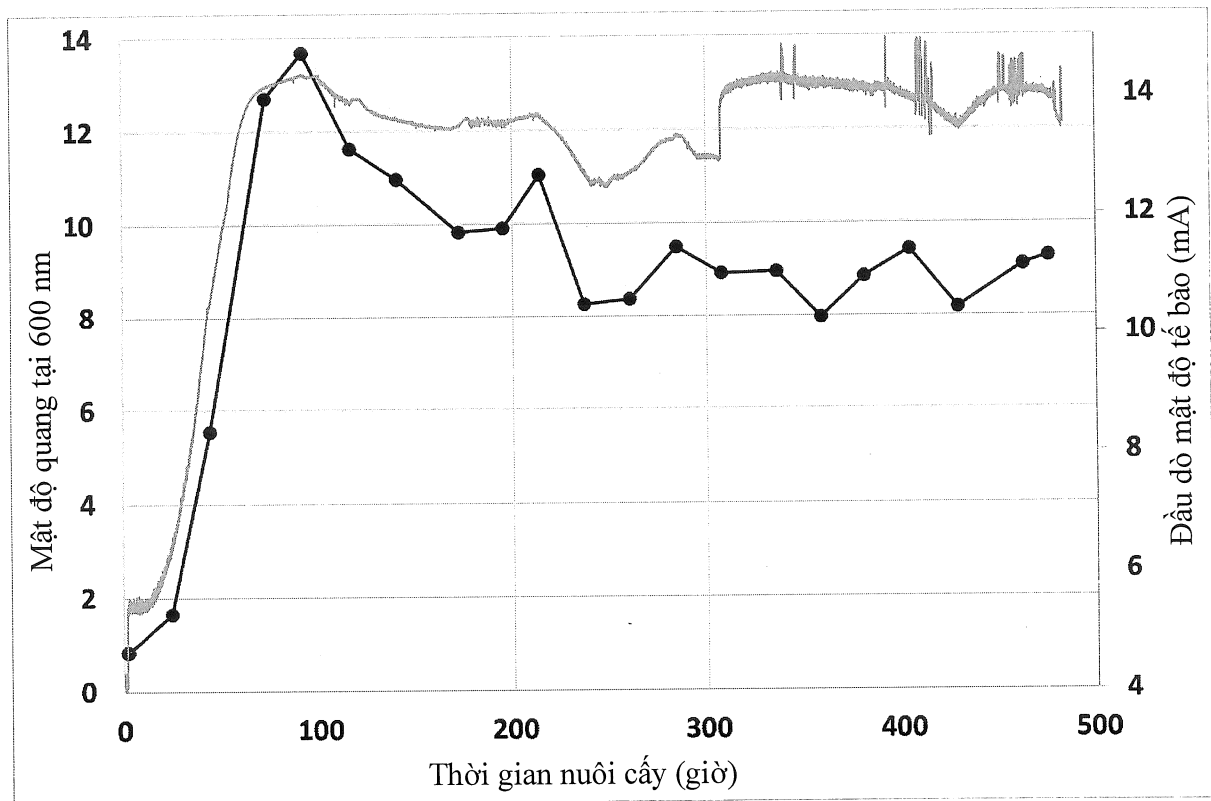
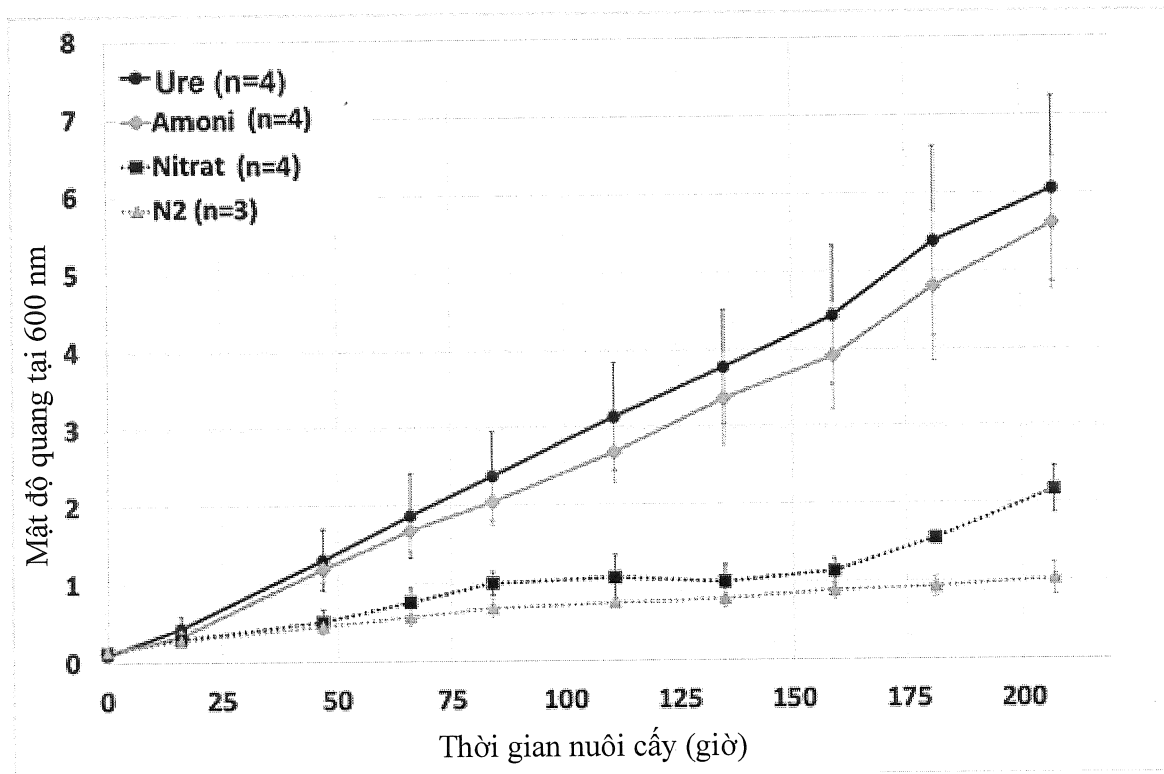


Fig. 2



## DANH MỤC TRÌNH TỰ

&lt;110&gt; Solar Foods

&lt;120&gt; CHÙNG VI KHUẨN ĐƯỢC PHÂN LẬP VTT-E-193585 VÀ MÔI TRƯỜNG NUÔI CẤY CHỨA CHÙNG NÀY

&lt;130&gt; Solar002

&lt;160&gt; 59

&lt;170&gt; PatentIn phiên bản 3.5

&lt;210&gt; 1

&lt;211&gt; 1482

&lt;212&gt; ADN

&lt;213&gt; Chưa biết

&lt;220&gt;

&lt;223&gt; Chủng mới

&lt;400&gt; 1

|                                                                    |      |
|--------------------------------------------------------------------|------|
| cttgagagtt tgatcctggc tcagagcgaa cgctggcggc aggcctaaca catgcaagtc  | 60   |
| gagcgcccag caatgggagc ggcagacggg tgagtaacgc gtggggatgt gccaatggt   | 120  |
| acggaataac ccagggaac ttggactaat accgtatgag cccttcgggg gaaagattta   | 180  |
| tcgccattgg atcaaccgc gtctgattag ctagttggtg gggtaacggc ccaccaaggc   | 240  |
| gacgatcagt agctgggtctg agaggatgat cagccacact gggactgaga cacggcccag | 300  |
| actcctacgg gaggcagcag tggggaatat tggacaatgg gcgcaagcct gatccagcca  | 360  |
| tgccgcgtgt gtgatgaagg ccttaggggt gtaaagcact ttcgccggtg aagataatga  | 420  |
| cggtaaccgg agaagaagcc cgggctaact tcgtgccagc agccgcggta atacgaaggg  | 480  |
| ggctagcgtt gctcggaatc actgggcgta aagcgcacgt aggcggatcg ttaagtcagg  | 540  |
| ggtgaaatcc tggagctcaa ctccagaact gcccttgata ctggcgacct tgagttcgag  | 600  |
| agaggttggt ggaactgcga gtgtagaggt gaaattcgta gatattcgca agaacaccag  | 660  |
| tggcgaaggc ggccaactgg ctcgatactg acgctgaggt gcgaaagcgt ggggagcaaa  | 720  |
| caggattaga taccctggta gtccacgccg taaacgatgg atgctagccg ttgggcagct  | 780  |
| tgctgttcag tggcgagct aacgcattaa gcatcccgcc tggggagtac ggtcgcaaga   | 840  |
| ttaaaactca aaggaattga cgggggcccc cacaagcggg ggagcatgtg gtttaattcg  | 900  |
| aagcaacgcg cagaacctta ccagcctttg acatggcagg acgatttcca gagatggatc  | 960  |
| tcttccagca atggacctgc acacaggtgc tgcatggctg tcgtcagctc gtgtcgtgag  | 1020 |

atgttgggtt aagtccccgca acgagcgcaa ccctcgcctc tagttgccag cattcagttg 1080  
 ggcaactctag agggactgcc ggtgataagc cgagaggaag gtggggatga cgtcaagtcc 1140  
 tcatggccct tacgggctgg gctacacacg tgctacaatg gtggtgacag tgggatgcga 1200  
 aaggcgacc tctagcaaat ctccaaaagc catctcagtt cggattgtac tctgcaactc 1260  
 gagtgcata agttggaatc gctagtaatc gtggatcagc atgccacggt gaatacgttc 1320  
 ccgggccttg tacacaccgc ccgtcacacc atgggagttg gctttaccgc aaggcgctgc 1380  
 gctaaccgc aaggaggca ggcgaccacg gtagggtcag cgactggggt gaagtcgtaa 1440  
 caaggtagcc gtaggggaac ctgcggctgg atcacctcct tt 1482

<210> 2  
 <211> 1467  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 2  
 atgggtgccg aagcaaccgt cgggcagatc acggacgcca agaagagata cgccgccggc 60  
 gtgctgaagt acgcccagat gggctactgg aacggcgact acgttcccaa ggacaccgac 120  
 ctcttgccgg tgttccgcat cccccccag gcgggcgtgg acccggtgga agccgccgcg 180  
 gcggtcgccg gcgaaagctc caccgctacc tggaccgtgg tgtggaccga ccggctcacc 240  
 gccgccgacg tctaccgcgc caaggcctac aagggtggagc cggtgccggg ccaggaaggc 300  
 cagtatttct gctacatcgc ctatgatctc gatttgttcg aggaaggctc catcgccaac 360  
 ctcacggcgt cgatcatcgg caacgtcttc tccttcaagc cgctgaaggc ggcgcggctg 420  
 gaggacatgc ggcttcccggt cgcctatgtg aagaccttcc gcggcccgcc caccggcatc 480  
 gtggtcgagc gcgagcgctt ggacaagttc ggccgcccc ttctgggcgc caccaccaag 540  
 ccgaagcttg gcctctcggg caagaattac ggccgcgtgg tctatgaggc cctcaagggc 600  
 ggctcgcact tcgtgaagga cgacgagaac atcaactcgc agcccttcat gcactggcgc 660  
 gatcgcttcc tctattgcat ggaggccgtc aacaaggccc aggccgagac cggcgaggtg 720  
 aaggggcact atctcaacat caccgccggg accatggagg agatgtaccg ccgcgccgag 780  
 ttcgccaagg aactgggctc cgtgggtggg atgggtggatc tcatcatcgg ctggaccgcc 840  
 atccagtcca tgtccaactg gtgccgcgag aacgacatga tcctgcacat gcaccgtgcg 900

```

ggccatggca cctacacgcg ccagaagagc cacggcgtct ccttcgcgt catcgccaag      960
tggctgcggc tcgccggcgt cgaccacctg cacaccggca ccgccgtggg caagctggaa      1020
ggcgacccca tgaccgtgca gggcttctac aatgtctgcc gcgagacgac gacgcagcag      1080
gacctacccc gcggcctggt cttcgagcag gactggggcg gcatccgcaa ggtgatgccg      1140
gtggcctccg gcggcatcca tgcgggccag atgcaccagc tcatcgacct gttcggcgag      1200
gacgtggtgc tccagttcgg cggcggcacc atcggccacc cggacggcat ccaggccggc      1260
gccaccgcca accgcgtggc gctggaaacc atgatcctcg cccgcaacga gggccgcgac      1320
atcaggaacg agggcccgga aatcctggtg gaagccgcca aatggtgccg tccgctgcgc      1380
gcggcgctcg atacctgggg cgaggtgacc ttcaactacg cctccaccga cacgtccgat      1440
tacgtgcca cgcgtccgt cgcctga                                           1467

```

<210> 3  
 <211> 488  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 3

```

Met Gly Ala Glu Ala Thr Val Gly Gln Ile Thr Asp Ala Lys Lys Arg
1           5           10           15

```

```

Tyr Ala Ala Gly Val Leu Lys Tyr Ala Gln Met Gly Tyr Trp Asn Gly
          20           25           30

```

```

Asp Tyr Val Pro Lys Asp Thr Asp Leu Leu Ala Val Phe Arg Ile Thr
          35           40           45

```

```

Pro Gln Ala Gly Val Asp Pro Val Glu Ala Ala Ala Val Ala Gly
          50           55           60

```

```

Glu Ser Ser Thr Ala Thr Trp Thr Val Val Trp Thr Asp Arg Leu Thr
65           70           75           80

```

```

Ala Ala Asp Val Tyr Arg Ala Lys Ala Tyr Lys Val Glu Pro Val Pro
          85           90           95

```

```

Gly Gln Glu Gly Gln Tyr Phe Cys Tyr Ile Ala Tyr Asp Leu Asp Leu
                                     3

```

|                                                                 |     |     |
|-----------------------------------------------------------------|-----|-----|
| 100                                                             | 105 | 110 |
| Phe Glu Glu Gly Ser Ile Ala Asn Leu Thr Ala Ser Ile Ile Gly Asn |     |     |
| 115                                                             | 120 | 125 |
| Val Phe Ser Phe Lys Pro Leu Lys Ala Ala Arg Leu Glu Asp Met Arg |     |     |
| 130                                                             | 135 | 140 |
| Leu Pro Val Ala Tyr Val Lys Thr Phe Arg Gly Pro Pro Thr Gly Ile |     |     |
| 145                                                             | 150 | 155 |
| Val Val Glu Arg Glu Arg Leu Asp Lys Phe Gly Arg Pro Leu Leu Gly |     |     |
|                                                                 | 165 | 170 |
| Ala Thr Thr Lys Pro Lys Leu Gly Leu Ser Gly Lys Asn Tyr Gly Arg |     |     |
|                                                                 | 180 | 185 |
| Val Val Tyr Glu Ala Leu Lys Gly Gly Leu Asp Phe Val Lys Asp Asp |     |     |
|                                                                 | 195 | 200 |
| Glu Asn Ile Asn Ser Gln Pro Phe Met His Trp Arg Asp Arg Phe Leu |     |     |
|                                                                 | 210 | 215 |
| Tyr Cys Met Glu Ala Val Asn Lys Ala Gln Ala Glu Thr Gly Glu Val |     |     |
| 225                                                             | 230 | 235 |
| Lys Gly His Tyr Leu Asn Ile Thr Ala Gly Thr Met Glu Glu Met Tyr |     |     |
|                                                                 | 245 | 250 |
| Arg Arg Ala Glu Phe Ala Lys Glu Leu Gly Ser Val Val Val Met Val |     |     |
|                                                                 | 260 | 265 |
| Asp Leu Ile Ile Gly Trp Thr Ala Ile Gln Ser Met Ser Asn Trp Cys |     |     |
|                                                                 | 275 | 280 |
| Arg Glu Asn Asp Met Ile Leu His Met His Arg Ala Gly His Gly Thr |     |     |
|                                                                 | 290 | 300 |
| Tyr Thr Arg Gln Lys Ser His Gly Val Ser Phe Arg Val Ile Ala Lys |     |     |
| 305                                                             | 310 | 315 |
| Trp Leu Arg Leu Ala Gly Val Asp His Leu His Thr Gly Thr Ala Val |     |     |

```
<400> 4
atgcgcatca cccaaggctc cttctccttc ctgccggacc tcaccgacac gcagatcaag      60
gcccgaggtgc aatattgcct ggaccagggc tgggcgggtct cggtggagca caccgacgat      120
```

```

ccccacccgc gcaacaccta ttgggagatg tggggcccg ccatgttcga tctgcgcgac      180
gcggccggcg tcttcggcga gatcgaagcc tgccggggccg ccaatcccga gcattatgtg      240
cgggtgaacg ctttcgattc cagccgcgga tgggagacga tccgcctgtc cttcatcggt      300
cagcggccca ccgtggaaga gggcttcgcg ctcgaccgca ccgaaggcaa gggccgcaac      360
cagagctacg ccatgcgcta ccgggcgcag ttcgcgccgc gctga                        405

