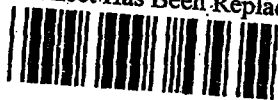


This Sheet Has Been Replaced



REPLACED SHEETS SG

☒ Description	: paged 1 to paged 112
☒ Claim(s)	: paged 113 to paged 122
☒ Drawing(s)	: paged ~ to paged ~
☒ Abstract	: paged 123 to paged 123

SPECIFICATION

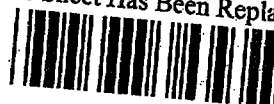
DNA CODING FOR DECARBAMYLASE
IMPROVED IN THERMOSTABILITY AND USE THEREOF

Technical Field

The present invention relates to a DNA fragment, and more particularly, it relates to a DNA fragment coding for a thermostability-improved enzyme capable of converting D-N-carbamoyl- α -amino acids into the corresponding D- α -amino acids (hereinafter referred to as decarbamylase). The present invention also relates to a process for preparing the DNA fragment, a vector comprising the DNA fragment, a transformant obtained by transformation with the vector, a decarbamylase improved in thermostability and its production process, and an improved process for producing amino acids using the decarbamylase. The present invention further relates to a method for increasing, by one or more, the number of mutations in amino acids related to the thermostability of a decarbamylase.

Background Art

Optically active D- α -amino acids are important compounds as intermediates of medicaments, and in particular, D-phenylglycine, D-parahydroxyphenylglycine and the like, which are intermediates for the production of semi-synthetic penicillins and semisynthetic sepharospolines, are useful compounds from an industrial point of view. As the production process for such D- α -amino acids, a process is known in which these compounds are obtained by the removal of carbamo-



yl groups of the corresponding D-N-carbamoyl- α -amino acids, wherein the removal of carbamoyl groups is effected by a chemical method (the specification of the Japanese Patent Publication No. 58-4707) or by the utilization of enzyme reactions of microorganisms (the specifications of the Japanese Patent Publication Nos. 57-18793, 63-20520 and 1-48758, and the Japanese Patent Application No. 2-407922).

However, because a mineral acid such as sulfuric acid is used at great amounts in the chemical method employed for the removal of carbamoyl groups, serious problems on the environment may occur in relation to the treatment thereof and the like. On the other hand, the method utilizing an enzyme reaction has disadvantages that the amount of enzyme produced is not satisfactory and there are some difficulties in its properties even if mass production is made possible, so that enzymes with reactivity to substrates and enzyme stability have not yet been found.

In general, some enzymes have poor stability, and when such an enzyme is prepared, a stabilizing agent is added or a means of preventing inactivation such as treatment at low temperatures is employed. When an enzyme is used in an actual reaction at ordinary or high temperatures, attention will be directed to the stability of the enzyme. In particular case where an enzyme is used on an industrial scale, its stability often has an effect upon the product cost. Also, in particular case where an enzyme is used on an industrial scale, as a means of allowing the enzyme reaction to proceed with advantage, the enzyme is repeatedly used in the reaction



as a so-called bioreactor such as an immobilized enzyme or immobilized bacterial cells, and even at this time, the stability of the enzyme will give a limitation to the number of use and have an important effect upon the product cost.

In the present invention, attention is paid to the thermostability as an index of the enzyme stability and an improvement in the thermostability of an enzyme is achieved using a genetic engineering means, whereby the stability of this enzyme is increased and this makes it possible to produce a stabilized enzyme having excellent advantages for industrial applications.

To solve such a problem, the present invention has an object to prepare a decarbamylase with high reactivity to D-N-carbamoyl- α -amino acids as the substrate and excellent stability by an improvement of currently available decarbamylases and to produce D- α -amino acids using this enzyme with high efficiency.

Summary of the Invention

That is, the present invention provides an improved process for producing D- α -amino acids, comprising converting D-N-carbamoyl- α -amino acids into the corresponding D- α -amino acids in an aqueous medium by the action of an enzyme and collecting the D- α -amino acids produced, said enzyme being produced by a transformant which has been obtained by the transformation of a host bacterial cell selected from the microorganisms of the genus *Escherichia*, *Pseudomonas*, *Flavobacterium*, *Bacillus*, *Serratia*, *Corynebacterium* or *Brevibacterium* with a gene which has been obtained by a process



comprising subjecting a DNA fragment containing a decarbamylase gene to chemical or enzyme mutagenesis, introducing a recombinant DNA having a vector DNA ligated with this DNA fragment into a host cell and screening a strain capable of producing an enzyme improved in thermostability.

Thus, the present invention also provides a DNA fragment for conducting the improved process for producing D- α -amino acids according to the present invention, its production process and an expression vector containing the fragment, a transformed microorganism obtained by transformation with the expression vector, as well as a decarbamylase improved in thermostability and its production process. The present invention further provides a method for increasing, by one or more, the number of mutations in amino acids related to the thermostability of decarbamylases.

Brief Description of the Drawings

The present invention will hereinafter be illustrated by reference to the accompanying drawings.

Figure 1 is a graph showing the thermostable activity of a decarbamylase after thermal treatment, which is obtained by the use of a strain capable of producing the decarbamylase improved in thermostability according to the present invention.

Figure 2 is a graph showing the results of stability obtained by a repeated continuous test on the thermostability using a resin having immobilized thereon a decarbamylase improved in thermostability according to the present invention.



Figure 3 shows restriction maps of plasmids pKK233-2 and pKK NE.

Figure 4 shows restriction maps of plasmids pAD108 and pAD1086.

Figure 5 shows restriction maps of plasmids pAD402 and pAD4021.

Figure 6 shows restriction maps of plasmids pAD429 and pAD455.

Figure 7 is a graph showing the relationship between the reaction temperature of a decarbamylase produced by E. coli JM109 pAD108 or E. coli JM109 pAD455 and its activity.

Figure 8 is a graph showing the stability against pH, of a decarbamylase produced by E. coli JM109 pAD416 or E. coli JM109 pAD455.

Figure 9 shows a process for preparing vector pUC NT for the introduction of a foreign gene and vector pNT4553 for the expression of a decarbamylase improved in thermostability.

Figure 10 shows a restriction map of plasmid pUC NT and pNT4553.

Detailed Description of the Invention

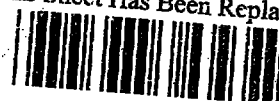
In general, it is known that the thermostability and stability of an enzyme are correlated with each other. It is also known that various enzymes produced by thermophiles are those which have high thermostability and also have high stability with almost no exception, and that enzymes having thermostability given by the substitution of amino acids according to a gene recombination technique also have excel-



lent stability (Kagawa et al., Saibo Kogaku, vol. 7, 509-571 (1988)).

As the method for making an enzyme thermostable, there can be employed a method in which an amino acid sequence of the enzyme protein is changed or a method in which the enzyme is modified by chemical treatment, enzyme reaction or the like. As the method for improving the thermostability by a change in the amino acid sequence of the enzyme protein, there can be employed a method in which bacterial cells are allowed to cause some mutation with a chemically mutagenic agent such as nitrosoguanidine (NTG) or by ultraviolet rays, followed by screening, or a method in which a gene of interest is obtained by a gene recombination technique and then allowed to cause some chemical or enzyme mutation, followed by returning it into the bacterial cell and then screening. When a gene is obtained and allowed to cause some mutation, it is more effective to subject a DNA fragment coding for a decarbamylase protein to mutagenesis in the form of a single strand. To make the fragment into a single-stranded form, there can be employed, for example, a method for incorporating the fragment into phage particles. As the single-stranded DNA, there can be used either a DNA strand containing the codons corresponding to the amino acid sequence of a decarbamylase protein or the complementary strand thereof. Moreover, as to the mutation sites, there can be employed a method for introducing random mutations or a method for introducing site-specific mutations.

As the agent for allowing a gene to cause some random



mutation and thereby introducing some mutation into the amino acid sequence of an enzyme protein, there can be used hydroxylamine hydrochloride, sodium nitrite, formic acid, hydrazine or the like.

For example, in the mutagenesis using hydroxylamine hydrochloride, a decarbamylase gene is first ligated with the double-stranded DNA of an M13 phage such as M13mp19, with which *E. coli* JM109 or the like is infected, and this bacterial culture is incubated to prepare phage particles for use in the mutagenesis. The mutation can be effected by the reaction using hydroxylamine hydrochloride in a concentration of 0.1 to 2 M, desirably 0.25 M, at a pH of 6.0 to 8.0, preferably 6.0, at a temperature of 37°C for a period of 1 to 24 hours. The phage particles undergoing such a mutagenic reaction are used for the infection of *E. coli* and recovered as a double-stranded recombinant DNA, followed by the transfer of such a mutant decarbamylase gene to a plasmid.

In the mutagenesis using sodium nitrite, formic acid, hydrazine or the like, there can be basically used a method as described by R.M. Myers, et al. (*Science*, 229, 242-247 (1985)). This method can be conducted by the preparation of a single-stranded DNA from the recombinant M13 phage having a decarbamylase gene incorporated therein and the subsequent mutagenesis. The mutation with sodium nitrite can be effected by the reaction using sodium nitrite in a concentration of 0.5 to 2 M, preferably 1 M, at a pH on the acidic side, preferably pH 4.3, at a temperature of 4°C to 37°C, preferably 25°C, for a period of 1 minute to 5 hours, preferably 30 min-



utes. The mutation with formic acid can be effected by the reaction using 12 M formic acid at 4°C to 37°C, preferably 15°C, for a period of 2 to 10 minutes. The mutation with hydrazine can be effected by the treatment using hydrazine in a concentration of 20% to 60% at 25°C for 3 to 10 minutes. These mutant single-stranded DNAs are made into double-stranded forms using E. coli DNA polymerase Klenow fragment, Sequenase^R and the like, followed by incorporation into a plasmid, which makes it possible to produce a mutant decarbamylase gene.

As the method for causing some random mutation using enzyme reaction, there can be employed a method using PCR (polymerase chain reaction) (D.W. Leung et al., A Journal of Methods in Cell and Molecular Biology, 1, 11-15 (1989)) or the like. This is a method in which TaqDNA polymerase is used for the synthesis of a gene from a synthetic DNA (DNA primer) having the sequences at both ends of the gene, at which time a mutant decarbamylase gene can be prepared by causing an error in the gene synthesis using TaqDNA polymerase under the reaction conditions such as higher concentrations of MgCl₂ and substrates (dNTPs) as compared with ordinary reaction or an extremely lowered concentration of only one substrate in four kinds of substrates.

As the method for introducing a mutation into a gene, which is specific to the site of a enzyme protein, there can be employed an in vitro method for introducing a site-specific mutation using an oligonucleotide, a method for replacing



a cassette at the mutation site, a method using PCR or the like. The in vitro method for introducing a site-specific mutation using an oligonucleotide can be conducted by preparing a single-stranded DNA from a recombinant M13 phage having a decarbamylase gene incorporated therein, and making the DNA into a double-stranded form using a synthetic DNA primer containing a sequence, after the mutation, of a gene part to be allowed to cause some mutation and using an enzyme such as a DNA polymerase Klenow fragment, followed by introduction into *E. coli*. These reactions can readily be effected using a commercially available kit, for example, "MutanTM-K" or "MutanTM-G" available from Takara Shuzo, or "Oligonucleotide-directed in vitro mutagenesis System, version 2.0" from Amersham Japan. The method for replacing a cassette at the mutation site can be conducted by replacing the whole restriction enzyme fragment containing the site to be allowed to cause a mutation, with a synthetic DNA containing such a mutation. The method using a PCR reaction can be conducted according to the method of W. Ito, et al., *Gene*, 102, 67-70 (1991) or the like. In this method, PCR reaction is effected using a synthetic DNA primer containing the mutation and a synthetic DNA corresponding to the end of the gene, after which the resulting DNA fragment is allowed to hybridize with the full-length DNA of the gene, followed by another PCR reaction for the enzyme elongation of the fragment to its full length.

In the present invention, the mutant decarbamylase



gene prepared in the above-described manner is ligated to a plasmid vector such as pUC19 or pKK-233-2, and the resulting recombinant vector is incorporated into *E. coli*, such as JM109, followed by colony formation on a plate containing an antibiotic and the like. Then, this colony is replicated onto a sterilized filter paper, and the plate is stored. The filter paper, after drying, is immersed in a solution containing lysozyme and Triton X-100 or repeatedly put in and out of an acetone-dry ice bath (freeze and thaw), thereby causing lysis. Then, this filter paper is immersed in a water bath thermostated at a temperature such as 65°C or 70°C for a constant period of time, for example, 5 minutes, and then dried, after which this filter paper is immersed in a reaction solution capable of making color development by decarbamylase activity. As the reaction solution, there can be used, for example, those which contain a carbamyl-D-amino acid as the substrate and also contains phenol, 4-aminoantipyrine, D-amino acid oxidase and peroxidase. The colony exhibiting color development can be separated as a thermostability-acquiring strain.

First, as an index indicating the thermostability of an enzyme, thermostable temperature is defined. The thermostable temperature is defined as the treatment temperature at which the activity of an enzyme is 50% inactivated by the thermal treatment for 10 minutes. As the source of a decarbamylase gene, microorganisms which are known to have a decarbamylase (e.g., microorganisms disclosed in the International Publication No. WO92/10579) can usually be used. The



thermostable temperature is different with the respective decarbamylases of these microorganisms. To attain excellent stability tolerable to repeated use by an improvement in the thermostability, various properties other than thermostability should, of course, be taken into consideration, and it is desirable to use a decarbamylase exhibiting high thermostable temperature, for example, in the range of 60° to 63°C, from the first. From this point of view, a decarbamylase produced by *Agrobacterium radiobacter* KNK712 (FERM BP-1900) is preferred because of its relatively higher thermostable temperature of about 62°C. The decarbamylase from KNK712 and the DNA fragment coding for this enzyme have the amino acid sequence and the DNA sequence, respectively, shown in Sequence Listing, SEQ ID NO: 1. The decarbamylase improved in thermostability obtained by the mutagenesis of this decarbamylase gene was found to contain a replacement of 57-histidine, 203-proline or 236-valine with a different amino acid, as shown in Sequence Listing, SEQ ID NOs: 2-16. Therefore, amino acids at at least three positions are related to the thermostability. High thermostability can be obtained not only by the replacement of one amino acid but also by the combined replacement of two or three amino acids. In the present invention, the thermostable temperature could be increased by at least 2°C, mostly about 5°C or higher, and in some cases, by 10°C or higher.

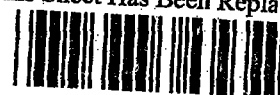
For example, it is possible to obtain a decarbamylase improved in thermostability by the substitution of leucine or tyrosine for 57-histidine; of leucine, asparagine, glutamic



acid, threonine or serine for 203-proline; or of alanine, threonine or serine for 236-valine, as the different amino acid (Sequence Listing, SEQ ID NOs: 2-30). As the DNA sequences coding for these decarbamylases improved in thermostability, there were obtained those which contain a substitution of TAT, CTT or CTA for CAT, i.e., 401- to 403-nucleotides corresponding to 57-histidine; of TCT, CTT, GAA, AAC or ACC for CCT, i.e., 839- to 841-nucleotides corresponding to 203-proline; of GCG, GCT, ACC, ACG, TCA, TCG or AGT for GTG, i.e., 938- to 940-nucleotides corresponding to 236-valine (Sequence Listing, SEQ ID NOs: 2-30). Once the thermostability-related amino acids are identified, the use of site-specific mutagenesis makes it possible to substitute many amino acids for these amino acids by the substitution of various bases. Thus, the DNA fragment coding for the decarbamylase improved in thermostability can be subjected to mutagenesis for a further improvement in the thermostability.

In case of *Agrobacterium radiobacter* KNK712 (FERM BP-1900), the derivative having two or three amino acid mutations in three amino acids which have been found to have a relation to the thermostability, i.e., 57-histidine, 203-proline and 236-valine (multiple mutant) is produced, for example, as follows.

In the gene structures of pAD108 (coding for a naturally occurring decarbamylase) and pAD402 (coding for a decarbamylase having Tyr substituted for 57-His), respectively, as shown in Figs. 4 and 5, SphI and SalI are close to each other in the vicinity of the decarbamylase gene; there-



fore, an NdeI restriction site is formed at the initiation codon of the decarbamylase gene by the PCR method. Thus, 57-amino acid is present on the NdeI-SacI DNA fragment of about 190 bp; 203-amino acid on the SalI-ClaI DNA fragment of about 170 bp; and 236-amino acid on the ClaI-SphI DNA fragment of about 75 bp.

The multiple mutant is prepared by the replacement of these mutation-free DNA fragments with DNA fragments having some mutation in the thermostability-related amino acids. The DNA fragment having some amino acid mutation is prepared from the transformant or expression vector as obtained in Example 3 or 8.

On the other hand, an NdeI-EcoRI DNA fragment of about 1.6 kb containing a decarbamylase gene (which is naturally occurring or contains one amino acid mutation) is obtained by cleaving pAD108, pAD402 or the like, which has been conferred an NdeI restriction site, and then incorporated into an appropriate vector having NdeI and EcoRI restriction sites, such as pKK NE (prepared from pKK233-2 shown in Fig. 3) described in Example 9 and shown in Fig. 3, thereby obtaining a plasmid (e.g., pAD 1086 shown in Fig. 4 from pAD108; and pAD4021 shown in Fig. 5 from pAD402). The DNA fragments of this plasmid, which have no mutation site for the above-mentioned three amino acids, are replaced with DNA fragments containing such amino acid mutation sites, thereby obtaining an expression vector for multiple mutants.

In general, the multiple mutants thus prepared have improved thermostability in an additive manner depending upon

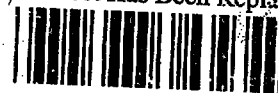


the degree of thermostability improvement obtained by a single mutation. For the decarbamylase having two or more amino acid mutations in three kinds of amino acids, it is possible to select any combination of the substituted amino acids as described above.

The above-described method is not limited to the case of *Agrobacterium radiobacter* KNK712 (FERM BP-1900), and it can be used, in general, for a further improvement in the thermostability of a decarbamylase enzyme. That is, according to the method for making an enzyme thermostable as described above, several kinds of decarbamylases improved in thermostability can be obtained by the replacement of an amino acid, from which the thermostability-related amino acids are identified, and still another replacement of this amino acid makes it possible to find other amino acids giving an improvement in the thermostability. Under the circumstances that DNA fragments corresponding to the respective thermostable decarbamylases have already been obtained, simultaneous mutations can be introduced into plural amino acid sites related to the thermostability improvement by the operations in the following steps 1) to 5) (multiple mutant), which afforded a decarbamylase additively improved in thermostability.

- 1) Restriction enzymes are found out, each of which can yield a DNA fragment comprising a DNA portion coding for one of the thermostability-related amino acids by cleavage for the DNA fragment.

- 2) A DNA fragment comprising a DNA portion coding for



one amino acid mutation site is obtained from the DNA fragments each coding for a thermostable decarbamylase having at least one amino acid mutation or from the vectors comprising the fragments with the corresponding restriction enzyme in step 1).

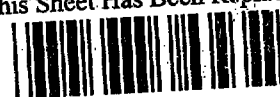
3) The corresponding DNA fragment but having no amino acid mutation is removed by cleaving a DNA fragment coding for a decarbamylase or a vector comprising the fragment with the same restriction enzyme as described above.

4) The mutation-containing DNA fragment obtained in step 2) is incorporated into the remaining DNA fragment after the cleavage for removing the corresponding DNA fragment in step 3) or into the vector comprising the remaining fragment.

