



(86) Date de dépôt PCT/PCT Filing Date: 2000/03/28

(87) Date publication PCT/PCT Publication Date: 2000/10/19

(85) Entrée phase nationale/National Entry: 2001/10/05

(86) N° demande PCT/PCT Application No.: EP 2000/002713

(87) N° publication PCT/PCT Publication No.: 2000/061792

(30) Priorité/Priority: 1999/04/10 (199 16 176.3) DE

(51) Cl.Int.⁷/Int.Cl.⁷ C12Q 1/18, C12Q 1/68

(71) Demandeur/Applicant:
BAYER AKTIENGESELLSCHAFT, DE

(72) Inventeurs/Inventors:
LABISCHINSKI, HARALD, DE;
WIELAND, BERND, DE;
SPALTMANN, FRANK, US;
FREIBERG, CHRISTOPH, DE;
EHLERT, KERSTIN, DE;
BROTZ, HEIKE, DE

(74) Agent: FETHERSTONHAUGH & CO.

(54) Titre : GENES BACTERIENS ESSENTIELS ET LEURS PROTEINES

(54) Title: NOVEL ESSENTIAL BACTERIAL GENES AND THEIR PROTEINS

(57) **Abrégé/Abstract:**

The invention relates to nucleic acids and to protein sequences derived from said nucleic acids which possess functions that are essential to the survival of Escherichia coli. The invention also relates to vectors and host cells that have been transformed with these vectors for expressing essential proteins of the type described above and to proteins produced with these vectors. Finally, the invention also relates to the use of the proteins produced with these vectors in screening methods for identifying antibacterial substances and to the antibacterial substances found with these proteins.



Novel essential bacterial genes and their proteins**A b s t r a c t**

The invention relates to nucleic acids and protein sequences derived therefrom, which have functions essential for the survival of Escherichia coli. It additionally encompasses vectors and host cells transformed with the latter for the expression of such essential proteins, and the proteins produced using these vectors. It further encompasses the use of the proteins produced using these vectors in screening methods for finding antibacterial substances, and of the antibacterial substances found using these proteins.

Novel essential bacterial genes and their proteins

5 Research and development of novel innovative antibiotics formerly took place in particular by modification and optimization of known classes of substances, such as, for example, β -lactams (penicillin-cephalosporins I-V) and testing of their effect in microorganisms in so-called MIC tests.

10 In recent years the search for novel active principles with the aid of biochemically, microbiologically well-characterized targets (points of attack) has become established in parallel. However, the selection of targets by these methods was based on an empirical procedure, which made the finding of novel points of attack very complicated.

15 Recently the knowledge of complete genome sequences of prokaryotes and eukaryotes in conjunction with methods of bioinformatics and molecular biology is opening up the possibility of systematic and rational selection of antibacterial points of attack.

20 Thus, analysis of various bacterial genomes has revealed a large number of genes of unknown function.

Unknown proteins from the Gram-negative model organism *Escherichia coli* (F.R. Blattner et al., Science 277: 1453-1462, 1997) usually have names which begin
25 with "Y": for example YQGF. However, some proteins of unknown function have names deviating from this system, because the genes coding for them are located in the vicinity of functionally characterized genes: for example KDTB (T. Clementz and C.R.H. Raetz, J. Biol. Chem. 266: 9687-9696, 1991; T. Clementz, J. Bacteriol. 174: 7750-7756, 1992).

30

In the setting of the discovery of antiinfective agents, genes of unknown function and proteins derived therefrom represent novel potential points of attack if it can be shown that they are essential for survival of microorganisms (F. Argoni et al., Nature Biotechnology, 16:851, 1998; D. T. Moir et al., Antimicrob. Agents Chemother. 43: 439-446, 1999; US 5821076; B.J. Akerly et al., Proc. Natl. Acad. Sci. USA 95: 8927-8932, 1998).

Knockout technologies can be employed for analyzing the essentiality of such gene sequences for an organism. Knockout strategies are based on replacing the intact (wild-type) gene sequence information in the chromosome of a bacterium by a gene sequence which has been deranged by insertion or deletion. Such a gene sequence thus identifies the coding gene product as essential for this cell from the ability to be capable of survival only with an intact complementary copy in the cell (on a plasmid or a second artificial copy in the chromosome).

Knockout experiments are carried out as homologous DNA sequence recombination events. Vectors able controllably to lose the ability to replicate the plasmid are used for this, such as in the case of temperature-sensitive replicons (pKO3, pMAK 700 or pSC101) or vectors unable to replicate in the desired target cell (E.coli): for example Gram-positive plasmids (pC194, pT181) in Gram-negative bacteria (E. coli). Integration vectors such as pMAK700 harbor not only sequences for temperature-sensitive replication but additionally markers such as antibiotic resistances (for example chloramphenicol) which can be used to select for plasmid-harboring cells.

The invention relates to genes with previously unknown function, to indicate the distribution thereof in the phylogenetically diverse bacterial kingdom and moreover to select in particular those genes which have no or, at the most, only a very remote homology in eukaryotes, and to demonstrate the essentiality of these genes for the ability of microorganisms to survive.

Another demand on the genes to be found is to select, through suitable choice of the search strategy, only those whose derived protein can, on the basis of their expected physicochemical properties, be expected to be soluble in water, so that they can advantageously be used in cell-free assays.

5

It is possible from this analysis to select gene sequences which would not have been selected as the point of attack for finding antibiotic active substances in the conventional way.

10

For this purpose, the gene products of said genes are expressed and purified within the scope of the invention. The novel proteins obtained in this way are used in assays for finding agonists and antagonists of these proteins. In addition, overexpression or underexpression mutants for the corresponding genes are also used in assays for finding agonists and antagonists.

15

The present invention relates to essential genes which encode the proteins from the group of YQGF, YHBC, YGGJ, YGBP, YCHB, YGBB, YJEE, KDTB and genes from other microorganisms which encode the corresponding orthologous gene products, and genes as indicated above for which it can be shown that their controllable switching off leads to inhibition of growth of the host organism, and genes as indicated above from *Escherichia coli* which encode gene products which are 95-100% identical and for which it can be shown that their controllable switching off leads to inhibition of growth of the host organism, and genes as indicated above from other microorganisms which encode orthologous gene products which have 42-100% identical or conservatively exchanged amino acids compared with the corresponding gene product from *E. coli* and for which it can be shown that their controllable switching off leads to inhibition of growth of the host organism, and genes which contain as component of their nucleic acid sequence at least one gene from the group of genes indicated above, in whole or in part, and the gene products of the genes indicated above.