```

<210> 5  
 <211> 134  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 5

Met Arg Ile Thr Gln Gly Ser Phe Ser Phe Leu Pro Asp Leu Thr Asp  
1 5 10 15

Thr Gln Ile Lys Ala Gln Val Gln Tyr Cys Leu Asp Gln Gly Trp Ala  
20 25 30

Val Ser Val Glu His Thr Asp Asp Pro His Pro Arg Asn Thr Tyr Trp  
35 40 45

Glu Met Trp Gly Pro Pro Met Phe Asp Leu Arg Asp Ala Ala Gly Val  
50 55 60

Phe Gly Glu Ile Glu Ala Cys Arg Ala Ala Asn Pro Glu His Tyr Val  
65 70 75 80

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

Ser Phe Ile Val Gln Arg Pro Thr Val Glu Glu Gly Phe Arg Leu Asp  
100 105 110

Arg Thr Glu Gly Lys Gly Arg Asn Gln Ser Tyr Ala Met Arg Tyr Arg  
115 120 125

Ala Gln Phe Ala Pro Arg  
130

<210> 6  
 <211> 1896  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 6  
 atgatgccat ctgagccgca cggcgcgggc atgccgcccc cacgggaagc ggccgcggtt 60  
 cccaccccc aggaggtgag cgcggtggtg gccgaggtgg tcgcggatgc cgtggcatcg 120  
 gtgggcggcg cacgcacccg gctcatggac atcgtccagc tggcccagca gcgtctcggc 180  
 catctctcgc aagagaccat ggcgccatt gccgcgcggc tcgccattcc gccggtggaa 240  
 gtggcggaca tgggtgtcctt ctacgccttc ctgaaccgcg cgcccaaggg ccgctaccac 300  
 atccgcctgt cgcgcagccc catctcgtcg atgaaggcg cgcgagcggt ggctgccgcc 360  
 ttctgccaga tcctcggcat cgccatgggc gagacctcgc aggatggcga cttcaccctg 420  
 gaatggacca acgacatcgg catggccgac caggagccgg ccgccctcgt caacggcacg 480  
 gtgatgacgc agctcgcgcc cggcgatgcg gccatcatcg tcggccggct gcgggccccat 540  
 cacgcgcca atgccctgcc gctgttcctt ggagccggcg tggccggctc cggcctgccc 600  
 catgcccgga tccgccccag cctggtgatg ccgggacagc ttctgttccg cgaggaccac 660  
 acgacgccgg gcgccggcat caaggcggca ctcgccctca ccccgacga agtgggtgcag 720  
 aaggtctccg ccgcgcgcct gcgcgggcgg ggtggcgccg gctttccac cggctctaaa 780  
 tggaagctct gccgccagtc gcccgccacc acccgccatg tgatctgcaa tgcggacgag 840  
 ggcgagcccg gcaccttcaa ggatcgctg ctgctcacgc aggcgccga cctcatgttc 900  
 gacggcatga ccatcgccgg ctacgccttg gggcgcggg agggcgtggt ctatctgcgc 960  
 ggcgagtacg cctatctgtg ggagcctctg catgcggtcc tgcgcgagcg ctatgggctc 1020  
 gggctcgccg gcgcgaacat cctgggacac gcgggcttcg acttcgacat ccgcatccag 1080  
 ctgggcgccc gcgcctatat ctgcggcgag gaatccgcgc tgggtggaatc gctggaaggc 1140  
 aagcgcggct cgccccgcga ccgccccccc ttccccaccg tgcgcggcca tctccagcag 1200  
 cccaccgccg tggacaatgt ggagaccttc gcctgcgcc cccgcatcct ggaggatggc 1260  
 gtggaggcgt tcgcgggcat cggcacgccc gaatccgcc gcacgaagct cctctcgggtg 1320  
 tcgggcgatt gcccgcgccc cggcgtgtat gaggtgccct tcggcctcac ggtgaacgcg 1380

ctgctcgacc ttgtcggcgc gccggacgcc gccttcgtgc agatgggtgg gccgtccggc 1440  
 caatgcgtgg cgccgaagga ttacggccgc cgcacgcct tcgaggacct gcccaccggc 1500  
 ggctcgggtga tgggtgttcgg cccggggcgc gacgtgctcg ccatggtgcg cgagttcgcg 1560  
 gatttcttcg ccggcgaatc ctgcggctgg tgcacgcct gccgggtggg caccaccttg 1620  
 ctcaaggaag agctggacaa gtcctcgcc aaccgcgcca ccctcgccga catccgcgcg 1680  
 ctggagaccc tggccacgac cgtctcccgc accagccgct gcggcctcgg ccagacggcg 1740  
 cccaaccca tcctttccac catgcgcaac ctgccggaag cctatgaggc gaggctgagg 1800  
 cccgaagact tcctgccctg ggcctcgctc gacgaggcgc tgaagcccgc catcgatc 1860  
 cagggccgcg cgcccgtgcc ggaggaagag gcatga 1896

<210> 7  
 <211> 631  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 7

Met Met Pro Ser Glu Pro His Gly Ala Gly Met Pro Pro Pro Arg Glu  
 1 5 10 15

Ala Ala Ala Val Pro Thr Pro Gln Glu Val Ser Ala Val Val Ala Glu  
 20 25 30

Val Val Ala Asp Ala Val Ala Ser Val Gly Gly Ala Arg Thr Arg Leu  
 35 40 45

Met Asp Ile Val Gln Leu Ala Gln Gln Arg Leu Gly His Leu Ser Glu  
 50 55 60

Glu Thr Met Ala Ala Ile Ala Ala Arg Leu Ala Ile Pro Pro Val Glu  
 65 70 75 80

Val Ala Asp Met Val Ser Phe Tyr Ala Phe Leu Asn Arg Ala Pro Lys  
 85 90 95

Gly Arg Tyr His Ile Arg Leu Ser Arg Ser Pro Ile Ser Leu Met Lys  
 100 105 110

Gly Ala Glu Ala Val Ala Ala Ala Phe Cys Gln Ile Leu Gly Ile Ala  
 115 120 125

Met Gly Glu Thr Ser Gln Asp Gly Asp Phe Thr Leu Glu Trp Thr Asn  
 130 135 140

Asp Ile Gly Met Ala Asp Gln Glu Pro Ala Ala Leu Val Asn Gly Thr  
 145 150 155 160

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

Leu Arg Ala His His Ala Pro Asn Ala Leu Pro Leu Phe Pro Gly Ala  
 180 185 190

Gly Val Ala Gly Ser Gly Leu Pro His Ala Arg Ile Arg Pro Ser Leu  
 195 200 205

Val Met Pro Gly Gln Leu Leu Phe Arg Glu Asp His Thr Thr Pro Gly  
 210 215 220

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

Lys Val Ser Ala Ala Arg Leu Arg Gly Arg Gly Gly Ala Gly Phe Pro  
 245 250 255

Thr Gly Leu Lys Trp Lys Leu Cys Arg Gln Ser Pro Ala Thr Thr Arg  
 260 265 270

His Val Ile Cys Asn Ala Asp Glu Gly Glu Pro Gly Thr Phe Lys Asp  
 275 280 285

Arg Val Leu Leu Thr Gln Ala Pro His Leu Met Phe Asp Gly Met Thr  
 290 295 300

Ile Ala Gly Tyr Ala Leu Gly Ala Arg Glu Gly Val Val Tyr Leu Arg  
 305 310 315 320

Gly Glu Tyr Ala Tyr Leu Trp Glu Pro Leu His Ala Val Leu Arg Glu  
 325 330 335

Arg Tyr Gly Leu Gly Leu Ala Gly Ala Asn Ile Leu Gly His Ala Gly  
 340 345 350

Phe Asp Phe Asp Ile Arg Ile Gln Leu Gly Ala Gly Ala Tyr Ile Cys  
 355 360 365

Gly Glu Glu Ser Ala Leu Val Glu Ser Leu Glu Gly Lys Arg Gly Ser  
 370 375 380

Pro Arg Asp Arg Pro Pro Phe Pro Thr Val Arg Gly His Leu Gln Gln  
 385 390 395 400

Pro Thr Ala Val Asp Asn Val Glu Thr Phe Ala Cys Ala Ala Arg Ile  
 405 410 415

Leu Glu Asp Gly Val Glu Ala Phe Ala Gly Ile Gly Thr Pro Glu Ser  
 420 425 430

Ala Gly Thr Lys Leu Leu Ser Val Ser Gly Asp Cys Pro Arg Pro Gly  
 435 440 445

Val Tyr Glu Val Pro Phe Gly Leu Thr Val Asn Ala Leu Leu Asp Leu  
 450 455 460

Val Gly Ala Pro Asp Ala Ala Phe Val Gln Met Gly Gly Pro Ser Gly  
 465 470 475 480

Gln Cys Val Ala Pro Lys Asp Tyr Gly Arg Arg Ile Ala Phe Glu Asp  
 485 490 495

Leu Pro Thr Gly Gly Ser Val Met Val Phe Gly Pro Gly Arg Asp Val  
 500 505 510

Leu Ala Met Val Arg Glu Phe Ala Asp Phe Phe Ala Gly Glu Ser Cys  
 515 520 525

Gly Trp Cys Thr Pro Cys Arg Val Gly Thr Thr Leu Leu Lys Glu Glu  
 530 535 540

Leu Asp Lys Leu Leu Ala Asn Arg Ala Thr Leu Ala Asp Ile Arg Ala  
 545 550 555 560

Leu Glu Thr Leu Ala Thr Thr Val Ser Arg Thr Ser Arg Cys Gly Leu  
565 570 575

Gly Gln Thr Ala Pro Asn Pro Ile Leu Ser Thr Met Arg Asn Leu Pro  
580 585 590

Glu Ala Tyr Glu Ala Arg Leu Arg Pro Glu Asp Phe Leu Pro Trp Ala  
595 600 605

Ser Leu Asp Glu Ala Leu Lys Pro Ala Ile Val Ile Gln Gly Arg Ala  
610 615 620

Pro Val Pro Glu Glu Glu Ala  
625 630

<210> 8

<211> 1491

<212> ADN

<213> Chưa biết

<220>

<223> chủng mới

<400> 8

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| atgagccggg gatccccga tgccgggaaa gaccgcacca tgagcgccac cgacggcacc   | 60  |
| accgcccccc gcaagatcgt catcgatccg gtgaccgcgc tggagggcca cggcaaggctc | 120 |
| accatccgcc tggatgaagc cggcgcgggtg gaggatgcgc gtttccacat cgtggagttc | 180 |
| cgcggtttcg agcggttcat ccagggccgg atgtactggg aagtgccct tatcatccag   | 240 |
| cggctgtgcg gcatctgccc ggtgagccac catctggcgg cggcgaaagc catggaccag  | 300 |
| gtggcgggcg tggaccgcgt accgcccacc gccgagaaac tgcgccggct gatgcattat  | 360 |
| gggcaggtgc tgcaatcaa cgctttgcac atcttccacc tcgcctcgcc cgacctctg    | 420 |
| ttcggtttcg acgcgccggc cgagcagcgc aacatcatcg ccgtgctcca gcgttatccg  | 480 |
| gagatcggca aatgggcgat cttcatcagg aagttcggcc aggaggtcat caaggccacc  | 540 |
| ggcgggcgca agatccatcc caccagcgcc attcccggcg ggggtcaacca gaacctcgcc | 600 |
| gtggaggacc gcgacgccct gcgcgccaag gtgggcgaga tcatcagctg gtgcatggcg  | 660 |
| gcgctggacc atcacaaggc ctatgtggcg gaaaaccggg cgctgcatga cagcttcgcc  | 720 |
| gccttcccct ccgccttcat gagcctcgtg gggccggatg gcggcatgga cctttatgac  | 780 |

```

ggcaccctgc gggatgatcga tgccgagggc gccccctca tcgaaggcgc gccgcccgcc      840
tcctaccgcg accacctcat cgaggaggtg cggccctgga gctatctgaa attcccccat      900
ctgcgcgcct tcggccgcga cgatggctgg tatcgggtcg gccccctcgc ccaggtcaat      960
tgcgccgcgt ccatcgacac gccccgcgcc gaggcggccc ggcgggactt catggccgag     1020
ggcgggcgga agccggtgca tgccaccctc gcttatcact gggcgcggtt catcgtgctg     1080
gtccattgcg cggagaagat cgaacagctg ctgttcgacg acgacctgca aggctgcgat     1140
ctgcgtgcgg agggcacccg gcgcggggaa ggcgtcgcct ggatcgaggc gccgcgcggc     1200
accctcatcc accattacga ggtggacgag aacgaccagg tgcgccgcgc caacctcatc     1260
gtctccacca cccacaataa cgaggccatg aaccgcgccg tgcggcaggt ggcgaagacg     1320
gacctttccg gtcgcgagat caccgaaggg ctgctgaacc atatcgaggt ggccatccgc     1380
gccttcgacc cctgcctgtc ctgcgccacc catgcgctgg gccagatgcc gctgatcgtg     1440
acgcttgaag atgcctccgg cgagagatc gccgcggag tgaaggaatg a                 1491

```

<210> 9  
 <211> 496  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 9

Met Ser Arg Gly Ser Pro Asp Ala Gly Lys Asp Arg Thr Met Ser Ala  
 1 5 10 15

Thr Asp Gly Thr Thr Ala Pro Arg Lys Ile Val Ile Asp Pro Val Thr  
 20 25 30

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

Ala Val Glu Asp Ala Arg Phe His Ile Val Glu Phe Arg Gly Phe Glu  
 50 55 60

Arg Phe Ile Gln Gly Arg Met Tyr Trp Glu Val Pro Leu Ile Ile Gln  
 65 70 75 80

Arg Leu Cys Gly Ile Cys Pro Val Ser His His Leu Ala Ala Ala Lys

85

90

95

Ala Met Asp Gln Val Ala Gly Val Asp Arg Val Pro Pro Thr Ala Glu  
 100 105 110

Lys Leu Arg Arg Leu Met His Tyr Gly Gln Val Leu Gln Ser Asn Ala  
 115 120 125

Leu His Ile Phe His Leu Ala Ser Pro Asp Leu Leu Phe Gly Phe Asp  
 130 135 140

Ala Pro Ala Glu Gln Arg Asn Ile Ile Ala Val Leu Gln Arg Tyr Pro  
 145 150 155 160

Glu Ile Gly Lys Trp Ala Ile Phe Ile Arg Lys Phe Gly Gln Glu Val  
 165 170 175

Ile Lys Ala Thr Gly Gly Arg Lys Ile His Pro Thr Ser Ala Ile Pro  
 180 185 190

Gly Gly Val Asn Gln Asn Leu Ala Val Glu Asp Arg Asp Ala Leu Arg  
 195 200 205

Ala Lys Val Gly Glu Ile Ile Ser Trp Cys Met Ala Ala Leu Asp His  
 210 215 220

His Lys Ala Tyr Val Ala Glu Asn Arg Ala Leu His Asp Ser Phe Ala  
 225 230 235 240

Ala Phe Pro Ser Ala Phe Met Ser Leu Val Gly Pro Asp Gly Gly Met  
 245 250 255

Asp Leu Tyr Asp Gly Thr Leu Arg Val Ile Asp Ala Glu Gly Ala Pro  
 260 265 270

Leu Ile Glu Gly Ala Pro Pro Ala Ser Tyr Arg Asp His Leu Ile Glu  
 275 280 285

Glu Val Arg Pro Trp Ser Tyr Leu Lys Phe Pro His Leu Arg Ala Phe  
 290 295 300

Gly Arg Asp Asp Gly Trp Tyr Arg Val Gly Pro Leu Ala Gln Val Asn  
 13

|                       |           |
|-----------------------|-----------|
| $\langle 210 \rangle$ | 10        |
| $\langle 211 \rangle$ | 720       |
| $\langle 212 \rangle$ | ADN       |
| $\langle 213 \rangle$ | Chưa biết |
| $\langle 220 \rangle$ |           |
| $\langle 223 \rangle$ | chúng mới |
| $\langle 400 \rangle$ | 10        |

atgagcgaga ccccttcac ctttaccgtg gacggcatcg cgggtcccggc ccccccggc 60  
 cagagcgtca tcgaggcgtg cgatgcggcg ggcattctata tcccgcgcct gtgccaccac 120  
 ccggacctgc cgccggcggg ccattgccgg gtgtgcacct gcatcatcga cgggcggccg 180  
 gccagcgctt gcaccatgcc cgccgccagg ggcattggtgg tggagaacga gacgcccgt 240  
 ttgtggcgg agcggcgcac gctgacgag atgctgttcg cggaaggcaa ccatttctgc 300  
 cagttctgcg aggcgagcgg cgattgcgaa ttgcaggcgc tgggctacct gttcggcatg 360  
 gtggccccgc ctttccccca tctgtggccg aagcggccgg tggatgccag ccatccggat 420  
 atctatatcg accacaatcg ctgcatcctg tgctcgcgt gcgtgcgcgc ctcgcgcacc 480  
 ctggacggca agtccgtgtt cggcttcgag gggcgcggca tcgagatgca tctggcgggtg 540  
 accggcgggc acctggacga cagcgccatc gccgccgccg acagggcgggt tgagatgtgc 600  
 ccggtgggct gcatcgtcct caagcgacc ggctaccgca cgccctatgg ccggcggcgc 660  
 tacgacgccc cgcccatcgg ctccgacatc accgcccggc gcggcggcgc gaaggactga 720

<210> 11  
 <211> 239  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 11

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

Ala Thr Pro Gly Gln Ser Val Ile Glu Ala Cys Asp Ala Ala Gly Ile  
20 25 30

Tyr Ile Pro Arg Leu Cys His His Pro Asp Leu Pro Pro Ala Gly His  
35 40 45

Cys Arg Val Cys Thr Cys Ile Ile Asp Gly Arg Pro Ala Ser Ala Cys  
50 55 60

Thr Met Pro Ala Ala Arg Gly Met Val Val Glu Asn Glu Thr Pro Ala  
65 70 75 80

Leu Leu Ala Glu Arg Arg Thr Leu Ile Glu Met Leu Phe Ala Glu Gly  
15

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

<210> 12  
 <211> 543  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

|                                                                     |     |
|---------------------------------------------------------------------|-----|
| <400> 12                                                            |     |
| atggccaagc ccaaactcgc cacctgcgcg ctggccggct gcttcggctg ccacatgtcc   | 60  |
| ttcctggaca tggacgagcg catcgtcgag ctcatcgacc tgggtggacct cgacgtctcg  | 120 |
| cccctcgacg acaagaaaaa cttcaccggc atggtggaaa tcggcctggt ggaaggcggc   | 180 |
| tgcgccgacg agcgccatgt gaagggtgctg cgcgagttcc gcgagaaatc ccgcatacctg | 240 |

gtggcgggtgg gcgcctgcgc catcaccggc ggcatcccgg cattgcgcaa cctcgccggc 300  
 ctcgacgaat gcctgaggga agcctacctc accggcccca cgggtggaagg cggcgggctc 360  
 attcccaacg acccggagct gccgctgctg ctggacaagg tctatccggt gcaggacttc 420  
 gtgaagatcg accatttcct gcccggctgc ccgccctcgg ccgacgccat ctgggcggct 480  
 ctgaaggcgc tgctgaccgg caccgagccg catctgccct acccgctttt caagtacgaa 540  
 tga 543

<210> 13  
 <211> 180  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chúng mới