5) The operations in steps 2) to 4) are repeated, if necessary.

According to this series of operations, it is possible to increase, by one or more, the number of mutations in the thermostability-related amino acids.

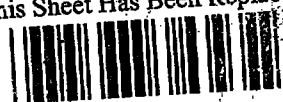
In case of *Agrobacterium radiobacter* KNK712 (FERM BP-1900), as the vector in step 2), any vector can be used so long as it has been prepared by giving an NdeI restriction site in an expression vector which is pUC19 having incorporated therein one of the DNA fragments shown in Sequence Listing, SEQ ID NOs: 2-30. As the vector containing a decarbamylase gene in step 3), various vectors, including pAD4021 and pAD1086, can be used which are obtained using pKK NE in the same operations as for pAD4021 from the vector prepared by giving an NdeI restriction site in any vector



selected from various examples of the expression vector as recited below, except for the vectors each containing a DNA fragment coding for a decarbamylase having already received an amino acid substitution for an improvement in the thermostability at every one of three thermostability-related sites (i.e., triple mutation-containing pAD426, pAD427, pAD454, pAD455 and pAD456). The DNA fragments obtained with restriction enzymes are NdeI-SacI, SalI-ClaI and ClaI-SphI in any case of step 2) or step 3).

The preparation of DNA fragments each coding for a decarbamylase, expression vectors obtainable by incorporating each of the DNA fragments into a vector, and transformants obtainable by incorporating each of the expression vectors into a host cell can be conducted using a gene manipulation technique as described in, for example, International Publication No. WO92/10579.

From the expression vector obtained by the incorporation of a DNA fragment shown in Sequence Listing, SEQ ID NOs: 2-34 into pUC19 or pKK NE, for example, the incorporation into pUC19 gives pAD402, pAD404, pAD406, pAD416, pAD428, pAD429, pAD431, pAD434, pAD435, pAD439, pAD441, pAD445, pAD447, pAD448, pAD450, pAD451, pAD452, pAD453, pAD454 or pAD456; the incorporation into pKK NE gives pAD421, pAD422, pAD423, pAD424, pAD425, pAD426, pAD427, pAD461 and pAD455; and the incorporation into pUC NT (Fig. 10, upper part) gives pNT4553 (Fig. 10, lower part). The replacement of amino acids at the thermostability-related sites makes no change in the above-described restriction sites for restriction en-



zymes; therefore, the expression vector obtained by incorporation into pUC19 has the same restriction map as that of pAD429 (Fig. 6, upper part) and the expression vector obtained by incorporation into pKK NE has the same restriction map as that of pAD455 (Fig. 6, lower part). These expression vectors were used for the transformation of *Escherichia coli* JM109 or HB101, which afforded the following transformants capable of producing a decarbamylase improved in thermostability.

E. coli JM109 pAD402 (FERM BP-3912),
E. coli JM109 pAD404 (FERM BP-3913),
E. coli JM109 pAD406 (FERM BP-3914),
E. coli JM109 pAD416 (FERM BP-3915),
E. coli JM109 pAD428,
E. coli JM109 pAD429 (FERM BP-4035),
E. coli JM109 pAD431,
E. coli JM109 pAD434,
E. coli JM109 pAD435,
E. coli JM109 pAD439,
E. coli JM109 pAD441,
E. coli JM109 pAD445,
E. coli JM109 pAD447,
E. coli JM109 pAD448,
E. coli JM109 pAD450,
E. coli JM109 pAD421,
E. coli JM109 pAD422,
E. coli JM109 pAD423,
E. coli JM109 pAD424 (FERM BP-4034),



E. coli JM109 pAD425,
E. coli JM109 pAD426,
E. coli JM109 pAD427,
E. coli JM109 pAD451,
E. coli JM109 pAD452,
E. coli JM109 pAD453,
E. coli JM109 pAD461,
E. coli JM109 pAD454,
E. coli JM109 pAD455 (FERM BP-4036),
E. coli JM109 pAD456,
E. coli JM109 pAD468,
E. coli JM109 pAD469,
E. coli JM109 pAD470 or
E. coli HB101 pNT4553 (FERM BP-4368).

For these decarbamylases of the improved type, the amount of such an enzyme of interest produced by a transformant can be increased by incorporating the corresponding gene downstream the strong promotor of a vector.

The transformant can be cultured on a conventional nutrient medium to express a transformant DNA which has been introduced thereinto. In case where the transformant DNA has been conferred with certain properties originating in the genetic or vector DNA, the medium may be supplemented with various agents depending upon the properties.

The transformant thus produced can be obtained as an enzyme source only by culture preparation on a conventional medium; if necessary, various treatments may be employed for enzyme induction, such as the addition of hydantoin com-



pounds, D-N-carbamoyl- α -amino acids, isopropyl-1-thio- β -D-galactoside (IPTG) or the like, and a temperature increase.

The medium to be used for the culture preparation of a transformant may usually be a conventional medium containing carbon sources, nitrogen sources and inorganic ions. If organic micronutrients such as vitamins and amino acids are added thereto, preferred results are often obtained. As the carbon sources, carbohydrates such as glucose and sucrose, organic acids such as acetic acid, alcohols and the like are suitably used. As the nitrogen sources, ammonia gas, ammonia water, ammonium salts and the like are used. As the inorganic salts, phosphate ion, magnesium ion, potassium ion, ferric ion and the like may be used.

The culture preparation may be conducted under aerobic conditions for 1 to 10 days while making an appropriate adjustment to the pH range of 4 to 8 and the temperature range of 25° to 45°C, which makes it possible to obtain desirable results. The enzyme produced by a transformant can be applied as a culture solution of the transformant, bacterial cells, treated bacterial cells, enzyme extracts from bacterial cells, immobilized bacterial cells or the like.

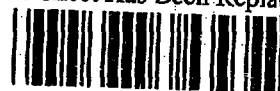
As the bacterial cells, there can be used any one of the forms such as an untreated culture solution after completion of the culture preparation, bacterial cells separated from the culture solution and washed bacterial cells. As the treated bacterial cells, there can be used lyophilized bacterial cells, acetone-dried bacterial cells, toluene- or surfactant-contacted bacterial cells, lysozyme-treated bacterial



cells, ultrasonicated bacterial cells, mechanically-disrupted bacterial cells, enzyme extracts having an enzyme activity to convert D-N-carbamoyl- α -amino acids into the corresponding D- α -amino acids by removal of the carbamoyl groups, which have been obtained from these treated bacterial cells; immobilization products of these bacterial cells; insolubilization products of the treated bacterial cells; immobilization products of the treated bacterial cells on a support for the immobilization of an enzyme protein (e.g., anion exchange resin), or the like. For the method for immobilization, reference is made to the specification of the Japanese Patent Laid-open Publication No. 63-185382.

As the support to be used for immobilization, various anion exchange resins having various amine-, ammonium salt- or diethanolamine-type functional groups are suitable, such as phenol-formaldehyde anion exchange resins, for example, Duolite A568 and DS17186 (Rohm & Haas; registered trade names), and polystyrene resins, for example, Amberlite IRA935, IRA945, IRA901 (Rohm & Haas; registered trade names), Lewatit OC1037 (Bayer; registered trade name) and Diaion EX-05 (Mitsubishi Chemical Industries; registered trade name). In addition to these, any other support such as DEAE-cellulose can also be used.

Further, to ensure firm and stable adsorption of an enzyme, a crosslinking agent is usually used, and a preferred example thereof is glutaraldehyde. As the enzyme to be used, there can be applied, in addition to purified enzymes, those which have been purified to various degrees, such as partial-

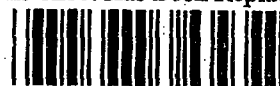


ly-purified enzymes, solutions containing disrupted bacterial cells and cell-free extracts.

For the preparation of an immobilized enzyme, conventional methods can be used, which comprises, for example, allowing an enzyme to be adsorbed from an enzyme solution on a support and then subjecting the support to crosslinking treatment.

D-N-carbamoyl- α -amino acids used as the substrate of an enzyme reaction in the present invention can be expressed by the formula: $R-CH(NHCONH_2)-COOH$, and their forms to be subjected to the reaction in actual cases are as follows: in case where the enzyme to be used has strict stereoselectivity on the D-N-carbamoyl- α -amino acids, they can be used as the D-form or a mixture of the D- and L-forms; and in case where the enzyme has loose stereoselectivity because of its action even on L-carbamoylamino acids, or in case where it is used as an enzyme mixture capable of acting even on the L-amino acids, it is preferred that they are used only in the D-form for the production of α -amino acids in the D-form.

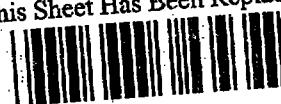
The substituent R can be selected within a wide range as described in the specifications of the Japanese Patent Publication Nos. 57-18793, 63-20520 and 1-48758. In particular, to provide useful compounds from an industrial point of view, such as intermediates for medicaments, it is preferred that R is phenyl, hydroxy-substituted phenyl, alkyl, substituted alkyl, aralkyl or thienyl. When R is hydroxy-substituted phenyl, it may be substituted with one or more hydroxy



groups at the o-, m- and/or p-positions, and the typical example thereof is p-hydroxyphenyl. The term alkyl refers to a group of 1 to 4 carbon atoms, the corresponding amino acid of which is D-alanine, D-valine, D-leucine, D-isoleucine or the like. The term substituted alkyl refers to an alkyl group of 1 to 4 carbon atoms, which is substituted with hydroxy, alkylthio, carboxyl, amino, phenyl, hydroxy-substituted phenyl, amide or the like, the corresponding amino acid of which is D-serine, D-threonine, D-methionine, D-cysteine, D-asparagine, D-glutamine, D-tyrosine, D-tryptophan, D-aspartic acid, D-glutamic acid, D-histidine, D-lysine, D-arginine, D-citrulline or the like. The term aralkyl refers to a group of 7 to 8 carbon atoms, such as benzyl or phenethyl, the corresponding amino acid of which is D-phenylalanine or the like.

As the aqueous medium, there can be used water, buffer or those which contain an organic solvent such as ethanol. If necessary, various additives such as nutrients needed for the growth of microorganisms, antioxidants, surfactants, coenzymes, hydroxylamine and metals can also be added to the aqueous medium.

In case where the bacterial cells of the above-described microorganism are cultured in an aqueous medium, during which the bacterial cells are brought into contact with D-N-carbamoyl- α -amino acids for the reaction, D-N-carbamoyl- α -amino acids and nutrients needed for the growth of the microorganism, such as carbon sources, nitrogen sources and inorganic ions, are contained in the aqueous medium used.



If organic micronutrients such as vitamins and amino acids are added thereto, preferred results are often obtained. As the nitrogen sources, carbohydrates such as glucose and sucrose, organic acids such as acetic acid, alcohols and the like are suitably used. As the nitrogen sources, ammonia gas, ammonia water, ammonium salts and the like are used. As the inorganic salts, phosphate ion, magnesium ion, potassium ion, ferric ion and the like may be used.

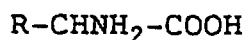
The culture preparation may be conducted under aerobic conditions for 1 to 10 days while making an appropriate adjustment to the pH range of 4 to 8 and the temperature range of 25°C to 45°C, which makes it possible that D-N-carbamoyl- α -amino acids are converted only into D- α -amino acids with high efficiency.

In contrast, when the reaction of an untreated culture solution of the above-described microorganism, cultured bacterial cells, treated bacterial cells, enzyme extracts, immobilization products of bacterial cells, insolubilization products of bacterial cells or immobilization products of enzyme proteins is effected in an aqueous medium containing D-N-carbamoyl- α -amino acids dissolved or suspended therein, the reaction system may be allowed to stand or stirred for a time, while the temperature is adjusted to an appropriate range of 10°C to 80°C and the pH is kept in the range of 4 to 9.5. In this manner, after the lapse of 5 to 100 hours, D- α -amino acids are produced in great amounts and accumulated in the aqueous medium. The D-N-carbamoyl- α -amino acids may be added in separated portions according as the reaction pro-



ceeds. The D- α -amino acids produced can be separated and purified by a conventional separation method.

The D- α -amino acids obtained herein can be expressed by the formula:



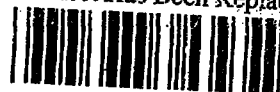
(wherein R is as defined above).

The following will describe the typical embodiments of the present invention. The detection and quantitative analysis of D- α -amino acids produced were conducted by high performance liquid chromatography (HPLC) or thin layer chromatography (TLC).

Example 1

Mutagenesis of *Agrabacterium radiobacter* KNK712 decarbamylase gene with hydroxylamine

Plasmid pAD108 having a KNK712 decarbamylase gene was digested with restriction enzymes HindIII and EcoRI, which was mixed and ligated with the digest of an M13mp18 double-stranded DNA with HindIII and EcoRI. This was transformed into *E. coli* JM109, which was mixed with 2 ml of H top agar medium (10 g/l bactotrypton, 8 g/l NaCl, 8 g/l bactoagar) containing 100 mM isopropyl- β -D-thiogalactopyranoside (IPTG) and 0.2% 5-bromo-4-chloro-3-indolyl- β -D-galactoside (X-Gal) added both in an amount of 40 μ l, and then plated on an H agar plate (10 g/l bactotrypton, 8 g/l NaCl, 12 g/l bactoagar), followed by incubation at 37°C and separation of a white plaque. This recombinant phage was cultured on 100 ml of 2YT medium (16 g/l bactotrypton, 10 g/l bactoyeast ex-

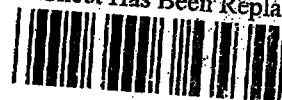


tract, 5 g/l NaCl), and a PEG-NaCl (20% polyethylene glycol 6000, 2.5 M NaCl) solution was added to the supernatant at a 1/5 volume, so that phage particles were precipitated and collected by centrifugation. These recombinant phage particles were treated in NH_2OH having a final concentration of 0.25 M (pH 6.0) at 37°C for 1 to 8 hours. The phage particles were precipitated with a PEG-NaCl solution, and then dissolved in sterilized water, with which *E. coli* JM109 was infected. The microorganism was cultured on 800 ml of 2YT medium, and a double-stranded DNA having some mutation was prepared by the alkaline-SDS method and CsCl ultracentrifugation method. This was digested with HindIII and EcoRI, and then incorporated into pUC19, which was transformed into *E. coli* JM109, and then plated on a 2YT agar plate containing 50 $\mu\text{g/ml}$ ampicillin. Thus, about 15,000 *E. coli* transformants each having a mutant decarbamylase gene were obtained.

Example 2

Mutagenesis of *Agrobacterium radiobacter* KNK712 decarbamylase gene with nitrous acid

The recombinant phage particles obtained by incorporating a KNK712 decarbamylase gene into M13mp18 as prepared in Example 1 were cultured on 800 ml of 2YT medium for preparation, and then subjected to phenol extraction and ethanol precipitation, which afforded a single-stranded phage DNA. This single-stranded DNA was treated with NaNO_2 having a final concentration of 0.9 M (pH 4.3) at 25°C for 30 minutes,



allowing a gene to cause some mutation. This was made into a double-stranded form by the use of Sequenase^R ver. 2.0 (United States Biochemical) and AMV reverse transcriptase (Life Science). This was digested with restriction enzymes HindIII and EcoRI, and incorporated into pUC19, which was transformed into E. coli JM109 and then plated on a 2YT plate (containing ampicillin). Thus, about 7600 E. coli transformants each having a mutant decarbamylase gene were obtained.

Example 3

Screening of decarbamylase-producing strain improved in thermostability

The colony of the E. coli transformants each having a mutant decarbamylase gene on the plate were replicated onto a filter paper (Toyo Roshi, 5C, ϕ 83 mm), which was soaked with 1.5 ml of lytic solution (20 mM Tris-HCl (pH 7.5), 10 mM EDTA, 2 mg/ml lysozyme, 1% Triton X-100), followed by a reaction at 37°C for 30 minutes, washing with water and drying. This filter paper was immersed in hot water at 65°C for thermal treatment, and after drying, it was soaked with 1 ml of color development reaction solution (30 mM K-phosphate buffer (pH 7.4), 0.3% carbamyl-D-phenylglycine, 0.25% phenol, 10 mg/ml D-amino acid oxidase (Sigma), 2.36 μ g/ml peroxidase (derived from horseradish, CALZYME Lab.), 0.1 mg/ml 4-aminoantipyrine), followed by a reaction at 37°C for 30 minutes. Colonies corresponding to the red-color developed spots were separated, as the strain improved in thermostability, from the original plate.



From 27,000 mutants as prepared by mutagenesis with hydroxylamine in Example 1 and 7600 mutants as prepared by mutagenesis with nitrous acid in Example 2, twelve and seven mutants improved in thermostability were obtained, respectively.

Example 4

Thermostability evaluation of decarbamylase improved in thermostability

A shaken culture containing one of four decarbamylase-producing strains improved in thermostability as obtained in Example 3 (three mutants obtained by mutagenesis with hydroxylamine and one mutant obtained by mutagenesis with nitrous acid), together with recombinant *E. coli* JM109 pAD108 (FERM BP-3184) having a KNK712 decarbamylase gene, was prepared in 10 ml of 2YT liquid medium (containing 50 µg/ml ampicillin and 1 mM IPTG) and incubated overnight. After harvested, the bacterial cells were washed with 0.1 M K-phosphate buffer (pH 7.0), and suspended in 1 ml of the same buffer, after which the suspension was disrupted with an ultrasonic disrupting apparatus (Tomy Seiko, model UR-20P) and the residue was removed by centrifugation. One part of this crude enzyme solution was subjected to thermal treatment at a temperature such as 55°C, 60°C, 65°C, 70°C and 75°C for 10 minutes, and some denatured insolubilized protein was removed by centrifugation. Then, the decarbamylase activity before and after the thermal treatment were measured. In the measurement, using 1% carbamyl-D-hydroxy-phenylglycine as the substrate, the reaction was effected in 0.1 M K-phosphate buffer (pH



7.0) at 40°C for 20 minutes, followed by protein denaturation with 5% TCA and removal thereof, after which the amount of D-hydroxyphenylglycine produced was determined by high performance liquid chromatography. The results are shown in Fig. 1.

Example 5

Gene analysis of decarbamylase improved in thermostability

The gene analysis was conducted for the mutant decarbamylase improved in thermostability, to presume the mutation sites of the decarbamylase protein. The plasmid having a gene for the decarbamylase improved in thermostability was reacted in a programmable incubator (ASTECH, model PC-700) using Taq Dye DeoxyTM Terminator Cycle Sequencing Kit (Applied Biosystems), followed by removal of excess Dye Deoxy using Bio Spin 30 (BIO-RAD). This sample was subjected to electrophoresis and data analysis with DNA sequencer model 373A (Applied Biosystems). As the result, the mutation sites were found as shown in Table 1. In every cases, the thermostability was improved by one amino acid mutation.