20

25

30

The scope of the invention also includes vectors which contain one or more of the abovementioned genes, and transformed microorganisms which contain one or more of the abovementioned genes, and the use of constructed, recombinant microorganisms in which one of the abovementioned genes is capable of regulatable expression for finding substances which bind to said gene products, and the use of the gene products of the genes indicated above in assays for finding substances which bind to said gene products, and the use of said gene products in assays, where said assays are based on the principle of affinity selection, and the use of said gene products of YQGF in assays, where the assay uses a flavin mononucleotide or a flavin adenine dinucleotide as substrate, and an indicator system which indicates the redox state of the protein, or the use of said gene products of YQGF in assays, where the assay measures the hydrolytic cleavage of a phosphodiester linkage or of another ester linkage, and the use of said gene products of YJEE in assays, where the assay uses an ATP/GTP as substrate, and the use of said gene products of KDTB in assays, where the assay used is a nucleoside derivate as substrate and measures its phosphorylation and/or dehydrogenation, and the use of said gene products of YGBP in assays, where the assay uses a nucleoside diphosphate derivative as substrate and measures conversion thereof, or the use of said gene products of YGBP, where the assay uses a CDP-sorbitol or another polyol substrate and measures pyrophosphorylysis thereof, or where the nucleotide transfer to sorbitol 1-phosphate or another polyol phosphate is measured, the use of said gene products of YGBB in assays, where the key step of the assay is the hydrolytic cleavage of the substrate, or the reduction of the substrate with NAD(P)H, and the use of said gene products of YCHB in assays, where the key step of the assay is either a phosphorylation or a dehydrogenation, and the use of said gene products of YHBC in assays, where the binding to DNA is measured, and the use of said gene products of YGGJ in assays, where the key step of the assay is a phosphorylation, and the use of said gene products or the use of parts of said gene products for finding or preparing antibodies or other proteins which bind to said gene products. The scope of the invention also includes the use of substances which bind said gene products for producing pharmaceuticals, and the use of substances which bind said gene products for

producing pharmaceuticals with antimicrobial activity, and the use of said gene products for finding substances which bind to these gene products, and the preparation and use of antisense constructs directed against the abovementioned genes. Processes for purifying said gene products with the aid of antibodies are also
5 within the scope of the invention.

The substances which bind to the abovementioned gene products and have been found using the processes detailed above, and the pharmaceuticals produced from these substances, can be used for the treatment of infections caused by pathogens in
10 humans and animals.

This preferably includes disorders caused by bacterial pathogens selected from the families of Bacillaceae, Bacteroidaceae, Chlamydiales, Enterobacteriaceae, Micrococcaceae, Mycobacteriaceae, Neisseriaceae, Pasteurellaceae,
15 Pseudomonadaceae, Rickettsiaceae, Spirillaceae, Spirochaetaceae, Streptococcaceae and Vibrionaceae.

The treatment is very particularly preferred on disorders caused by one or more bacterial pathogens selected from the genera Porphyromonas gingivalis, Chlamydia trachomatis, Escherichia coli, Yersinia pestis, Staphylococcus aureus,
20 Mycobacterium tuberculosis, Neisseria gonorrhoeae, Neisseria meningitidis, Bordetella pertussis, Haemophilus influenzae, Pseudomonas aeruginosa, Rickettsia prowazekii, Campylobacter jejuni, Helicobacter pylori, Borrelia burgdorferi, Treponema pallidum, Enterococcus faecalis, Streptococcus mutans, Streptococcus
25 pneumoniae, Streptococcus pyogenes, Vibrio cholerae, Bacillus subtilis.

The substances which bind to the abovementioned gene products and have been found using the processes detailed above, and the pharmaceuticals produced from these substances, can be used in humans and animals for the treatment of disorders
30 and, in particular, for the treatment of infections which affect the blood; the cardiovascular system; the central nervous system and its appendages; the bones, the

bone marrow; the muscles and fasciae, the teeth; the joints and their appendages and cavities; the internal organs, their cavities, internal and external membranes and appendages; the gastrointestinal tract and its cavities, internal and external membranes and appendages; the skin, skin appendages and soft tissues and their
5 appendages; the urogenital system and its appendages; localized and/or generalized inflammations and/or general bacteriemia and sepsis as may occur in the context of systemic disease, trauma, polytrauma, and/or after surgical intervention; opportunistic infections and sepsis also in conjunction with other diseases such as diseases of the blood, viral diseases, consumptive diseases and/or antineoplastic
10 and/or immunosuppressant therapy.

It is possible to treat by way of example diseases occurring after infection with *Porphyromonas gingivalis* in the region of the throat, nose, ear, head and neck, in the region of wounds in the mouth after surgical and/or dental treatment, and/or after
15 bites of humans and animals; *Escherichia coli* in the region of urogenital system including the urinary tract, the bladder, the kidney, the renal pelvis, and urosepsis and sepsis after surgical interventions and other traumas; *Staphylococcus aureus* as pathogen of inflammatory diseases of the respiratory tract and of the nasopharyngeal space, of the lung, of the bone, of the skin, of the appendages of the skin and soft
20 tissues, of the heart and of the heart valves, and toxic diseases caused by the pathogens, such as hemolysis of blood, epidermolysis of the skin, enterocolitis of the gastrointestinal tract, toxic shock syndrome; *Mycobacterium tuberculosis* as tuberculosis and/or tuberculoid diseases of the lung, of the skin, and other organs and organ systems; *Neisseria gonorrhoeae* and the pathological state of gonorrhea
25 resulting therefrom; *Neisseria meningitidis* as pathogen of acute purulent meningitis; *Bordetella pertussis* as pathogen of whooping cough; *Haemophilus influenzae* as pathogen of meningitis, epiglottitis, osteomyelitis, pneumonia, septicemia, laryngotracheitis, pharyngitis, sinusitis, otitis media, septic arthritis, acute endocarditis, cellulitis; *Pseudomonas aeruginosa* and inflammation, resulting
30 therefrom, of wounds, of the urogenital tract, of the endocardium, of the respiratory tract, of the gastrointestinal tract; *Rickettsia prowazekii* and typhus resulting

therefrom, *Chlamydia trachomatis* and conjunctivitis granulosa trachomatosa of the eye resulting therefrom; *Yersinia pestis* and plague resulting therefrom; *Campylobacter jejuni* and inflammatory diseases of the gastrointestinal tract resulting therefrom, such as enteritis, colitis, proctitis, and peptic ulcer; *Helicobacter pylori* and chronic inflammatory diseases of the gastrointestinal tract resulting therefrom such as gastritis and duodenal ulcer, and peptic ulcer; *Borrelia burgdorferi* and Lyme disease and erythema migrans disorder; *Treponema pallidum* and the pathological state of syphilis resulting therefrom; *Enterococcus faecalis* and disorder, resulting therefrom, of the urinary tract, of the urogenital system, of the endocardium, and wound infection; *Streptococcus mutans* and endocarditis and/or inflammation of the heart valves resulting therefrom; *Streptococcus pneumoniae* and inflammatory diseases, resulting therefrom, of the upper and lower respiratory tract such as pneumonia, otitis media, sinusitis, and meningitis, peritonitis, and inflammations of the endocardium and of the pericardium, and of caries; *Streptococcus pyogenes* and inflammatory diseases of the upper respiratory tract, resulting therefrom, such as streptococcal pharyngitis, scarlet fever, otitis media, and pyoderma and erysipelas of the skin and sepsis; *Vibrio cholerae* and the pathological state of cholera resulting therefrom; *Bacillus subtilis* and the pathological state, resulting therefrom, of opportunistic infections, disseminated dispersion and secondary infections associated with other bacterial infections such as otitis, mastoiditis, urinary tract infection, bacteremia, meningitis, endocarditis.

Bioinformatic analysis of various bacterial genomes:

The amino acid sequences of the 4289 proteins of *Escherichia coli* K12 MG1655 with the Genbank/EMBL entry U00096 were compared with the sequences of the complete protein sets of the Gram-positive model organism *Bacillus subtilis* (Genbank/EMBL accession no. AL009126) and of the Gram-negative pathogen *Haemophilus influenzae* (Genbank/EMBL accession no. L42023) and with all other amino acid sequences in the public database (Genpept, Swiss-Prot).

Comparative analyses were carried out with the FASTA program (FASTA vers. 2; Pearson and Lipman 1988).

5 All *E. coli* proteins with significant [expected value $E(N) < 10^{-4}$] sequence homology with proteins of *H. influenzae* and *B. subtilis* were selected.