<400> 13

Met Ala Lys Pro Lys Leu Ala Thr Cys Ala Leu Ala Gly Cys Phe Gly  
 1 5 10 15

Cys His Met Ser Phe Leu Asp Met Asp Glu Arg Ile Val Glu Leu Ile  
 20 25 30

Asp Leu Val Asp Leu Asp Val Ser Pro Leu Asp Asp Lys Lys Asn Phe  
 35 40 45

Thr Gly Met Val Glu Ile Gly Leu Val Glu Gly Gly Cys Ala Asp Glu  
 50 55 60

Arg His Val Lys Val Leu Arg Glu Phe Arg Glu Lys Ser Arg Ile Leu  
 65 70 75 80

Val Ala Val Gly Ala Cys Ala Ile Thr Gly Gly Ile Pro Ala Leu Arg  
 85 90 95

Asn Leu Ala Gly Leu Asp Glu Cys Leu Arg Glu Ala Tyr Leu Thr Gly  
 100 105 110

Pro Thr Val Glu Gly Gly Gly Leu Ile Pro Asn Asp Pro Glu Leu Pro  
 115 120 125

Leu Leu Leu Asp Lys Val Tyr Pro Val Gln Asp Phe Val Lys Ile Asp

130

135

140

His Phe Leu Pro Gly Cys Pro Pro Ser Ala Asp Ala Ile Trp Ala Ala  
 145 150 155 160

Leu Lys Ala Leu Leu Thr Gly Thr Glu Pro His Leu Pro Tyr Pro Leu  
 165 170 175

Phe Lys Tyr Glu  
 180

<210> 14  
 <211> 1013  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 14  
 tccagaccgc ggcaacattg ctccatgtgc tgggcaccct ggccggccgc tggccccata 60  
 ccctcgcgct ccagcccggc ggggtgacct gaagcgccga ccagcacgac cgcattgcgc 120  
 tgctcgcgac gctgaaggcg gtgcgggcgg cgctggaaga gacctgttc ggcgcgctt 180  
 tggaagaggt ggcgccctg gacggcgccg ccgctgtgga ggcttggcg gccaacggcc 240  
 cggaagggga tttccgcctg ttcctggaga tcgcccga cctggagctg gaccggctcg 300  
 gccgcgcgca cgaccgcttt ctctccttcg gcgcctacgc ccaggacgag gggcgcttt 360  
 atggcgccgg caccttcgag gccgggacgg cgggagggt cgatcccaac gccatcaccg 420  
 aggaccacgc cttcgccgc atggaggacc gcgcggcgcc ccattgcgcc tttgacggct 480  
 ccaccttccc cgatgccgac gacaccgagg gctacacctg gtgcaaggcg ccgcgcttg 540  
 ccggcctgcc cttcgagacc ggcgcttcg cccggcaggt ggtggcgggc catccgctcg 600  
 cccgggacct cgtgacgcgg gaaggcggca ctgtgcgag ccgctgggtc ggccggctgc 660  
 tggaaccgc gcgcacctg atcgccatgg agggctgggt gaaggactg cggcccgaag 720  
 ggccctggtg cggccagggc cacctgcccc aggaaggccg cgccttcggc ctcaccgagg 780  
 cggcgcgcg ggcgctcggc cactggatgg tggaggagaa gggccgcat gcccgtacc 840  
 agatcatcgc cccaccacc tggaacttct ccccccgcga cggcgcgggc ctgcccggcc 900  
 cgctggagac ggccctggtg ggcgcgccc tgccggcagg agagacgac cccgtgagcg 960

tgcagcacat cgtgcgctcc ttcgacccgt gcatggcttg cactgtgcat tga

1013

<210> 15  
 <211> 485  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 15

Met Ser Ala Glu Thr Arg Arg Leu Val Val Gly Pro Phe Asn Arg Val  
 1 5 10 15

Glu Gly Asp Leu Glu Val Arg Leu Asp Val Gln Asp Gly Arg Val Gln  
 20 25 30

Gln Ala Phe Val Ser Ser Pro Leu Phe Arg Gly Phe Glu Arg Ile Leu  
 35 40 45

Glu Gly Arg Asp Pro Arg Asp Ala Leu Val Ile Ala Pro Arg Ile Cys  
 50 55 60

Gly Ile Cys Ser Val Ser Gln Ser His Ala Ala Ala Leu Ala Leu Ala  
 65 70 75 80

Gly Leu Gln Gly Ile Ala Pro Thr His Asp Gly Arg Ile Ala Thr Asn  
 85 90 95

Leu Ile Val Ala Ala Glu Asn Val Ala Asp His Leu Thr His Phe His  
 100 105 110

Val Phe Phe Met Pro Asp Phe Ala Arg Ala Val Tyr Glu Asp Arg Pro  
 115 120 125

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

Val Arg Arg Ala Leu Gln Thr Arg Ala Thr Leu Leu His Val Leu Gly  
 145 150 155 160

Thr Leu Ala Gly Arg Trp Pro His Thr Leu Ala Leu Gln Pro Gly Gly  
 165 170 175

Val Thr Arg Ser Ala Asp Gln His Asp Arg Met Arg Leu Leu Ala Thr  
 180 185 190

Leu Lys Ala Val Arg Ala Ala Leu Glu Glu Thr Leu Phe Gly Ala Pro  
 195 200 205

Leu Glu Glu Val Ala Ala Leu Asp Gly Ala Ala Ala Val Glu Ala Trp  
 210 215 220

Arg Ala Asn Gly Pro Glu Gly Asp Phe Arg Leu Phe Leu Glu Ile Ala  
 225 230 235 240

Ala Asp Leu Glu Leu Asp Arg Leu Gly Arg Ala His Asp Arg Phe Leu  
 245 250 255

Ser Phe Gly Ala Tyr Ala Gln Asp Glu Gly Arg Leu Tyr Gly Ala Gly  
 260 265 270

Thr Phe Glu Ala Gly Thr Ala Gly Gly Leu Asp Pro Asn Ala Ile Thr  
 275 280 285

Glu Asp His Ala Phe Ala Arg Met Glu Asp Arg Ala Ala Pro His Ala  
 290 295 300

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

Thr Trp Cys Lys Ala Pro Arg Leu Ala Gly Leu Pro Phe Glu Thr Gly  
 325 330 335

Ala Phe Ala Arg Gln Val Val Ala Gly His Pro Leu Ala Arg Asp Leu  
 340 345 350

Val Thr Arg Glu Gly Gly Thr Val Arg Ser Arg Val Val Gly Arg Leu  
 355 360 365

Leu Glu Thr Ala Arg Thr Leu Ile Ala Met Glu Gly Trp Val Lys Glu  
 370 375 380

Leu Arg Pro Glu Gly Pro Trp Cys Ala Gln Gly His Leu Pro Gln Glu  
 385 390 395 400

Gly Arg Ala Phe Gly Leu Thr Glu Ala Ala Arg Gly Ala Leu Gly His  
 405 410 415

Trp Met Val Val Glu Lys Gly Arg Ile Ala Arg Tyr Gln Ile Ile Ala  
 420 425 430

Pro Thr Thr Trp Asn Phe Ser Pro Arg Asp Gly Ala Gly Leu Pro Gly  
 435 440 445

Pro Leu Glu Thr Ala Leu Val Gly Ala Pro Val Arg Gln Gly Glu Thr  
 450 455 460

Thr Pro Val Ser Val Gln His Ile Val Arg Ser Phe Asp Pro Cys Met  
 465 470 475 480

Val Cys Thr Val His  
 485

<210> 16  
 <211> 858  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 16  
 acgggggagg aagcccgcg catcttcgac gccatccttg ccggcggttat cgtcctcgac 60  
 gccctgtgcg tggaaggcg gctgctgcg gggccgaacg gcaccgggcg cttccatgtg 120  
 ctggcgggca cggacacccc caccatcgac tgggcgcggc agctcgccgg catggcgcg 180  
 cacgtggttg cggtgggcac ctgcgccgcc tatgggggcg tgacggcggc gggcatcaac 240  
 cccaccgatg cctgcggcct ccagttcgac ggacgccgga aggggtggggc gctggggggcg 300  
 gacttccgct cccgctcggg gcttccggtc atcaatgttg ccggctgccc caccatccc 360  
 aactgggtga cggaaccct gatgctgctc gcctgcggcc tgctgggcga ggccgacctc 420  
 gacgtctatg gccgcccgcg cttctatgcg gacctgctgg tgcatcacgg ctgcccgcgc 480  
 aacgaatact atgaatacaa ggcgagcgcc gagaagatga gcgacctcgg ctgcatgatg 540  
 gagcatctgg gctgcctcgg caccagggc cagccgact gcaacacgcg cctttggaat 600  
 ggcgagggct cgtgcacccg cggcgggtat gcctgcatca actgcacggc gccggaattc 660

gaggagccgg gccacgcctt cctggagacg cccaagatcg gcggcatccc catcggcctg 720  
 cccaccgaca tgcccaaggc ctggttcacg gccttgtcct ccctcgccaa ggcggcgacg 780  
 ccggagcggc tgcgcaagaa cgcggtgtcc gaccatgtgg tcacgccgcc cgccgtcaag 840  
 gacatcaagc ggcgatga 858

<210> 17  
 <211> 335  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 17

Met Ser Thr Pro Phe Ser Val Leu Trp Leu Gln Ser Gly Gly Cys Gly  
 1 5 10 15

Gly Cys Thr Met Ser Leu Leu Cys Ala Glu Ala Pro Asp Leu Ala Thr  
 20 25 30

Thr Leu Asp Ala Ala Gly Ile Gly Phe Leu Trp His Pro Ala Leu Ser  
 35 40 45

Glu Glu Thr Gly Glu Glu Ala Arg Ala Ile Phe Asp Ala Ile Leu Ala  
 50 55 60

Gly Val Ile Val Leu Asp Ala Leu Cys Val Glu Gly Ala Leu Leu Arg  
 65 70 75 80

Gly Pro Asn Gly Thr Gly Arg Phe His Val Leu Ala Gly Thr Asp Thr  
 85 90 95

Pro Thr Ile Asp Trp Ala Arg Gln Leu Ala Gly Met Ala Arg His Val  
 100 105 110

Val Ala Val Gly Thr Cys Ala Ala Tyr Gly Gly Val Thr Ala Ala Gly  
 115 120 125

Ile Asn Pro Thr Asp Ala Cys Gly Leu Gln Phe Asp Gly Arg Arg Lys  
 130 135 140

Gly Gly Ala Leu Gly Ala Asp Phe Arg Ser Arg Ser Gly Leu Pro Val  
 22

|                                                                 |  |     |  |     |  |     |
|-----------------------------------------------------------------|--|-----|--|-----|--|-----|
| 145                                                             |  | 150 |  | 155 |  | 160 |
| Ile Asn Val Ala Gly Cys Pro Thr His Pro Asn Trp Val Thr Glu Thr |  |     |  |     |  |     |
|                                                                 |  | 165 |  | 170 |  | 175 |
| Leu Met Leu Leu Ala Cys Gly Leu Leu Gly Glu Ala Asp Leu Asp Val |  |     |  |     |  |     |
|                                                                 |  | 180 |  | 185 |  | 190 |
| Tyr Gly Arg Pro Arg Phe Tyr Ala Asp Leu Leu Val His His Gly Cys |  |     |  |     |  |     |
|                                                                 |  | 195 |  | 200 |  | 205 |
| Pro Arg Asn Glu Tyr Tyr Glu Tyr Lys Ala Ser Ala Glu Lys Met Ser |  |     |  |     |  |     |
|                                                                 |  | 210 |  | 215 |  | 220 |
| Asp Leu Gly Cys Met Met Glu His Leu Gly Cys Leu Gly Thr Gln Ala |  |     |  |     |  |     |
|                                                                 |  | 225 |  | 230 |  | 240 |
| His Ala Asp Cys Asn Thr Arg Leu Trp Asn Gly Glu Gly Ser Cys Thr |  |     |  |     |  |     |
|                                                                 |  | 245 |  | 250 |  | 255 |
| Arg Gly Gly Tyr Ala Cys Ile Asn Cys Thr Ala Pro Glu Phe Glu Glu |  |     |  |     |  |     |
|                                                                 |  | 260 |  | 265 |  | 270 |
| Pro Gly His Ala Phe Leu Glu Thr Pro Lys Ile Gly Gly Ile Pro Ile |  |     |  |     |  |     |
|                                                                 |  | 275 |  | 280 |  | 285 |
| Gly Leu Pro Thr Asp Met Pro Lys Ala Trp Phe Ile Ala Leu Ser Ser |  |     |  |     |  |     |
|                                                                 |  | 290 |  | 295 |  | 300 |
| Leu Ala Lys Ala Ala Thr Pro Glu Arg Leu Arg Lys Asn Ala Val Ser |  |     |  |     |  |     |
|                                                                 |  | 305 |  | 310 |  | 315 |
| Asp His Val Val Thr Pro Pro Ala Val Lys Asp Ile Lys Arg Arg     |  |     |  |     |  |     |
|                                                                 |  | 325 |  | 330 |  | 335 |

<210> 18  
 <211> 870  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 18

gtgaccgagc gcctgtccga cgtcaacgcc cgcacgcct cggcgggca gctctcatcg 60  
 gtcacacagg ccatgcgggg cattgcggcg gcgcggggcg gggaggcgcg gggcggctc 120  
 gacggcatcc gcgcctatgc gcagaccatc gccgaggcca tcggccatgt gctcgccgtg 180  
 ctgcccagg aggcccgcg ccggtcctcc gggcaccggc atcggggcca tgcggtcac 240  
 gccctgtgcg cggagcaggg ctttgccggc gtcttcaac agcgggtgct ggacgaggcc 300  
 gcccggctgc tgaccggcgg ggcggggccg gccgagctgc tgctggtggg cgaccggggc 360  
 ctgatggtgg ccgcgagcg ggggctcgat gtctcctggt cggcgcccat ggtggcccat 420  
 gcggggcagg cctcggcgct ggcggaccgc atcagcgagg agctctaccg gcggatcgat 480  
 gcgggacggg tgacgcgggt gtcgggtggt cacgccgagc ccgccgcgtc cgccgccatc 540  
 gagacggtgg tgaaagtgct ggtgccgttc gacttcgccc gcttccccct ggcgcgggtg 600  
 gcatccgcc cgctcatgac catgccgcc ccgcggctgc tggcccagct gtcggaggaa 660  
 tatgtgttcg ccgagctgtg cgaggcgctc accttgcct tcgcggcgga gaacgaggcc 720  
 cgcacgcggg ccatgatcgc gcccgcgcc aatgtggccg ataccctgga gggcctcgtc 780  
 ggccgcgcc ggcagatgcg ccaggaggag ataccaacg agatcatcga gctggaaggc 840  
 ggcgccggca gcgccggca tgcggattga 870

<210> 19  
 <211> 289  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 19

Met Thr Glu Arg Leu Ser Asp Val Asn Ala Arg Ile Ala Ser Val Arg  
 1 5 10 15

Gln Leu Ser Ser Val Ile Thr Ala Met Arg Gly Ile Ala Ala Ala Arg  
 20 25 30

Ala Arg Glu Ala Arg Gly Arg Leu Asp Gly Ile Arg Ala Tyr Ala Gln  
 35 40 45

Thr Ile Ala Glu Ala Ile Gly His Val Leu Ala Val Leu Pro Glu Glu  
 50 55 60

Ala Arg Ala Arg Ser Ser Gly His Arg His Arg Gly His Ala Val Ile  
65 70 75 80

Ala Leu Cys Ala Glu Gln Gly Phe Ala Gly Val Phe Asn Glu Arg Val  
85 90 95

Leu Asp Glu Ala Ala Arg Leu Leu Thr Gly Gly Ala Gly Pro Ala Glu  
100 105 110

Leu Leu Leu Val Gly Asp Arg Gly Leu Met Val Ala Arg Glu Arg Gly  
115 120 125

Leu Asp Val Ser Trp Ser Val Pro Met Val Ala His Ala Gly Gln Ala  
130 135 140

Ser Ala Leu Ala Asp Arg Ile Ser Glu Glu Leu Tyr Arg Arg Ile Asp  
145 150 155 160

Ala Gly Arg Val Thr Arg Val Ser Val Val His Ala Glu Pro Ala Ala  
165 170 175

Ser Ala Ala Ile Glu Thr Val Val Lys Val Leu Val Pro Phe Asp Phe  
180 185 190

Ala Arg Phe Pro Leu Ala Arg Val Ala Ser Ala Pro Leu Met Thr Met  
195 200 205

Pro Pro Pro Arg Leu Leu Ala Gln Leu Ser Glu Glu Tyr Val Phe Ala  
210 215 220

Glu Leu Cys Glu Ala Leu Thr Leu Ser Phe Ala Ala Glu Asn Glu Ala  
225 230 235 240

Arg Met Arg Ala Met Ile Ala Ala Arg Ala Asn Val Ala Asp Thr Leu  
245 250 255

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

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

Asp

<210> 20  
 <211> 1536  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 20  
 atgagcacgg gcgcgcaagc gagcgaggat tggctcacc ggagccgggc ggccctggcc 60  
 gggacgcgcc tttcccagca atcccaatcg gtgggcccggg tggaggagat ggccgacggc 120  
 atcgcccgcg tctccggcct gccggatgtg cggctcgacg agcttctcac cttcgagggc 180  
 ggccagaccg gctatgccct caccctcgat cgcaccgaga tcgccgtggt gctgctggat 240  
 gacgcctccg gcgtggaggc gggcgcccgg gtgttcggca ccggcgagggt ggtgaagggt 300  
 ccggtggggc cggggctgct gggccgcata gtcgaccccc tcggccggcc catggaccgc 360  
 tccgagccgg tgggtggcgca ggcgcaccat cccatcgagc ggccggcgcc ggccatcatc 420  
 gcccgcgacc tgggtctcgca gccggttcag accggcacgc tgggtggtgga tgcgctgttc 480  
 tccctcggcc ggggccagcg cgagctcatc atcggcgacc gggctaccgg caagaccgcc 540  
 atcgcggtgg acaccatcat cagccagaag cattcggaca tcgtgtgcat ctacgtggcg 600  
 gtgggccagc gcgccgccgc cgtggagcgg gtgggtggagg cgggtgcgcgc ccacggggcg 660  
 atcgagcgct gcatcttcgt ggtcgcctcg gccgccgcct cgccagggct gcaatggatc 720  
 gcgccgttcg ccggcatgac catggcggaa tatttccgcg acaacggcca gcatgcgctc 780  
 atcatcatcg atgatctcac caagcatgcg gccacccatc gcgagctggc gctgctcacc 840  
 cagagccgc cgggccgcga ggcctatccc ggcgacatct tctatgtgca cggccgcctt 900  
 ctggagcggg ccgccaagct ctccgccgag ctgggcgggtg gctcgctcac ggccctgccc 960  
 atcgcggaga cggacgcggg aaacctctcc gcctatatcc ccaccaacct catctccatc 1020  
 accgatgggc agatcgtgct ggattcgcgg ctgttcgcgg ccaaccagcg cccggcgggtg 1080  
 gatgtggggc tctccgtgag ccgggtgggc ggcaaggcgc agcatccgc gcttcggggc 1140  
 gtgtccgggc gcatccggct cgattattcc cagttcctgg agctggaaat gttcaccgc 1200  
 ttcggcggca tcaccgatac ccgctgaag gcgcagatca cccggggcga gcgcatccgc 1260

gcgctgctca cccagccgcg cttttccacc ctgcgccttc aggacgaggt ggcgctgctg 1320  
gccgcgctgg cggaggggggt gttcgacact ttggccccgg ggctgatggg cgccgtgcgt 1380  
gccccgattc cggcccagct ggatgcgcag gtgaaggacg tggcctcggc cctcgccgag 1440  
ggcaagggtgc tggaggaggg cttgcacgcc cgtctcgtgg cggccgtgcg ggccgtcgcg 1500  
gcgggacgtgg ccgcgaccgc gaaggccggg ccgtga 1536