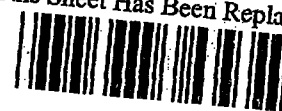


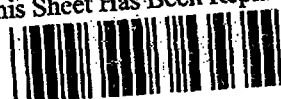
Table 1

Mutant	Thermostable temperature(t)	DNA mutation	Amino acid mutation
(712)	61.8	none	none
Escherichia coli			
JM109 pAD402	67.3	401C→T	57His→Tyr
JM109 pAD404	68.0	840C→T	203Pro→Leu
JM109 pAD406	66.5	839C→T	203Pro→Ser
JM109 pAD416	71.4	939T→C	236Val→Ala

Example 6

Immobilization of decarbamylase improved in thermostability on resin

A shaken culture containing one of four decarbamylase-producing strains improved in thermostability, together with recombinant E. coli JM109 pAD108 (FERM BP-3184) having a KNK712 decarbamylase gene, was prepared in 1 liter of 2YT medium (containing 50 µg/ml ampicillin and 1 mM IPTG) and incubated overnight. After harvested, the bacterial cells were washed with 0.1 M K-phosphate buffer (pH 7.0), and suspended in 100 ml of the same buffer, after which the suspension was disrupted with an ultrasonic disrupting apparatus (BRANSON, Sonifier model 250) and the residue was removed by centrifugation to yield a crude enzyme solution. To this, Duolite A-568 (Rohm & Haas) equilibrated with 0.1 M



K-phosphate buffer (pH 7.0) was added at a ratio of 1 g resin to 40 mg protein, and the mixture was stirred under a nitrogen seal at 4°C for 20 hours to ensure the adsorption. This enzyme-adsorbing resin was washed with 0.1 M K-phosphate buffer (pH 7.0) and 10 mM dithiothreitol (DTT), and reacted in 0.2% glutaraldehyde and 0.1 M K-phosphate buffer (pH 7.0) at 4°C for 10 minutes for the protein crosslinking, which afforded an immobilized enzyme-containing resin. The activities of the resins thus obtained are shown in Table 2.

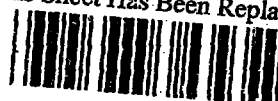
Table 2

Mutant	E. coli	E. coli	E. coli	E. coli	712
	JM109	JM109	JM109	JM109	
	pAD402	pAD404	pAD406	pAD416	
Activity (u/g resin)	12.4	9.4	10.0	14.6	9.6

Example 7

Repeated continuous reaction using immobilized decarbamylase

The mutant decarbamylase improved in thermostability was evaluated by the repeated continuous reaction using the immobilized enzyme-containing resin as obtained in Example 6. The reaction was effected (100 ml of reaction mixture) using 50 units of the immobilized decarbamylase and 3% carbamyl-D-HPG as the substrate at 40°C with stirring under a stream of nitrogen gas while the pH was adjusted to 7.0. The sample was taken for the activity measurement after 10 and 60 min-



utes, and the reaction was continued for 23.5 hours in total. The reaction mixture was removed by suction, after which another fresh reaction mixture was charged and allowed to react in the same manner as described above; such an operation was repeated 15 times, and the change in the activity of the immobilized decarbamylase was examined. The results are shown in Fig. 2. All the decarbamylases improved in thermostability were found to have improved stabilities when used in the reaction as the immobilized enzyme-containing resin, as compared with those before the mutagenesis.

Example 8

Amino acid replacement at thermostability-related sites

Various amino acids were substituted for three sites found to be related to thermostability, i.e., histidine which is 57-amino acid, proline which is 203-amino acid, and valine which is 236-amino acid, to prepare various derivatives. The preparation of these derivatives was conducted by the use of a method utilizing polymerase chain reaction (PCR) (W. Ito, et al., Gene, 102, 67-70 (1991)). For each of the thermostability-related sites, various synthetic primers were prepared with a mixed solution of A, T, G and C by the DNA synthesizer model 391 (Applied Biosystems), said primers having about 10-base sequence portions which were the same as the complimentary gene sequences on both sides of the gene portion corresponding to the amino acid at each site and also having any combination of A, T, G and C in the 3-base portion corresponding to the amino acid to be replaced, so that any



kind of amino acid was incorporated into this 3-base portion after the replacement. With the use of pAD108 as the template, and using this primer and M13RV primer (Takara Shuzo), the PCR was effected by the programmable incubator model PC-700 (ASTECH). A DNA fragment containing the full-length decarbamylase gene fragment (1785 bases) and short-length sequences on both sides and having a mutation at every HindIII restriction sites so as not to be cleaved was prepared by the PCR reaction using pAD108 and MUTF3 and M13M4 primers (both, Takara Shuzo). These two kinds of PCR products were mixed together, and heated at 94°C for 10 minutes, after which the mixture was gradually cooled for annealing, which afforded a combined form of these two kinds of DNAs. Using Taq DNA polymerase, single-stranded DNA portions were made into double-stranded forms. Using this DNA and M13M4 and M13RV primers, the PCR reaction was effected, and the DNA produced was simultaneously digested with HindIII and EcoRI. The digest was ligated with pUC19 which had been digested in the same manner as described above, and transformed into E. coli. According to the method as described in Example 3, only the strains improved in thermostability were screened, and it was found that derivatives having a substitution of leucine for 57-histidine, derivatives having a substitution of asparagine, glutamic acid, threonine, alanine, isoleucine or histidine for 203-proline, and derivatives having a substitution of threonine or serine for 236-valine had improved thermostability, as shown in Tables 3 and 4.

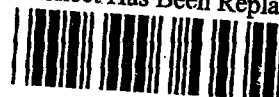


Table 3

Mutant	Thermostable temperature (°C)	DNA mutation	Amino acid mutation
(712)	61.8	none	none
Escherichia coli			
JN109 pAD434	67.5	402 A → T 403 T → A	57 His → Leu
JN109 pAD435	67.0	402 A → T	57 His → Leu
JN109 pAD431	67.0	839 C → A 840 C → A 841 T → C	203 Pro → Asn
JN109 pAD429	70.0	839 C → G 840 C → A 841 T → A	203 Pro → Glu
JN109 pAD445	67.5	839 C → A 841 T → C	203 Pro → Thr
JN109 pAD468	67.7	839 C → G	203 Pro → Ala
JN109 pAD469	67.2	839 C → A 840 C → T	203 Pro → Ile
JN109 pAD470	65.2	840 C → A	203 Pro → His

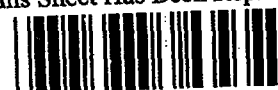


Table 4

Mutant	Thermostable temperature(°C)	DNA mutation	Amino acid mutation
(712)	61.8	none	none
Escherichia coli			
JM109 pAD428	70.0	939 T → C 940 G → C	236 Val → Ala
JM109 pAD439	71.5	939 T → C 940 G → T	236 Val → Ala
JM109 pAD441	71.5	938 G → A 939 T → G 940 G → T	236 Val → Ser
JM109 pAD447	72.0	938 G → T 939 T → C 940 G → A	236 Val → Ser
JM109 pAD448	72.0	938 G → T 939 T → C	236 Val → Ser
JM109 pAD450	69.5	938 G → A 939 C → C	236 Val → Thr



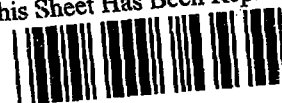
Example 9

Multiple mutation at thermostability-related site

Derivatives (multiple mutants) having a combination of two or three amino acid mutations which had been found to improve the thermostability, at three sites which had been found to be related to the thermostability, i.e., histidine which is 57-amino acid, proline which is 203-amino acid, and valine which is 236-amino acid, were prepared as follows.

The multiple mutants were prepared by replacing the mutation-free DNA fragments of a native gene with the restriction enzyme-digested DNA fragments each containing some mutation site of a mutant decarbamylase gene which had been improved in thermostability by a single amino acid substitution, as obtained in Example 3 or 8. First, because pAD108, pAD402 and other plasmids, which had the same restriction sites, each have two SalI and two SphI restriction sites for use in the replacement of the above-described DNA fragments, it is necessary to make these sites into a single restriction site. Therefore, as a new vector for gene incorporation, pKK NE was prepared, into which a decarbamylase gene fragment was incorporated, thereby constructing an expression vector to be used for the DNA fragment replacement.

Plasmid pKK NE was prepared as follows. Plasmid pKK233-2 (Pharmacia) shown in Fig. 3 was digested with EcoRI and NdeI, and subjected to agarose gel electrophoresis to separate a 2.7-kb DNA fragment, after which the cohesive end of the DNA fragment was changed to the blunt end with a DNA blunting kit (Takara Shuzo), followed by ligation and trans-



formation into *E. coli* JM109.

The *Nco*I restriction site of the vector thus prepared was changed to the *Nde*I restriction site using the PCR method. The resulting plasmid was digested with *Hind*III, and the cohesive end thereof was changed to the blunt end, followed by ligation with pEcoRI linker (Takara Shuzo), which afforded pKK NE as shown in Fig. 3. Next, into the initiation codon portion of the decarbamylase gene in plasmid pAD108 or pAD402, the *Nde*I restriction site was generated by the PCR method. The resulting *Nde*I-EcoRI DNA fragment of 1.6 kb was then incorporated into pKK NE, which afforded expression vectors pAD1086 (derived from pAD108; shown in Fig. 4) and pAD4021 (derived from pAD402; shown in Fig. 5).

The 57-amino acid mutation is positioned on the *Nde*I-SacI DNA fragment of about 190 bp; the 203-amino acid mutation on the *Sal*I-ClaI DNA fragment of about 170 bp (this DNA fragment is hereinafter referred to as fragment A); and the 236-amino acid mutation on the *Cla*I-SphI DNA fragment of about 75 bp (hereinafter referred to as fragment B). Therefore, fragment A of pAD4021 was removed, into which site the same fragment of pAD404 or pAD406 was incorporated, resulting in pAD421 or pAD422, respectively. Also, fragment B of pAD1086, pAD4021, pAD421 or pAD422 was replaced with the same fragment of pAD416, resulting in pAD4161, pAD423, pAD426 or pAD427, respectively. Further, fragment A of pAD4161 was replaced with the same fragment of pAD404, pAD406 or pAD429, resulting in pAD424, pAD425 or pAD461, respectively; fragment A of pAD402 or pAD423 was replaced with the same frag-



ment of pAD429, resulting in pAD451 or pAD455 (shown in Fig. 6), respectively; fragment B of pAD402, pAD429 (shown in Fig. 6), pAD451 or pAD421 was replaced with the same fragment of pAD447, resulting in pAD452, pAD453, pAD454 or pAD456, respectively.

The expression vectors prepared in the above-described manner were separately transformed into *E. coli* JM109, from which the extracts of disrupted bacterial cells were prepared and examined for the thermostability.

As shown in Tables 5 and 6, for the multiple mutants, it was found that the thermostability was additively improved according to the degree of an improvement in the thermostability attained by a single mutation. For the decarbamylase produced by *E. coli* JM109 pAD455 (FERM PB-4036) having highest thermostability, it was found that the thermostability was improved even by about 19°C, as compared with the decarbamylase before the mutagenesis.



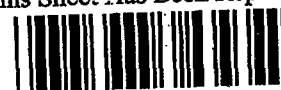
Table 5

Mutant	Amino acid mutation	Thermostable temperature(°C)
(712)	none	61.8
Escherichia coli		
JM109 pAD421	57His→Tyr 203Pro→Leu	71.5
JM109 pAD422	57His→Tyr 203Pro→Ser	70.0
JM109 pAD423	57His→Tyr 236Val→Ala	75.4
JM109 pAD424	203Pro→Leu 236Val→Ala	77.5
JM109 pAD425	203Pro→Ser 236Val→Ala	75.9
JM109 pAD426	57His→Tyr 203Pro→Leu 236Val→Ala	78.8
JM109 pAD427	57His→Tyr 203Pro→Ser 236Val→Ala	77.8



Table 6

Mutant	Amino acid mutation	Thermostable temperature(°C)
(712)	none	61.8
Escherichia coli		
JM109 pAD451	57His→Tyr 203Pro→Glu	74.0
JM109 pAD452	57His→Tyr 236Val→Ser	75.3
JM109 pAD453	203Pro→Glu 236Val→Ser	76.0
JM109 pAD461	203Pro→Glu 236Val→Ala	79.0
JM109 pAD454	57His→Tyr 203Pro→Glu 236Val→Ser	80.4
JM109 pAD455	57His→Tyr 203Pro→Glu 236Val→Ala	80.8
JM109 pAD456	57His→Tyr 203Pro→Leu 236Val→Ser	78.5



Example 10

Reaction temperature characteristics of decarbamylase improved in thermostability

A shaken culture containing a thermostability-improved decarbamylase-producing strain, *E. coli* JM109 pAD108 (FERM BP-3184) or *E. coli* JM109 pAD455 (FERM BP-4036), was prepared in 10 ml of 2YT medium and incubated overnight. After harvested, the bacterial cells were washed with 0.1 M K-phosphate buffer (pH 7.0), and then suspended in 1 ml of the same buffer, after which the suspension was disrupted with an ultrasonic disrupting apparatus (Tomy Seiko, model UR-20P) and the residue was removed by centrifugation to yield a crude enzyme solution. This crude enzyme solution was 10-fold diluted with a solution obtained by the addition of 5 mM dithiothreitol to the same buffer. Several 1 ml substrate solutions (1% carbamyl-D-p-hydroxyphenylglycine, 0.1% K-phosphate buffer (pH 6.5)) were kept at various temperatures of 30°C to 85°C for 3 minutes, respectively, to which 100 µl of the diluted crude enzyme solution (or when the temperature was in the range of 50°C to 80°C for higher thermostability (the temperature was in the range of 50°C to 65°C for the crude enzyme solution of *E. coli* JM109 pAD108), 100 µl of the enzyme solution which had been further 2- to 3-fold diluted) was added, followed by a reaction at the respective temperatures for 20 minutes. Then, the reaction was allowed to stop by adding 250 µl of 20% trichloroacetic acid solution, followed by centrifugation, after which the supernatant was analyzed by high performance liquid chromatography



(Nakalai tesque, Cosmosil 5C18-AR column). Taking the activity at 40°C as 100%, the relative activity at each temperature is shown in Fig. 7. As can be seen from Fig. 7, the enzyme produced by *E. coli* JM109 pAD455 exhibits the highest activity at a temperature around 75°C and has extremely improved stability, indicating that it was improved into an enzyme capable of acting at higher temperatures, as compared with the enzyme produced by *E. coli* JM109 pAD108.

Example 11

pH stability of decarbamylase improved in thermostability

A shaken culture containing a thermostability-improved decarbamylase-producing strain, *E. coli* JM109 pAD416 (FERM BP-3915) or *E. coli* JM109 pAD455 (FERM BP-4036), was prepared in 10 ml of 2YT medium and incubated overnight. After harvested, the bacterial cells were washed with 10 mM K-phosphate buffer (pH 7.0; containing 0.5 mM dithiothreitol), and suspended in 1 ml of the same buffer, after which the suspension was disrupted with a small-sized ultrasonic disrupting apparatus and the residue was removed by centrifugation to yield a crude enzyme solution. Various buffers for the respective pHs were prepared from 0.1 M K-phosphate buffer (pH 5.5, 6, 7, 8), Tris·HCl buffer (pH 7.5, 8, 9) and sodium carbonate buffer (pH 9, 10, 11). To 800 µl of each of these buffers, 200 µl of the crude enzyme solution was added, and the mixture was incubated at 40°C for 12.5 hours. Then, 100 µl of the mixture was added to 1 ml of substrate solution (1% carbamyl-D-p-hydroxyphenylglycine, 0.1% K-phosphate



buffer (pH 7.0)), followed by a reaction at 40°C, after which the reaction mixture was analyzed in the same manner as described in Example 10. Taking the activity at pH 7.0 as 100%, the relative activity of the treated sample at each pH is shown in Fig. 8.

Example 12

Preparation and expression of expression vector for decarbamylase improved in thermostability

(1) Preparation of vector for expression of foreign gene

First, a vector was prepared for the incorporation of a thermostability-improved decarbamylase gene and expression thereof (Fig. 9). From pUC19, HindIII-Cfr10I fragment of 1.3 kb was prepared by the PCR method, at which time as the PCR primer, a primer having a sequence designed so that an NdeI restriction site was formed at the initiation codon site of a lacZ gene inside the HindIII restriction site and the sequence could be further digested with HindIII and Cfr10I, was used for the reaction. This DNA fragment obtained by the PCR was digested with HindIII and Cfr10I, followed by ligation with 1.4-kb fragment obtained by removal of the corresponding 1.3-kb fragment from pUC19, resulting in plasmid pUC·Nde which corresponds pUC19 having an NdeI restriction site added thereto. Next, from pTrc99A (commercially available from Pharmacia), an EcoRI-EcoRV fragment of 0.6 kb was prepared by the PCR method, at which time as the PCR primer, a primer having a sequence designed so that an NdeI restriction site was substituted for the NcoI restriction site



inside the EcoRI restriction site and the sequence could be further digested with EcoRI and EcoRV, was used for the reaction. This DNA fragment obtained by the PCR was digested with EcoRI and EcoRV, followed by ligation with 3.6-kb fragment obtained by removal of the corresponding 0.6-kb fragment from pTrc99A, resulting in plasmid pTrcNde which corresponds pTrc99A having an NdeI restriction site added thereto. Then, a 2.0-kb fragment obtained by the digestion of pUC Nde with NdeI and SspI was ligated with a 0.6-kb fragment obtained by the digestion of pTrc Nde with NdeI and SspI, which afforded expression vector plasmid pUCNT for the expression of foreign genes (Fig. 10).

(2) Incorporation of thermostability-improved decarbamylase gene into vector

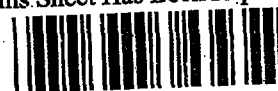
From pAD455 having a gene for one kind of decarbamylases improved in thermostability, only the gene portion coding for the decarbamylase was prepared by the PCR method, at which time as the PCR primer, a primer having a sequence designed so that an NdeI restriction site was formed in the initiation codon portion of the gene and a PstI restriction site was made just after the termination codon of the gene, was used for the reaction. A 0.9-kb DNA fragment obtained by this PCR was digested with NdeI and PstI, followed by ligation with pUCNT which was also digested with NdeI and PstI, which afforded expression vector plasmid pNT4553 (Fig. 10) for the decarbamylase improved in thermostability.

(3) Expression of thermostability-improved decarbamylase expression vector



The pNT4553 as prepared above was transformed into *Escherichia coli* HB101 by the calcium chloride method. A shaken culture of this transformant *E. coli* HB101 pNT4553 (FERM BP-4368) was prepared on a 2YT medium and incubated at 37°C for 16 hours. After the bacterial cells were harvested, a crude enzyme solution was prepared by the method as shown in (Example 10), and the decarbamylase activity was measured to be 5.6 unites/ml per culture solution.

According to the present invention, DNA fragments each coding for a decarbamylase protein improved in thermostability can be obtained, from which a decarbamylase having high reactivity to D-N-carbamoyl- α -amino acids and excellent stability can be produced, and with the use of this enzyme, D- α -amino acids can be produced with high efficiency.