10 Sequences which show a homology with an expected value $E(N)$ of $< 10^{-4}$ for yeast proteins (H.W. Mewes et al., Nature, suppl. Vol. 387: 9, 1997), and all proteins with known function or with homologies for known proteins from other organisms were excluded.

Also excluded were the proteins from *E. coli* for which homologous protein sequences occur with the *E. coli* species itself [expected value $E(N) < 10^{-4}$].

15 In addition, sequences with more than one predicted transmembrane domains [program: "Peptidestructure", Wisconsin Sequence Analysis Package (Vers. 9, Genetics Computer Group, Madison WI, USA)] were excluded.

Results of the bioinformatic analysis of various bacterial genomes:
20 27 protein sequences were selected from the sequences of potential targets obtained in this way:

25 KDTB, SMPB, YBAB, YBAD, YBAK, YBAX, YBEA, YBEY, YCHB, YFGB, YGAG, YGBB, YGBP, YGGJ, YGGV, YHBC, YHBJ, YHBY, YHIN, YIBK, YIDA, YIGZ, YJEE, YJEQ, YQGF, YRAL, YRFL.

"Knockout" construction and essentiality tests using temperature-sensitive plasmids

Material and methods

30

1. Bacterial strains and plasmids

Table 1: E. coli strains used

E. coli strain	Genotype
E. coli W3110 F	lambda ⁻ IN(rrnD-rrnE)1 rph-1
E. coli MC1061	hsdR ⁻ mcrB ⁻ araD ⁻ 139 ⁻ delta(araABC-leu)7679 delta(lacX74) galU galK rpsL thi
E. coli MG1655	lambda ⁻ F ⁻
E. coli XL1-BLUE	supE44 hsdR17 recA1 endA1 gyrA46 thi relA1 lac ⁻ F'[traD36 proAB ⁺ lacI ^q lacZ delta(M15)]
E. coli JM101	supE thi-1 ((lac-proAB)[F' traD36 proAB lacI ^q Z(M15)])
E. coli TOP10	F ⁻ mrcA ⁻ delta(mrr-hsdRMS-mcrBC)Phi80lacZdeltaM15 delta(lacX74) deoR recA1 araD139 delta(araA-leu)7697 galU galK rpsL endA1 nupG
E. coli M15 [pREP4]	nal ^s str ^s rif ^s lacI ⁻ ara ⁻ gal ⁻ mtl ⁻ F ⁻ recA ⁺ uvr ⁺ lon ⁺ pREP4 (ref. Quiagen, Hilden, Germany)

Table 2: Plasmids used

Plasmid	Reference
pMAK705	C.M. Hamilton et al. J.Bacteriol. 171: 4617-4622 (1989)
pKO3	A.J. Link et al. J. Bacteriol. 179: 6228-6237 (1997)
pAL761	pKO3 + NotI 1739 bp fragment with 890 bp of the region downstream from araD including 108 bp of the 3' end of araD, with a 21 bp linker (5'- AGGCTAGCTTTAGCTCGAGGG-3') which contains the NheI and XhoI cleavage sites, and with 828 bp of the region from 14 bp downstream of

	araB. The araB and araD regions originate from E. coli MG1655.
pCR-blunt	Invitrogen (Groningen, The Netherlands)
pBAD/HisA, HisB	Invitrogen (Groningen, The Netherlands)
PBluescript KS +/-	Stratagene, (Heidelberg, Germany)
pUC4K	Contains kanamycin resistance cassette (Genebank/EMBL Acc. No. X06404)
pBR322	Contains tetracycline resistance cassette (Life Technologies, GibcoBRL, Eggenstein, Germany)

Nutrient media used

LB (Luria-Bertani) medium

- 5 10 g of tryptone, 5 g of yeast extract, 10 g of NaCl in 1 l of H₂O

M9 minimal medium

6 g of Na₂HPO₄, 3 g of KH₂PO₄, 0.5 g of NaCl, 1 g of NH₄Cl in 1 l of H₂O

- 10 Table 3: Antibiotic concentrations in the media used

Antibiotic	Concentration
Chloramphenicol	20 µg/ml
Tetracycline	10 µg/ml
Kanamycin	25 µg/ml
Ampicillin	100 µg/ml

Molecular biology techniques

- 15 Basic molecular biology techniques

Basic molecular biology techniques such as plasmid isolation, cutting with restriction endonucleases, modification of DNA (dephosphorylation, filling-in reaction, ligation), preparation of competent *E. coli* cells and transformation were carried out using conventional protocols known to the skilled workers described, for example, in

5 J. Sambrook et al., *Molecular Cloning*, and Ausubel et al., *Current Protocols in Molecular Biology*.

Polymerase chain reaction (PCR)

All the PCR primers used had calculated melting temperatures of 58°C to 62°C.

10

The following DNA polymerases were used for cloning purposes: Pwo and Taq DNA polymerase (Boehringer, Mannheim, Germany), Akku-Taq DNA polymerase (Sigma-Aldrich Chemie GmbH, Deisenhofen, Germany) and Pfu DNA polymerase (Stratagene, Heidelberg, Germany).

15

Genes and their flanking regions were amplified from 100 ng of chromosomal DNA from *E. coli* W3110 F in a reaction volume of 25 to 100 μ l. The primer concentrations were in each case 1 μ M and the dNTP concentrations were in each case 250 μ M. Twenty-five to thirty reaction cycles at 94°C for 30-60 s, at 48-55°C

20 for 30-60 s and at 72°C for 1-3 min were carried out with a final 5-minute step at 72°C [ref.: *PCR Protocols: a Guide to Method and Application*, Ed. M. Innis et al., Academic Press (1990)].

Assembling PCRs

25 Amplified 5'- and 3'-flanking gene regions were fused by mixing in each case 1 μ l of the two preceding PCR mixtures as template for a second PCR with the primers located distally. This PCR was carried out under the conditions described above.

Diagnostic colony PCRs

30 Diagnostic colony PCRs were carried out with a mixture of Taq DNA polymerase (Boehringer) and ThermoSequenase (Amersham).

One colony in each case was resuspended in 50 μ l of H₂O. 5 μ l of this were used for the PCR, which was carried out in a volume of 50 μ l under the conditions described above.

5

Overexpression of essential proteins in *E. coli*

Various vector systems were used for overexpression of essential proteins in *E. coli*: pBAD/His (Invitrogen), pQE vectors (Quiagen).

10 pBAD/His plasmids (pUC vector derivatives) which were constructed for regulated, dose-dependent expression and purification of recombinant proteins in *E. coli* were used:

15 The pBAD/His vector system includes the sequence which codes for six histidines, including a linker for fusion with the N or C terminus of the required recombinant protein. Maximal expression of the soluble, recombinant protein is achieved by specific activation of the araBAD promoter with arabinose. The proteins were purified by affinity chromatography on nickel-NTA agarose.

20 For purification of the fusion proteins it is necessary first to cultivate the cells containing the pBAD construct to an optical density (OD_{600nm}) of 0.3 to 0.6. The specific protein production is then started with arabinose [0.002-0.2% (w/v)] and incubation is continued for 2-3h. After completion of the incubation, the cells are harvested by centrifugation, and the cell pellet is frozen. Cell disruption and the
25 purification procedure were carried out in accordance with "The QIAexpressionist, A handbook for high-level expression and purification of 6xHis-tagged proteins, Quiagen, Hilden, Germany 1997". The purification steps were checked with polyacrylamide gels (and subsequent staining with Coomassie blue).