<210> 21  
<211> 511  
<212> PRT  
<213> Chưa biết

<220>  
<223> chủng mới

<400> 21

Met Ser Thr Gly Ala Gln Ala Ser Glu Asp Trp Leu Thr Arg Ser Arg  
1 5 10 15

Ala Ala Leu Ala Gly Thr Arg Leu Ser Gln Gln Ser Gln Ser Val Gly  
20 25 30

Arg Val Glu Glu Met Ala Asp Gly Ile Ala Arg Val Ser Gly Leu Pro  
35 40 45

Asp Val Arg Leu Asp Glu Leu Leu Thr Phe Glu Gly Gly Gln Thr Gly  
50 55 60

Tyr Ala Leu Thr Leu Asp Arg Thr Glu Ile Ala Val Val Leu Leu Asp  
65 70 75 80

Asp Ala Ser Gly Val Glu Ala Gly Ala Arg Val Phe Gly Thr Gly Glu  
85 90 95

Val Val Lys Val Pro Val Gly Pro Gly Leu Leu Gly Arg Ile Val Asp  
100 105 110

Pro Leu Gly Arg Pro Met Asp Arg Ser Glu Pro Val Val Ala Gln Ala  
115 120 125

His His Pro Ile Glu Arg Pro Ala Pro Ala Ile Ile Ala Arg Asp Leu  
130 135 140

Val Ser Gln Pro Val Gln Thr Gly Thr Leu Val Val Asp Ala Leu Phe  
 145 150 155 160

Ser Leu Gly Arg Gly Gln Arg Glu Leu Ile Ile Gly Asp Arg Ala Thr  
 165 170 175

Gly Lys Thr Ala Ile Ala Val Asp Thr Ile Ile Ser Gln Lys His Ser  
 180 185 190

Asp Ile Val Cys Ile Tyr Val Ala Val Gly Gln Arg Ala Ala Ala Val  
 195 200 205

Glu Arg Val Val Glu Ala Val Arg Ala His Gly Ala Ile Glu Arg Cys  
 210 215 220

Ile Phe Val Val Ala Ser Ala Ala Ala Ser Pro Gly Leu Gln Trp Ile  
 225 230 235 240

Ala Pro Phe Ala Gly Met Thr Met Ala Glu Tyr Phe Arg Asp Asn Gly  
 245 250 255

Gln His Ala Leu Ile Ile Ile Asp Asp Leu Thr Lys His Ala Ala Thr  
 260 265 270

His Arg Glu Leu Ala Leu Leu Thr His Glu Pro Pro Gly Arg Glu Ala  
 275 280 285

Tyr Pro Gly Asp Ile Phe Tyr Val His Ala Arg Leu Leu Glu Arg Ala  
 290 295 300

Ala Lys Leu Ser Ala Glu Leu Gly Gly Gly Ser Leu Thr Ala Leu Pro  
 305 310 315 320

Ile Ala Glu Thr Asp Ala Gly Asn Leu Ser Ala Tyr Ile Pro Thr Asn  
 325 330 335

Leu Ile Ser Ile Thr Asp Gly Gln Ile Val Leu Asp Ser Arg Leu Phe  
 340 345 350

Ala Ala Asn Gln Arg Pro Ala Val Asp Val Gly Leu Ser Val Ser Arg  
 355 360 365

Val Gly Gly Lys Ala Gln His Pro Ala Leu Arg Ala Val Ser Gly Arg  
370 375 380

Ile Arg Leu Asp Tyr Ser Gln Phe Leu Glu Leu Glu Met Phe Thr Arg  
385 390 395 400

Phe Gly Gly Ile Thr Asp Thr Arg Val Lys Ala Gln Ile Thr Arg Gly  
405 410 415

Glu Arg Ile Arg Ala Leu Leu Thr Gln Pro Arg Phe Ser Thr Leu Arg  
420 425 430

Leu Gln Asp Glu Val Ala Leu Leu Ala Ala Leu Ala Glu Gly Val Phe  
435 440 445

Asp Thr Leu Ala Pro Gly Leu Met Gly Ala Val Arg Ala Arg Ile Pro  
450 455 460

Ala Gln Leu Asp Ala Gln Val Lys Asp Val Ala Ser Ala Leu Ala Glu  
465 470 475 480

Gly Lys Val Leu Glu Glu Gly Leu His Ala Arg Leu Val Ala Ala Val  
485 490 495

Arg Ala Val Ala Ala Asp Val Ala Ala Thr Ala Lys Ala Gly Pro  
500 505 510

<210> 22

<211> 741

<212> ADN

<213> Chưa biết

<220>

<223> chủng mới

<400> 22

|                                                                     |     |
|---------------------------------------------------------------------|-----|
| atgcagatcg actggtggac gctgggcctg cagacgggtca acgtcctcgt tctcatctgg  | 60  |
| ctcctgagcc gcttcctgtt caagccggtg gcgcgaggtca tcgcgagcgc cgtgcccag   | 120 |
| atcgagaagc tgggtggagga tgcgcgcgcc gccaaaggccg ccgcccaggc cgagcgggac | 180 |
| acggcgaagg cggaggaggc ggccttgcc gccgagcgcg gcgcccgcgc ggcggcggtc    | 240 |
| gccaaaggagg cggaggcgca gaaggcggca ttgctggccg ccgccaagac cgaggccgag  | 300 |

gccctgcacg cggccgcgga agcggccatc gtccgggcg cggcgagcga ggaggaagcc 360  
 gccgccgacc gcgccagccg ccttgccgtg gacatcgccg ccaagctgct ggaccggctg 420  
 cccgacgacg cccgggtcgc gggcttcacg gatggcctcg ccgaggggct tgaagccctg 480  
 cccgaggcga gccgggcggt gatcggcgtc gacggcgcg cagtgcgcgt gacggccgcg 540  
 cgcgccctta tgccggcgga ggaggaggcc tgccgcacgc ggctctcca ggcgctgggc 600  
 cgtccggtga cgctggccgt gaccatcgac cccgccctca tcgccggcct ggagatggag 660  
 acgccccacg cgggtggtgc caattccttc aaggccgac tcgaccgcgt caccgcggcg 720  
 ctcacccatc atgggacctg a 741

<210> 23  
 <211> 246  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 23

Met Gln Ile Asp Trp Trp Thr Leu Gly Leu Gln Thr Val Asn Val Leu  
1 5 10 15

Val Leu Ile Trp Leu Leu Ser Arg Phe Leu Phe Lys Pro Val Ala Gln  
20 25 30

Val Ile Ala Gln Arg Arg Ala Glu Ile Glu Lys Leu Val Glu Asp Ala  
35 40 45

Arg Ala Ala Lys Ala Ala Ala Glu Ala Glu Arg Asp Thr Ala Lys Ala  
50 55 60

Glu Glu Ala Arg Leu Ala Ala Glu Arg Gly Ala Arg Met Ala Ala Val  
65 70 75 80

Ala Lys Glu Ala Glu Ala Gln Lys Ala Ala Leu Leu Ala Ala Ala Lys  
85 90 95

Thr Glu Ala Glu Ala Leu His Ala Ala Ala Glu Ala Ala Ile Val Arg  
100 105 110

Ala Arg Ala Ser Glu Glu Glu Ala Ala Ala Asp Arg Ala Ser Arg Leu  
30

|                                                                 |     |     |
|-----------------------------------------------------------------|-----|-----|
| 115                                                             | 120 | 125 |
| Ala Val Asp Ile Ala Ala Lys Leu Leu Asp Arg Leu Pro Asp Asp Ala |     |     |
| 130                                                             | 135 | 140 |
| Arg Val Ala Gly Phe Ile Asp Gly Leu Ala Glu Gly Leu Glu Ala Leu |     |     |
| 145                                                             | 150 | 155 |
| Pro Glu Ala Ser Arg Ala Val Ile Gly Val Asp Gly Ala Pro Val Arg |     |     |
|                                                                 | 165 | 170 |
| Val Thr Ala Ala Arg Ala Leu Met Pro Ala Glu Glu Glu Ala Cys Arg |     |     |
|                                                                 | 180 | 185 |
| Thr Arg Leu Ser Gln Ala Leu Gly Arg Pro Val Thr Leu Ala Val Thr |     |     |
|                                                                 | 195 | 200 |
| Ile Asp Pro Ala Leu Ile Ala Gly Leu Glu Met Glu Thr Pro His Ala |     |     |
|                                                                 | 210 | 215 |
| Val Val Arg Asn Ser Phe Lys Ala Asp Leu Asp Arg Val Thr Ala Ala |     |     |
|                                                                 | 225 | 230 |
| Leu Thr His His Gly Thr                                         |     |     |
|                                                                 | 245 |     |

&lt;210&gt; 24

&lt;211&gt; 243

&lt;212&gt; ADN

&lt;213&gt; Chưa biết

&lt;220&gt;

&lt;223&gt; chủng mới

&lt;400&gt; 24

atgactgtcg agatgggtcag catcttcgcg gcggcgctcg ccgtctcctt cggcgccatc 60

gggcccggccc tgggcgaggg ccgggcggtg gccgcggcca tggacgcat cgcccggcag 120

ccggaggcgg ccggaacctt gtcgcgcacg ctcttcgtcg gcctcgccat gatcgagacc 180

atggcgatct actgcctggt gatcgcgctc ctggtgctct tcgccaatcc gttcgtgaag 240

tga 243

&lt;210&gt; 25

<211> 80  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 25

Met Thr Val Glu Met Val Ser Ile Phe Ala Ala Ala Leu Ala Val Ser  
 1 5 10 15

Phe Gly Ala Ile Gly Pro Ala Leu Gly Glu Gly Arg Ala Val Ala Ala  
 20 25 30

Ala Met Asp Ala Ile Ala Arg Gln Pro Glu Ala Ala Gly Thr Leu Ser  
 35 40 45

Arg Thr Leu Phe Val Gly Leu Ala Met Ile Glu Thr Met Ala Ile Tyr  
 50 55 60

Cys Leu Val Ile Ala Leu Leu Val Leu Phe Ala Asn Pro Phe Val Lys  
 65 70 75 80

<210> 26  
 <211> 699  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 26

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| atgggctcgc cgctgatacct cgaacccctg ttccatatcg ggcccgtgcc catcaccgcg | 60  |
| ccggtgggtgg tcacctgggt catcatggcc gccttcattg ggctggcgcg gctcatcacc | 120 |
| cgggaagcttt ccaccgatcc caccgggacc caggcggcgg tggaaacggt gctgaccgcc | 180 |
| atcgattccc agatcgccga caccatgcag gccgatcccg cgccttatcg cgcgctcatc  | 240 |
| ggcaccatct tcctttatgt gctgggtggcc aactggtcct cgctcatccc gggcatcgag | 300 |
| ccgcccacgg cgcatatcga gaccgatgcg gcgctcgctt tcatcgtgtt cgccgccacc  | 360 |
| atcgggttcg gggtgaagac aaggggtgtg aagggtatc tcgccacctt cgccgaaccc   | 420 |
| tcctgggtga tgatcccgt caatgtggtg gagcagatca cccggacctt ctcgctcatc   | 480 |
| gtg'gcctgt tcggcaacat catgagcggg gtgttcgtgg tcggcatcat cctgtccctc  | 540 |

gccgggctgc tgggtgcccacat cccctcatg gcgctcgatc tcctgaccgg cgccgtgcag 600  
gcctacatct tcgcggtgct ggcctgcgtg ttcacggcg cggccattgg cgaggcgccg 660  
gcaaagcccc aatcgaagga gccagggaaa acatcatga 699

<210> 27  
<211> 232  
<212> PRT  
<213> Chưa biết

<220>  
<223> chủng mới

<400> 27

Met Gly Ser Pro Leu Ile Leu Glu Pro Leu Phe His Ile Gly Pro Val  
1 5 10 15

Pro Ile Thr Ala Pro Val Val Val Thr Trp Leu Ile Met Ala Ala Phe  
20 25 30

Ile Gly Leu Ala Arg Leu Ile Thr Arg Lys Leu Ser Thr Asp Pro Thr  
35 40 45

Arg Thr Gln Ala Ala Val Glu Thr Val Leu Thr Ala Ile Asp Ser Gln  
50 55 60

Ile Ala Asp Thr Met Gln Ala Asp Pro Ala Pro Tyr Arg Ala Leu Ile  
65 70 75 80

Gly Thr Ile Phe Leu Tyr Val Leu Val Ala Asn Trp Ser Ser Leu Ile  
85 90 95

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

Ala Phe Ile Val Phe Ala Ala Thr Ile Gly Phe Gly Leu Lys Thr Arg  
115 120 125

Gly Val Lys Gly Tyr Leu Ala Thr Phe Ala Glu Pro Ser Trp Val Met  
130 135 140

Ile Pro Leu Asn Val Val Glu Gln Ile Thr Arg Thr Phe Ser Leu Ile  
145 150 155 160

Val Arg Leu Phe Gly Asn Ile Met Ser Gly Val Phe Val Val Gly Ile  
 165 170 175

Ile Leu Ser Leu Ala Gly Leu Leu Val Pro Ile Pro Leu Met Ala Leu  
 180 185 190

Asp Leu Leu Thr Gly Ala Val Gln Ala Tyr Ile Phe Ala Val Leu Ala  
 195 200 205

Cys Val Phe Ile Gly Ala Ala Ile Gly Glu Ala Pro Ala Lys Pro Gln  
 210 215 220

Ser Lys Glu Pro Gly Lys Thr Ser  
 225 230

<210> 28  
 <211> 456  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 28  
 gtgagcgcgc cgctgcacct caccatcacc acgccggccg ccgttctggt ggaccgtgcc 60  
 gacatcgtgg ccctgcgtgc cgaggacgag agcggcagct tcggcatcct gcccggccat 120  
 gcggatttcc tgaccgttct ggaggcctgc gtggtgctgc tcaaggatgg ggccgacggc 180  
 gtgcattatt gtgctctcag tgggtggcgtg ctgtcggtcg aggagggccg gcgcatcgcc 240  
 atgcctgcc gtcagggcac ggtgagcgac gacctggtcg ccctggaagg ggcggtggac 300  
 gccatgcgtt cggcggagag cgatgccgac aagcgggccc gggaggagca gatgcgcctt 360  
 catgcccacg ccgtgcgcca gtcctgcac tatctgcggc ccggccgggc cggcggcgtg 420  
 gcgccggccg ccgcgccgga ggaggggccc tcatga 456

<210> 29  
 <211> 151  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 29

Met Ser Ala Pro Leu His Leu Thr Ile Thr Thr Pro Ala Ala Val Leu  
1 5 10 15

Val Asp Arg Ala Asp Ile Val Ala Leu Arg Ala Glu Asp Glu Ser Gly  
20 25 30

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

Ala Cys Val Val Arg Phe Lys Asp Gly Ala Asp Gly Val His Tyr Cys  
50 55 60

Ala Leu Ser Gly Gly Val Leu Ser Val Glu Glu Gly Arg Arg Ile Ala  
65 70 75 80

Ile Ala Cys Arg Gln Gly Thr Val Ser Asp Asp Leu Val Ala Leu Glu  
85 90 95

Gly Ala Val Asp Ala Met Arg Ser Ala Glu Ser Asp Ala Asp Lys Arg  
100 105 110

Ala Arg Val Glu Gln Met Arg Leu His Ala His Ala Val Arg Gln Leu  
115 120 125

Leu His Tyr Leu Arg Pro Gly Arg Ala Gly Gly Val Ala Pro Ala Ala  
130 135 140

Ala Pro Glu Glu Gly Pro Ser  
145 150

<210> 30

<211> 1494

<212> ADN

<213> Chưa biết

<220>

<223> chủng mới

<400> 30

atggcagcgg cagatgagga ggcgcaatcg gccgccggcc ccgcctcggg ccgggtggtg 60

gccgtgcgcg gcgcggtgat cgacatcgcc ttgcccagc ctccgctgcc gccgctggac 120

gagccccttc tcatcaccga cggccggggc ggcacggtgc tgggtggaggt gcagagccat 180

atggatcggc acacggtgcg cgccatcgcc cttcaggcca ccaccggcct cagccggggg 240  
 ctggaggcgg cgcggtggg cgggccggtg aaggtgccgg tgggagacca tgtgctcggc 300  
 cgcctcctgg atgtcaccgg cgccatcggc gacaagggcg ggccgctgcc ggccgacgtg 360  
 cccacgcggc cgatccacca cgcgccgcca tccttcgccg cgcagggcgg cacgtccgat 420  
 ctgtttcgca ccggcatcaa ggtcatcgac ctctggcgcc ccctcgcca gggcggcaag 480  
 gcggccatgt tcggcggggc cggcgtgggc aagaccgtgc tggatgatgga gctgatccac 540  
 gccatggtgg cgagctacaa gggcatctcg gtgtttgccg gcgtggggga gcgctcccgc 600  
 gagggccacg agatgctgct ggacatgacc gattccggcg tgctcgaccg caccgttctg 660  
 gtctatggcc agatgaacga gcccccggg gcccgctggc gggtgcccac gacggcgctg 720  
 accatcgccg aatatttccg cgacgagaag caccagaacg tcctgctgct gatggacaac 780  
 atcttccgct tcgtccaggc gggggcggag gtctccggcc ttttgggccg tccgccctcc 840  
 cgggtgggat accagccgac gctggcgagc gaggtggcgg cgctccagga acgcatcacc 900  
 tccgtgggcg aggcctcggt gaccgccatc gaggcggtct acgtgccggc ggatgacttc 960  
 accgatcccg ccgtgaccac catcgccgcc cacgtggatt ccatggtggt gctctcccgc 1020  
 gccatggcgg cggagggcac gtatccggcg gtggaccca tctcctctc gtcggtgctg 1080  
 ctcgaccgc tcatcgtggg ggacgagcat gcgcgcgtcg ccaacgaggt gcgccggacc 1140  
 atcgagcatt atcgcgagct tcaggatgtg atctcgctgc tgggcatgga ggaattgggc 1200  
 accgaggatc gccgcatcgt ggagcgggcg cgccggctcc agcgcttcct caccagccc 1260  
 ttcacggtca ccgaggcctt caccggcgtg cccggccgct cgggtggccat cgccgacacc 1320  
 atcgccggct gcaggatgat cctgtccggc gcctgcgacg actggcagga aagcgccctc 1380  
 tacatggtgg gcaccatcga cgaggcccgc cagaaggagg aggccgctcg cgccaaggcg 1440  
 gggcagggcg ccccgccgg gacggcagcc gagacggcgg aggccgcccc gtga 1494