"Sequence Listing"

SEQ ID NO: 1

SEQUENCE LENGTH: 1785

SEQUENCE TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

ORIGINAL SOURCE

ORGANISM: Agrobacterium radiobacter

STRAIN: KNK712 (FERM BP-1900)

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCCGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						

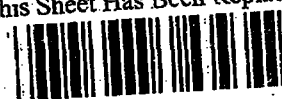
155

150						155						160								
GCC Ala	GCG Ala	AAA Lys 165	ATG Met	GGG Gly	ATG Met	TTC Phe	ATC Ile 170	TGC Cys	AAC Asn	GAT Asp	CGC Arg	CGC Arg 175	TGG Trp	CCT Pro	GAA Glu	766				
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	814				
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	CCT Pro	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	862				
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	910				
GGG Gly	GCC Ala	TGG Trp	TCC Ser 230	GCG Ala	GCC Ala	GCG Ala	GGC Gly	AAG Lys 235	GTG Val	GGC Gly	ATG Met	GAG Glu 240	GAG Glu	AAC Asn	TGC Cys	958				
ATG Met	CTG Leu 245	CTC Leu	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GCG Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	1006				
GCT Ala 260	CTC Leu	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	1054				
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	1102				
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu 300	ATC Ile	GCG Ala	GAA Glu	CTC Leu				1141				
TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201	CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261	GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACCTCGA	1321
CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAACCTG	TAGCCCATAT	1381	ATAGATTTC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441	GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501
CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561	GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621	ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681
ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCCGC	1741	TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785							

SEQ ID NO: 2
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD402 (FERM BP-3912)

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCGGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCTCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG TAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp Tyr Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTT GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

GCC Ala	GCG Ala	AAA Lys 165	ATG Met	GGG Gly	ATG Met	TTC Phe	ATC Ile 170	TGC Cys	AAC Asn	GAT Asp	CGC Arg 175	TGG Trp	CCT Pro	GAA Glu	766																																																												
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	814																																																											
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	CCT Pro	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	862																																																											
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	910																																																											
GGG Gly	GCC Ala	TGG Trp	TCC Ser 230	GCG Ala	GCC Ala	GCG Ala	GGC Gly	AAG Lys 235	GTG Val	GGC Gly	ATG Met	GAG Glu	GAG Glu 240	AAC Asn	TGC Cys	958																																																											
ATG Met	CTG Leu	CTC Leu 245	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GCG Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	1006																																																											
GCT Ala	CTC Leu 260	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	1054																																																											
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	1102																																																											
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu 300	ATC Ile	GCG Ala	GAA Glu	CTC Leu				1141																																																											
TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201	CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261	GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321	CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAAGTG	TAGCCCATAT	1381	ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441	GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501	CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561	GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621	ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681	ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCGGC	1741	TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC	1785



- 49 -

SEQ ID NO: 3
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD404 (FERM BP-3913)

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCGGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
				Thr Arg		
				1		
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

GCC Ala	GCG Ala	AAA Lys 165	ATG Met	GGG Gly	ATG Met	TTC Phe	ATC Ile 170	TGC Cys	AAC Asn	GAT Asp	CGC Arg	CGC Arg 175	TGG Trp	CCT Pro	GAA Glu	766																																																												
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	814																																																												
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	CTT Leu	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	862																																																												
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	910																																																												
GGG Gly	GCC Ala	TGG Trp	TCC Ser 230	GCG Ala	GCC Ala	GCG Ala	GGC Gly	AAG Lys 235	GTG Val	GGC Gly	ATG Met	GAG Glu 240	GAG Glu	AAC Asn	TGC Cys	958																																																												
ATG Met	CTG Leu	CTC Leu 245	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GCG Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	1006																																																												
GCT Ala	CTC Leu 260	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	1054																																																												
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	1102																																																												
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu	ATC Ile 300	GCG Ala	GAA Glu	CTC Leu				1141																																																												
TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201	CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261	GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAATCGA	1321	CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAAGT	TAGCCCATAT	1381	ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441	GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501	CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561	GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621	ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681	ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCTGGC	1741	TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785



SEQ ID NO: 4
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD406 (FERM BP-3914)

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCCGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

GCC Ala	GCG Ala	AAA Lys 165	ATG Met	GGG Gly	ATG Met	TTC Phe	ATC Ile 170	TGC Cys	AAC Asn	GAT Asp	CGC Arg	CGC Arg 175	TGG Trp	CCT Pro	GAA Glu	766																																																												
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	814																																																												
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	TCT Ser	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	862																																																												
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	910																																																												
GGG Gly	GCC Ala	TGG Trp	TCC Ser 230	GCG Ala	GCC Ala	GCG Ala	GGC Gly	AAG Lys 235	GTG Val	GGC Gly	ATG Met	GAG Glu	GAG Glu 240	AAC Asn	TGC Cys	958																																																												
ATG Met	CTG Leu	CTC Leu 245	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GCG Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	1006																																																												
GCT Ala	CTC Leu 260	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	1054																																																												
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	1102																																																												
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu 300	ATC Ile	GCG Ala	GAA Glu	CTC Leu				1141																																																												
TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201	CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261	GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAATCTGA	1321	CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAATGT	TAGCCCATAT	1381	ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441	GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501	CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561	GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621	ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681	ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCCGC	1741	TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785



- 53 -

SEQ ID NO: 5
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD416 (FERM BP-3915)

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCCGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA 766
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu
165 170 175

GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC 814
Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly
180 185 190

TAC	AAC	ACG	CCG	ACC	CAC	AAT	CCC	CCT	GTT	CCC	CAG	CAC	GAC	CAC	CTG	862
Tyr	Asn	Thr	Pro	Thr	His	Asn	Pro	Pro	Val	Pro	Gln	His	Asp	His	Leu	
195					200					205					210	

ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC 910
Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn
215 220 225

GGG GCC TGG TCC GCG GCC GCG GGC AAG GCG GGC ATG GAG GAG AAC TGC 958
Gly Ala Trp Ser Ala Ala Ala Gly Lys Ala Gly Met Glu Glu Asn Cys
230 235 240

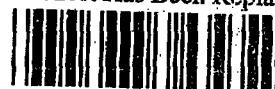
ATG CTG CTC GGC CAC TCC TGC ATC GTG GCG CCG ACC GGG GAA ATC GTC 1006
Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val
245 250 255

GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT 1054
Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp
260 265 270

CTC	GAT	CGC	TGC	CGG	GAA	CTG	CGT	GAA	CAC	ATC	TTC	AAC	TTC	AAG	CAG	1102
Leu	Asp	Arg	Cys	Arg	Glu	Leu	Arg	Glu	His	Ile	Phe	Asn	Phe	Lys	Gln	
275					280					285					290	

CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC 1141
His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu
295 300

TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201
CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATCG	1261
GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321
CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAAGTG	TAGCCCATAT	1381
ATAGATTTC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441
GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501
CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	TGCGGAGGAA	1561
GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACCTGAGAA	TGCATCGCC	1621
ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTC AACG	1681
ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCCGC	1741
TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785



SEQ ID NO: 6
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD428

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCGGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCTCG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA	CTT GCA GTG	GGA CAA CAA	GGT CCG ATC	GCG CGC GCG	GAG	286
Gln Met Ile	Leu Ala Val	Gly Gln Gln	Gly Pro Ile	Ala Arg Ala	Glu	
	5	10		15		
ACA CGC GAA	CAG GTC GTC	GTT CGT CTT	CTC GAC ATG	CTG ACG AAA	GCC	334
Thr Arg Glu	Gln Val Val	Val Arg Leu	Leu Asp Met	Leu Thr Lys	Ala	
	20	25	30			
GCG AGC CGG	GGC GCG AAT	TTC ATT GTC	TTC CCC GAA	CTC GCG CTT	ACG	382
Ala Ser Arg	Gly Ala Asn	Phe Ile Val	Phe Pro Glu	Leu Ala Leu	Thr	
	35	40	45	50		
ACC TTC TTC	CCG CGC TGG	CAT TTC ACC	GAC GAG GCC	GAG CTC GAT	AGC	430
Thr Phe Phe	Pro Arg Trp	His Phe Thr	Asp Glu Ala	Glu Leu Asp	Ser	
	55	60	65			
TTC TAT GAG	ACC GAA ATG	CCC GGC CCG	GTG GTC CGT	CCA CTC TTT	GAG	478
Phe Tyr Glu	Thr Glu Met	Pro Gly Pro	Val Val Arg	Pro Leu Phe	Glu	
	70	75	80			
AAG GCC GCG	GAA CTC GGG	ATC GGC TTC	AAT CTG GGC	TAC GCT GAA	CTC	526
Lys Ala Ala	Glu Leu Gly	Ile Gly Phe	Asn Leu Gly	Tyr Ala Glu	Leu	
	85	90	95			
GTC GTC GAA	GGC GGC GTC	AAG CGT CGC	TTC AAC ACG	TCC ATT TTG	GTG	574
Val Val Glu	Gly Gly Val	Lys Arg Arg	Phe Asn Thr	Ser Ile Leu	Val	
	100	105	110			
GAT AAG TCA	GGC AAG ATC	GTC GGC AAG	TAT CGT AAG	ATC CAT TTG	CCG	622
Asp Lys Ser	Gly Lys Ile	Val Gly Lys	Tyr Arg Lys	Ile His Leu	Pro	
	115	120	125	130		
GGT CAC AAG	GAG TAC GAG	GCC TAC CGG	CCG TTC CAG	CAT CTT GAA	AAG	670
Gly His Lys	Glu Tyr Glu	Ala Tyr Arg	Pro Phe Gln	His Leu Glu	Lys	
	135	140	145			
CGT TAT TTC	GAG CCG GGC	GAT CTC GGC	TTC CCG GTC	TAT GAC GTC	GAC	718
Arg Tyr Phe	Glu Pro Gly	Asp Leu Gly	Phe Pro Val	Tyr Asp Val	Asp	
	150	155	160			

GCC Ala	CGC Ala	AAA Lys 165	ATG Met	GGG Gly	ATG Met	TTC Phe	ATC Ile 170	TGC Cys	AAC Asn	GAT Asp	CGC Arg	CGC Arg 175	TGG Trp	CCT Pro	GAA Glu	766																																																											
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	814																																																											
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	CCT Pro	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	862																																																											
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	910																																																											
GGG Gly	GCC Ala	TGG Trp	TCC Ser 230	GCG Ala	GCC Ala	GCG Ala	GGC Gly	AAG Lys 235	GCC Ala	GGC Gly	ATG Met	GAG Glu	GAG Glu 240	AAC Asn	TGC Cys	958																																																											
ATG Met	CTG Leu	CTC Leu 245	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GCG Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	1006																																																											
GCT Ala	CTC Leu 260	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	1054																																																											
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	1102																																																											
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu	ATC Ile 300	GCG Ala	GAA Glu	CTC Leu				1141																																																											
TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201	CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261	GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321	CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAACGT	TAGCCCATAT	1381	ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441	GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501	CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561	GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621	ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681	ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTTCGGC	1741	TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC	1785



SEQ ID NO: 7
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD429

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCGGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCTCG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG						286
Gln Met Ile Leu Ala val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
	5		10		15	
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC						334
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
	20		25		30	
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG						382
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
	35		40		45	50
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC						430
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser						
		55		60		65
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG						478
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
	70		75		80	
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC						526
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
	85		90		95	
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG						574
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
	100		105		110	
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG						622
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
	115		120		125	130
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG						670
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
		135		140		145
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC						718
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
	150		155		160	

GCC Ala	GCG Ala	AAA Lys 165	ATG Met	GGG Gly	ATG Met	TTC Phe	ATC Ile 170	TGC Cys	AAC Asn	GAT Asp	CGC Arg 175	CGC Arg	TGG Trp	CCT Pro	GAA Glu	766																																																											
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	814																																																											
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	GAA Glu	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	862																																																											
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	910																																																											
GGG Gly	GCC Ala	TGG Trp 230	TCC Ser	GCG Ala	GCC Ala	GCG Ala	GGC Gly	AAG Lys 235	GTG Val	GGC Gly	ATG Met	GAG Glu 240	GAG Glu	AAC Asn	TGC Cys	958																																																											
ATG Met	CTG Leu 245	CTC Leu	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GCG Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	1006																																																											
GCT Ala 260	CTC Leu	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	1054																																																											
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	1102																																																											
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu 300	ATC Ile	GCG Ala	GAA Glu	CTC Leu				1141																																																											
TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201	CCAGGGTGAC	GCTGGTGGAA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261	GGGTGCGCCG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321	CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAACGT	TAGCCCATAT	1381	ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441	GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501	CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561	GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621	ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681	ACGCTCGGGA	TATCCAAGAG	CCACTGTAC	TCCATCCTCA	AGACGCTGAC	GCATTTCGGC	1741	TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC	1785

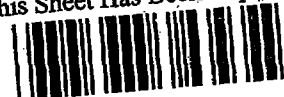


- 59 -

SEQ ID NO: 8
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD431

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCCGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

GCC Ala	GCG Ala	AAA Lys 165	ATG Met	GGG Gly	ATG Met	TTC Phe	ATC Ile 170	TGC Cys	AAC Asn	GAT Asp	CGC Arg	CGC Arg 175	TGG Trp	CCT Pro	GAA Glu	766																																																												
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	814																																																												
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	AAC Asn	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	862																																																												
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	910																																																												
GGG Gly	GCC Ala	TGG Trp	TCC Ser 230	GCG Ala	GCC Ala	GCG Ala	GGC Gly	AAG Lys 235	GTG Val	GGC Gly	ATG Met	GAG Glu	GAG Glu 240	AAC Asn	TGC Cys	958																																																												
ATG Met	CTG Leu	CTC Leu 245	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GCG Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	1006																																																												
GCT Ala	CTC Leu 260	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	1054																																																												
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	1102																																																												
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu	ATC Ile 300	GCG Ala	GAA Glu	CTC Leu				1141																																																												
TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201	CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261	GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321	CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAACGT	TAGCCCATAT	1381	ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441	GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501	CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561	GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACCTGAGAA	TGCGATCGCC	1621	ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681	ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTGCGC	1741	TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785



SEQ ID NO: 9
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD434

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCCGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG CTA TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp Leu Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA 766
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu
165 170 175

GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC 814
Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly
180 185 190

TAC	AAC	ACG	CCG	ACC	CAC	AAT	CCC	CCT	GTT	CCC	CAG	CAC	GAC	CAC	CTG	862
Tyr	Asn	Thr	Pro	Thr	His	Asn	Pro	Pro	Val	Pro	Gln	His	Asp	His	Leu	
195					200					205					210	

ACG	TCC	TTC	CAC	CAT	CTC	CTA	TCG	ATG	CAG	GCC	GGG	TCT	TAT	CAG	AAC	910
Thr	Ser	Phe	His	His	Leu	Leu	Ser	Met	Gln	Ala	Gly	Ser	Tyr	Gln	Asn	
				215					220					225		

GGG GCC TGG TCC GCG GCC GCG GGC AAG GTG GGC ATG GAG GAG AAC TGC 958
Gly Ala Trp Ser Ala Ala Ala Gly Lys Val Gly Met Glu Glu Asn Cys
230 235 240

ATG	CTG	CTC	GGC	CAC	TCC	TGC	ATC	GTG	GCG	CCG	ACC	GGG	GAA	ATC	GTC	1006
Met	Leu	Leu	Gly	His	Ser	Cys	Ile	Val	Ala	Pro	Thr	Gly	Glu	Ile	Val	
		245					250					255				

GCT	CTC	ACT	ACG	ACG	CTG	GAA	GAC	GAG	GTG	ATC	ACC	GCC	GCC	GTC	GAT	1054
Ala	Leu	Thr	Thr	Thr	Leu	Glu	Asp	Glu	Val	Ile	Thr	Ala	Ala	Val	Asp	
260						265					270					

CTC	GAT	CGC	TGC	CGG	GAA	CTG	CGT	GAA	CAC	ATC	TTC	AAC	TTC	AAG	CAG	1102
Leu	Asp	Arg	Cys	Arg	Glu	Leu	Arg	Glu	His	Ile	Phe	Asn	Phe	Lys	Gln	
275					280					285					290	

CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC 1141
His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu
295 300

TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201
CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261
GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321
CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAAGCTG	TAGCCCATAT	1381
ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441
GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501
CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561
GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621
ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681
ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTTCGGC	1741
TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785



- 63 -

SEQ ID NO: 10
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD435

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCCGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CCG GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG CTT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp Leu Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA 766
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu
165 170 175

GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC 814
Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly
180 185 190

TAC	AAC	ACG	CCG	ACC	CAC	AAT	CCC	CCT	GTT	CCC	CAG	CAC	GAC	CAC	CTG	862
Tyr	Asn	Thr	Pro	Thr	His	Asn	Pro	Pro	Val	Pro	Gln	His	Asp	His	Leu	
195					200					205					210	

ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC 910
Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn
215 220 225

GGG GCC TGG TCC GCG GCC GCG GGC AAG GTG GGC ATG GAG GAG AAC TGC 958
Gly Ala Trp Ser Ala Ala Ala Gly Lys Val Gly Met Gln Glu Asn Cys
230 235 240

ATG	CTG	CTC	GGC	CAC	TCC	TGC	ATC	GTG	GCG	CCG	ACC	GGG	GAA	ATC	GTC	1006
Met	Leu	Leu	Gly	His	Ser	Cys	Ile	Val	Ala	Pro	Thr	Gly	Glu	Ile	Val	
		245					250					255				

GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT 1054
Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp
260 265 270

CTC	GAT	CGC	TGC	CGG	GAA	CTG	CGT	GAA	CAC	ATC	TTC	AAC	TTC	AAG	CAG	1102
Leu	Asp	Arg	Cys	Arg	Glu	Leu	Arg	Glu	His	Ile	Phe	Asn	Phe	Lys	Gln	
275					280					285					290	

CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC 1141
His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu
295 300

TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201
CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261
GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321
CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAAC TG	TAGCCCATAT	1381
ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441
GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501
CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561
GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621
ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681
ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCGGC	1741
TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785



SEQ ID NO: 11
SEQUENCE LENGTH: 1785
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE
ORGANISM: Escherichia coli
STRAIN: JM109 pAD439

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCCGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA 766
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu
165 170 175

GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC 814
Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly
180 185 190

TAC	AAC	ACG	CCG	ACC	CAC	AAT	CCC	CCT	GTT	CCC	CAG	CAC	GAC	CAC	CTG	862
Tyr	Asn	Thr	Pro	Thr	His	Asn	Pro	Pro	Val	Pro	Gln	His	Asp	His	Leu	
195					200					205					210	

ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC 910
Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn
215 220 225

GGG GCC TGG TCC GCG GCC GCG GGC AAG GCT GGC ATG GAG GAG AAC TGC 958
Gly Ala Trp Ser Ala Ala Ala Gly Lys Ala Gly Met Glu Glu Asn Cys
230 235 240

ATG CTG CTC GGC CAC TCC TGC ATC GTG GCG CCG ACC GGG GAA ATC GTC 1006
Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val
245 250 255

GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT 1054
Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp
260 265 270

CTC	GAT	CGC	TGC	CGG	GAA	CTG	CGT	GAA	CAC	ATC	TTC	AAC	TTC	AAG	CAG	1102
Leu	Asp	Arg	Cys	Arg	Glu	Leu	Arg	Glu	His	Ile	Phe	Asn	Phe	Lys	Gln	
275					280					285					290	

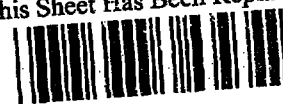
CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC 1141
His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu
295 300

TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCAACA	TCCAGGCGCG	1201
CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTACG	GGCGATCCGC	AAGGAGATGC	1261
GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321
CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAAGTC	TAGCCCATAT	1381
ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441
GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501
CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561
GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621
ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681
ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCCGG	1741
TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785

SEQ ID NO: 12
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD441

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCCGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCTCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

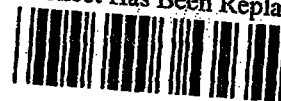
GCC Ala	CGC Ala	AAA Lys 165	ATG Met	GGG Gly	ATG Met	TTC Phe	ATC Ile 170	TGC Cys	AAC Asn	GAT Asp	CGC Arg 175	TGG Trp	CCT Pro	GAA Glu	766																																																												
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	814																																																											
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	CCT Pro	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	862																																																											
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	910																																																											
GGG Gly	GCC Ala	TGG Trp 230	TCC Ser	GCG Ala	GCC Ala	GCG Ala	GGC Gly	AAG Lys 235	AGT Ser	GGC Gly	ATG Met	GAG Glu 240	GAG Glu	AAC Asn	TGC Cys	958																																																											
ATG Met	CTG Leu	CTC Leu 245	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GCG Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	1006																																																											
GCT Ala	CTC Leu 260	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	1054																																																											
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	1102																																																											
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu	ATC Ile 300	GCG Ala	GAA Glu	CTC Leu				1141																																																											
TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201	CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261	GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321	CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAACGT	TAGCCCATAT	1381	ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441	GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501	CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561	GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621	ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681	ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCGGC	1741	TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC	1785