30 The pQE vector system includes the sequence which codes for six histidines, including a linker for fusion to an N or C terminus of the required recombinant

protein. Maximal expression of the soluble, recombinant protein is achieved by specific activation of the phage T5 promoter system and induction with IPTG. The proteins were purified by affinity chromatography on nickel-NTA agarose.

- 5 Cell cultivation and purification of the fusion proteins from cells containing the pQE constructs were carried out in accordance with the protocol from "The QIAexpressionist, A handbook for high-level expression and purification of 6xHis-tagged proteins, Quiagen, Hilden, Germany 1997".

10 Cloning strategy

Knockout plasmids:

The essentiality of the open reading frames of the selected genes was tested by using knockout plasmids having the temperature-sensitive origin of replication of pSC101.

- 15 This means that these plasmids are capable of autonomous replication in *E. coli* at 30°C but not at 42-44°C. If DNA regions which are homologous with the *E. coli* chromosome are cloned into such plasmids, it is possible to integrate these plasmids specifically into the gene locus which is to be investigated in the *E. coli* chromosome at 42-44°C by means of homologous recombination. The subsequent excision of the
- 20 plamid at 30°C can lead to replacement of the wild-type sequence by the sequence originally cloned in the plasmid.

- The flanking regions of the investigated genes in the region of in each case 400-800 bp were cloned into pMAK705 or pKO3. In place of the investigated gene, either
- 25 a tetracycline or a kanamycin resistance cassette was integrated into the respective plasmids. Two routes were followed for the cloning strategy:

Integration of the resistance cassette into the gene:

- The gene of interest, including its flanking regions (400-800 bp), was amplified by
- 30 PCR from chromosomal *E. coli* W3110 F DNA and cloned into a cloning vector (pCR-blunt). By restriction digestion and subsequent ligations, most of the coding

region of the gene of interest was replaced by an antibiotic cassette, or the gene was interrupted by such a cassette (tetracycline or kanamycin resistance cassette). The cassette was then cloned, with the sequences originally flanking the gene, into the knockout plasmid pMAK705 or pKO3.

5

Cloning of in-frame deletions with integration of a resistance cassette:

DNA regions with lengths of 500 to 800 bp neighboring the 5' and 3' ends of the particular gene were amplified in two separate PCR reactions from chromosomal E. coli W3110 F DNA. The primers necessary for this were selected so that the ends of the PCR products covering the 5' and 3' ends of the gene additionally had a 21 bp tag with one or two restriction cleavage sites. Accordingly, the two PCR products contained two 21 bp-long extensions complementary to one another. These were used to assemble the two products in a second PCR step with the aid of the primers located on the outside.

15

It was possible to clone the fusion PCR product into pKO3 via restriction cleavage sites incorporated at the 5' ends of the primers located on the outside. The fusion PCR product contained the flanking regions (each 400-800 bp) and exactly 18 nucleotides of the 5' end and 36 nucleotides of the 3' end of the investigated gene, with the 5' and 3' ends of the gene being interrupted by a 21 bp tag. A so-called in-frame deletion of the gene was generated in this way. A kanamycin resistance cassette was integrated into the 21 bp tag via a restriction cleavage site.

20

Temperature shift experiments (integration and excision)

For integration of pMAK705 and pKO3 constructs, E. coli JM101 and E. coli MC1061, respectively, were transformed with the plasmids by standard protocols and incubated on LB agar plates with chloramphenicol at 43-44°C.

25

To make it possible to excise the plasmids, the 43/44°C clones (cointegrates) were then transferred to 30°C LB. The routes followed for this differed according to the vector system (pMAK705 or pKO3):

30

Temperature shifts with pMAK705:

43/44°C clones were transferred into 100 ml of LB-chloramphenicol liquid medium and incubated at 30°C for 16h. After two passages in fresh medium with renewed incubation at 30° for 16 h in each case, the plasmids were isolated from 1 ml of the culture grown last. The prepared plasmids were investigated by restriction digestion for the presence of the wild-type copy of the gene in the plasmid. The culture in which positive plasmids were found was isolated on LB chloramphenicol plates at 30°C.

10

A number of colonies was picked and cultured for the purpose of plasmid preparation. The prepared plasmids were investigated by restriction digestion for the presence of the wild-type copy of the gene in the plasmid. The presence of the antibiotic resistance cassette in the appropriate chromosomal locus was verified by PCR for positive clones. A chromosomal knockout for the investigated gene was generated in *E. coli* in this way, with simultaneous complementation by a functional gene copy on the plasmid pMAK705.

15

Temperature shifts with pKO3:

The cointegrates of the pKO3 derivatives in *E. coli* MC1061 were cut out while, at the same time, eliminating the pKO3 plasmid from the cell by a described method [A.J. Link et al., *J. Bacteriol.* 179: 6228-6237 (1997)]. This entailed three 43/44°C colonies being diluted in Luria-Bertani (LB) medium and plated out on LB-5% (w/v) sucrose-kanamycin plates and incubated overnight. Then 50 to 100 of the clones resulting from each integration clone were again transferred to LB-sucrose-kanamycin and LB-chloramphenicol plates (replica plating). Chloramphenicol-sensitive but kanamycin-resistant colonies were finally analyzed by PCR using primers flanking the gene. It was possible to show on the basis of the size of the PCR product whether the particular clone had a kanamycin resistance cassette in place of the wild-type gene. In this way, a chromosomal knockout for the investigated gene was generated in *E. coli* without functional complementation.

20

25

30

Proofs of essentiality:

The essentiality of the genes was shown in various ways:

- 5 Proof of essentiality by means of temperature shifts with the pMAK705 vector system:

Aliquots were taken from 5 ml of overnight LB-chloramphenicol cultures of the respective complemented knockout strains obtained at the end of the pMAK705 experiments. These aliquots were adjusted to a uniform OD₆₀₀ of 0.1. Dilutions
10 thereof were plated out on LB plates, in each case with and without chloramphenicol, and incubated at 30°C and 43/44°C overnight.

The clones form a distinctly reduced number of colonies on LB-chloramphenicol plates at 43/44°C compared with the number of colonies on LB+/-chloramphenicol at
15 30°C, because most of the clones lose the pMAK705 construct at 43/44°C and are thus no longer able to develop resistance to chloramphenicol. The number of colonies is reduced by a factor of at least 10². Under these conditions, the clones are regarded as incapable of survival.

- 20 Criterion for chromosomal knockout in an essential gene:

Clones which, according to the counting of colonies on LB without chloramphenicol at 43/44°C, formed at least 10² fewer colonies than at 30°C on LB (with or without chloramphenicol) and not more than 10 times more colonies than on LB plates with chloramphenicol at 43/44°C had, according to our criteria, a chromosomal knockout
25 in an essential gene. The clones were not viable if they lost the plasmid at 43/44°C on LB without chloramphenicol (measured in "number of colonies capable of survival").

By contrast, all the clones which formed a similar number (+/- factor of 10) of colonies at 43/44°C on LB without chloramphenicol as on LB (with and without
30 chloramphenicol) at 30°C had a chromosomal knockout in a nonessential gene. Deletion of the plasmid with the complementing functional copy from the E. coli

cells on LB without chloramphenicol at 43/44°C does not lead to a reduction in the viability of the clones (measured in “number of colonies capable of survival”).

5 Proof of essentiality by generation of knockouts exclusively in the presence of a complementing gene copy (pKO3 system):

The clones which were generated with the pKO3 vector system and which were chloramphenicol-sensitive and contained a kanamycin-resistance cassette in place of the wild-type gene had a chromosomal knockout in a nonessential gene. Plasmid-encoded complementation no longer exists.