<210> 31  
 <211> 497  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 31

Met Ala Ala Ala Asp Glu Glu Ala Gln Ser Ala Ala Gly Pro Ala Ser

|     |     |     |     |
|-----|-----|-----|-----|
| 1   | 5   | 10  | 15  |
| Gly | Arg | Val | Val |
| 20  | Ala | Val | Arg |
| Gly | Ala | Val | Ile |
| 25  | Asp | Ile | Ala |
| 30  | Phe | Ala |     |
| Gln | Pro | Pro | Leu |
| 35  | Pro | Pro | Leu |
| 40  | Asp | Asp | Ala |
| 45  | Leu | Leu | Ile |
| Thr | Asp | Gly |     |
| Arg | Gly | Gly | Thr |
| 50  | Val | Leu | Val |
| 55  | Glu | Val | Gln |
| 60  | Ser | His | Met |
| Asp | Arg | His |     |
| Thr | Val | Arg | Ala |
| 65  | Ile | Ala | Leu |
| 70  | Gln | Ala | Thr |
| 75  | Thr | Gly | Leu |
| 80  | Ser | Arg | Gly |
| Leu | Glu | Ala | Ala |
| 85  | Arg | Val | Gly |
| 90  | Gly | Pro | Val |
| 95  | Lys | Val | Pro |
| Val | Gly | Asp |     |
| His | Val | Leu | Gly |
| 100 | Arg | Leu | Leu |
| 105 | Asp | Val | Thr |
| 110 | Gly | Ala | Ile |
| Gly | Asp | Lys |     |
| Gly | Gly | Pro | Leu |
| 115 | Pro | Ala | Asp |
| 120 | Val | Pro | Thr |
| 125 | Arg | Pro | Ile |
| His | His | Ala |     |
| Pro | Pro | Ser | Phe |
| 130 | Ala | Ala | Gln |
| 135 | Gly | Gly | Thr |
| 140 | Ser | Asp | Leu |
| Phe | Arg | Thr |     |
| Gly | Ile | Lys | Val |
| 145 | Ile | Asp | Leu |
| 150 | Leu | Ala | Pro |
| 155 | Leu | Ala | Gln |
| 160 | Gly | Gly | Lys |
| Ala | Ala | Met | Phe |
| 165 | Gly | Gly | Ala |
| 170 | Gly | Val | Gly |
| 175 | Lys | Thr | Val |
| Leu | Val | Met |     |
| Glu | Leu | Ile | His |
| 180 | Ala | Met | Val |
| 185 | Ala | Ser | Tyr |
| 190 | Lys | Gly | Ile |
| Ser | Val | Phe |     |
| Ala | Gly | Val | Gly |
| 195 | Glu | Arg | Ser |
| 200 | Arg | Glu | Gly |
| 205 | His | Glu | Met |
| Leu | Leu | Asp |     |
| Met | Thr | Asp | Ser |
| 210 | Gly | Val | Leu |
| 215 | Asp | Arg | Thr |
| 220 | Val | Leu | Val |
| Tyr | Gly | Gln |     |
| Met | Asn | Glu | Pro |
| 37  | Pro | Gly | Ala |
| Arg | Trp | Arg | Val |
| Pro | Met | Thr | Ala |
| Leu |     |     |     |

38

|                                                                 |     |         |
|-----------------------------------------------------------------|-----|---------|
| 450                                                             | 455 | 460     |
| Thr Ile Asp Glu Ala Arg Gln Lys Glu Glu Ala Ala Arg Ala Lys Ala |     |         |
| 465                                                             | 470 | 475 480 |
| Gly Gln Gly Ala Pro Ala Gly Thr Ala Ala Glu Thr Ala Glu Ala Ala |     |         |
|                                                                 | 485 | 490 495 |

Pro

<210> 32  
 <211> 1437  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 32  
 atggcgaaca aggtcggacg catcaccacg atcatcggcg ccgtcgtcga cgtgcagttc 60  
 gacgggcatc tgccggcgat tctcaacgcg atcgagacca ccaaccaggg caaccggctg 120  
 gtgctcgaag tggctcagca tctcggcgag aacaccgtgc gctgcatcgc catggatgcc 180  
 actgaaggcc tgggtgcgtgg ccaggaggtg gccgacaccg atgcgccccat ccaggtgccc 240  
 gtgggcgccc ccaccctcgg ccgcatcatg aacgtgatcg gcgagccggt ggacgagctg 300  
 ggccccatcg agggcgaagc gctgcgcggc atccatcagc cggccccctc ctatgcggag 360  
 caggccacgg aagctgagat cctcgtcacc ggcatcaagg tgggtggatct gctggcgccc 420  
 tattccaagg gcggcaaggt gggcctgttc ggcggcgccc gcgtgggcaa gaccgtgctc 480  
 atcatggagc tgatcaacaa cgtggccaag gcgcacggcg gctattccgt gttcgccggc 540  
 gtgggtgagc gcacccgcga gggcaacgac ctctaccacg agatgatcga gtccaacgtg 600  
 aacaaggacc cgcacgagaa caatggctcg gcggccggtt ccaagtgcgc cctgggtctat 660  
 ggccagatga acgagccgcc cggcgcccgcc gcccgcggtg ccctcaccgg cctcaccgtc 720  
 gccgagcatt tccgcgacca gggccaggac gtgctgttct tcgtggacaa catcttccgc 780  
 ttcacccagg cgggctccga ggtgtcggcg cttctcggcc gcaccccctc ggcgggtgggc 840  
 taccagccga cgctggccac cgacatgggc cagctgcagg agcgcacac caccaccacc 900  
 aagggtcca tcacctcggg gcaggccatc tacgtgccgg cggacgatct gaccgatccg 960

gcgccggccg cctccttcgc ccatctggac gccaccacgg tgctgtcgcg ctccatcgcg 1020  
 gagaagggca tctacccggc ggtggatccg ctggactcca cctcgcgcat gctgtctccc 1080  
 gccatcctcg gcgacgagca ctacaacacc gcgcgccagg tgcagcagac cctgcagcgc 1140  
 tacaaggcgc tccaggacat catcgccatc ctgggcatgg acgaactctc cgaagaggac 1200  
 aagctcaccg tggcccgcgc ccgcaagatc gagcgcttcc tctcccagcc cttccacgtg 1260  
 gccgaggtgt tcaccgggtc gcccggaag ctggtcgacc tcgccgacac catcaagggc 1320  
 ttcaagggcc tgggtggacgg caagtacgac tacctgcccg agcaggcctt ctacatggtg 1380  
 ggccaccatcg aagaagccat cgagaagggc aagaagctgg cggccgaggc ggcctga 1437

<210> 33  
 <211> 478  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 33

Met Ala Asn Lys Val Gly Arg Ile Thr Gln Ile Ile Gly Ala Val Val  
1 5 10 15

Asp Val Gln Phe Asp Gly His Leu Pro Ala Ile Leu Asn Ala Ile Glu  
20 25 30

Thr Thr Asn Gln Gly Asn Arg Leu Val Leu Glu Val Ala Gln His Leu  
35 40 45

Gly Glu Asn Thr Val Arg Cys Ile Ala Met Asp Ala Thr Glu Gly Leu  
50 55 60

Val Arg Gly Gln Glu Val Ala Asp Thr Asp Ala Pro Ile Gln Val Pro  
65 70 75 80

Val Gly Ala Ala Thr Leu Gly Arg Ile Met Asn Val Ile Gly Glu Pro  
85 90 95

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

Gln Pro Ala Pro Ser Tyr Ala Glu Gln Ala Thr Glu Ala Glu Ile Leu  
40

|                                                                 |     |     |
|-----------------------------------------------------------------|-----|-----|
| 115                                                             | 120 | 125 |
| Val Thr Gly Ile Lys Val Val Asp Leu Leu Ala Pro Tyr Ser Lys Gly |     |     |
| 130                                                             | 135 | 140 |
| Gly Lys Val Gly Leu Phe Gly Gly Ala Gly Val Gly Lys Thr Val Leu |     |     |
| 145                                                             | 150 | 155 |
| Ile Met Glu Leu Ile Asn Asn Val Ala Lys Ala His Gly Gly Tyr Ser |     |     |
|                                                                 | 165 | 170 |
| Val Phe Ala Gly Val Gly Glu Arg Thr Arg Glu Gly Asn Asp Leu Tyr |     |     |
|                                                                 | 180 | 185 |
| His Glu Met Ile Glu Ser Asn Val Asn Lys Asp Pro His Glu Asn Asn |     |     |
|                                                                 | 195 | 200 |
| Gly Ser Ala Ala Gly Ser Lys Cys Ala Leu Val Tyr Gly Gln Met Asn |     |     |
|                                                                 | 210 | 215 |
| Glu Pro Pro Gly Ala Arg Ala Arg Val Ala Leu Thr Gly Leu Thr Val |     |     |
|                                                                 | 225 | 230 |
| Ala Glu His Phe Arg Asp Gln Gly Gln Asp Val Leu Phe Phe Val Asp |     |     |
|                                                                 | 245 | 250 |
| Asn Ile Phe Arg Phe Thr Gln Ala Gly Ser Glu Val Ser Ala Leu Leu |     |     |
|                                                                 | 260 | 265 |
| Gly Arg Ile Pro Ser Ala Val Gly Tyr Gln Pro Thr Leu Ala Thr Asp |     |     |
|                                                                 | 275 | 280 |
| Met Gly Gln Leu Gln Glu Arg Ile Thr Thr Thr Thr Lys Gly Ser Ile |     |     |
|                                                                 | 290 | 295 |
| Thr Ser Val Gln Ala Ile Tyr Val Pro Ala Asp Asp Leu Thr Asp Pro |     |     |
|                                                                 | 305 | 310 |
| Ala Pro Ala Ala Ser Phe Ala His Leu Asp Ala Thr Thr Val Leu Ser |     |     |
|                                                                 | 325 | 330 |
| Arg Ser Ile Ala Glu Lys Gly Ile Tyr Pro Ala Val Asp Pro Leu Asp |     |     |

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

<210> 34  
 <211> 876  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 34  
 atggcgagtc tgaaggacct gagaaaccgc attgcctcgg tgaaggcgac gcagaagatc 60  
 accaaggcga tgcagatggc cgccgcggcg aagctgcgtc gcgcccaggc ggcggctgaa 120  
 gcggcccgct cctatgcgga acgcatggag acggtgctcg gaaatcttgc ctccggcatg 180  
 gtggtgggcg cgcaggcgcc tgttctcatg accgggacgg gcaagagcga caccacctg 240  
 ctgctggtgt gcaccggcga gcgcggcctg tgcggcgctt tcaactcgtc catcgtgcgc 300  
 ttcgcccgcg agcgggcgca gctgctgctg gccgagggca agaaggtgaa aatcctgtgc 360

gtgggccgca agggccacga gcagctgcgc cgcattctacc cggacaacat catcgacgtg 420  
 gtggacctgc gcgcggtgcg caacatcggc ttcaaggagg cgcacgccat cgcccgaag 480  
 gtgctggccc tgctcgatga aggcgcattc gacgtctgca cgctcttcta ctcccacttc 540  
 aggagcgtga tcgcccaggt gccgacggcc cagcagctca ttccggccac cttcgacgag 600  
 cggccggccg tcgccgatgc gccggtctat gaatatgagc cggaggagga ggagatcctc 660  
 gccgagctgc tgccgcgcaa cgtggcggtg cagatcttca aggccctcct cgagaaccag 720  
 gcttctttct atggctccca gatgagcgcc atggacaacg ccacgcgcaa tgcgggagag 780  
 atgatcaaga agcagacgtt cacctacaac cgtacccgcc aggccatgat cacgaaggaa 840  
 ctcacgcaga tcattctcgg cgccgaggcc gtctga 876

<210> 35  
 <211> 291  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 35

Met Ala Ser Leu Lys Asp Leu Arg Asn Arg Ile Ala Ser Val Lys Ala  
1 5 10 15

Thr Gln Lys Ile Thr Lys Ala Met Gln Met Val Ala Ala Lys Leu  
20 25 30

Arg Arg Ala Gln Ala Ala Ala Glu Ala Ala Arg Pro Tyr Ala Glu Arg  
35 40 45

Met Glu Thr Val Leu Gly Asn Leu Ala Ser Gly Met Val Val Gly Ala  
50 55 60

Gln Ala Pro Val Leu Met Thr Gly Thr Gly Lys Ser Asp Thr His Leu  
65 70 75 80

Leu Leu Val Cys Thr Gly Glu Arg Gly Leu Cys Gly Ala Phe Asn Ser  
85 90 95

Ser Ile Val Arg Phe Ala Arg Glu Arg Ala Gln Leu Leu Leu Ala Glu  
100 105 110

Gly Lys Lys Val Lys Ile Leu Cys Val Gly Arg Lys Gly His Glu Gln  
 115 120 125

Leu Arg Arg Ile Tyr Pro Asp Asn Ile Ile Asp Val Val Asp Leu Arg  
 130 135 140

Ala Val Arg Asn Ile Gly Phe Lys Glu Ala Asp Ala Ile Ala Arg Lys  
 145 150 155 160

Val Leu Ala Leu Leu Asp Glu Gly Ala Phe Asp Val Cys Thr Leu Phe  
 165 170 175

Tyr Ser His Phe Arg Ser Val Ile Ala Gln Val Pro Thr Ala Gln Gln  
 180 185 190

Leu Ile Pro Ala Thr Phe Asp Glu Arg Pro Ala Val Ala Asp Ala Pro  
 195 200 205

Val Tyr Glu Tyr Glu Pro Glu Glu Glu Glu Ile Leu Ala Glu Leu Leu  
 210 215 220

Pro Arg Asn Val Ala Val Gln Ile Phe Lys Ala Leu Leu Glu Asn Gln  
 225 230 235 240

Ala Ser Phe Tyr Gly Ser Gln Met Ser Ala Met Asp Asn Ala Thr Arg  
 245 250 255

Asn Ala Gly Glu Met Ile Lys Lys Gln Thr Leu Thr Tyr Asn Arg Thr  
 260 265 270

Arg Gln Ala Met Ile Thr Lys Glu Leu Ile Glu Ile Ile Ser Gly Ala  
 275 280 285

Glu Ala Val  
 290

<210> 36  
 <211> 1530  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

&lt;400&gt; 36

|                                                                    |      |
|--------------------------------------------------------------------|------|
| atggacattc gagccgctga aatctctgcc atcctgaaag agcagatcca gaatttcggc  | 60   |
| caggaggcgg aagtctccga ggtgggtcag gttctgtccg tgggtgacgg catcgcgcg   | 120  |
| gtctacggcc tcgacaacgt ccaggcgggc gagatggtcg agttcgagaa cggcacgcg   | 180  |
| ggcatggcg tgaacctcga gctcgacaat gtcggcatcg tgatcttcgg ttccgaccgc   | 240  |
| gagatcaagg aaggccagac cgtcaagcgg accggcgcca tcgtggacgc ccccgtcggc  | 300  |
| aagggcctgc tcggccgctg cgtggacgct ctcggcaacc cgatcgacgg caagggccc   | 360  |
| atcatgttca ccgagcgtcg ccgggtcgac gtgaaggcgc cgggcatcat cccgcgaag   | 420  |
| tcggtgcacg agcccatgca gaccggcctg aaggccatcg atgcgctcat ccccatcggc  | 480  |
| cgcggccagc gcgagctcat catcggcgac cgccagaccg gcaagaccgc cgtggcgctc  | 540  |
| gactcgatcc tgaaccagaa gcccatcaac caggcgacg acgagaaggc caagctctac   | 600  |
| tgcgtctatg tcgcggtggg ccagaagcgt tccactgtcg cgcagttcgt gaaggtgctc  | 660  |
| gaggagcacg gcgcgctgga atattccatc gtcgtcgccg ccaccgcctc ggacgcggcc  | 720  |
| cccatgcagt tcctggcgcc gttcaccggc accgccatgg gcgagtatct ccgcgacaac  | 780  |
| ggcatgcacg ccctcatcat ccatgatgac ctgtccaagc aggcctggc ctaccgccag   | 840  |
| atgtcgctgc tgctgcgccg cccgccgggc cgcgaggcct atcccggcga tgtgttctac  | 900  |
| ctgcactccc gcctcttgga gcgcgccgcc aagctcaatg acgagcacgg cgccggctcg  | 960  |
| ctgaccgccc tgccggtgat cgagaccag gccaacgacg tgtcggccta catcccgacc   | 1020 |
| aacgtgatct ccatcaccga cggtcagatc ttccttgaat ccgatctgtt ctaccagggc  | 1080 |
| atccgcccgg cgggtgaacgt gggcctgtcg gtgtcgcgcg tgggctcttc ggcccagatc | 1140 |
| aaggcgatga agcaggtggc cggcaagatc aaggcgagc tcgcccagta tcgagagctg   | 1200 |
| gcggccttcg ccagttcgg ttccggacctg gacgcggcca cccagaagct gctgaaccgc  | 1260 |
| ggcgcccgcc tcaccgagct gctgaagcag agccagttct cgccctcaa ggtggaggag   | 1320 |
| caggtggcgg tgatctatgc cggcaccaat ggctatctcg atccgctgcc ggtctccaag  | 1380 |
| gtgcgcgagt tcgagcaggg tctgctcctg tcgctgcgct cgcagcatcc ggagatcctg  | 1440 |
| gacgccatcc gcacgtccaa ggagctttcc aaggacaccg ccgagaagct gacgaaggcc  | 1500 |
| atcgacgcct tcgccaagag cttctcctga                                   | 1530 |