SEQ ID NO: 13
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD445

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCGGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTC	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG						286
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
	5		10		15	
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC						334
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
	20		25		30	
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG						382
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
	35		40		45	50
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC						430
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser						
		55		60		65
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG						478
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
	70		75		80	
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC						526
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
	85		90		95	
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG						574
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
	100		105		110	
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG						622
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
	115		120		125	130
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG						670
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
		135		140		145
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC						718
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
	150		155		160	

GCC Ala	GCG Ala	AAA Lys 165	ATG Met	GGG Gly	ATG Met	TTC Phe	ATC Ile 170	TGC Cys	AAC Asn	GAT Asp	CGC Arg	CGC Arg 175	TGG Trp	CCT Pro	GAA Glu	766																																																											
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	814																																																											
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	ACC Thr	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	862																																																											
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	910																																																											
GGG Gly	GCC Ala	TGG Trp	TCC Ser 230	GCG Ala	GCC Ala	GCG Ala	GGC Gly	AAG Lys 235	GTG Val	GGC Gly	ATG Met	GAG Gln 240	GAG Glu	AAC Asn	TGC Cys	958																																																											
ATG Met	CTG Leu	CTC Leu 245	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GCG Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	1006																																																											
GCT Ala	CTC Leu 260	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	1054																																																											
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	1102																																																											
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu	ATC Ile 300	GCG Ala	GAA Glu	CTC Leu				1141																																																											
TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201	CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261	GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321	CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAACTG	TAGCCCATAT	1381	ATAGATTTC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441	GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501	CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561	GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621	ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681	ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCCGC	1741	TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC	1785



- 71 -

SEQ ID NO: 14
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD447

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCCGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCTCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA 766
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu
165 170 175

GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC 814
Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly
180 185 190

TAC	AAC	ACG	CCG	ACC	CAC	AAT	CCC	CCT	GTT	CCC	CAG	CAC	GAC	CAC	CTG	862
Tyr	Asn	Thr	Pro	Thr	His	Asn	Pro	Pro	Val	Pro	Gln	His	Asp	His	Leu	
195					200					205					210	

ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC 910
Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn
215 220 225

GGG GCC TGG TCC GCG GCC GCG GGC AAG TCA GGC ATG GAG GAG AAC TGC 958
Gly. Ala Trp Ser Ala Ala Ala Gly Lys Ser Gly Met Glu Glu Asn Cys
230 235 240

ATG CTG CTC GGC CAC TCC TGC ATC GTG GCG CCG ACC GGG GAA ATC GTC 1006
Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val
245 250 255

GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT 1054
Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp
260 265 270

CTC	GAT	CGC	TGC	CGG	GAA	CTG	CGT	GAA	CAC	ATC	TTC	AAC	TTC	AAG	CAG	1102
Leu	Asp	Arg	Cys	Arg	Glu	Leu	Arg	Glu	His	Ile	Phe	Asn	Phe	Lys	Gln	
275					280					285					290	

CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC 1141
His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu
295 300

TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTACACA	TCCAGGCGCG	1201
CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261
GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321
CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAAGTG	TAGCCCATAT	1381
ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441
GTGCCCAGAT	CAAAGAAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501
CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561
GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621
ATCGTCGATT	ACATCAACGG	TACACGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681
ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTTCGGC	1741
TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785



- 73 -

SEQ ID NO: 15
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD448

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCGGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
				Thr Arg		
				1		
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA 766
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu
165 170 175

GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC 814
Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly
180 185 190

TAC	AAC	ACG	CCG	ACC	CAC	AAT	CCC	CCT	GTT	CCC	CAG	CAC	GAC	CAC	CTG	862
Tyr	Asn	Thr	Pro	Thr	His	Asn	Pro	Pro	Val	Pro	Gln	His	Asp	His	Leu	
195					200					205					210	

ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC 910
Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn
215 220 225

GGG GCC TGG TCC GCG GCC GCG GGC AAG TCG GGC ATG GAG GAG AAC TGC 958
Gly Ala Trp Ser Ala Ala Ala Gly Lys Ser Gly Met Glu Glu Asn Cys
230 235 240

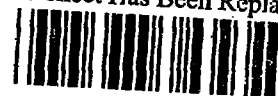
ATG CTG CTC GGC CAC TCC TGC ATC GTG GCG CCG ACC GGG GAA ATC GTC 1006
Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val
245 250 255

GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT 1054
Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp
260 265 270

CTC	GAT	CGC	TGC	CGG	GAA	CTG	CGT	GAA	CAC	ATC	TTC	AAC	TTC	AAG	CAG	1102
Leu	Asp	Arg	Cys	Arg	Glu	Leu	Arg	Glu	His	Ile	Phe	Asn	Phe	Lys	Gln	
275					280					285					290	

CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC 1141
His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu
295 300

TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201
CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261
GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321
CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAAC TG	TAGCCCATAT	1381
ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441
GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501
CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561
GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621
ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681
ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTTCGGC	1741
TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785



SEQ ID NO: 16
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD450

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCCGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

GCC Ala	GCG Ala	AAA Lys 165	ATG Met	GGG Gly	ATG Met	TTC Phe	ATC Ile 170	TGC Cys	AAC Asn	GAT Asp	CGC Arg	CGC Arg 175	TGG Trp	CCT Pro	GAA Glu	766																																																												
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	814																																																												
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	CCT Pro	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	862																																																												
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	910																																																												
GGG Gly	GCC Ala	TGG Trp	TCC Ser 230	GCG Ala	GCC Ala	GCG Ala	GGC Gly	AAG Lys 235	ACG Thr	GGC Gly	ATG Met	GAG Glu 240	GAG Glu	AAC Asn	TGC Cys	958																																																												
ATG Met	CTG Leu	CTC Leu 245	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GCG Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	1006																																																												
GCT Ala	CTC Leu 260	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	1054																																																												
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	1102																																																												
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu 300	ATC Ile	GCG Ala	GAA Glu	CTC Leu				1141																																																												
TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201	CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261	GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACCTCGA	1321	CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAACCTG	TAGCCCATAT	1381	ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441	GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501	CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561	GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621	ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681	ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTGCGC	1741	TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785



- 77 -

SEQ ID NO: 17
 SEQUENCE LENGTH: 1559
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD421

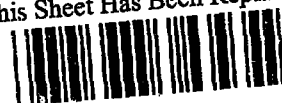
	CATA TG ACA CGT	12
	Thr Arg	
	1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	60	
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu		
5 10 15		
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	108	
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala		
20 25 30		
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	156	
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr		
35 40 45 50		
ACC TTC TTC CCG CGC TGG TAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	201	
Thr Phe Phe Pro Arg Trp Tyr Phe Thr Asp Glu Ala Glu Leu Asp Ser		
55 60 65		
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	252	
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu		
70 75 80		
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	300	
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu		
85 90 95		
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	348	
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val		
100 105 110		
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	396	
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro		
115 120 125 130		
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	444	
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys		
135 140 145		
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	492	
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp		
150 155 160		
GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA	540	
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu		



- 78 -

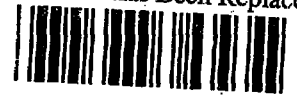
165	170	175	
GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly 180 185 190			588
TAC AAC ACG CCG ACC CAC AAT CCC CTT GTT CCC CAG CAC GAC CAC CTG Tyr Asn Thr Pro Thr His Asn Pro Leu Val Pro Gln His Asp His Leu 195 200 205 210			636
ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn 215 220 225			684
GGG GCC TGG TCC GCG GCC GCG GGC AAG GTG GGC ATG GAG GAG AAC TGC Gly Ala Trp Ser Ala Ala Ala Gly Lys Val Gly Met Glu Glu Asn Cys 230 235 240			732
ATG CTG CTC GGC CAC TCC TGC ATC GTG GCG CCG ACC GGG GAA ATC GTC Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val 245 250 255			780
GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp 260 265 270			828
CTC GAT CGC TGC CGG GAA CTG CGT GAA CAC ATC TTC AAC TTC AAG CAG Leu Asp Arg Cys Arg Glu Leu Arg Glu His Ile Phe Asn Phe Lys Gln 275 280 285 290			876
CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu 295 300			915
TGAGGTTGCC GAAAAGCATG TGTGTCGTTG TTCTCGGCGC CTGGGTCACA TCCAGGCGCG CCAGGGTGAC GCTGGTGGAA TAGTACCACG ACCGCTTCAG GGCGATCCGC AAGGAGATGC GGGTCGCCGG AGCGGCAAAG CCCGACATTC GTTTCGCACC GACGGCCGTC GTGAACTCGA CAGTCCGCGA GAAGGGCGTA TTGCGCGGCC TGGACCTGTA CGTGGAAGTGT TAGCCCATAT ATAGATTTCC AAAGAGTTTC GGCGAGGCGC GGCGCGCCTA GCCCCATGTG AGCGAGAACC GTGCCCAGAT CAAAGAATGA GACCGACGCG CCGGCCGCGG CAAAGGATGA TCCTCAGGGT CGGATCTATC GCTCCGCCCT GAAGCAGGAG GGCGCACGCT GGCTGCTGAC GGCGGAGGAA GGGTTGCTGG CAAAGCCCAA GCCGCCCGGC CTTGTTCCGG CACTTGAGAA TGCGATCGCC ATCGTCGATT ACATCAACGG TACACCGCCC CATATCGCGT CCCTGGCGGA GCTTTCAACG ACGCTCGGGA TATCCAAGAG CCACTGTCAC TCCATCCTCA AGACGCTGAC GCATTTCTGGC TGGCTGAAAT TCGACAATCG CTCAAAGAGC TACGAGCTGA ATTC			975 1035 1095 1155 1215 1275 1335 1395 1455 1515 1559

This Sheet Has Been Replaced



SEQ ID NO: 18
SEQUENCE LENGTH: 1559
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE
ORGANISM: Escherichia coli
STRAIN: JM109 pAD422

	CATA TG ACA CGT	12
	Thr Arg	
	1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	60	
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu		
5 10 15		
ACA CGC GAA CAG GTC GTC GTC CGT CTT CTC GAC ATG CTG ACG AAA GCC	108	
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala		
20 25 30		
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	156	
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr		
35 40 45 50		
ACC TTC TTC CCG CGC TGG TAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	201	
Thr Phe Phe Pro Arg Trp Tyr Phe Thr Asp Glu Ala Glu Leu Asp Ser		
55 60 65		
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	252	
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu		
70 75 80		
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	300	
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu		
85 90 95		
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	348	
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val		
100 105 110		
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	396	
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro		
115 120 125 130		
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	444	
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys		
135 140 145		
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	492	
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp		
150 155 160		
GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA	540	
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu		



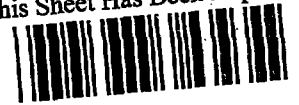
- 80 -

165	170	175	
GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly 180 185 190			588
TAC AAC ACG CCG ACC CAC AAT CCC TCT GTT CCC CAG CAC GAC CAC CTG Tyr Asn Thr Pro Thr His Asn Pro Ser Val Pro Gln His Asp His Leu 195 200 205 210			636
ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn 215 220 225			684
GGG GCC TGG TCC GCG GCC GCG GGC AAG GTG GGC ATG GAG GAG AAC TGC Gly Ala Trp Ser Ala Ala Ala Gly Lys Val Gly Met Glu Glu Asn Cys 230 235 240			732
ATG CTG CTC GGC UAC TCC TGC ATC GTG GCG TCG ACC GGG GAA ATC GTC Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val 245 250 255			780
GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp 260 265 270			828
CTC GAT CGC TGC CGG GAA CTG CGT GAA CAC ATC TTC AAC TTC AAG CAG Leu Asp Arg Cys Arg Glu Leu Arg Glu His Ile Phe Asn Phe Lys Gln 275 280 285 290			876
CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu 295 300			915
TGAGGTTGCC GAAAAGCATG TGTGTCGTTG TTCTCGGCGC CTGGGTCACA TCCAGGCGCG CCAGGGTGAC GCTGGTGGA TAGTACCACG ACCGCTTCAG GCGATCCGC AAGGAGATGC GGGTCGCCGG AGCGGCAAAG CCCGACATTC GTTTCGCACC GACGGCCGTC GTGAACCTGA CAGTCCGCGA GAAGGGCGTA TTGCGCGGCC TGGACCTGTA CGTGGAACCTG TAGCCCATAT ATAGATTTC AAAGAGTTTC GGCGAGGCGC GGCGCGCCTA GCCCCATGTG AGCGAGAACC GTGCCCAGAT CAAAGAATGA GACCGACGCG CCGGCCGCGG CAAAGGATGA TCCTCAGGGT CGGATCTATC GCTCCGCCCT GAAGCAGGAG GGCGCACGCT GGCTGCTGAC GCGGAGGAA GGGTTGCTGG CAAAGCCCAA GCCGCCCGGC CTTGTTCCGG CACTTGAGAA TGCGATCGCC ATCGTCGATT ACATCAACGG TACACCGCCC CATATCGCGT CCCTGGCGGA GCTTTCAACG ACGCTCGGGA TATCCAAGAG CCACTGTCAC TCCATCCTCA AGACGCTGAC GCATTTCCGGC TGGCTGAAAT TCGACAATCG CTCAAAGAGC TACGAGCTGA ATTC			975 1035 1095 1155 1215 1275 1335 1395 1455 1515 1559



SEQ ID NO: 19
 SEQUENCE LENGTH: 1559
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD423

	CATA	TG	ACA	CGT	
	Thr	Arg			
	1				
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG					60
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu					
	5			10	15
ACA CGC GAA CAG GTC GTC GTC CGT CTT CTC GAC ATG CTG ACG AAA GCG					108
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala					
	20			25	30
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG					156
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr					
	35			40	45
ACC TTC TTC CCG CGC TGG TAT TTC ACC GAC GAG GCC GAG CTC GAT AGC					201
Thr Phe Phe Pro Arg Trp Tyr Phe Thr Asp Glu Ala Glu Leu Asp Ser					
		55		60	65
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG					252
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu					
		70		75	80
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC					300
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu					
		85		90	95
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG					348
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val					
	100			105	110
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG					396
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro					
	115			120	125
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG					444
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys					
		135		140	145
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC					492
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp					
		150		155	160
GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA					540
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu					



- 82 -

165	170	175	
GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly 180 185 190			588
TAC AAC ACG CCG ACC CAC AAT CCC CCT GTT CCC CAG CAC GAC CAC CTG Tyr Asn Thr Pro Thr His Asn Pro Pro Val Pro Gln His Asp His Leu 195 200 205 210			636
ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn 215 220 225			684
GGG GCC TGG TCC GCG GCC GCG GGC AAG GCG GGC ATG GAG GAG AAC TGC Gly Ala Trp Ser Ala Ala Ala Gly Lys Ala Gly Met Glu Glu Asn Cys 230 235 240			732
ATG CTG CTC GGC CAC TCC TGC ATC GTG GCG CCG ACC GGG GAA ATC GTC Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val 245 250 255			780
GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp 260 265 270			828
CTC GAT CGC TGC CGG GAA CTG CGT GAA CAC ATC TTC AAC TTC AAG CAG Leu Asp Arg Cys Arg Glu Leu Arg Glu His Ile Phe Asn Phe Lys Gln 275 280 285 290			876
CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu 295 300			915
TGAGGTTGCC GAAAAGCATG TGTGTCGTTG TTCTCGGCGC CTGGGTCACA TCCAGGCGCG CCAGGGTGAC GCTGGTGGAA TAGTACCACG ACCGCTTCAG GGCGATCCGC AAGGAGATGC GGGTCGCCGG AGCGGCAAAG CCCGACATTC GTTTCGCACC GACGGCCGTC GTGAACTCGA CAGTCCGCGA GAAGGGCGTA TTGCGCGGCC TGGACCTGTA CGTGGAAGTGTAGCCCATAT ATAGATTTCC AAAGAGTTTC GGCGAGGCGC GGCGCGCCTA GCCCCATGTG AGCGAGAACC GTGCCCAGAT CAAAGAATGA GACCGACGCG CCGGCCGCGG CAAAGGATGA TCCTCAGGGT CGGATCTATC GCTCCGCCCT GAAGCAGGAG GGCGCACGCT GGCTGCTGAC GGCGGAGGAA GGGTTGCTGG CAAAGCCCAA GCCGCCCGGC CTTGTTCCGG CACTTGAGAA TGCGATCGCC ATCGTCGATT ACATCAACGG TACACCGCCC CATATCGCGT CCCTGGCGGA GCTTTCAACG ACGCTCGGGA TATCCAAGAG CCACTGTCAC TCCATCCTCA AGACGCTGAC GCATTTTCGGC TGGCTGAAAT TCGACAATCG CTCAAAGAGC TACGAGCTGA ATTC			975 1035 1095 1155 1215 1275 1335 1395 1455 1515 1559



- 83 -

SEQ ID NO: 20
 SEQUENCE LENGTH: 1559
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD424 (FERM BP-4034)

	CATA	TG	ACA	CGT	12
	Thr	Arg			
	1				
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG					60
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu					
5 10 15					
ACA CGC GAA CAG GTC GTC GTC CGT CTT CTC GAC ATG CTG ACG AAA GCC					108
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala					
20 25 30					
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG					156
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr					
35 40 45 50					
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC					201
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser					
55 60 65					
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG					252
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu					
70 75 80					
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC					300
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu					
85 90 95					
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG					348
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val					
100 105 110					
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG					396
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro					
115 120 125 130					
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG					444
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys					
135 140 145					
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC					492
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp					
150 155 160					
GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA					540
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu					



- 84 -

165	170	175	
GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly 180 185 190			588
TAC AAC ACG CCG ACC CAC AAT CCC CTT GTT CCC CAG CAC GAC CAC CTG Tyr Asn Thr Pro Thr His Asn Pro Leu Val Pro Gln His Asp His Leu 195 200 205 210			636
ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn 215 220 225			684
GGG GCC TGG TCC GCG GCC GCG GGC AAG GCG GGC ATG GAG GAG AAC TGC Gly Ala Trp Ser Ala Ala Ala Gly Lys Ala Gly Met Glu Glu Asn Cys 230 235 240			732
ATG CTG CTC GGC CAC TCC TGC ATC GTG GCG CCG ACC GGG GAA ATC GTC Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val 245 250 255			780
GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp 260 265 270			828
CTC GAT CGC TGC CGG GAA CTG CGT GAA CAC ATC TTC AAC TTC AAG CAG Leu Asp Arg Cys Arg Glu Leu Arg Glu His Ile Phe Asn Phe Lys Gln 275 280 285 290			876
CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu 295 300			915
TGAGGTTGCC GAAAAGCATG TGTGTCGTTG TTCTCGGCGC CTGGGTCACA TCCAGGCGCG CCAGGGTGAC GCTGGTGGAA TAGTACCACG ACCGCTTCAG GGCGATCCGC AAGGAGATGC GGGTCGCCGG AGCGGCAAAG CCCGACATTC GTTTCGCACC GACGGCCGTC GTGAACTCGA CAGTCCGCGA GAAGGGCGTA TTGCGCGGCC TGGACCTGTA CGTGGAACCTG TAGCCCATAT ATAGATTTCC AAAGAGTTTC GGCGAGGCGC GGCGCGCCTA GCCCCATGTG AGCGAGAACC GTGCCCAGAT CAAAGAATGA GACCGACGCG CCGGCCGCGG CAAAGGATGA TCCTCAGGGT CGGATCTATC GCTCCGCCCT GAAGCAGGAG GGCGCACGCT GGCTGCTGAC GGCGGAGGAA GGGTTGCTGG CAAAGCCCAA GCCGCCCGGC CTTGTTCCGG CACTTGAGAA TGCGATCGCC ATCGTCGATT ACATCAACGG TACACCGCCC CATATCGCGT CCCTGGCGGA GCTTTCAACG ACGCTCGGGA TATCCAAGAG CCACTGTCAC TCCATCCTCA AGACGCTGAC GCATTTCGGC TGGCTGAAAT TCGACAATCG CTCAAAGAGC TACGAGCTGA ATTC			975 1035 1095 1155 1215 1275 1335 1395 1455 1515 1559