10

If all the colonies remained chloramphenicol-resistant it was not possible to generate a knockout in the chromosome with simultaneous loss of the plasmid.

15 For further examination of the essentiality of these genes, it was attempted to generate a knockout of the corresponding genes in the chromosome with simultaneous functional complementation.

20 The *E. coli* strain MG1655 which differs from MC1061 in having a complete araBAD locus was used for this purpose. In the first place, the araBAD genes of *E. coli* MG1655 were replaced by the functional copy of the particular gene to be investigated. The plasmid pAL761 was used for this purpose. After transformation of the pAL761 derivatives and following integration and excision steps, the replacement of the araBAD genes by the functional copy of a gene was demonstrated by PCR on the clones unable to grow on M9 minimal medium with 0.2% (w/v) arabinose.

25

Positive clones were subsequently transformed with the pKO3 derivative of the appropriate gene. The subsequent cointegration and excision were carried out as described above, with the exception that the integration and excision step was carried out in the presence of 0.2% (w/v) arabinose.

30

5 PCR was used to diagnose the genotype of deleted chromosomal gene locus with kanamycin cassette among the chloramphenicol-sensitive colonies. If the resulting clones showed a dependence on arabinose for growth (i.e. no single colony formation on LB agar plates without arabinose in contrast to the growth on LB agar plates with arabinose), the corresponding genes were regarded as essential.

Proof of essentiality by means of in-frame deletions (pKO3 system):

10 In some cases it was not possible to generate a replacement of the genes by appropriate interrupted or deleted copies with antibiotic resistance cassettes even in the presence of a complementation. The reason for this may be the polar effects caused by the cassettes, which lead to impaired expression of essential genes located downstream from the investigated gene. For this reason, the in-frame deletion plasmids, described above, of the corresponding genes were used without kanamycin cassette for the following experiments.

15

In-frame deletion clones were formed into *E. coli* MC1061. Integration, excision and elimination of pKO3, and subsequent replica plating, were carried out as described above, with the modification that kanamycin selection was no longer used. Viewed statistically, 50% of the chloramphenicol-sensitive colonies ought to harbor the gene deletion. If no clone among 50 colonies tested by PCR harbored the deletion, further experiments followed to confirm the essentiality of the investigated gene.

20

For this purpose, a deletion of the gene in the chromosome was generated with simultaneous complementation as described above, and dependency on arabinose for growth was shown.

25

Proof of essentiality on complemented knockouts without arabinose-dependent growth:

30

It was attempted with the complemented knockout strains showing no arabinose-dependent growth to replace the functional copy of the gene in the araBAD locus by a

kanamycin cassette. For this purpose, the corresponding strains were transformed with pAL767, a pKO3 derivative with kanamycin-resistance cassette and araBAD-flanking regions.

- 5 If, after integration, excision and elimination of the pAL767, 100% of the resulting clones (200 clones inspected) were both kanamycin- and chloramphenicol-resistant, it was shown that pAL767 were still present integrated in the chromosome. It was thus impossible to delete the intact copy of the gene from the araBAD locus. This finding underlines the essentiality of the relevant gene. If it was possible in any
- 10 kanamycin-resistant and chloramphenicol-sensitive clones to show successful replacement of the functional gene copy by the kanamycin resistance cassette by PCR, the gene was nonessential.

15 Results of proofs of essentiality:

Table 4: Genes whose essentiality was assessed by the abovementioned methods:

Knockout with pMAK705 system:	
Essential:	YJEE, KDTB
Nonessential:	YHBY, YBAB, YBEY, YHIN, YJEQ, SMPB

20

Knockout with pKO3 system:	
Essential	YQGF, YGBP, YHBC, YGGJ, YGBB, YCHB
Nonessential:	YBAK, YBAX, YGGV, YBAD, YIGZ, YRFI, YRAL, YFGB, YIDA, YIBK, YGAG, YHBJ, YBEA

Accordingly, eight of the 27 investigated genes were identified as essential. The protein sequences derived from these genes were used to carry out a homology search comparing with nucleotide sequences of completely and incompletely deciphered bacterial genomes using the TBLASTN program.

5

It is possible with this method to distinguish all significantly homologous protein sequences from sequences which agree by chance through the choice of a threshold value, the expected value $P(N)$. It was possible with use of a stringent threshold value (expected value $P(N) < 10^{-4}$) to find gene products with a significant sequence
10 homology in a large number of disease-relevant bacteria.

15

The homologs of a bacterial species with the greatest similarity which were selected from those found in this way were those having the maximum number of identical and/or conservatively exchanged amino acids compared with the corresponding
E. coli protein (Table 5). Such proteins are referred to as orthologous proteins (gene products) because the skilled worker knows that they carry out the same function as in Escherichia coli.

20

The orthologous proteins listed in Table 5 have at least 21% identical or 42% identical or conservatively exchanged amino acids, i.e. 21% identity/42% similarity.

Table 5:

Distribution of the orthologous proteins of the identified essential gene products of E. coli in pathogenic bacteria. Results of a TBLASTN search with the E. coli protein sequences found to be essential compared with nucleotide sequences of completely or incompletely deciphered bacterial genomes as deposited, for example, in the publicly accessible genome database of the National Center for Biotechnology Information (NCBI): criterion for significant homology is an expected value $P(N) < 10^{-4}$. The identity and similarity of the orthologs by comparison with the corresponding E. coli protein has been indicated as %identity/%similarity. Completely sequenced genomes are marked by (*). Lack of homology in incompletely deciphered genomes is indicated by (?).

Le A 33 616-Foreign Countries

- 22 -

Organism/ ORF	YJEE	KDTB	YQGF	YHBC	YGGJ	YGBP	YCHB	YGBB
<i>E. coli</i>	100/100	100/100	100/100	100/100	100/100	100/100	100/100	100/100
<i>Haemophilus influenzae</i> *	55/74	52/73	70/81	55/73	63/76	53/71	58/72	64/72
<i>Yersinia pestis</i>	73/80	73/87	76/91	90/97	72/83	68/84	73/81	75/80
<i>Vibrio cholerae</i>	67/80	61/73	64/80	67/86	61/69	56/72	62/72	64/73
<i>Pseudomonas aeruginosa</i>	58/70	61/80	50/67	49/69	43/59	46/61	52/65	60/67
<i>Bordetella pertussis</i>	44/61	52/74	47/65	39/59	35/52	32/45	47/60	48/56
<i>Neisseria meningitidis</i>	42/64	37/62	41/55	40/60	36/53	43/59	40/56	48/60
<i>Neisseria gonorrhoeae</i>	48/64	36/64	41/54	51/65	36/52	46/61	39/57	48/60
<i>Rickettsia prowazekii</i> *	31/54	-	31/50	43/67	28/49	-	-	-
<i>Campylobacter jejuni</i>	33/58	48/72	34/51	41/64	26/48	25/45	26/48	43/58
<i>Helicobacter pylori</i> *	34/52	49/74	29/43	35/58	28/51	28/42	-	37/54
<i>Porphyromonas gingivalis</i>	35/56	41/59	30/46	?	30/49	29/43	32/52	46/60
<i>Streptococcus pneumoniae</i>	32/50	37/62	38/57	35/61	28/49	26/45	?	?
<i>Streptococcus pyogenes</i>	35/55	42/61	33/56	39/61	26/47	?	?	?
<i>Streptococcus mutans</i>	30/52	40/58	?	39/57	28/49	?	30/49	?
<i>Enterococcus faecalis</i>	?	44/70	36/57	40/62	26/44	21/44	32/52	44/59
<i>Staphylococcus aureus</i>	39/61	46/69	27/52	36/62	32/52	?	?	?
<i>Mycobacterium tuberculosis</i> *	35/50	44/67	28/43	34/58	24/44	33/48	36/52	32/44
<i>Bacillus subtilis</i> *	37/53	43/67	33/55	38/63	32/50	37/57	29/49	56/63
<i>Chlamydia trachomatis</i> *	41/55	-	27/43	-	-	31/48	34/53	29/44
<i>Borrelia burgdorferi</i> *	39/55	35/59	-	-	23/44	-	-	-
<i>Treponema pallidum</i> *	35/52	38/57	-	-	-	35/44	29/49	33/47

It is possible on the basis of sequence motifs in the amino acid sequence of the eight essential gene products and/or their spatial structure to be expected from the amino acid sequence to predict the following functions:

YJEE probably represents an ATP- or GTP-binding protein. The protein may have ATPase/GTPase or ATP/GTP synthase activity.