&lt;210&gt; 37

<211> 509  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 37

Met Asp Ile Arg Ala Ala Glu Ile Ser Ala Ile Leu Lys Glu Gln Ile  
 1 5 10 15

Gln Asn Phe Gly Gln Glu Ala Glu Val Ser Glu Val Gly Gln Val Leu  
 20 25 30

Ser Val Gly Asp Gly Ile Ala Arg Val Tyr Gly Leu Asp Asn Val Gln  
 35 40 45

Ala Gly Glu Met Val Glu Phe Glu Asn Gly Thr Arg Gly Met Ala Leu  
 50 55 60

Asn Leu Glu Leu Asp Asn Val Gly Ile Val Ile Phe Gly Ser Asp Arg  
 65 70 75 80

Glu Ile Lys Glu Gly Gln Thr Val Lys Arg Thr Gly Ala Ile Val Asp  
 85 90 95

Ala Pro Val Gly Lys Gly Leu Leu Gly Arg Val Val Asp Ala Leu Gly  
 100 105 110

Asn Pro Ile Asp Gly Lys Gly Pro Ile Met Phe Thr Glu Arg Arg Arg  
 115 120 125

Val Asp Val Lys Ala Pro Gly Ile Ile Pro Arg Lys Ser Val His Glu  
 130 135 140

Pro Met Gln Thr Gly Leu Lys Ala Ile Asp Ala Leu Ile Pro Ile Gly  
 145 150 155 160

Arg Gly Gln Arg Glu Leu Ile Ile Gly Asp Arg Gln Thr Gly Lys Thr  
 165 170 175

Ala Val Ala Leu Asp Ser Ile Leu Asn Gln Lys Pro Ile Asn Gln Gly  
 180 185 190

Asp Asp Glu Lys Ala Lys Leu Tyr Cys Val Tyr Val Ala Val Gly Gln  
 195 200 205

Lys Arg Ser Thr Val Ala Gln Phe Val Lys Val Leu Glu Glu His Gly  
 210 215 220

Ala Leu Glu Tyr Ser Ile Val Val Ala Ala Thr Ala Ser Asp Ala Ala  
 225 230 235 240

Pro Met Gln Phe Leu Ala Pro Phe Thr Gly Thr Ala Met Gly Glu Tyr  
 245 250 255

Phe Arg Asp Asn Gly Met His Ala Leu Ile Ile His Asp Asp Leu Ser  
 260 265 270

Lys Gln Ala Val Ala Tyr Arg Gln Met Ser Leu Leu Leu Arg Arg Pro  
 275 280 285

Pro Gly Arg Glu Ala Tyr Pro Gly Asp Val Phe Tyr Leu His Ser Arg  
 290 295 300

Leu Leu Glu Arg Ala Ala Lys Leu Asn Asp Glu His Gly Ala Gly Ser  
 305 310 315 320

Leu Thr Ala Leu Pro Val Ile Glu Thr Gln Ala Asn Asp Val Ser Ala  
 325 330 335

Tyr Ile Pro Thr Asn Val Ile Ser Ile Thr Asp Gly Gln Ile Phe Leu  
 340 345 350

Glu Ser Asp Leu Phe Tyr Gln Gly Ile Arg Pro Ala Val Asn Val Gly  
 355 360 365

Leu Ser Val Ser Arg Val Gly Ser Ser Ala Gln Ile Lys Ala Met Lys  
 370 375 380

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

Ala Ala Phe Ala Gln Phe Gly Ser Asp Leu Asp Ala Ala Thr Gln Lys  
 405 410 415

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

Phe Ser Pro Leu Lys Val Glu Glu Gln Val Ala Val Ile Tyr Ala Gly  
 435 440 445

Thr Asn Gly Tyr Leu Asp Pro Leu Pro Val Ser Lys Val Arg Glu Phe  
 450 455 460

Glu Gln Gly Leu Leu Leu Ser Leu Arg Ser Gln His Pro Glu Ile Leu  
 465 470 475 480

Asp Ala Ile Arg Thr Ser Lys Glu Leu Ser Lys Asp Thr Ala Glu Lys  
 485 490 495

Leu Thr Lys Ala Ile Asp Ala Phe Ala Lys Ser Phe Ser  
 500 505

<210> 38

<211> 555

<212> ADN

<213> Chưa biết

<220>

<223> chủng mới

<400> 38

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| gtggcggaag cgcgtggtc aggcattggcg ggacgctatg cgaccgcgct gttcgagctg  | 60  |
| gcggacgaag ccggtgccat cgattccgctc caggcggatc ttgatcgctt gtccggcctt | 120 |
| ctggccgaga gcgcggatct ggcgcggtg gtcaagagcc cggctctcac cgccgagcag   | 180 |
| cagctcggcg cgatggcggc cattctcgat caagcaggca tttccggcct tgcgggcaaa  | 240 |
| ttcgtgaagc tgggtggcgca gaaccgccgc ctgttcgcac tgccgcgcat gattgccgaa | 300 |
| tacgccgtcc tgggtggccc gaagaaggcg gagacctcg cgagcgtgac cgttgccacc   | 360 |
| cccctgagcg atgagcatct ggccacgctc aaggcggccc tggctgaaaa gaccggcaag  | 420 |
| gacgtgaagc tcgacgtcac cgtcgatccg tccatcctcg gtggtctcat cgtgaagctc  | 480 |
| ggctcgcgca tggtcgatgc ttccctgaag accaaactca attctatccg gcatgcgatg  | 540 |
| aaagaggtcc gctga                                                   | 555 |

<210> 39

<211> 184  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 39

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

Leu Phe Glu Leu Ala Asp Glu Ala Gly Ala Ile Asp Ser Val Gln Ala  
 20 25 30

Asp Leu Asp Arg Leu Ser Gly Leu Leu Ala Glu Ser Ala Asp Leu Ala  
 35 40 45

Arg Leu Val Lys Ser Pro Val Phe Thr Ala Glu Gln Gln Leu Gly Ala  
 50 55 60

Met Ala Ala Ile Leu Asp Gln Ala Gly Ile Ser Gly Leu Ala Gly Lys  
 65 70 75 80

Phe Val Lys Leu Val Ala Gln Asn Arg Arg Leu Phe Ala Leu Pro Arg  
 85 90 95

Met Ile Ala Glu Tyr Ala Val Leu Val Ala Arg Lys Lys Gly Glu Thr  
 100 105 110

Ser Ala Ser Val Thr Val Ala Thr Pro Leu Ser Asp Glu His Leu Ala  
 115 120 125

Thr Leu Lys Ala Ala Leu Ala Glu Lys Thr Gly Lys Asp Val Lys Leu  
 130 135 140

Asp Val Thr Val Asp Pro Ser Ile Leu Gly Gly Leu Ile Val Lys Leu  
 145 150 155 160

Gly Ser Arg Met Val Asp Ala Ser Leu Lys Thr Lys Leu Asn Ser Ile  
 165 170 175

Arg His Ala Met Lys Glu Val Arg  
 180

<210> 40  
 <211> 492  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 40  
 atgaccgaaa tggaactggc tgagctctgg gtcgccatcg ccttcctggt tttcgtaggc 60  
 ctctgatct atgcgggcg ccaccgcgc atcgtctccg ccctggattc ccgcggctcg 120  
 cgcacgcct cggaactgga ggaggcccgt cggctcaagg aagaggcca gaagctggtg 180  
 gccgaattca agcgcaagca gcgcgaggcc gaggccgagg ccgaatccat cgtcaccggc 240  
 gccaaaggccg aggccgagcg cctcgccgcc gaggccaagg cgaagatcga ggatttcgtc 300  
 acccgccgca ccaagatggc cgaggacaag atcgcccagg ccgagcatca ggctctggcg 360  
 gacgtgaagt ccatcgccgc cgaggcggcg gccaaaggcg ccgaggtgat cctcggcgcc 420  
 caggccaccg gcgcggtggc ggagcgtctg ctgtcgggcg ccatctccga ggtcaagacc 480  
 aagctcaact ga 492

<210> 41  
 <211> 163  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 41

Met Thr Glu Met Glu Leu Ala Glu Leu Trp Val Ala Ile Ala Phe Leu  
 1 5 10 15

Val Phe Val Gly Leu Leu Ile Tyr Ala Gly Ala His Arg Ala Ile Val  
 20 25 30

Ser Ala Leu Asp Ser Arg Gly Ser Arg Ile Ala Ser Glu Leu Glu Glu  
 35 40 45

Ala Arg Arg Leu Lys Glu Glu Ala Gln Lys Leu Val Ala Glu Phe Lys  
 50 55 60

Arg Lys Gln Arg Glu Ala Glu Ala Glu Ala Glu Ser Ile Val Thr Gly  
 50

|                 |                 |             |             |             |  |    |
|-----------------|-----------------|-------------|-------------|-------------|--|----|
| 65              |                 | 70          |             | 75          |  | 80 |
| Ala Lys Ala Glu | Ala Glu Arg Leu | Ala Ala Glu | Ala Lys Ala | Lys Ile     |  |    |
|                 | 85              |             | 90          | 95          |  |    |
| Glu Asp Phe Val | Thr Arg Arg Thr | Lys Met Ala | Glu Asp Lys | Ile Ala     |  |    |
|                 | 100             | 105         | 110         |             |  |    |
| Gln Ala Glu His | Gln Ala Leu     | Ala Asp Val | Lys Ser Ile | Ala Ala Glu |  |    |
|                 | 115             | 120         | 125         |             |  |    |
| Ala Ala Ala Lys | Ala Ala Glu     | Val Ile Leu | Gly Ala Gln | Ala Thr Gly |  |    |
|                 | 130             | 135         | 140         |             |  |    |
| Ala Val Ala Glu | Arg Leu Leu     | Ser Gly Ala | Ile Ser Glu | Val Lys Thr |  |    |
| 145             | 150             | 155         | 160         |             |  |    |
| Lys Leu Asn     |                 |             |             |             |  |    |

<210> 42  
 <211> 729  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 42  
 atgatgattg catggaagcg gaccttcgca gtcgtgacct tcggggccgc cctgatggcc 60  
 atgccgctcg cgggcgtggt cgcagctgag acttctcccg ctccggcggc agtggcgag 120  
 gccgatcatg cggtgcccac cgaggcggcc ggccagggca ccgccgatgc ggcccatgcc 180  
 gccgcgccgg gcgaggccgc ccatggtggc gcggccaagc acgaaacca tttcccgccc 240  
 ttcgacggca ccaccttcgc ctcccagttg ctgtggctcg ccgtcacctt cggcctgctt 300  
 tactacctca tgagcaaggt cacgctgccg cgcacgccc gcacacctga agagcgccac 360  
 gaccgcatcg ccgatgatct ggaggaagcc tccaagcatc gcgccgagag cgaggccgcc 420  
 cagcgggcct atgagaaggc gctgagcgag gccgcgcga aggccattc catcgccgcg 480  
 gaaacccgcg accgccttgc cgcccacgcc gacaccaacc gcaaggcgct ggagagcgag 540  
 ctcaccgcca agctgcaggc ggccgaggag cgcacgcca ccaccaagag cgaagccctc 600

acccatgtgc gcggcatcgc ggtggacgcc acccaatcca tcgtctccac cctcatcggt 660  
 gtcgcgccccg cggcggccga cgtggaaaaa gcggtggacg gcgccctgtc ccagcacggc 720  
 caggcctga 729

<210> 43  
 <211> 242  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chúng mới

<400> 43

Met Met Ile Ala Trp Lys Arg Thr Phe Ala Val Val Thr Phe Gly Ala  
 1 5 10 15

Ala Leu Met Ala Met Pro Val Ala Gly Val Val Ala Ala Glu Thr Ser  
 20 25 30

Pro Ala Pro Ala Ala Val Ala Gln Ala Asp His Ala Val Pro Thr Glu  
 35 40 45

Ala Ala Gly Gln Gly Thr Ala Asp Ala Ala His Ala Ala Ala Pro Gly  
 50 55 60

Glu Ala Ala His Gly Gly Ala Ala Lys His Glu Thr His Phe Pro Pro  
 65 70 75 80

Phe Asp Gly Thr Thr Phe Ala Ser Gln Leu Leu Trp Leu Ala Val Thr  
 85 90 95

Phe Gly Leu Leu Tyr Tyr Leu Met Ser Lys Val Thr Leu Pro Arg Ile  
 100 105 110

Gly Arg Ile Leu Glu Glu Arg His Asp Arg Ile Ala Asp Asp Leu Glu  
 115 120 125

Glu Ala Ser Lys His Arg Ala Glu Ser Glu Ala Ala Gln Arg Ala Tyr  
 130 135 140

Glu Lys Ala Leu Ser Glu Ala Arg Ala Lys Ala His Ser Ile Ala Ala  
 145 150 155 160

Glu Thr Arg Asp Arg Leu Ala Ala His Ala Asp Thr Asn Arg Lys Ala  
 165 170 175

Leu Glu Ser Glu Leu Thr Ala Lys Leu Gln Ala Ala Glu Glu Arg Ile  
 180 185 190

Ala Thr Thr Lys Ser Glu Ala Leu Thr His Val Arg Gly Ile Ala Val  
 195 200 205

Asp Ala Thr Gln Ser Ile Val Ser Thr Leu Ile Gly Val Ala Pro Ala  
 210 215 220

Ala Ala Asp Val Glu Lys Ala Val Asp Gly Ala Leu Ser Gln His Gly  
 225 230 235 240

Gln Ala

<210> 44  
 <211> 228  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 44  
 atggaagcgg aagctggaaa gttcatcggt gccggcctcg cctgcctcgg catgggtctc 60  
 gctggcgtcg gcgtcggtaa catcttcggt aacttcctct ccggcgccct gcgcaacccg 120  
 tccgctgccg acggccagtt cgcccgcgcc ttcacggcg ccgccctcgc ggaaggtctc 180  
 ggcatcttct cgctgggtcgt tgcgctcgtc ctgctgttcg tggcctga 228

<210> 45  
 <211> 75  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 45

Met Glu Ala Glu Ala Gly Lys Phe Ile Gly Ala Gly Leu Ala Cys Leu  
 1 5 10 15

Gly Met Gly Leu Ala Gly Val Gly Val Gly Asn Ile Phe Gly Asn Phe  
 20 25 30

Leu Ser Gly Ala Leu Arg Asn Pro Ser Ala Ala Asp Gly Gln Phe Ala  
 35 40 45

Arg Ala Phe Ile Gly Ala Ala Leu Ala Glu Gly Leu Gly Ile Phe Ser  
 50 55 60

Leu Val Val Ala Leu Val Leu Leu Phe Val Ala  
 65 70 75

<210> 46  
 <211> 753  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 46  
 atgaccgtcg atccgatcca ccagttcgag atcaagcgct acgtggatct gctgaacgtc 60  
 ggcggtgtcc agttctcctt caccaacgca acgggtgttca tgattggcat cgtcctgggtg 120  
 attttcttct tcctgacttt cgcgacacgc ggctgcaccc ttgtgccggg ccggatgcag 180  
 tcggcgggcg agctgagcta cgagttcatc gccaaagatgg tgcgcgacgc ggccggcagc 240  
 gagggaatgg tgttctttcc cttcgtcttc tcgctcttca tgttcgtgct ggtggcgaac 300  
 gtattggggc tcatccccta caccttcacg gtgaccgccc acctcatcgt caccgccgcc 360  
 ctggcgggcga cggatgacct caccgtcatc atctacggct tcgtgcggca cggcaccac 420  
 ttcctgcacc tgttcgtgcc gtcgggcgtg ccgggcttcc tcctgccctt cctcgtgggtg 480  
 atcgaggtgg tgctgttcct gtcgcggccc atcagcctct cgctgcgtct gttcgccaac 540  
 atgctggcgg gccacatcg cctcaagggtg ttgccttct tcgtcgtggg actggcctcg 600  
 gccggcgcgga tcggctggtt cggcgccacc ctgcccttct tcatgatcgt ggcgctcacc 660  
 gcgctggagc tgctgggtggc ggtgctgcag gcctacgtgt tcgcggtgct gacctcgatc 720  
 tacctcaacg acgcatcca tcccgccac tga 753

<210> 47  
 <211> 250  
 <212> PRT

<213> Chưa biết

<220>

<223> chủng mới

<400> 47

Met Thr Val Asp Pro Ile His Gln Phe Glu Ile Lys Arg Tyr Val Asp  
1 5 10 15

Leu Leu Asn Val Gly Gly Val Gln Phe Ser Phe Thr Asn Ala Thr Val  
20 25 30

Phe Met Ile Gly Ile Val Leu Val Ile Phe Phe Phe Leu Thr Phe Ala  
35 40 45

Thr Arg Gly Arg Thr Leu Val Pro Gly Arg Met Gln Ser Ala Ala Glu  
50 55 60

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

Glu Gly Met Val Phe Phe Pro Phe Val Phe Ser Leu Phe Met Phe Val  
85 90 95

Leu Val Ala Asn Val Leu Gly Leu Ile Pro Tyr Thr Phe Thr Val Thr  
100 105 110

Ala His Leu Ile Val Thr Ala Ala Leu Ala Ala Thr Val Ile Leu Thr  
115 120 125

Val Ile Ile Tyr Gly Phe Val Arg His Gly Thr His Phe Leu His Leu  
130 135 140

Phe Val Pro Ser Gly Val Pro Gly Phe Leu Leu Pro Phe Leu Val Val  
145 150 155 160