SEQ ID NO: 21
 SEQUENCE LENGTH: 1559
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD425

	CATA TG ACA CGT	12
	Thr Arg	
	1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	60	
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu		
5 10 15		
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	108	
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala		
20 25 30		
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	156	
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr		
35 40 45 50		
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	201	
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser		
55 60 65		
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	252	
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu		
70 75 80		
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	300	
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu		
85 90 95		
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	348	
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val		
100 105 110		
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	396	
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro		
115 120 125 130		
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	444	
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys		
135 140 145		
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	492	
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp		
150 155 160		
GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA	540	
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu		



- 86 -

165	170	175	
GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly 180 185 190			588
TAC AAC ACG CCG ACC CAC AAT CCC TCT GTT CCC CAG CAC GAC CAC CTG Tyr Asn Thr Pro Thr His Asn Pro Ser Val Pro Gln His Asp His Leu 195 200 205 210			636
ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn 215 220 225			684
GGG GCC TGG TCC GCG GCC GCG GGC AAG GCG GGC ATG GAG GAG AAC TGC Gly Ala Trp Ser Ala Ala Ala Gly Lys Ala Gly Met Glu Glu Asn Cys 230 235 240			732
ATG CTG CTC GGC CAC TCC TGC ATC GTG GCG CCG ACC GGG GAA ATC GTC Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val 245 250 255			780
GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp 260 265 270			828
CTC GAT CGC TGC CGG GAA CTG CGT GAA CAC ATC TTC AAC TTC AAG CAG Leu Asp Arg Cys Arg Glu Leu Arg Glu His Ile Phe Asn Phe Lys Gln 275 280 285 290			876
CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu 295 300			915
TGAGGTTGCC GAAAAGCATG TGTGTCGTTG TTCTCGGCGC CTGGGTCACA TCCAGGCGCG 975			
CCAGGGTGAC GCTGGTGGA TAGTACCACG ACCGCTTCAG GGCGATCCGC AAGGAGATGC 1035			
GGGTGCGCGG AGCGGCAAAG CCCGACATTC GTTTCGCACC GACGGCCGTC GTGAACTCGA 1095			
CAGTCCGCGA GAAGGGCGTA TTGCGCGGCC TGGACCTGTA CGTGGAAC TGAGCCCATAT 1155			
ATAGATTTCC AAAGAGTTTC GGCGAGGCGC GGCGCGCCTA GCGCCATGTG AGCGAGAACC 1215			
GTGCCCAGAT CAAAGAATGA GACCGACGCG CCGGCCGCGG CAAAGGATGA TCCTCAGGGT 1275			
CGGATCTATC GCTCCGCCCT GAAGCAGGAG GGCGCACGCT GGCTGCTGAC GGCGGAGGAA 1335			
GGGTGCTG GAAAGCCCAA GCCGCCCGGC CTTGTTCCGG CACTTGAGAA TGCGATCGCC 1395			
ATCGTCGATT ACATCAACGG TACACCGCCC CATATCGCGT CCCTGGCGGA GCTTTCAACG 1455			
ACGCTCGGGA TATCCAAGAG CCACTGTCAC TCCATCCTCA AGACGCTGAC GCATTTCCGGC 1515			
TGGCTGAAAT TCGACAATCG CTCAAAGAGC TACGAGCTGA ATTC 1559			



- 87 -

SEQ ID NO: 22
 SEQUENCE LENGTH: 1559
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD426

	CATA TG ACA CGT	12
	Thr Arg	
	1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	60	
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu		
5 10 15		
ACA CGC GAA CAG GTC GTC GGT CGT CTT TTC GAG ATG CTG ACC AAA GCC	108	
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala		
20 25 30		
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	156	
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr		
35 40 45 50		
ACC TTC TTC CCG CGC TGG TAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	201	
Thr Phe Phe Pro Arg Trp Tyr Phe Thr Asp Glu Ala Glu Leu Asp Ser		
55 60 65		
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	252	
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu		
70 75 80		
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	300	
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu		
85 90 95		
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	348	
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val		
100 105 110		
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	396	
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro		
115 120 125 130		
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	444	
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys		
135 140 145		
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	492	
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp		
150 155 160		
GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA	540	
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu		

	165						170						175							
GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly																			588	
180						185						190								
TAC AAC ACG CCG ACC CAC AAT CCC CTT GTT CCC CAG CAC GAC CAC CTG Tyr Asn Thr Pro Thr His Asn Pro Leu Val Pro Gln His Asp His Leu																			636	
195						200						205						210		
ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn																			684	
215						220						225								
GGG GCC TGG TCC GCG GCC GCG GGC AAG GCG GGC ATG GAG GAG AAC TGC Gly Ala Trp Ser Ala Ala Ala Gly Lys Ala Gly Met Glu Glu Asn Cys																			732	
230						235						240								
ATG CTG CTC GGC CAC TCC TGC ATC GTG GCG CCG ACC GGG GAA ATC GTC Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val																			780	
245						250						255								
GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp																			828	
260						265						270								
CTC GAT CGC TGC CGG GAA CTG CGT GAA CAC ATC TTC AAC TTC AAG CAG Leu Asp Arg Cys Arg Glu Leu Arg Glu His Ile Phe Asn Phe Lys Gln																			876	
275						280						285						290		
CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu																			915	
295						300														
TGAGGTTGCC GAAAAGCATG TGTGTCGTTG TTCTCGGCGC CTGGGTCACA TCCAGGCGCG	975																			
CCAGGGTGAC GCTGGTGGA TAGTACCACG ACCGCTTCAG GGCGATCCGC AAGGAGATGC	1035																			
GGGTCGCCGG AGCGGCAAAG CCCGACATTC GTTTCGCACC GACGGCCGTC GTGAACTCGA	1095																			
CAGTCCGCGA GAAGGGCGTA TTGCGCGGCC TGGACCTGTA CGTGGAActg TAGCCCAtat	1155																			
ATAgatttcc AAAGagtttc GGcGaggcgC GgcGcgCcta gCCCCatgtG AgCGagaacc	1215																			
GTGcccagat CAAGAATGa GAccGacGcG cCGGCCGCGG CAAGgATga TCctCaGGGT	1275																			
cGgatctatc Gctccgccct GaagcaGGag ggCGcacGct GGctGctGAc GGcggaggaa	1335																			
GGGTTgcttg CAAAGCCCA GCCGCCCGGC CTTGTTCcGG CaCttGagaa TGcGatCGCc	1395																			
aTCGtCGAtt acATCAACGG TaCACCGCCC CatATCGcGT CCctGGCGGA GcTTTTCAACG	1455																			
ACGCTCGGGA TatCCAAGag CCACTGtCAC tCCATCctCa agAcGctGAc GCATTTCGGC	1515																			
TGGCTGAAAT TCGACAATCG CTCAAAGAGC TACGAGCTGA ATTc	1559																			



- 89 -

SEQ ID NO: 23
 SEQUENCE LENGTH: 1559
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD427

	CATA	TG	ACA	CGT	
	Thr	Arg			
	1				
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG					60
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu					
5 10 15					
ACA CGC GAA CAG GTC CTC GTC GGT CTC CTC GAC ATG CTG ACG AAA GCC					108
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala					
20 25 30					
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG					156
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr					
35 40 45 50					
ACC TTC TTC CCG CGC TGG TAT TTC ACC GAC GAG GCC GAG CTC GAT AGC					201
Thr Phe Phe Pro Arg Trp Tyr Phe Thr Asp Glu Ala Glu Leu Asp Ser					
55 60 65					
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG					252
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu					
70 75 80					
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC					300
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu					
85 90 95					
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG					348
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val					
100 105 110					
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG					396
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro					
115 120 125 130					
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG					444
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys					
135 140 145					
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC					492
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp					
150 155 160					
GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA					540
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu					



- 90 -

165	170	175	
GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly 180 185 190			588
TAC AAC ACG CCG ACC CAC AAT CCC TCT GTT CCC CAG CAC GAC CAC CTG Tyr Asn Thr Pro Thr His Asn Pro Ser Val Pro Gln His Asp His Leu 195 200 205 210			636
ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn 215 220 225			684
GGG GCC TGG TCC GCG GCC GCG GGC AAG GCG GGC ATG GAG GAG AAC TGC Gly Ala Trp Ser Ala Ala Ala Gly Lys Ala Gly Met Glu Glu Asn Cys 230 235 240			732
ATG CTG CTC GGC CAC TCC TGC ATC GTG GCG CCG ACC GGG GAA ATC GTC Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val 245 250 255			780
GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp 260 265 270			828
CTC GAT CGC TGC CGG GAA CTG CGT GAA CAC ATC TTC AAC TTC AAG CAG Leu Asp Arg Cys Arg Glu Leu Arg Glu His Ile Phe Asn Phe Lys Gln 275 280 285 290			876
CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu 295 300			915
TGAGGTTGCC GAAAAGCATG TGTGTCGTTG TTCTCGGCGC CTGGGTCACA TCCAGGCGCG CCAGGGTGAC GCTGGTGGA TAGTACCACG ACCGCTTCAG GGCGATCCGC AAGGAGATGC GGGTGCGCGG AGCGGCAAAG CCCGACATTC GTTTCGCACC GACGGCCGTC GTGAACTCGA CAGTCCGCGA GAAGGGCGTA TTGCGCGGCC TGGACCTGTA CGTGGAACCTG TAGCCCATAT ATAGATTTCC AAAGAGTTTC GGCGAGGCGC GGCGCGCCTA GCGCCATGTG AGCGAGAACC GTGCCCAGAT CAAAGAATGA GACCGACGCG CCGGCCGCGG CAAAGGATGA TCCTCAGGGT CGGATCTATC GCTCCGCCCT GAAGCAGGAG GGCGCACGCT GGCTGCTGAC GGCGGAGGAA GGGTTGCTGG CAAAGCCCAA GCCGCCCGGC CTTGTTCCGG CACTTGAGAA TGCGATCGCC ATCGTCGATT ACATCAACGG TACACCGCCC CATATCGCGT CCCTGGCGGA GCTTTCAACG ACGCTCGGGA TATCCAAGAG CCACTGTCAC TCCATCCTCA AGACGCTGAC GCATTTCCGG TGGCTGAAAT TCGACAATCG CTCAAAGAGC TACGAGCTGA ATTC			975 1035 1095 1155 1215 1275 1335 1395 1455 1515 1559



- 91 -

SEQ ID NO: 24
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD451

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCGGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG CGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG TAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp Tyr Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						



- 92 -

GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA 766
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu
165 170 175

GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC 814
Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly
180 185 190

TAC	AAC	ACG	CCG	ACC	CAC	AAT	CCC	GAA	GTT	CCC	CAG	CAC	GAC	CAC	CTG	862
Tyr	Asn	Thr	Pro	Thr	His	Asn	Pro	Glu	Val	Pro	Gln	His	Asp	His	Leu	
195					200					205					210	

ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC 910
Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn
215 220 225

GGG GCC TGG TCC GCG GCC GCG GGC AAG GTG GGC ATG GAG GAG AAC TGC 958
Gly Ala Trp Ser Ala Ala Ala Gly Lys Val Gly Met Glu Glu Asn Cys
230 235 240

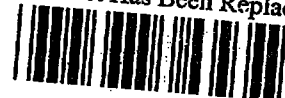
ATG CTG CTC GGC CAC TCC TGC ATC GTG GCG CCG ACC GGG GAA ATC GTC 1006
Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val
245 250 255

GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT 1054
Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp
260 265 270

CTC	GAT	CGC	TGC	CGG	GAA	CTG	CGT	GAA	CAC	ATC	TTC	AAC	TTC	AAG	CAG	1102
Leu	Asp	Arg	Cys	Arg	Glu	Leu	Arg	Glu	His	Ile	Phe	Asn	Phe	Lys	Gln	
275					280					285					290	

CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC 1141
His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu
295 300

TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCA	TCCAGGCGCG	1201
CCAGGGTGAC	GCTGGTGGAA	TAGTACCACG	ACCGCTTCAG	GGCGATCCG	AAGGAGATGC	1261
GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321
CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGACCTGTA	CGTGGAAC TG	TAGCCCATAT	1381
ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441
GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501
CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561
GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621
ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681
ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCGGC	1741
TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785



SEQ ID NO: 25
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD452

GTCGACGGCG GGCTCGCGCG AGAGCTTGTC AAGCAGCGCA AATTCCGGTT CCGCTCCGGT 60
 TGACAGATCA AAAATTTTAC GCCTGTTATT GTCGTGCTGC ATGTAATATT TCGTACTTTA 120
 TGTAGAATTT GCATTGCGCC GCGAGTCACA AAGCCGGTTT TCGGCGATGT GTTTCACAAC 180
 GTTTTCCCGG CCGCTGGGCC GGACATCACC TAGGAAGGAG CAGAGGTTCA TG ACA CGT 238
 Thr Arg
 1

CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG 286
 Gln Met Ile Leu Ala Val Gly Gln Gln Gly Phe Ile Ala Arg Ala Glu
 5 10 15

ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC 334
 Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala
 20 25 30

GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG 382
 Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr
 35 40 45 50

ACC TTC TTC CCG CGC TGG TAT TTC ACC GAC GAG GCC GAG CTC GAT AGC 430
 Thr Phe Phe Pro Arg Trp Tyr Phe Thr Asp Glu Ala Glu Leu Asp Ser
 55 60 65

TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG 478
 Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu
 70 75 80

AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC 526
 Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu
 85 90 95

GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG 574
 Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val
 100 105 110

GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG 622
 Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro
 115 120 125 130

GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG 670
 Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys
 135 140 145

CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC 718
 Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp
 150 155 160



GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA 766
 Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu
 165 170 175

GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC 814
 Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly
 180 185 190

TAC AAC ACG CCG ACC CAC AAT CCC CCT GTT CCC CAG CAC GAC CAC CTG 862
 Tyr Asn Thr Pro Thr His Asn Pro Pro Val Pro Gln His Asp His Leu
 195 200 205 210

ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC 910
 Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn
 215 220 225

GGG GCC TGG TCC GCG GCC GCG GGC AAG TCA GGC ATG GAG GAG AAC TGC 958
 Gly Ala Trp Ser Ala Ala Ala Gly Lys Ser Gly Met Glu Glu Asn Cys
 230 235 240

ATG CTG CTC GGC CAC TCC TGC ATC GTG GCG CCG ACC GGG GAA ATC GTC 1006
 Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val
 245 250 255

GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT 1054
 Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp
 260 265 270

CTC GAT CGC TGC CGG GAA CTG CGT GAA CAC ATC TTC AAC TTC AAG CAG 1102
 Leu Asp Arg Cys Arg Glu Leu Arg Glu His Ile Phe Asn Phe Lys Gln
 275 280 285 290

CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC 1141
 His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu
 295 300

TGAGGTTGCC GAAAAGCATG TGTGTCGTTG TTCTCGGCGC CTGGGTCACA TCCAGGCGCG 1201
 CCAGGGTGAC GCTGGTGGAA TAGTACCACG ACCGCTTCAG GGCATCCGC AAGGAGATGC 1261
 GGGTCGCCCG AGCGGCAAAG CCCGACATTC GTTTCGCACC GACGGCCGTC GTGAACTCGA 1321
 CAGTCCGCGA GAAGGGCGTA TTGCGCGGCC TGGACCTGTA CGTGGAACCTG TAGCCCATAT 1381
 ATAGATTTCC AAAGAGTTTC GGCGAGGCGC GGCGCGCCTA GCCCCATGTG AGCGAGAACC 1441
 GTGCCCAGAT CAAAGAATGA GACCGACGCG CCGGCCGCGG CAAAGGATGA TCCTCAGGGT 1501
 CGGATCTATC GCTCCGCCCT GAAGCAGGAG GGCGCACGCT GGCTGCTGAC GGCGGAGGAA 1561
 GGGTTGCTGG CAAAGCCCAA GCCGCCCGGC CTTGTTCCGG CACTTGAGAA TGCGATCGCC 1621
 ATCGTCGATT ACATCAACGG TACACCGCCC CATATCGCGT CCCTGGCGGA GCTTTCAACG 1681
 ACGCTCGGGA TATCCAAGAG CCACTGTCAC TCCATCCTCA AGACGCTGAC GCATTTCGGC 1741
 TGGCTGAAAT TCGACAATCG CTCAAAGAGC TACGAGCTGA ATTC 1785



- 95 -

SEQ ID NO: 26
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD453

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCCGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG						286
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
	5		10		15	
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC						334
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
	20		25		30	
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG						382
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
	35		40		45	50
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC						430
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser						
		55		60		65
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG						478
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
		70		75		80
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC						526
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
	85		90		95	
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG						574
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
	100		105		110	
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG						622
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
	115		120		125	130
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG						670
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
		135		140		145
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC						718
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
	150		155		160	

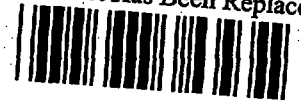
[illegible]



- 97 -

SEQ ID NO: 27
 SEQUENCE LENGTH: 1559
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD461

	CATA	TG	ACA	CGT	
			Thr	Arg	
			1		
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG					60
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu					
	5			15	
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GGC					108
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala					
	20			30	
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG					156
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr					
	35			45	50
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC					201
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser					
		55		60	65
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG					252
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu					
		70		75	80
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC					300
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu					
		85		90	95
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG					348
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val					
	100			110	
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG					396
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro					
	115			125	130
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG					444
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys					
		135		140	145
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC					492
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp					
		150		155	160
GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA					540
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu					