5 KDTB is probably a sugar- and/or nucleoside and/or nucleoside derivative-binding protein with phosphorylation or dehydrogenase activity.

YQGF possibly represents a flavine derivative-binding protein. It may bind flavomononucleotide (FMN) or flavin adenine dinucleotide (FAD) and transfer electrons to another substrate/protein. YQGF possibly represents a hydrolysis which
10 cleaves phosphodiester linkages or other ester linkages.

YHBC is a putative DNA-binding protein with modulatory/regulatory function. YGGJ is a possible phosphotransferase. YGBP is probably a nucleoside diphosphate derivative phosphorylase or pyrophosphorylase. YGBP probably cleaves CDP-
15 sorbitol or another polyol nucleoside diphosphate with the aid of pyrophosphate or transfers the nucleotid CMP to sorbitol 1-phosphate or another polyol phosphate. YCHB is a putative kinase or a putative oxidoreductase. YGBB is a putative hydrolase or NAD(P)H-dependent reductase.

Examples

Example I on YJEE

Construction of the integration vector with the gene sequence for yjeE

5

The integration plasmid pMAK705 (5.5 kb) it is possible to integrate by homologous recombination into the chromosome of E. coli JM101. Besides a temperature-sensitive origin of replication from pSC101, the plasmid contains a chloramphenicol resistance gene for selection of the plasmid in the host organism.

10

The coding sequence for yjeE and 420 base pairs (bp) upstream and 413 bp downstream of the coding sequence for yjeE were amplified with PCR (Perkin Elmer 480 cyclor). 100 ng of chromosomal DNA from E.coli W3110 F were amplified with Pfu DNA polymerase in a reaction volume of 100 µl in 25 cycles at 94°C for 45 sec, at 48°C for 45 sec and at 72°C for 3 min, and a final incubation at 72°C for 5 minutes. The amplification primers used were primer A (5'-CCT GCT GGC AAT CAA TCC CGA TA-3') and primer B (5'-AGG CGG TGG CGG CAC ATCGGC GTT-3'). The amplification product (1482 bp) was purified using Qiaex (Qiagen, Hilden, Germany) and then subcloned into the vector pCR-blunt using the "pCR - blunt cloning kit" (Invitrogen Corporation, Carlsbad,CA). A tetracycline resistance cassette was cloned into the BpmI restriction cleavage site with pCR - blunt/yjee. The tetracycline cassette was isolated as 1400 bp EcoRI/AvaI fragment from the plasmid pBR322, and blunt ends were produced with Klenow polymerase.

20

25 The plasmid pCR - blunt/yjeE::tet produced in this way was digested with KpnI/XbaI restriction enzymes. A DNA band 2.8 kb in size was isolated and cloned into the vector pMAK705 pretreated with KpnI/XbaI (host organism E.coli JM101).

The resulting construct (pMAK705yjeE) was used for the integration experiments.

30

Preparation and analysis of the yjeE knockout

Cells with the plasmid pMAK705yjeE were cultured in LB medium at 30°C overnight and, the next day, plated out on LB test plates with chloramphenicol and incubated at 43°C for 16 h. Integration of the plasmid into the chromosome of E.coli JM101 was identified by selection of chloramphenicol-resistant colonies able to grow at 43°C. After identification of the integration (at 43°C), the cells were incubated in 100 ml of LB medium at 30°C for 16 h. These cells were transferred twice into fresh medium and incubated further at 30°C for 16 h each time.

After this series of incubations, quick plasmid preparations were carried out with Qiaprep (Qiagen, Hilden, Germany) on 1 ml of each of the last incubation mixtures. The prepared plasmids were examined by restriction digestion for the presence of the wild-type copy of the gene in the plasmid. Cells from the mixtures which contained positive clones (yjeE gene from the wild-type chromosome on the plasmid) were isolated on LB plates with chloramphenicol at 30°C.

The plasmid DNA of the isolated clones was again tested for intact yjeE gene from the chromosome, and positive colonies were identified. Integration of the yjeE copy inactivated by insertion of the tetracycline cassette in the chromosome was tested by PCR.

The clones examined in this way were then employable for essentiality of the target yjeE for survival of the E.coli cells. The cells were incubated on LB test plates with and without chloramphenicol at 30° and 43°C.

Table 6: Cell counts on incubation +/-chloramphenicol at 30°C/43°C

	+Chloramphenicol	-Chloramphenicol
30°C	1×10^7	0.7×10^7
43°C	1×10^3	1.6×10^3

As is evident from the table, the cells are able to grow at 30°C both with and without chloramphenicol. At 43°C, the growth is drastically, i.e. by a factor of 10^4 , reduced both with chloramphenicol and without chloramphenicol. This clearly shows that the presence of the plasmid with the yjeE-complementing sequence is necessary for growth of the E. coli cell at 43°C without chloramphenicol, and thus that yjeE is essential for survival of these cells.

Example: Demonstration of essentiality for ygbP

Cloning of in-frame deletions with and without integration of a resistance cassette:

DNA regions neighboring the 5' and 3' ends of the ygbP gene were amplified from 100 ng of chromosomal E.-coli W3110 F DNA in a reaction volume of 25 μ l in two separate PCR reactions. The primer concentrations were in each case 1 μ M, and the dNTP concentrations were in each case 250 μ M. Twenty-five reaction cycles at 94°C for 30 s, at 52°C for 30 s and at 72°C for 1-2 min, with a final 5-minute step at 72°C, were carried out. The PCR product which was prepared with the primers YGBP1A (5'-cgcgatccCCACATGGTCACTGCCTGG-3') and YGBP1B (5'-cccatccactaaactgcagctCAAATGAGTGGTTGCCATGTT-3') comprises 18 bp of the 5' end of ygbP and 776 bp of the chromosomal region upstream from ygbP. The PCR product prepared with the primers YGBP2A (5'-agctgcagtttagtgatgggTACCTACCCGAACCATCC-3') and YGBP2B (5'-cgcgatccACCTGGCAGCCTTCCAGTTG-3') comprises 36 bp of the 3' end of ygbP and 807 bp of the chromosomal region downstream from ygbP.

The internal primers YGBP1B and YGBP2A additionally contained at their 5' ends 21mer tags complementary to one another (small letters in the abovementioned sequences). These were used to assemble the two PCR products in a second PCR step using the primers YGBP1A and YGBP2B located on the outside. For this purpose, 1 μ l of each of the two preceding PCR mixtures were mixed as template for a second PCR with YGBP1A and YGBP2B. This PCR was carried out under the

conditions described above. The PCR product of the expected size was purified by an agarose gel electrophoresis using QIAEX (QIAGEN, Hilden, Germany).