Ile Glu Val Val Ser Phe Leu Ser Arg Pro Ile Ser Leu Ser Leu Arg  
165 170 175

Leu Phe Ala Asn Met Leu Ala Gly His Ile Ala Leu Lys Val Phe Ala  
180 185 190

Phe Phe Val Val Gly Leu Ala Ser Ala Gly Ala Ile Gly Trp Phe Gly  
55

195

200

205

Ala Thr Leu Pro Phe Phe Met Ile Val Ala Leu Thr Ala Leu Glu Leu  
 210 215 220

Leu Val Ala Val Leu Gln Ala Tyr Val Phe Ala Val Leu Thr Ser Ile  
 225 230 235 240

Tyr Leu Asn Asp Ala Ile His Pro Gly His  
 245 250

<210> 48  
 <211> 390  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 48  
 atgtccgagc cgaatgatcc atcccgcagg gacggtgcga aggcgaaaga cgagacgcag 60  
 gactcccggc ccggtgaggc ggatcttgct cggcgcctcg atgcgctcgg cacctccatc 120  
 ggtcagggtca agtccagaag cggggagccc gcggcgacgc cgcgcaagga cacctcctcg 180  
 gcctccggcg cggccctggc gtttcggctg ggcgccgagt ttgtttcagg cgtgctgggtg 240  
 ggctcgctca tcggctacgg gttggattat gcgtttgcga tttcgccctg ggggctgata 300  
 gccttcacgc tgatcggctt tgccgccggc gtcctgaaca tgctgcgcgt ggcgaacagc 360  
 gatgccaaag gccacagcgc ggacaggtga 390

<210> 49  
 <211> 129  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 49

Met Ser Glu Pro Asn Asp Pro Ser Arg Arg Asp Gly Ala Lys Ala Lys  
 1 5 10 15

Asp Glu Thr Gln Asp Ser Arg Pro Gly Glu Ala Asp Leu Ala Arg Arg  
 20 25 30

Leu Asp Ala Leu Gly Thr Ser Ile Gly Gln Val Lys Ser Arg Ser Gly  
 35 40 45

Glu Pro Ala Ala Thr Pro Arg Lys Asp Thr Ser Ser Ala Ser Gly Ala  
 50 55 60

Ala Leu Ala Phe Arg Leu Gly Ala Glu Phe Val Ser Gly Val Leu Val  
 65 70 75 80

Gly Ser Leu Ile Gly Tyr Gly Leu Asp Tyr Ala Phe Ala Ile Ser Pro  
 85 90 95

Trp Gly Leu Ile Ala Phe Thr Leu Ile Gly Phe Ala Ala Gly Val Leu  
 100 105 110

Asn Met Leu Arg Val Ala Asn Ser Asp Ala Lys Arg His Ser Ala Asp  
 115 120 125

Arg

<210> 50  
 <211> 1929  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 50  
 atgagttcgc tctccgccac tattcaacag gtcttcaacg agccgggctg cgccaagaac 60  
 cagaataagt ccgaggcgga gaagaagaag ggctgcacca agcagctgca acccggcgga 120  
 gcggccggcg gctgcgcgtt cgacggcgcg aagatcgcg tccagccctt gaccgacgtc 180  
 gccacactgg tgcacggccc catgcctgc gaaggcaatt cctgggacaa tcgtggcgcc 240  
 aagtcctccg gctcgaacat ctggcgacc ggcttcacca cggacatcaa cgaaaccgac 300  
 gtggtgttcg gcggcgagaa gcgtctgttc aagtccatca aggaaatcat cgagaagtac 360  
 gaccgcggg ccgtcttcgt ctatcagacc tgcgtccccg ccatgatcgg cgacgacatc 420  
 gacgcgggtgt gcaaggcggc cagggagaag ttcggaaagc cggatgaccc gatcaattcc 480  
 cccggcttcg tggggccgaa gaatctcggc aacaagctcg ccggcgaggc gctcctcgac 540

catgtgatcg gcaccgagga gcccgattac acgacggcct acgacatcaa catcatcggc 600  
gaatacaatc tctccggcga gttgtggcag gtgaagccgc tgctggacga gctgggcatc 660  
cgcatcctcg cctgcatctc cggcgacggg aagtacaagg atgtggcgctc ctcccaccgc 720  
gccaaggcgg cgatgatggt gtgctccaag gccatgatca acgtggcccc caagatggag 780  
gagcgctacg acatccccctt cttcgaaggc tccttctacg gcatcgagga tagctccgat 840  
tccctgcgcg agattgcgcg catgctcatc gagaagggcg ccgatccgga gctgatggac 900  
cgcaccgagg cgctgattga gcgggaagag aagaaggcgt gggacgccat cgccgcctac 960  
aagccccgct tcaaggacaa gaagggtgtg ctcatcaccg gcggcgtaga atcctggctg 1020  
gtggtggcag cgctccagga agccggcctc gaactggtgg gcacctcggt gaagaagtcc 1080  
accaaggagg acaaggagcg catcaaggaa ctgatgggcc aggacgcca catgatcgac 1140  
gacatgacgc cccgcgaaat gtacaagatg ctgaaggacg ccaaggcgga catcatgctc 1200  
tcgggcgggc gctcgcaatt catcgcgctc aaggccgcca tgccctggct cgacatcaac 1260  
caggagcgcc accacgccta tatgggctat gtgggcatgg tgaagctggt cgaggagatc 1320  
gacaaggcgc tctacaatcc cgtgtgggaa cagggtgcga agcccgcccc gtgggaaaat 1380  
ccggaagaca cctggcaggc ccgtgcgctc gccgaaatgg aggcggaggc cgccgcgctc 1440  
gccgccgatc cgggtgcgcgc ggaagagggt cgccggtcca agaagatctg caattgcaag 1500  
agcgtcgacc tcggaaccat tgaggacgcc atcaaggctc acgcgctgac caccgtggag 1560  
ggtgtgcgag agcacaccaa tgcctcgga ggctgcggag cctgcagcgg gcggatcgag 1620  
gagatcttcg aggccgtggg cgttgtcgcc gccccgcctc ccgcggaggc cgccccgtct 1680  
ccgcaggaga tcgcgcccga tccgctcgct gcggaggaaa agcgccgcgc caagaaggcc 1740  
tgcggctgca aggaggtagc ggtcggcacc attgaggatg ccatccgcgc caagggtctg 1800  
cgaaacatcg cggagggtgcg tgcggccacc gatgccaaca ccggctgcgg caattgccag 1860  
gagcgggtgg agggcatcct cgaccgggtt ctcgccgagg cggcctcaga actccaggcg 1920  
gcggaatag 1929

<210> 51  
<211> 642  
<212> PRT  
<213> Chưa biết

<220>  
<223> chủng mới

&lt;400&gt; 51

Met Ser Ser Leu Ser Ala Thr Ile Gln Gln Val Phe Asn Glu Pro Gly  
 1 5 10 15

Cys Ala Lys Asn Gln Asn Lys Ser Glu Ala Glu Lys Lys Lys Gly Cys  
 20 25 30

Thr Lys Gln Leu Gln Pro Gly Gly Ala Ala Gly Gly Cys Ala Phe Asp  
 35 40 45

Gly Ala Lys Ile Ala Leu Gln Pro Leu Thr Asp Val Ala His Leu Val  
 50 55 60

His Gly Pro Ile Ala Cys Glu Gly Asn Ser Trp Asp Asn Arg Gly Ala  
 65 70 75 80

Lys Ser Ser Gly Ser Asn Ile Trp Arg Thr Gly Phe Thr Thr Asp Ile  
 85 90 95

Asn Glu Thr Asp Val Val Phe Gly Gly Glu Lys Arg Leu Phe Lys Ser  
 100 105 110

Ile Lys Glu Ile Ile Glu Lys Tyr Asp Pro Pro Ala Val Phe Val Tyr  
 115 120 125

Gln Thr Cys Val Pro Ala Met Ile Gly Asp Asp Ile Asp Ala Val Cys  
 130 135 140

Lys Ala Ala Arg Glu Lys Phe Gly Lys Pro Val Ile Pro Ile Asn Ser  
 145 150 155 160

Pro Gly Phe Val Gly Pro Lys Asn Leu Gly Asn Lys Leu Ala Gly Glu  
 165 170 175

Ala Leu Leu Asp His Val Ile Gly Thr Glu Glu Pro Asp Tyr Thr Thr  
 180 185 190

Ala Tyr Asp Ile Asn Ile Ile Gly Glu Tyr Asn Leu Ser Gly Glu Leu  
 195 200 205

Trp Gln Val Lys Pro Leu Leu Asp Glu Leu Gly Ile Arg Ile Leu Ala  
 59

|                                                                 |     |         |
|-----------------------------------------------------------------|-----|---------|
| 210                                                             | 215 | 220     |
| Cys Ile Ser Gly Asp Gly Lys Tyr Lys Asp Val Ala Ser Ser His Arg |     |         |
| 225                                                             | 230 | 235 240 |
| Ala Lys Ala Ala Met Met Val Cys Ser Lys Ala Met Ile Asn Val Ala |     |         |
|                                                                 | 245 | 250 255 |
| Arg Lys Met Glu Glu Arg Tyr Asp Ile Pro Phe Phe Glu Gly Ser Phe |     |         |
|                                                                 | 260 | 265 270 |
| Tyr Gly Ile Glu Asp Ser Ser Asp Ser Leu Arg Glu Ile Ala Arg Met |     |         |
|                                                                 | 275 | 280 285 |
| Leu Ile Glu Lys Gly Ala Asp Pro Glu Leu Met Asp Arg Thr Glu Ala |     |         |
|                                                                 | 290 | 295 300 |
| Leu Ile Glu Arg Glu Glu Lys Lys Ala Trp Asp Ala Ile Ala Ala Tyr |     |         |
| 305                                                             | 310 | 315 320 |
| Lys Pro Arg Phe Lys Asp Lys Lys Val Leu Leu Ile Thr Gly Gly Val |     |         |
|                                                                 | 325 | 330 335 |
| Lys Ser Trp Ser Val Val Ala Ala Leu Gln Glu Ala Gly Leu Glu Leu |     |         |
|                                                                 | 340 | 345 350 |
| Val Gly Thr Ser Val Lys Lys Ser Thr Lys Glu Asp Lys Glu Arg Ile |     |         |
|                                                                 | 355 | 360 365 |
| Lys Glu Leu Met Gly Gln Asp Ala His Met Ile Asp Asp Met Thr Pro |     |         |
|                                                                 | 370 | 375 380 |
| Arg Glu Met Tyr Lys Met Leu Lys Asp Ala Lys Ala Asp Ile Met Leu |     |         |
| 385                                                             | 390 | 395 400 |
| Ser Gly Gly Arg Ser Gln Phe Ile Ala Leu Lys Ala Ala Met Pro Trp |     |         |
|                                                                 | 405 | 410 415 |
| Leu Asp Ile Asn Gln Glu Arg His His Ala Tyr Met Gly Tyr Val Gly |     |         |
|                                                                 | 420 | 425 430 |
| Met Val Lys Leu Val Glu Glu Ile Asp Lys Ala Leu Tyr Asn Pro Val |     |         |
|                                                                 |     | 60      |

435

440

445

Trp Glu Gln Val Arg Lys Pro Ala Pro Trp Glu Asn Pro Glu Asp Thr  
 450 455 460

Trp Gln Ala Arg Ala Leu Ala Glu Met Glu Ala Glu Ala Ala Leu  
 465 470 475 480

Ala Ala Asp Pro Val Arg Ala Glu Glu Val Arg Arg Ser Lys Lys Ile  
 485 490 495

Cys Asn Cys Lys Ser Val Asp Leu Gly Thr Ile Glu Asp Ala Ile Lys  
 500 505 510

Ala His Ala Leu Thr Thr Val Glu Gly Val Arg Glu His Thr Asn Ala  
 515 520 525

Ser Gly Gly Cys Gly Ala Cys Ser Gly Arg Ile Glu Glu Ile Phe Glu  
 530 535 540

Ala Val Gly Val Val Ala Ala Pro Pro Pro Ala Glu Ala Ala Pro Ser  
 545 550 555 560

Pro Gln Glu Ile Ala Pro Asp Pro Leu Ala Ala Glu Glu Lys Arg Arg  
 565 570 575

Ala Lys Lys Ala Cys Gly Cys Lys Glu Val Ala Val Gly Thr Ile Glu  
 580 585 590

Asp Ala Ile Arg Ala Lys Gly Leu Arg Asn Ile Ala Glu Val Arg Ala  
 595 600 605

Ala Thr Asp Ala Asn Thr Gly Cys Gly Asn Cys Gln Glu Arg Val Glu  
 610 615 620

Gly Ile Leu Asp Arg Val Leu Ala Glu Ala Ala Ser Glu Leu Gln Ala  
 625 630 635 640

Ala Glu

<211> 1503  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 52  
 atgagtgtcg cacagtccca gagcgtcgcc gagatcaagg cgcgcaacaa ggaactcatc 60  
 gaagaggtcc tcaaggtcta tcccgagaag accgccaagc gccgcgccaa gcacctgaac 120  
 gtccacgaag ccggcaagtc cgactgcggc gtgaagtcca acatcaagtc catcccgggc 180  
 gtgatgacca tccgcggttg cgcttatgcc ggctccaagg gtgtggtgtg ggggtccatc 240  
 aaggacatga tccacatctc ccacggcccc gtgggctgcg gccagtatag ctgggcccgc 300  
 cgccgcaact actatatcgg caccgaccggc atcgacacct tcgtgacgat gcagttcacc 360  
 tccgacttcc aggagaagga catcgtcttc ggcggcgaca agaagctcgc caagatcatg 420  
 gacgagatcc aggagctggt cccgctgaac aacggcatca ccgttcagtc cgagtgcgcc 480  
 atcggcctca tcggcgacga catcgaggcc gtctccaagc agaagtccaa ggagtatgag 540  
 ggcaagacca tcgtgccggt gcgctgcgag ggcttccgcg gcgtgtccca gtccctgggc 600  
 caccacatcg ccaacgacgc catccgcgat tgggtgttcg acaagatcgc gcccgacgcc 660  
 gagccgcgct ttgagccgac cccgtacgac gtcgccatca tcggcgacta caatatcggc 720  
 ggtgacgcct ggtcgtcccg taccctctg gaggagatgg gcctgcgct gatcgcccag 780  
 tgggtccggcg acggttcgct cgctgagctg gaggccaccc cgaaggccaa gctcaacgtg 840  
 ctgcactgct accgctccat gaactacatc tcgcgccaca tggaagagaa gtacggtatc 900  
 ccgtggtgcg agtacaactt cttcggtcct tccaagatcg ccgagtcct gcgcaagatc 960  
 gccagctact tcgacgacaa gatcaaggaa ggcgcggagc gcgtcatcgc caagtatcag 1020  
 ccgctcatgg atgcggtgat cgcgaagtat cgtccccgcc tcgagggcaa gaccgtgatg 1080  
 ctgtacgtgg gcggcctgcg tccccgtcac gtcacggcg cctacgagga cctgggcatg 1140  
 gaagtggctc gcacgggcta cgagttcgcc cataacgacg actaccagcg caccgcccag 1200  
 cactacgtca aggatggcac catcatctat gacgacgtga ccggctacga gttcgagaag 1260  
 ttcgtcgaga agatccagcc ggacctggtc gggtcgggca tcaaggaaaa gtacgtcttc 1320  
 cagaagatgg gcgtgccgtt ccgccagatg cactcctggg actactcggg cccgtaccac 1380  
 ggctatgacg gcttcgcgat cttcgcgcgc gacatggaca tggccatcaa cagccccgtg 1440

tggaagatga cccaggctcc gtggaagagc gtccccaagc cgacgatgct cgcggctgaa 1500  
tga 1503

<210> 53  
<211> 500  
<212> PRT  
<213> Chưa biết

<220>  
<223> chủng mới

<400> 53

Met Ser Val Ala Gln Ser Gln Ser Val Ala Glu Ile Lys Ala Arg Asn  
1 5 10 15

Lys Glu Leu Ile Glu Glu Val Leu Lys Val Tyr Pro Glu Lys Thr Ala  
20 25 30

Lys Arg Arg Ala Lys His Leu Asn Val His Glu Ala Gly Lys Ser Asp  
35 40 45

Cys Gly Val Lys Ser Asn Ile Lys Ser Ile Pro Gly Val Met Thr Ile  
50 55 60

Arg Gly Cys Ala Tyr Ala Gly Ser Lys Gly Val Val Trp Gly Pro Ile  
65 70 75 80

Lys Asp Met Ile His Ile Ser His Gly Pro Val Gly Cys Gly Gln Tyr  
85 90 95

Ser Trp Ala Ala Arg Arg Asn Tyr Tyr Ile Gly Thr Thr Gly Ile Asp  
100 105 110

Thr Phe Val Thr Met Gln Phe Thr Ser Asp Phe Gln Glu Lys Asp Ile  
115 120 125

Val Phe Gly Gly Asp Lys Lys Leu Ala Lys Ile Met Asp Glu Ile Gln  
130 135 140

Glu Leu Phe Pro Leu Asn Asn Gly Ile Thr Val Gln Ser Glu Cys Pro  
145 150 155 160

Ile Gly Leu Ile Gly Asp Asp Ile Glu Ala Val Ser Lys Gln Lys Ser  
63

165

170

175

Lys Glu Tyr Glu Gly Lys Thr Ile Val Pro Val Arg Cys Glu Gly Phe  
 180 185 190

Arg Gly Val Ser Gln Ser Leu Gly His His Ile Ala Asn Asp Ala Ile  
 195 200 205

Arg Asp Trp Val Phe Asp Lys Ile Ala Pro Asp Ala Glu Pro Arg Phe  
 210 215 220

Glu Pro Thr Pro Tyr Asp Val Ala Ile Ile Gly Asp Tyr Asn Ile Gly  
 225 230 235 240