- 98 -

165	170	175	
GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly 180 185 190			588
TAC AAC ACG CCG ACC CAC AAT CCC GAA GTT CCC CAG CAC GAC CAC CTG Tyr Asn Thr Pro Thr His Asn Pro Glu Val Pro Gln His Asp His Leu 195 200 205 210			636
ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn 215 220 225			684
GGG GCC TGG TCC GCG GCC GCG GGC AAG GCG GGC ATG GAG GAG AAC TGC Gly Ala Trp Ser Ala Ala Ala Gly Lys Ala Gly Met Glu Glu Asn Cys 230 235 240			732
ATG CTG CTC GGC CAC TCC TGC ATC GTG GCG CCG ACC GGG GAA ATC GTC Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val 245 250 255			780
GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp 260 265 270			828
CTC GAT CGC TGC CGG GAA CTG CGT GAA CAC ATC TTC AAC TTC AAG CAG Leu Asp Arg Cys Arg Glu Leu Arg Glu His Ile Phe Asn Phe Lys Gln 275 280 285 290			876
CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu 295 300			915
TGAGGTTGCC GAAAAGCATG TGTGTCGTTG TTCTCGGCGC CTGGGTCACA TCCAGGCGCG CCAGGGTGAC GCTGGTGGAA TAGTACCACG ACCGCTTCAG GGCGATCCGC AAGGAGATGC GGGTCGCCGG AGCGGCAAAG CCCGACATTC GTTTCGCACC GACGGCCGTC GTGAACTCGA CAGTCCGCGA GAAGGGCGTA TTGCGCGGCC TGGACCTGTA CGTGGAACGT TAGCCCATAT ATAGATTTCC AAAGAGTTTC GGCAGGCGC GGCAGCCTA GCGGATGTG AGCGAGAACC GTGCCCAGAT CAAAGAATGA GACCGACGCG CCGGCCGCGG CAAAGGATGA TCCTCAGGGT CGGATCTATC GCTCCGCCCT GAAGCAGGAG GGCACACGCT GGCTGCTGAC GCGGAGGAA GGGTTGCTGG CAAAGCCCAA GCCGCCCGG CTTGTTCCGG CACTTGAGAA TCGATCGCC ATCGTCGATT ACATCAACGG TACACCGCCC CATATCGCGT CCCTGGCGGA GCTTTCAACG ACGCTCGGGA TATCCAAGAG CCACTGTCAC TCCATCCTCA AGACGCTGAC GCATTTCCGGC TGGCTGAAAT TCGACAATCG CTCAAAGAGC TACGAGCTGA ATTC			975 1035 1095 1155 1215 1275 1335 1395 1455 1515 1559



SEQ ID NO: 28
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD454

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCCGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCTCG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG TAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp Tyr Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

GCC Ala	GCG Ala	AAA Lys 165	ATG Met	GGG Gly	ATG Met	TTC Phe	ATC Ile 170	TGC Cys	AAC Asn	GAT Asp	CGC Arg	CGC Arg 175	TGG Trp	CCT Pro	GAA Glu	766																																																											
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	814																																																											
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	GAA Glu	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	862																																																											
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	910																																																											
GGG Gly	GCC Ala	TGG Trp	TCC Ser 230	GCG Ala	GCC Ala	GCG Ala	GGC Gly	AAG Lys 235	TCA Ser	GGC Gly	ATG Met	GAG Glu	GAG Glu 240	AAC Asn	TGC Cys	958																																																											
ATG Met	CTG Leu	CTC Leu 245	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GCG Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	1006																																																											
GCT Ala	CTC Leu 260	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	1054																																																											
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	1102																																																											
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu	ATC Ile 300	GCG Ala	GAA Glu	CTC Leu				1141																																																											
TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201	CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261	GGGTCGCCCG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAAGTCGA	1321	CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAAGTG	TAGCCCATAT	1381	ATAGATTTC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441	GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501	CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561	GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621	ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681	ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCCGC	1741	TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC	1785



SEQ ID NO: 29
 SEQUENCE LENGTH: 1559
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD455 (FERM BP-4036)

																CATA TG ACA CGT	12
																Thr Arg	
																1	
CAG	ATG	ATA	CTT	GCA	GTG	GGA	CAA	CAA	GGT	CCG	ATC	GCG	CGC	GCG	GAG		60
Gln	Met	Ile	Leu	Ala	Val	Gly	Gln	Gln	Gly	Pro	Ile	Ala	Arg	Ala	Glu		
		5					10					15					
ACA	CGC	GAA	CAG	GTC	GTC	GTT	CGT	CTT	CTC	GAC	ATG	CTG	ACG	AAA	GCC		108
Thr	Arg	Glu	Gln	Val	Val	Val	Arg	Leu	Leu	Asp	Met	Leu	Thr	Lys	Ala		
	20					25				30							
GCG	AGC	CGG	GGC	GCG	AAT	TTC	ATT	GTC	TTC	CCC	GAA	CTC	GCG	CTT	ACG		156
Ala	Ser	Arg	Gly	Ala	Asn	Phe	Ile	Val	Phe	Pro	Glu	Leu	Ala	Leu	Thr		
	35				40					45					50		
ACC	TTC	TTC	CCG	CGC	TGG	TAT	TTC	ACC	GAC	GAG	GCC	GAG	CTC	GAT	AGC		201
Thr	Phe	Phe	Pro	Arg	Trp	Tyr	Phe	Thr	Asp	Glu	Ala	Glu	Leu	Asp	Ser		
				55					60					65			
TTC	TAT	GAG	ACC	GAA	ATG	CCC	GGC	CCG	GTG	GTC	CGT	CCA	CTC	TTT	GAG		252
Phe	Tyr	Glu	Thr	Glu	Met	Pro	Gly	Pro	Val	Val	Arg	Pro	Leu	Phe	Glu		
			70				75						80				
AAG	GCC	GCG	GAA	CTC	GGG	ATC	GGC	TTC	AAT	CTG	GGC	TAC	GCT	GAA	CTC		300
Lys	Ala	Ala	Glu	Leu	Gly	Ile	Gly	Phe	Asn	Leu	Gly	Tyr	Ala	Glu	Leu		
		85				90					95						
GTC	GTC	GAA	GGC	GGC	GTC	AAG	CGT	CGC	TTC	AAC	ACG	TCC	ATT	TTG	GTG		348
Val	Val	Glu	Gly	Gly	Val	Lys	Arg	Arg	Phe	Asn	Thr	Ser	Ile	Leu	Val		
	100					105					110						
GAT	AAG	TCA	GGC	AAG	ATC	GTC	GGC	AAG	TAT	CGT	AAG	ATC	CAT	TTG	CCG		396
Asp	Lys	Ser	Gly	Lys	Ile	Val	Gly	Lys	Tyr	Arg	Lys	Ile	His	Leu	Pro		
	115				120					125					130		
GGT	CAC	AAG	GAG	TAC	GAG	GCC	TAC	CGG	CCG	TTC	CAG	CAT	CTT	GAA	AAG		444
Gly	His	Lys	Glu	Tyr	Glu	Ala	Tyr	Arg	Pro	Phe	Gln	His	Leu	Glu	Lys		
				135					140					145			
CGT	TAT	TTC	GAG	CCG	GGC	GAT	CTC	GGC	TTC	CCG	GTC	TAT	GAC	GTC	GAC		492
Arg	Tyr	Phe	Glu	Pro	Gly	Asp	Leu	Gly	Phe	Pro	Val	Tyr	Asp	Val	Asp		
			150					155					160				
GCC	GCG	AAA	ATG	GGG	ATG	TTC	ATC	TGC	AAC	GAT	CGC	CGC	TGG	CCT	GAA		540
Ala	Ala	Lys	Met	Gly	Met	Phe	Ile	Cys	Asn	Asp	Arg	Arg	Trp	Pro	Glu		



- 102 -

165	170	175	
GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly 180 185 190			588
TAC AAC ACG CCG ACC CAC AAT CCC GAA GTT CCC CAG CAC GAC CAC CTG Tyr Asn Thr Pro Thr His Asn Pro Glu Val Pro Gln His Asp His Leu 195 200 205 210			636
ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn 215 220 225			684
GGG GCC TGG TCC GCG GCC GCG GGC AAG GCG GGC ATG GAG GAG AAC TGC Gly Ala Trp Ser Ala Ala Ala Gly Lys Ala Gly Met Glu Glu Asn Cys 230 235 240			732
ATG CTG CTC GGC CAC TCC TGC ATC GTG GCG CCG ACC GGG GAA ATC GTC Met Leu Leu Gly His Ser Cys Ile Val Ala Pro Thr Gly Glu Ile Val 245 250 255			780
GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp 260 265 270			828
CTC GAT CGC TGC CGG GAA CTG CGT GAA CAC ATC TTC AAC TTC AAG CAG Leu Asp Arg Cys Arg Glu Leu Arg Glu His Ile Phe Asn Phe Lys Gln 275 280 285 290			876
CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu 295 300			915
TGAGGTTGCC GAAAAGCATG TGTGTCGTTG TTCTCGGCGC CTGGGTCACA TCCAGGCGCG 975			
CCAGGGTGAC GCTGGTGGAA TAGTACCACG ACCGCTTCAG GGCGATCCGC AAGGAGATGC 1035			
GGGTCGCCGG AGCGGCAAAG CCCGACATTC GTTTCGCACC GACGGCCGTC GTGAACTCGA 1095			
CAGTCCGCGA GAAGGGCGTA TTGCGCGGCC TGGACCTGTA CGTGGAAC TGACCCATAT 1155			
ATAGATTTCC AAAGAGTTTC GGCGAGGCGC GGCGCGCCTA GCCCCATGTG AGCGAGAACC 1215			
GTGCCCAGAT CAAAGAATGA GACCGACGCG CCGGCCGCGG CAAAGGATGA TCCTCAGGGT 1275			
CGGATCTATC GCTCCGCCCT GAAGCAGGAG GGCGCACGCT GGCTGCTGAC GGCGGAGGAA 1335			
GGGTTGCTGG CAAAGCCCAA GCCGCCCGGC CTTGTTCCGG CACTTGAGAA TGCGATCGCC 1395			
ATCGTCGATT ACATCAACGG TACACCGCCC CATATCGCGT CCCTGGCGGA GCTTTCAACG 1455			
ACGCTCGGGA TATCCAAGAG CCACTGTCAC TCCATCCTCA AGACGCTGAC GCATTTCGGC 1515			
TGGCTGAAAT TCGACAATCG CTCAAAGAGC TACGAGCTGA ATTC 1559			



- 103 -

SEQ ID NO: 30
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD456

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCGGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG Ala CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG TAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp Tyr Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

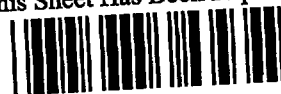
GCC Ala	GCG Ala	AAA Lys 165	ATG Met	GGG Gly	ATG Met	TTC Phe	ATC Ile 170	TGC Cys	AAC Asn	GAT Asp	CGC Arg	CGC Arg 175	TGG Trp	CCT Pro	GAA Glu	766																																																												
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	814																																																												
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	CTT Leu	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	862																																																												
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	910																																																												
GGG Gly	GCC Ala	TGG Trp	TCC Ser 230	GCG Ala	GCC Ala	GCG Ala	GGC Gly	AAG Lys 235	TCA Ser	GGC Gly	ATG Met	GAG Glu	GAG Glu 240	AAC Asn	TGC Cys	958																																																												
ATG Met	CTG Leu	CTC Leu 245	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GCG Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	1006																																																												
GCT Ala	CTC Leu 260	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	1054																																																												
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	1102																																																												
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu	ATC Ile 300	GCG Ala	GAA Glu	CTC Leu				1141																																																												
TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGGTCACA	TCCAGGCGCG	1201	CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261	GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321	CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAACGT	TAGCCCATAT	1381	ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441	GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501	CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561	GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACCTTGAGAA	TGCGATCGCC	1621	ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681	ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCCGC	1741	TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785



SEQ ID NO: 31
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD468

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCCGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG TAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp Tyr Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

GCC Ala	CGC Ala	AAA Lys 165	ATG Met	GGG Gly	ATG Met	TTC Phe	ATC Ile 170	TGC Cys	AAC Asn	GAT Asp	CGC Arg	CGC Arg 175	TGG Trp	CCT Pro	GAA Glu	766																																																												
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	814																																																												
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	GCT Ala	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	862																																																												
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	910																																																												
GGG Gly	GCC Ala	TGG Trp	TCC Ser 230	CGC Ala	GCC Ala	GGC Ala	GGC Gly	AAC Lys 235	TCA Ser	GGC Gly	ATG Met	GAG Glu 240	GAG Glu	AAC Asn	TGC Cys	958																																																												
ATG Met	CTG Leu	CTC Leu 245	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GGC Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	1006																																																												
GCT Ala	CTC Leu 260	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	1054																																																												
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	1102																																																												
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu 300	ATC Ile	GCG Ala	GAA Glu	CTC Leu				1141																																																												
TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201	CCAGGGTGAC	GCTGGTGGAA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261	GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321	CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAACGT	TAGCCCATAT	1381	ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441	GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501	CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561	GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621	ATCGTCGATT	ACATCAACGG	TACACGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681	ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCGGC	1741	TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785



- 107 -

SEQ ID NO: 32
 SEQUENCE LENGTH: 1785
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: JM109 pAD469

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCGGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTCACAAC	180
GTTCCTCCCG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
				Thr Arg		
				1		
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC CCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG TAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp Tyr Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

- 108 -

GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA 766
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu
165 170 175

GCC TGG CGG GTG ATG GGC CTC AGG GGC GCC GAG ATC ATC TGC GGC GGC 814
Ala Trp Arg Val Met Gly Leu Arg Gly Ala Glu Ile Ile Cys Gly Gly
180 185 190

TAC	AAC	ACG	CCG	ACC	CAC	AAT	CCC	ATT	GTT	CCC	CAG	CAC	GAC	CAC	CTG	862
Tyr	Asn	Thr	Pro	Thr	His	Asn	Pro	Ile	Val	Pro	Gln	His	Asp	His	Leu	
195					200					205					210	

ACG TCC TTC CAC CAT CTC CTA TCG ATG CAG GCC GGG TCT TAT CAG AAC 910
Thr Ser Phe His His Leu Leu Ser Met Gln Ala Gly Ser Tyr Gln Asn
215 220 225

GGG GCC TGG TCC GCG GCC GCG GGC AAG TCA GGC ATG GAG GAG AAC TGC 958
Gly Ala Trp Ser Ala Ala Ala Gly Lys Ser Gly Met Glu Glu Asn Cys
230 235 240

ATG	CTG	CTC	GGC	CAC	TCC	TGC	ATC	GTG	GCG	CCG	ACC	GGG	GAA	ATC	GTC	1006
Met	Leu	Leu	Gly	His	Ser	Cys	Ile	Val	Ala	Pro	Thr	Gly	Glu	Ile	Val	
		245					250					255				

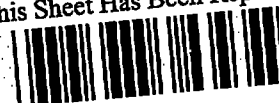
GCT CTC ACT ACG ACG CTG GAA GAC GAG GTG ATC ACC GCC GCC GTC GAT 1054
Ala Leu Thr Thr Thr Leu Glu Asp Glu Val Ile Thr Ala Ala Val Asp
260 265 270

CTC	GAT	CGC	TGC	CGG	GAA	CTG	CGT	GAA	CAC	ATC	TTC	AAC	TTC	AAG	CAG	1102
Leu	Asp	Arg	Cys	Arg	Glu	Leu	Arg	Glu	His	Ile	Phe	Asn	Phe	Lys	Gln	
275					280					285					290	

CAT CGT CAG CCC CAG CAC TAT GGT CTG ATC GCG GAA CTC 1141
His Arg Gln Pro Gln His Tyr Gly Leu Ile Ala Glu Leu
295 300

TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCA	TCCAGGCGCG	1201
CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261
GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACCTGA	1321
CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAACTG	TAGCCCATAT	1381
ATAGATTTC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441
GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501
CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561
GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACCTTGAGAA	TGCGATCGCC	1621
ATCGTCGATT	ACATCAACGG	TACACCGCCC	CATATCGCGT	CCCTGGCGGA	GCTTTC AACG	1681
ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCCGC	1741
TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC		1785

This Sheet Has Been Replaced



SEQ ID NO: 33
SEQUENCE LENGTH: 1785
SEQUENCE TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE
ORGANISM: Escherichia coli
STRAIN: JM109 pAD470

GTCGACGGCG	GGCTCGCGCG	AGAGCTTGTC	AAGCAGCGCA	AATTCGGGTT	CCGCTCCGGT	60
TGACAGATCA	AAAATTTTAC	GCCTGTTATT	GTCGTGCTGC	ATGTAATATT	TCGTACTTTA	120
TGTAGAATTT	GCATTGCGCC	GCGAGTCACA	AAGCCGGTTT	TCGGCGATGT	GTTTCACAAC	180
GTTTTCCCGG	CCGCTGGGCC	GGACATCACC	TAGGAAGGAG	CAGAGGTTCA	TG ACA CGT	238
					Thr Arg	
					1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	286					
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu						
5 10 15						
ACA CGC GAA CAG GTC GTC GTT CGT CTT CTC GAC ATG CTG ACG AAA GCC	334					
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala						
20 25 30						
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	382					
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr						
35 40 45 50						
ACC TTC TTC CCG CGC TGG TAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	430					
Thr Phe Phe Pro Arg Trp Tyr Phe Thr Asp Glu Ala Glu Leu Asp Ser						
55 60 65						
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	478					
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu						
70 75 80						
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	526					
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu						
85 90 95						
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	574					
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val						
100 105 110						
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	622					
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro						
115 120 125 130						
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	670					
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys						
135 140 145						
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	718					
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp						
150 155 160						

GCC Ala	GCG Ala	AAA Lys 165	ATG Met	GGG Gly	ATG Met	TTC Phe	ATC Ile 170	TGC Cys	AAC Asn	GAT Asp	CGC Arg	CGC Arg 175	TGG Trp	CCT Pro	GAA Glu	766																																																											
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	814																																																											
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	CAT His	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	862																																																											
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	910																																																											
GGG Gly	GCC Ala	TGG Trp	TCC Ser 230	GCG Ala	GCC Ala	GCG Ala	GGC Gly	AAG Lys 235	TCA Ser	GGC Gly	ATG Met	GAG Glu	GAG Glu	AAC Asn	TGC Cys	958																																																											
ATG Met	CTG Leu	CTC Leu 245	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GCG Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	1006																																																											
GCT Ala	CTC Leu 260	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	1054																																																											
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	1102																																																											
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu	ATC Ile 300	GCG Ala	GAA Glu	CTC Leu				1141																																																											
TGAGGTTGCC	GAAAAGCATG	TGTGTCGTTG	TTCTCGGCGC	CTGGGTCACA	TCCAGGCGCG	1201	CCAGGGTGAC	GCTGGTGGA	TAGTACCACG	ACCGCTTCAG	GGCGATCCGC	AAGGAGATGC	1261	GGGTCGCCGG	AGCGGCAAAG	CCCGACATTC	GTTTCGCACC	GACGGCCGTC	GTGAACTCGA	1321	CAGTCCGCGA	GAAGGGCGTA	TTGCGCGGCC	TGGACCTGTA	CGTGGAACGT	TAGCCCATAT	1381	ATAGATTTCC	AAAGAGTTTC	GGCGAGGCGC	GGCGCGCCTA	GCCCCATGTG	AGCGAGAACC	1441	GTGCCCAGAT	CAAAGAATGA	GACCGACGCG	CCGGCCGCGG	CAAAGGATGA	TCCTCAGGGT	1501	CGGATCTATC	GCTCCGCCCT	GAAGCAGGAG	GGCGCACGCT	GGCTGCTGAC	GGCGGAGGAA	1561	GGGTTGCTGG	CAAAGCCCAA	GCCGCCCGGC	CTTGTTCCGG	CACTTGAGAA	TGCGATCGCC	1621	ATCGTCGATT	ACATCAACGG	TACACGCCCC	CATATCGCGT	CCCTGGCGGA	GCTTTCAACG	1681	ACGCTCGGGA	TATCCAAGAG	CCACTGTCAC	TCCATCCTCA	AGACGCTGAC	GCATTTCGGC	1741	TGGCTGAAAT	TCGACAATCG	CTCAAAGAGC	TACGAGCTGA	ATTC	1785



SEQ ID NO: 34
 SEQUENCE LENGTH: 926
 SEQUENCE TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE
 ORGANISM: Escherichia coli
 STRAIN: HB101 pNT4553 (FERM BP-4368)