5 It was possible to clone the fusion PCR product into the BamHI-cut, dephosphorylated pKO3 via the BamHI cleavage site which were incorporated at the 5' ends of the primers YGBP1A and YGBP2B located on the outside (small letters in the abovementioned sequences). The plasmid pAL759 which contained the in-frame deletion of ygbP was generated in this way. The kanamycin-resistance cassette
10 obtained by a PstI digestion from pUC4K was cloned into pAL759 via the PstI cleavage site in the 21 bp tag (5'-ctgcag-3') . The resulting construct was called pAL759a.

Temperature shifts with pAL759a:

15 E. coli MC1061 was transformed with plasmid pAL759a and incubated on LB agar plates with chloramphenicol at 44°C.

Three of the resulting 44°C clones (cointegrates) were diluted in LB medium, plated out on LB-5% (w/v) sucrose-kanamycin plates and incubated overnight. 100 of the
20 resulting sucrose-kanamycin clones for each integration clone were again transferred to LB-sucrose-kanamycin and LB-chloramphenicol plates (replica plating) and incubated at 30°C overnight. 100% of the kanamycin-resistant clones were also chloramphenicol-resistant. It was thus not possible to carry out a replacement of ygbP by a kanamycin cassette successfully.

25

Temperature shifts with pAL759:

E. coli MC1061 was transformed with the plasmid pAL759 and incubated on LB agar plates with chloramphenicol at 44°C.

30

Three of the resulting 44°C clones (cointegrates) were diluted in LB medium, plated out on LB-5% (w/v) sucrose plates and incubated overnight.

100 of the resulting sucrose clones for each integration clone were transferred again
5 to LP-sucrose and LB-chloramphenicol plates (replica plating) and incubated at 30°C overnight. In each case 50 chloramphenicol-sensitive colonies were finally analyzed by PCR using YGBP1A and YGBP2B as primers. It was possible to show on the basis of the size of the PCR product that none of the tested clones had the deletion of ygbP (PCR product with ygbP deletion: 1637 bp). It was possible in all the colonies
10 to amplify the wild-type gene locus of ygbP (2294 bp). It was thus not possible to carry out replacement of ygbP by a deletion successfully.

Generation of a ygbP deletion in the presence of a chromosomal complementation:

15 It was attempted in the next step to generate a deletion of ygbP in the chromosome with simultaneous functional complementation. The E. coli strain MG1655 which differs from MC1061 in having a complete araBAD locus was used for this purpose. Firstly the araBAD genes of E. coli MG1655 were replaced by the functional copy of ygbP. The plasmid pAL763 was used for this purpose. The plasmid pAL763 was
20 generated by amplifying the complete YGBP gene, including its ribosome binding site (18 bp upstream from the 5' end) and including 8 bp downstream from the 3' end of ygbP, from chromosomal E. coli W3110 F DNA using the primers YGBPR (5'-atcgctagcATCAGCCCGGGAATTAACATG-3') and YGBPT (5'-atcctcgagAATTCGCATTATGTATTCTCCTG-3'). The primers contain at their 5'
25 termini the restriction cleavage sites for XhoI (YGBPT) and NheI (YGBPR). The PCR product was cloned into the vector pAL761 via the cleavage sites. After transformation of E. coli MG1655 with pAL763 and following integration and excision steps (see above), replacement of the araBAD genes by the functional copy of ygbP was demonstrated by PCR on the clones unable to grow on M9 minimal
30 medium with 0.2% (w/v) arabinose. The diagnostic PCR was carried out with the primers YGBPR (see above) and ARA2B (atcgcgccgcAAAGCCGTGCTCGCGC).

The primer ARA2B binds in the 3' region 865 bp downstream from the araD gene so that, together with the primer YGBPR, a 1655 bp product was produced with positive clones. Positive clones were subsequently transformed with the in-frame deletion plasmid pAL759. The subsequent cointegration and excision were carried out as described above, with the exception that the integration and excision step was carried out in the presence of 0.2% (w/v) arabinose. The genotype of the deleted chromosomal gene locus was diagnosed among the chloramphenicol-sensitive colonies by PCR. The ygbP deletion clone which contained an arabinose-regulatable ygbP copy in the araBAD locus showed a dependence on arabinose for growth. It was not possible for single colonies to form on LB agar plates without arabinose, and single colonies were obtained on LB agar plates with arabinose. The ygbP gene is accordingly essential.

Example 3

Preparation and purification of recombinant YJEE protein

Expression and purification of YJEE protein was carried out with the pBAD/His
5 vector system (invitrogen). The sequence of yjeE was amplified from 100 ng of
chromosomal DNA from E. coli W3110F with Pfu DNA polymerase in a 100 μ l
reaction mixture in 25 cycles at 94°C for 1 min, at 52°C for 1 min, at 72°C for
3 min. The primers employed were YJEE-START (5'-TGA AGA TCT AAT CGA
GTA ATT CCG CTC CCT GAT-3') and YJEE-STOP (5'-TCA GAA TTC TTA
10 ACC GGC TAA ACG CGC CAG CAA CAA TC-3'). The 5' end of YJEE-START
contains the sequence for the restriction enzyme BglII. The 5' end of YJEE-STOP
contains the sequence for EcoRI. The PCR mixture was fractionated in an agarose
gel, and a band at 450 bp was cut out. The band was then isolated using Qiaex
(Qiagen, Hilden, Germany) and incubated with the restriction enzymes BglII and
15 EcoRI at 37°C for 3 h.

The DNA fragment treated in this way was cloned into the vector pBAD/HisB which
had been treated with BglII/EcoRI [host organism: E.coli strain TOP10 (invitrogen)].
Successful cloning was shown by plasmid isolation of the resulting transformands
20 and restriction with EcoRI. The clone with the plasmid pBAD/HisB-YjeE30 was
employed for expression and purification.

The culturing and purification of YjeE from E. coli TOP 10 with plasmid
pBAD/HisB-YjeE30 was carried out by the protocol for the purification of histidine-
25 tagged proteins on Ni-NTA agarose (The QIAexpressionist, A handbook for high-
level expression and purification of 6xHis-tagged proteins, Quiagen, Hilden,
Germany 1997). The induction of YjeE expression was started with 0.02% arabinose
(w/v) at an optical density of the culture of $OD_{600} = 0.5$. Various concentrations of
imidazole (50, 100, 150 and 200 mM) were used to elute the bound protein. The
30 highest yield of protein was obtained with 150 mM. The purification was checked by
polyacrylamide gel fractionation and subsequent Coomassie staining. It was possible

to isolate 10 mg of protein with a purity of >90% from 1 l of culture medium (LB medium).

Example 4

5 Screening of libraries of low molecular weight chemical compounds with targets of unknown function based on affinity selection and mass spectrometry

10 To find inhibitors for the target proteins with an unknown function which is, however, essential for bacterial survival (such as, for example, YJEE) it is possible to employ screening methods which test substance libraries for affinity for the protein. One possibility for screening is affinity selection from substance mixtures with subsequent detection of the ligands in a mass spectrometer. It is necessary to use for this defined substance mixtures from which individual substances can be identified by mass detection. Substance mixtures prepared by combinatorial syntheses are
15 therefore particularly suitable for this screening method.