Gly Asp Ala Trp Ser Ser Arg Ile Leu Leu Glu Glu Met Gly Leu Arg  
 245 250 255

Val Ile Ala Gln Trp Ser Gly Asp Gly Ser Leu Ala Glu Leu Glu Ala  
 260 265 270

Thr Pro Lys Ala Lys Leu Asn Val Leu His Cys Tyr Arg Ser Met Asn  
 275 280 285

Tyr Ile Ser Arg His Met Glu Glu Lys Tyr Gly Ile Pro Trp Cys Glu  
 290 295 300

Tyr Asn Phe Phe Gly Pro Ser Lys Ile Ala Glu Ser Leu Arg Lys Ile  
 305 310 315 320

Ala Ser Tyr Phe Asp Asp Lys Ile Lys Glu Gly Ala Glu Arg Val Ile  
 325 330 335

Ala Lys Tyr Gln Pro Leu Met Asp Ala Val Ile Ala Lys Tyr Arg Pro  
 340 345 350

Arg Leu Glu Gly Lys Thr Val Met Leu Tyr Val Gly Gly Leu Arg Pro  
 355 360 365

Arg His Val Ile Gly Ala Tyr Glu Asp Leu Gly Met Glu Val Val Gly  
 370 375 380

Thr Gly Tyr Glu Phe Ala His Asn Asp Asp Tyr Gln Arg Thr Ala Gln  
 64

|             |            |             |            |
|-------------|------------|-------------|------------|
| <210>       | 54         |             |            |
| <211>       | 1377       |             |            |
| <212>       | ADN        |             |            |
| <213>       | Chưa biết  |             |            |
| <220>       |            |             |            |
| <223>       | chủng mới  |             |            |
| <400>       | 54         |             |            |
| atggccaccg  | tttccgtctc | caagaaggcc  | tgcgcggtca |
| acccccctcaa | gatgagccag |             | 60         |
| ccggtgggcg  | gcgcgctcgc | cttcatgggc  | gtgcgcaagg |
| ccatgccgct  | gctgcacggc |             | 120        |
| tcgcagggct  | gcacctcctt | cggcctgggtg | ctgttcgtgc |
| gccacttcaa  | ggaagccatc |             | 180        |
| cccatgcaga  | ccaccgccat | gagcgaggtg  | gcgacggttc |
| tgggcggcct  | tgagaatgtg |             | 240        |
| gagcaggcca  | ttctcaacat | ctacaatcgc  | accaagccgg |
| agatcatcgg  | catctgctcc |             | 300        |
| accggcgtca  | ccgagaccaa | gggcgatgat  | gtcgacggct |
| acatcaagct  | gatccgggac |             | 360        |
| aagtatcccc  | agctggccga | cttcccgctg  | gtctatgtct |
| ccacccccga  | tttcaaggac |             | 420        |
| gccttccagg  | acggttggga | gaagaccgtg  | gcgaagatgg |
| tggaggcgct  | ggtgaagccc |             | 480        |

gccgccgaca agcagaagga caagacccgc gtcaacgtcc tgcccggctg ccacctcacg 540  
 cccggcgatc tggatgagat gcggaccatc ttcgaggatt tcgggctcac accctatttc 600  
 ctgccggatc tggccggctc gctggatggg catatcccg aggacttctc gcccaccacc 660  
 atcggcggca tcggcatcga tgagatcgcc accatgggcg aggcgggcca caccatctgc 720  
 atcggcgcgc agatgcgccg ggcgggag gagcatggaga agaagaccgg cattcccttc 780  
 aagctgttcg agcgcctgtg cggcctggag gcgaacgacg cttcatcat gcacctgtcg 840  
 cagatctccg gccggccggt gccggtgaag tatcgccggc agcggggcca gctggtggat 900  
 gccatgctgg acggccactt ccatctgggc ggtcgcaagg tggccatggg ggcggagccg 960  
 gacctgctct acgacgtggg ctcttctctg cacgagatgg gcgcccacat ctttccgcg 1020  
 gtcaccacca cccagtcgcc ggtgctggcg cgcctgcctg ccgaggaggt gcttatcggc 1080  
 gacctggagg atctggagac ccaggcgaag gcgcgcggat gcgatctcct gctcacccat 1140  
 tcccatgggc gccaggcggc ggagcgcctc cacatcccct tctaccggat cggcattccc 1200  
 atgtttgacc ggctgggggc ggggcatctg ttgtcggtgg gctatcgcg caccgcgac 1260  
 ctcatcttcc atctcgccaa ctttgtgac gccgaccacg aggaaaatca cgagccgacg 1320  
 cccgacacct gggccaccgg ccatggcgag catgccgccg cccccacttc ccattga 1377

<210> 55  
 <211> 458  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chúng mới

<400> 55

Met Ala Thr Val Ser Val Ser Lys Lys Ala Cys Ala Val Asn Pro Leu  
1 5 10 15

Lys Met Ser Gln Pro Val Gly Gly Ala Leu Ala Phe Met Gly Val Arg  
20 25 30

Lys Ala Met Pro Leu Leu His Gly Ser Gln Gly Cys Thr Ser Phe Gly  
35 40 45

Leu Val Leu Phe Val Arg His Phe Lys Glu Ala Ile Pro Met Gln Thr  
50 55 60

Thr Ala Met Ser Glu Val Ala Thr Val Leu Gly Gly Leu Glu Asn Val  
65 70 75 80

Glu Gln Ala Ile Leu Asn Ile Tyr Asn Arg Thr Lys Pro Glu Ile Ile  
85 90 95

Gly Ile Cys Ser Thr Gly Val Thr Glu Thr Lys Gly Asp Asp Val Asp  
100 105 110

Gly Tyr Ile Lys Leu Ile Arg Asp Lys Tyr Pro Gln Leu Ala Asp Phe  
115 120 125

Pro Leu Val Tyr Val Ser Thr Pro Asp Phe Lys Asp Ala Phe Gln Asp  
130 135 140

Gly Trp Glu Lys Thr Val Ala Lys Met Val Glu Ala Leu Val Lys Pro  
145 150 155 160

Ala Ala Asp Lys Gln Lys Asp Lys Thr Arg Val Asn Val Leu Pro Gly  
165 170 175

Cys His Leu Thr Pro Gly Asp Leu Asp Glu Met Arg Thr Ile Phe Glu  
180 185 190

Asp Phe Gly Leu Thr Pro Tyr Phe Leu Pro Asp Leu Ala Gly Ser Leu  
195 200 205

Asp Gly His Ile Pro Glu Asp Phe Ser Pro Thr Thr Ile Gly Gly Ile  
210 215 220

Gly Ile Asp Glu Ile Ala Thr Met Gly Glu Ala Ala His Thr Ile Cys  
225 230 235 240

Ile Gly Ala Gln Met Arg Arg Ala Gly Glu Ala Met Glu Lys Lys Thr  
245 250 255

Gly Ile Pro Phe Lys Leu Phe Glu Arg Leu Cys Gly Leu Glu Ala Asn  
260 265 270

Asp Ala Phe Ile Met His Leu Ser Gln Ile Ser Gly Arg Pro Val Pro  
275 280 285

Val Lys Tyr Arg Arg Gln Arg Gly Gln Leu Val Asp Ala Met Leu Asp  
 290 295 300

Gly His Phe His Leu Gly Gly Arg Lys Val Ala Met Gly Ala Glu Pro  
 305 310 315 320

Asp Leu Leu Tyr Asp Val Gly Ser Phe Leu His Glu Met Gly Ala His  
 325 330 335

Ile Leu Ser Ala Val Thr Thr Thr Gln Ser Pro Val Leu Ala Arg Leu  
 340 345 350

Pro Ala Glu Glu Val Leu Ile Gly Asp Leu Glu Asp Leu Glu Thr Gln  
 355 360 365

Ala Lys Ala Arg Gly Cys Asp Leu Leu Leu Thr His Ser His Gly Arg  
 370 375 380

Gln Ala Ala Glu Arg Leu His Ile Pro Phe Tyr Arg Ile Gly Ile Pro  
 385 390 395 400

Met Phe Asp Arg Leu Gly Ala Gly His Leu Leu Ser Val Gly Tyr Arg  
 405 410 415

Gly Thr Arg Asp Leu Ile Phe His Leu Ala Asn Leu Val Ile Ala Asp  
 420 425 430

His Glu Glu Asn His Glu Pro Thr Pro Asp Thr Trp Ala Thr Gly His  
 435 440 445

Gly Glu His Ala Ala Ala Pro Thr Ser His  
 450 455

<210> 56  
 <211> 1560  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 56  
 atgccacaaa atgctgacaa tgtgctcgat cacttcgagc tcttcctggg tcccgaatac

60

cagcagatgc tggccaataa gaaaaagatg ttcgagaacc cccgcatcc ggccgaagtc 120  
 gagcgcgtgc gggaaatgggc gaagactcct gaatacaagg agctgaactt cgcccgcgag 180  
 gcgctcaccg tgaatccggc caaggcttgt cagccgctgg gcgcggtgtt cgtcgccgtc 240  
 ggcttcgaga gcacgatccc cttcgtgcac ggctcgcagg gttgcgtcgc gtattaccgc 300  
 tcgcacctct cccgccactt caaggagccg tcctcctgcg tctcctcgtc catgaccgag 360  
 gatgcggcgg tgttcggcgg cctcaacaac atgattgacg gcctcgccaa cacctacaac 420  
 atgtacaagc cgaagatgat cgccgtctcc accacctgca tggcggaagt catcggcgac 480  
 gatctgaacg ctttcatcaa gaccgcgaag gaaaagggtc cggttcggc cgaatacgac 540  
 gtgcccttcg cccacacccc ggcttcgtc ggagccatg tcaccggcta cgacaatgcg 600  
 ctcaagggca tcctcgagca cttctgggac ggcaaggccg gcaccgcgcc gaagctggag 660  
 cgcgttccca acgagaagat caacttcata ggaggcttcg acggctacac cgtcggcaac 720  
 actcgcgaag tgaagcgcac cttcgaggcg ttggcgccg attacaccat cctcgccgac 780  
 aattccgaag tgttcgacac cccgaccgac ggaggttcc gcatgtatga cggcggcacg 840  
 accctggagg acgcggcgaa cgcggtgcac gccaaaggcca ccatctccat gcaggaatac 900  
 tgcacggaga agaccctgcc catgatcgcc ggtcatggcc aggacgtggt cgccctcaac 960  
 caccctgtgg gcgtgggcgg caccgacaag ttctcatgg agatcgccc cctcaccggc 1020  
 aaggagatcc ccgaggagct gaccgcgag cgcggccgtc tcgtggacgc tatcgcgac 1080  
 tcttccgcgc acatccacgg caagaagttc gccatctacg gcgatccgga tctgtgcctg 1140  
 ggctcgcgcg cgttcctgct ggagctgggc gccagaccga cccatgtgct ggccaccaac 1200  
 ggaccaaga agtgggcccga gaaggttcag gaactgttcg actcttcgcc gttcggcgcc 1260  
 aactgcaagg tctatcccgg caaggacctg tggcacatgc gctcgtcct gttcgtggag 1320  
 ccggtggatt tcatcatcgg caacacctac ggcaagtatc tcgagcgca cacgggcacc 1380  
 ccgctgatcc gtatcggctt cccggtgttc gaccgtcacc accaccaccg ccgtccggtg 1440  
 tggggctatc agggcggcat gaacgtcctg atcacgatcc tcgacaagat ctttgacgag 1500  
 atcgaccgca acaccaacgt gccggccaag accgactact cgttcgacat cattcgttga 1560

<210> 57

<211> 519

<212> PRT

<213> Chưa biết

&lt;220&gt;

&lt;223&gt; chủng mới

&lt;400&gt; 57

Met Pro Gln Asn Ala Asp Asn Val Leu Asp His Phe Glu Leu Phe Arg  
 1 5 10 15

Gly Pro Glu Tyr Gln Gln Met Leu Ala Asn Lys Lys Lys Met Phe Glu  
 20 25 30

Asn Pro Arg Asp Pro Ala Glu Val Glu Arg Val Arg Glu Trp Ala Lys  
 35 40 45

Thr Pro Glu Tyr Lys Glu Leu Asn Phe Ala Arg Glu Ala Leu Thr Val  
 50 55 60

Asn Pro Ala Lys Ala Cys Gln Pro Leu Gly Ala Val Phe Val Ala Val  
 65 70 75 80

Gly Phe Glu Ser Thr Ile Pro Phe Val His Gly Ser Gln Gly Cys Val  
 85 90 95

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

Cys Val Ser Ser Ser Met Thr Glu Asp Ala Ala Val Phe Gly Gly Leu  
 115 120 125

Asn Asn Met Ile Asp Gly Leu Ala Asn Thr Tyr Asn Met Tyr Lys Pro  
 130 135 140

Lys Met Ile Ala Val Ser Thr Thr Cys Met Ala Glu Val Ile Gly Asp  
 145 150 155 160

Asp Leu Asn Ala Phe Ile Lys Thr Ala Lys Glu Lys Gly Ser Val Pro  
 165 170 175

Ala Glu Tyr Asp Val Pro Phe Ala His Thr Pro Ala Phe Val Gly Ser  
 180 185 190

His Val Thr Gly Tyr Asp Asn Ala Leu Lys Gly Ile Leu Glu His Phe  
 195 200 205

Trp Asp Gly Lys Ala Gly Thr Ala Pro Lys Leu Glu Arg Val Pro Asn  
 210 215 220

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

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

Ile Leu Ala Asp Asn Ser Glu Val Phe Asp Thr Pro Thr Asp Gly Glu  
 260 265 270

Phe Arg Met Tyr Asp Gly Gly Thr Thr Leu Glu Asp Ala Ala Asn Ala  
 275 280 285

Val His Ala Lys Ala Thr Ile Ser Met Gln Glu Tyr Cys Thr Glu Lys  
 290 295 300

Thr Leu Pro Met Ile Ala Gly His Gly Gln Asp Val Val Ala Leu Asn  
 305 310 315 320

His Pro Val Gly Val Gly Gly Thr Asp Lys Phe Leu Met Glu Ile Ala  
 325 330 335

Arg Leu Thr Gly Lys Glu Ile Pro Glu Glu Leu Thr Arg Glu Arg Gly  
 340 345 350

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

Lys Phe Ala Ile Tyr Gly Asp Pro Asp Leu Cys Leu Gly Leu Ala Ala  
 370 375 380

Phe Leu Leu Glu Leu Gly Ala Glu Pro Thr His Val Leu Ala Thr Asn  
 385 390 395 400

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

Pro Phe Gly Ala Asn Cys Lys Val Tyr Pro Gly Lys Asp Leu Trp His  
 420 425 430

Met Arg Ser Leu Leu Phe Val Glu Pro Val Asp Phe Ile Ile Gly Asn  
 435 440 445

Thr Tyr Gly Lys Tyr Leu Glu Arg Asp Thr Gly Thr Pro Leu Ile Arg  
 450 455 460

Ile Gly Phe Pro Val Phe Asp Arg His His His His Arg Arg Pro Val  
 465 470 475 480

Trp Gly Tyr Gln Gly Gly Met Asn Val Leu Ile Thr Ile Leu Asp Lys  
 485 490 495

Ile Phe Asp Glu Ile Asp Arg Asn Thr Asn Val Pro Ala Lys Thr Asp  
 500 505 510

Tyr Ser Phe Asp Ile Ile Arg  
 515

<210> 58  
 <211> 621  
 <212> ADN  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 58  
 gtggagtccg gtggtcctga gccgggctg ggctgcgccg gccgcggcgt gatcacctcc 60  
 atcaacttcc tggaggagaa cggcgcctac gaggacatcg actatgtgtc ctacgacgtg 120  
 ctgggcgacg tgggtgtcgg cggcttcgcc atgcccattcc gcgagaacaa ggcgaggaa 180  
 atctacatcg tgatgtccgg cgagatgatg gccatgtatg cggccaacaa catctccaag 240  
 ggcatcctga agtatgcaa ttccggcggc gtgcgcctgg gcgggctggt ctgcaacgag 300  
 cgccagaccg acaaggagct ggagctggcg gaggctctgg cgaagaagct cggcaccgag 360  
 ctgatctact tcgtgccgcg cgacaacatc gtgcagcatg ccgagctgcg ccgatgaca 420  
 gtgatcgagt atgcgcccga ttccggccag gccagcact accggaacct ggccgagaag 480  
 gtgcacgcca acaagggcaa cggcatcatc ccgaccccga tcacatgga cgagctggaa 540  
 gacatgctca tggagcacgg catcatgaag gccgtggacg agagccagat cggcaagacc 600  
 gccgccgagc tcgccgtctg a 621

<210> 59  
 <211> 206  
 <212> PRT  
 <213> Chưa biết

<220>  
 <223> chủng mới

<400> 59

Met Glu Ser Gly Gly Pro Glu Pro Gly Val Gly Cys Ala Gly Arg Gly  
 1 5 10 15

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

Ile Asp Tyr Val Ser Tyr Asp Val Leu Gly Asp Val Val Cys Gly Gly  
 35 40 45

Phe Ala Met Pro Ile Arg Glu Asn Lys Ala Gln Glu Ile Tyr Ile Val  
 50 55 60

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

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

Val Cys Asn Glu Arg Gln Thr Asp Lys Glu Leu Glu Leu Ala Glu Ala  
 100 105 110

Leu Ala Lys Lys Leu Gly Thr Glu Leu Ile Tyr Phe Val Pro Arg Asp  
 115 120 125

Asn Ile Val Gln His Ala Glu Leu Arg Arg Met Thr Val Ile Glu Tyr  
 130 135 140

Ala Pro Asp Ser Ala Gln Ala Gln His Tyr Arg Asn Leu Ala Glu Lys  
 145 150 155 160

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

Asp Glu Leu Glu Asp Met Leu Met Glu His Gly Ile Met Lys Ala Val  
 73

180

185

190

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asp | Glu | Ser | Gln | Ile | Gly | Lys | Thr | Ala | Ala | Glu | Leu | Ala | Val |
|     | 195 |     |     |     |     |     | 200 |     |     |     |     | 205 |     |