	CATA TG ACA CGT	12
	Thr Arg	
	1	
CAG ATG ATA CTT GCA GTG GGA CAA CAA GGT CCG ATC GCG CGC GCG GAG	60	
Gln Met Ile Leu Ala Val Gly Gln Gln Gly Pro Ile Ala Arg Ala Glu		
5 10 15		
ACA CGC GAA CAG GTC GTC GTC CGT CTT CTC GAC ATG CTG ACG AAA GCC	108	
Thr Arg Glu Gln Val Val Val Arg Leu Leu Asp Met Leu Thr Lys Ala		
20 25 30		
GCG AGC CGG GGC GCG AAT TTC ATT GTC TTC CCC GAA CTC GCG CTT ACG	156	
Ala Ser Arg Gly Ala Asn Phe Ile Val Phe Pro Glu Leu Ala Leu Thr		
35 40 45 50		
ACC TTC TTC CCG CGC TGG CAT TTC ACC GAC GAG GCC GAG CTC GAT AGC	201	
Thr Phe Phe Pro Arg Trp His Phe Thr Asp Glu Ala Glu Leu Asp Ser		
55 60 65		
TTC TAT GAG ACC GAA ATG CCC GGC CCG GTG GTC CGT CCA CTC TTT GAG	252	
Phe Tyr Glu Thr Glu Met Pro Gly Pro Val Val Arg Pro Leu Phe Glu		
70 75 80		
AAG GCC GCG GAA CTC GGG ATC GGC TTC AAT CTG GGC TAC GCT GAA CTC	300	
Lys Ala Ala Glu Leu Gly Ile Gly Phe Asn Leu Gly Tyr Ala Glu Leu		
85 90 95		
GTC GTC GAA GGC GGC GTC AAG CGT CGC TTC AAC ACG TCC ATT TTG GTG	348	
Val Val Glu Gly Gly Val Lys Arg Arg Phe Asn Thr Ser Ile Leu Val		
100 105 110		
GAT AAG TCA GGC AAG ATC GTC GGC AAG TAT CGT AAG ATC CAT TTG CCG	396	
Asp Lys Ser Gly Lys Ile Val Gly Lys Tyr Arg Lys Ile His Leu Pro		
115 120 125 130		
GGT CAC AAG GAG TAC GAG GCC TAC CGG CCG TTC CAG CAT CTT GAA AAG	444	
Gly His Lys Glu Tyr Glu Ala Tyr Arg Pro Phe Gln His Leu Glu Lys		
135 140 145		
CGT TAT TTC GAG CCG GGC GAT CTC GGC TTC CCG GTC TAT GAC GTC GAC	492	
Arg Tyr Phe Glu Pro Gly Asp Leu Gly Phe Pro Val Tyr Asp Val Asp		
150 155 160		
GCC GCG AAA ATG GGG ATG TTC ATC TGC AAC GAT CGC CGC TGG CCT GAA	540	
Ala Ala Lys Met Gly Met Phe Ile Cys Asn Asp Arg Arg Trp Pro Glu		

- 112 -

165					170					175						
GCC Ala	TGG Trp 180	CGG Arg	GTG Val	ATG Met	GGC Gly	CTC Leu 185	AGG Arg	GGC Gly	GCC Ala	GAG Glu	ATC Ile 190	ATC Ile	TGC Cys	GGC Gly	GGC Gly	588
TAC Tyr 195	AAC Asn	ACG Thr	CCG Pro	ACC Thr	CAC His 200	AAT Asn	CCC Pro	GAA Glu	GTT Val	CCC Pro 205	CAG Gln	CAC His	GAC Asp	CAC His	CTG Leu 210	636
ACG Thr	TCC Ser	TTC Phe	CAC His	CAT His 215	CTC Leu	CTA Leu	TCG Ser	ATG Met	CAG Gln 220	GCC Ala	GGG Gly	TCT Ser	TAT Tyr	CAG Gln 225	AAC Asn	684
GGG Gly	GCC Ala	TGG Trp	TCC Ser 230	GCG Ala	GCC Ala	GCG Ala	GGC Gly	AAG Lys 235	GCG Ala	GGC Gly	ATG Met	GAG Glu	GAG Glu 240	AAC Asn	TGC Cys	732
ATG Met	CTG Leu	CTC Leu 245	GGC Gly	CAC His	TCC Ser	TGC Cys	ATC Ile 250	GTG Val	GCG Ala	CCG Pro	ACC Thr	GGG Gly 255	GAA Glu	ATC Ile	GTC Val	780
GCT Ala 260	CTC Leu	ACT Thr	ACG Thr	ACG Thr	CTG Leu	GAA Glu 265	GAC Asp	GAG Glu	GTG Val	ATC Ile	ACC Thr 270	GCC Ala	GCC Ala	GTC Val	GAT Asp	828
CTC Leu 275	GAT Asp	CGC Arg	TGC Cys	CGG Arg	GAA Glu 280	CTG Leu	CGT Arg	GAA Glu	CAC His	ATC Ile 285	TTC Phe	AAC Asn	TTC Phe	AAG Lys	CAG Gln 290	876
CAT His	CGT Arg	CAG Gln	CCC Pro	CAG Gln 295	CAC His	TAT Tyr	GGT Gly	CTG Leu	ATC Ile 300	GCG Ala	GAA Glu	CTC Leu				915
TGAGGCTGCAG																926



CLAIMS

1. A process for preparing a DNA fragment coding for an enzyme protein capable of converting an N-carbamyl-D- α -amino acid by removal of its carbamyl group into the corresponding D- α -amino acid (hereinafter referred to as decarbamylase) improved in thermostability, characterized in that a DNA fragment coding for a microorganism-derived decarbamylase protein is subjected to mutagenesis.
2. A preparation process according to claim 1, wherein the DNA fragment coding for a microorganism-derived decarbamylase protein is subjected to mutagenesis in the form of a single strand.
3. A preparation process according to claim 1 or 2, wherein the mutagenesis causes a random mutation at mutation sites.
4. A preparation process according to claim 3, wherein the random mutation is caused by chemical mutagenesis.
5. A preparation process according to any one of claims 2 to 4, wherein the DNA fragment coding for a microorganism-derived decarbamylase protein is incorporated into a phage and subjected to mutagenesis in the form of a single strand.
6. A preparation process according to any one of claims 2 to 5, wherein the single-stranded DNA fragment is a DNA strand containing the codons corresponding to the amino acid sequence of the decarbamylase protein.
7. A preparation process according to any one of



claims 2 to 5, wherein the single-stranded DNA fragment is the complementary strand of the DNA strand containing the codons corresponding to the amino acid sequence of the decarbamylase protein.

8. A preparation process according to claim 5, wherein the mutagenesis is effected with hydroxylamine.

9. A preparation process according to claim 5, wherein the mutagenesis is effected with nitrous acid.

10. A preparation process according to any one of claims 1 to 9, wherein the microorganism is *Agrobacterium radiobacter* KNK712 (FERM BP-1900).

11. A preparation process according to claim 1 or 2, wherein the mutagenesis is site-specific mutagenesis at mutation sites.

12. A preparation process according to claim 11, wherein polymerase chain reaction (PCR) is used in the site-specific mutagenesis.

13. A preparation process according to claim 11 or 12, wherein the microorganism is *Agrobacterium radiobacter* KNK712 (FERM BP-1900).

14. A process for producing a decarbamylase improved in thermostability, characterized in that a microorganism transformed with a recombinant plasmid containing a DNA fragment coding for the decarbamylase protein improved in thermostability, which is prepared by the preparation process according to claim 1, is cultured.

15. A production process according to claim 14, wherein the thermostable temperature of the microorganism-



derived decarbamylase is 63°C or lower, and the thermostable temperature of the decarbamylase improved in thermostability is 65°C or higher.

16. A production process according to claim 14, wherein a decarbamylase improved in thermostability by 2°C or higher in the thermostable temperature is used.

17. A production process according to claim 14, wherein the decarbamylase improved in thermostability is a decarbamylase containing a substitution of another amino acid for at least one amino acid selected from 57-histidine, 203-proline and 236-valine of a decarbamylase derived from *Agrobacterium radiobacter* KNK712 (FERM BP-1900).

18. A production process according to claim 17, wherein another amino acid is tyrosine or leucine for 57-histidine, leucine, serine, asparagine, glutamic acid, alanine, isoleucine, histidine or threonine for 203-proline, and alanine, threonine or serine for 236-valine.

19. A production process according to any one of claims 14 to 18, wherein the transformant microorganism is *Escherichia coli* (*E. coli*) JM109 pAD402 (FERM BP-3912), *E. coli* JM109 pAD404 (FERM BP-3913), *E. coli* JM109 pAD406 (FERM BP-3914), *E. coli* JM109 pAD416 (FERM BP-3915), *E. coli* JM109 pAD428, *E. coli* JM109 pAD429 (FERM BP-4035), *E. coli* JM109 pAD431, *E. coli* JM109 pAD434, *E. coli* JM109 pAD435, *E. coli* JM109 pAD439, *E. coli* JM109 pAD441, *E. coli* JM109 pAD445, *E. coli* JM109 pAD447, *E. coli* JM109 pAD448, *E. coli* JM109 pAD450, *E. coli* JM109 pAD421, *E. coli* JM109 pAD422, *E. coli* JM109 pAD423, *E. coli* JM109 pAD424 (FERM BP-4034),



E. coli JM109 pAD425, E. coli JM109 pAD426, E. coli JM109 pAD427, E. coli JM109 pAD451, E. coli JM109 pAD452, E. coli JM109 pAD453, E. coli JM109 pAD461, E. coli JM109 pAD454, E. coli JM109 pAD455 (FERM BP-4036), E. coli JM109 pAD456, E. coli JM109 pAD468, E. coli JM109 pAD469, E. coli JM109 pAD470 or E. coli HB101 pNT4553 (FERM BP-4368).

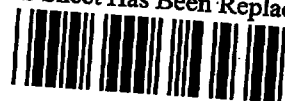
20. A process for producing a D- α -amino acid, characterized in that an N-carbamyl-D- α -amino acid is converted into the corresponding D- α -amino acid in an aqueous medium by the action of a decarbamylase with a thermostable temperature of 65°C or higher, and the resulting D- α -amino acid is collected.

21. A production process according to claim 20, wherein the decarbamylase is allowed to act in an immobilized form.

22. A production process according to claim 21, wherein the immobilized decarbamylase is repeatedly used.

23. A production process according to claim 20, wherein the decarbamylase with a thermostable temperature of 65°C or higher is a decarbamylase containing a substitution of another amino acid for at least one amino acid selected from 57-histidine, 203-proline and 236-valine of a decarbamylase derived from Agrobacterium radiobacter KNK712 (FERM BP-1900).

24. A production process according to claim 23, wherein another amino acid is tyrosine or leucine for 57-histidine, leucine, serine, asparagine, glutamic acid, alanine, isoleucine, histidine or threonine for 203-proline,



and alanine, threonine or serine for 236-valine.

25. A production process according to any one of claims 20 to 23, wherein the decarbamylase with a thermostable temperature of 65°C or higher is the transformant microorganism *E. coli* JM109 pAD402 (FERM BP-3912), *E. coli* JM109 pAD404 (FERM BP-3913), *E. coli* JM109 pAD406 (FERM BP-3914), *E. coli* JM109 pAD416 (FERM BP-3915), *E. coli* JM109 pAD428, *E. coli* JM109 pAD429 (FERM BP-4035), *E. coli* JM109 pAD431, *E. coli* JM109 pAD434, *E. coli* JM109 pAD435, *E. coli* JM109 pAD439, *E. coli* JM109 pAD441, *E. coli* JM109 pAD445, *E. coli* JM109 pAD447, *E. coli* JM109 pAD448, *E. coli* JM109 pAD450, *E. coli* JM109 pAD421, *E. coli* JM109 pAD422, *E. coli* JM109 pAD423, *E. coli* JM109 pAD424 (FERM BP-4034), *E. coli* JM109 pAD425, *E. coli* JM109 pAD426, *E. coli* JM109 pAD427, *E. coli* JM109 pAD451, *E. coli* JM109 pAD452, *E. coli* JM109 pAD453, *E. coli* JM109 pAD461, *E. coli* JM109 pAD454, *E. coli* JM109 pAD455 (FERM BP-4036), *E. coli* JM109 pAD456, *E. coli* JM109 pAD468, *E. coli* JM109 pAD469, *E. coli* JM109 pAD470 or *E. coli* HB101 pNT4553 (FERM BP-4368).

26. A process for producing a D- α -amino acid, characterized in that an N-carbamyl-D- α -amino acid is converted into the corresponding D- α -amino acid in an aqueous medium by the action of a decarbamylase with an improved thermostable temperature by 2°C or higher, and the resulting D- α -amino acid is collected.

27. A decarbamylase characterized in that the thermostable temperature thereof is improved by 2°C or higher.

28. A decarbamylase characterized in that at least



one amino acid of a decarbamylase with a thermostable temperature of 65°C or lower is replaced with another amino acid, so that the thermostable temperature thereof is increased to 65°C or higher.

29. A decarbamylase according to claim 28, wherein the number of replaced amino acids is in the range of 1 to 3.

30. A decarbamylase containing a substitution of another amino acid for at least one amino acid selected from 57-histidine, 203-proline and 236-valine of a decarbamylase derived from *Agrobacterium radiobacter* KNK712 (FERM BP-1900).

31. A decarbamylase according to claim 30, wherein at least 57-histidine is replaced with another amino acid.

32. A decarbamylase according to claim 31, wherein another amino acid is tyrosine or leucine.

33. A decarbamylase according to claim 30, wherein at least 203-proline is replaced with another amino acid.

34. A decarbamylase according to claim 33, wherein another amino acid is leucine, glutamic acid, serine, asparagine, threonine, alanine, isoleucine or histidine.

35. A decarbamylase according to claim 30, wherein at least 236-valine is replaced with another amino acid.

36. A decarbamylase according to claim 35, wherein another amino acid is alanine, threonine or serine.

37. A decarbamylase having an amino acid sequence as shown in Sequence Listing, SEQ ID NOs: 2-34.

38. A DNA fragment coding for a decarbamylase protein improved in thermostability which is attained by substitution of another base for at least one base of a DNA fragment



coding for a microorganism-derived decarbamylase protein so that at least one of the corresponding amino acids is replaced.

39. A DNA fragment according to claim 38, coding for a decarbamylase protein improved in thermostability which is attained by substitution of another base for at least one of the bases constituting a codon corresponding to histidine which is 57-amino acid, proline which is 203-amino acid, or valine which is 236-amino acid, of a DNA fragment coding for a decarbamylase protein derived from *Agrobacterium radiobacter* KNK712 (FERM BP-1900) so that at least one of these amino acids is replaced with another amino acid.

40. A DNA fragment according to claim 39, coding for a decarbamylase protein in which at least one of cytosine, adenine and thymine (CAT) of a base sequence coding for histidine which is the amino acid of amino acid number 57 is replaced with another base so that the histidine is replaced with tyrosine or leucine.

41. A DNA fragment according to claim 39, coding for a decarbamylase protein in which at least one of cytosine, cytosine and thymine (CCT) of a base sequence coding for proline which is the amino acid of amino acid number 203 is replaced with another base so that the proline is replaced with serine, leucine, asparagine, glutamic acid, threonine, alanine, isoleucine or histidine.

42. A DNA fragment according to claim 39, coding for a decarbamylase protein in which at least one of guanine, thymine and guanine (GTG) of a base sequence coding for



valine which is the amino acid of amino acid number 236 is replaced with another base so that the valine is replaced with alanine, serine or threonine.

43. A DNA fragment according to claim 40, wherein the base sequence CAT of claim 40 is replaced with TAT, CTT or CTA.

44. A DNA fragment according to claim 41, wherein the base sequence CCT of claim 41 is replaced with TCT, CTT, AAC, GAA, ACC, GCT, ATT or CAT.

45. A DNA fragment according to claim 42, wherein the base sequence GTG of claim 42 is replaced with GCG, GCT, ACC, ACG, TCA, TCG or AGT.

46. A base sequence comprising a base sequence composed of 233- to 1141-bases of a DNA fragment as shown in Sequence Listing, SEQ ID NOs: 2-16, 24-26, 28 and 33, and a base sequence composed of 7- to 915-bases of a DNA fragment as shown in Sequence Listing, SEQ ID NOs: 17-23, 27, 29 and 34.

47. An expression vector comprising a DNA fragment according to claim 38 to 45.

48. A transformant microorganism transformed with an expression vector according to claim 47.

49. An expression vector which is pAD402, pAD404, pAD406, pAD416, pAD428, pAD429, pAD431, pAD434, pAD435, pAD439, pAD441, pAD445, pAD447, pAD448, pAD450, pAD421, pAD422, pAD423, pAD424, pAD425, pAD426, pAD427, pAD451, pAD452, pAD453, pAD461, pAD454, pAD455, pAD456, pAD468, pAD469, pAD470 or pNT4553.



50. A transformant microorganism which is *E. coli* JM109 pAD402 (FERM BP-3912), *E. coli* JM109 pAD404 (FERM BP-3913), *E. coli* JM109 pAD406 (FERM BP-3914), *E. coli* JM109 pAD416 (FERM BP-3915), *E. coli* JM109 pAD428, *E. coli* JM109 pAD429, (FERM BP-4035), *E. coli* JM109 pAD431, *E. coli* JM109 pAD434, *E. coli* JM109 pAD435, *E. coli* JM109 pAD439, *E. coli* JM109 pAD441, *E. coli* JM109 pAD445, *E. coli* JM109 pAD447, *E. coli* JM109 pAD448, *E. coli* JM109 pAD450, *E. coli* JM109 pAD421, *E. coli* JM109 pAD422, *E. coli* JM109 pAD423, *E. coli* JM109 pAD424 (FERM BP-4034), *E. coli* JM109 pAD425, *E. coli* JM109 pAD426, *E. coli* JM109 pAD427, *E. coli* JM109 pAD451, *E. coli* JM109 pAD452, *E. coli* JM109 pAD453, *E. coli* JM109 pAD461, *E. coli* JM109 pAD454, *E. coli* JM109 pAD455 (FERM BP-4036), *E. coli* JM109 pAD456, *E. coli* JM109 pAD468, *E. coli* JM109 pAD469, *E. coli* JM109 pAD470 or *E. coli* HB101 pNT4553 (FERM BP-4368).

51. A method for increasing, by one or more, the number of mutations in amino acids related to the thermostability of a decarbamylase, characterized in that it comprises:

- 1) obtaining DNA fragments coding for a decarbamylase improved in thermostability or vectors comprising the fragments by repeated use of the process according to claim 1, said decarbamylases each containing one amino acid mutation at the respective amino acid sites;

- 2) finding out restriction enzymes, each of which can yield a DNA fragment comprising a DNA portion coding for one of the thermostability-related amino acids by cleavage for



This Sheet Has Been Replaced



- 122 -

the DNA fragment;

3) obtaining a DNA fragment comprising a DNA portion coding for one amino acid mutation site from the DNA fragments or vectors comprising the fragments in step 1) with the corresponding restriction enzyme in step 2);

4) removing the corresponding DNA fragment but having no amino acid mutation by cleaving a DNA fragment coding for a decarbamylase or a vector comprising the fragment with the same restriction enzyme as described above;

5) incorporating the DNA fragment obtained in step 3) into the remaining DNA fragment after the cleavage for removing the corresponding DNA fragment in step 4) or into the vector comprising the remaining fragment; and

6) repeating the operations in steps 3) to 5), if necessary.

52. A vector plasmid pUC NT for the expression of a foreign gene.