A liquid chromatography electrospray mass spectrometer is used as screening apparatus, the HPLC column being replaced by an ultrafiltration chamber. The ultrafiltration chamber consists of a filtration unit in which the filter disk is replaced
20 by an ultrafiltration membrane (exclusion molecular weight 10 kDa). Mass spectra are produced using a mass spectrometer (from Hewlett-Packard; Palo Alto, CA; "5989B MS Engine Quadrupole Mass Spectrometer"). It is operated as described (van Breemen et al., Pulsed Ultrafiltration Mass Spectrometry. A New Method for Screening Combinatorial [lacuna]: Anal. Chem., 69 [11], 2159-2164, 1997).

25

An equimolar mixture of 96 compounds (about 200 μ M each) in 10 μ l of DMSO is added to 90 μ l of 2 μ M protein (for example YJEE) which is present in a 20 mM ammonium acetate buffer. The protein/substance mixture is incubated at room temperature for at least 15 min. Injection of the mixture into the ultrafiltration
30 chamber is followed by washing with water at a flow rate of 50 l/min for 8 min. The mobile phase is then changed to methanol/water (50:50 v/v) in order to dissociate

possible ligand-protein complexes and filter out the ligands. The ligands are detected by electrospray mass spectrometry. The substances indentified from a mixture of 96 as possible ligands on the basis of their mass peaks are employed again as single substances in the affinity test in order to verify their affinity for the target protein (for
5 example YJEE).

Example 5

Substances picked up in the affinity selection are subjected to further tests, such as the MIC test.

10

For this purpose, the MIC values are determined by the microdilution method in BH medium. Each test substance is dissolved in the nutrient medium. Series of test substance concentrations are made up in the microtiter plate by serial dilution. Overnight cultures of the pathogens are used for inoculation after dilution 1:250 in
15 nutrient medium. 100 ml of each inoculation solution are added to 100 ml of the diluted active substance-containing nutrient solutions.

The microtiter plates are incubated at 37°C and read after about 20 hours or after 3 to 5 days. The MIC (mg/ml) indicates the lowest concentration of active substance
20 at which no growth is evident.

Another method is determination of the minimum inhibitory concentration (MIC) in the liquid dilution test. Overnight cultures of the test organisms (S. aureus 133) in Isosensitest broth are diluted 1:1000 in fetal calf serum/FCS) or Isosensitest broth
25 and incubated with dilutions of the test substances (1:2 dilution stages).

The cultures are incubated at 37°C for 18 to 24 hours. The lowest concentration of each substance at which there is no longer any visible bacterial growth is defined as the MIC.

30

References:

- Blattner et al. Science 277:1453-1462
- F. Argoni et al., Nature Biotechnology, 16:851 (1998)
- H.W. Mewes et al. Nature, suppl. Vol. 387, p 9 (1997)
- 5 Hamilton et al. J.Bacteriol. 171: 4617-4622 (1989)
- Link et al. J. Bacteriol. 179: 6228-6237 (1997)
- US 5821076 Identification of Essential Survival Genes
- B.J. Akerly et al. Proc. Natl. Acad. Si. USA 95: 8927-8932, 1998
- T. Clementz und C. R. H. Raetz: J. Biol. Chem. 266: 9687-9696, 1991
- 10 T. Clementz, J. Bact. 1992 174: 7750-7756, 1992
- D. T. Moir et al., Antimicrob. Agents Chemother. 43: 439-446, 1999;

SEQUENCE LISTING

<110> Bayer AG

<120> Novel essential genes and their proteins

<130> Essential Genes of EColi

<140>

<141>

<160> 28

<170> PatentIn Vers. 2.0

<210> 1

<211> 153

<212> PRT

<213> Escherichia coli

<220>

<223> YJEE

<400> 1

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

Leu Gly Glu Arg Val Ala Lys Ala Cys Asp Gly Ala Thr Val Ile Tyr
20 25 30

Leu Tyr Gly Asp Leu Gly Ala Gly Lys Thr Thr Phe Ser Arg Gly Phe
35 40 45

Leu Gln Ala Leu Gly His Gln Gly Asn Val Lys Ser Pro Thr Tyr Thr
50 55 60

Leu Val Glu Pro Tyr Thr Leu Asp Asn Leu Met Val Tyr His Phe Asp
65 70 75 80

Leu Tyr Arg Leu Ala Asp Pro Glu Glu Leu Glu Phe Met Gly Ile Arg
85 90 95

Asp Tyr Phe Ala Asn Asp Ala Ile Cys Leu Val Glu Trp Pro Gln Gln
100 105 110

Gly Thr Gly Val Leu Pro Asp Pro Asp Val Glu Ile His Ile Asp Tyr
115 120 125

Gln Ala Gln Gly Arg Glu Ala Arg Val Ser Ala Val Ser Ser Ala Gly
130 135 140

Glu Leu Leu Leu Ala Arg Leu Ala Gly
145 150

<210> 2

<211> 462

<212> DNA

<213> Escherichia coli

<220>

<223> yjeE Position 3299..3760 in accession number

AE000489; U00096

<400> 2

```
atgatgaatc gagtaattcc gctccctgat gagcaggcaa cattagacct gggcgagcgg 60
gtagcgaaag cctgcgatgg cgcaaccgta atctatctgt atggcgattt aggcgcaggt 120
aaaaccacct ttagccgggg ctttttacag gctctgggtc atcagggtaa tgtcaaaagc 180
cccacttata cgctggtcga accctatacg ctcgacaact taatgggtcta tcactttgat 240
ttgtaccgcc ttgccgatcc cgaggagctg gagtttatgg ggatccgcga ttattttgcc 300
aacgatgcca tctgcctggt ggagtggcca caacaaggta caggtgttct tcctgaccgc 360
gatgtcgaaa tacacattga ttatcaggca caaggccgtg aggcgcgcgt gagtgcgggt 420
tcctctgcgg gtgaattggt gctggcgcgt ttagccgggt aa 462
```

<210> 3

<211> 159

<212> PRT

<213> Escherichia coli

<220>

<223> KDTB

<400> 3

```
Met Gln Lys Arg Ala Ile Tyr Pro Gly Thr Phe Asp Pro Ile Thr Asn
  1              5              10              15

Gly His Ile Asp Ile Val Thr Arg Ala Thr Gln Met Phe Asp His Val
              20              25              30

Ile Leu Ala Ile Ala Ala Ser Pro Ser Lys Lys Pro Met Phe Thr Leu
  35              40              45

Glu Glu Arg Val Ala Leu Ala Gln Gln Ala Thr Ala His Leu Gly Asn
  50              55              60

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

Gln His Ala Thr Val Leu Ile Arg Gly Leu Arg Ala Val Ala Asp Phe
              85              90              95

Glu Tyr Glu Met Gln Leu Ala His Met Asn Arg His Leu Met Pro Glu
              100              105              110

Leu Glu Ser Val Phe Leu Met Pro Ser Lys Glu Trp Ser Phe Ile Ser
              115              120              125

Ser Ser Leu Val Lys Glu Val Ala Arg His Gln Gly Asp Val Thr His
              130              135              140

Phe Leu Pro Glu Asn Val His Gln Ala Leu Met Ala Lys Leu Ala
145              150              155
```

<210> 4

<211> 480

<212> DNA

<213> Escherichia coli

<220>

<223> kdtB Position 1720..2199 of accession number
AE000441; U00096

<400> 4

```

atgcaaaaac gggcgattta tccgggtact ttcgatccca ttaccaatgg tcatatcgat 60
atcgtgacgc gcgccacgca gatgttcgat cacgttattc tggcgattgc cgccagcccc 120
agtaaaaaac cgatgtttac cctggaagag cgtgtggcac tggcacagca ggcaaccgcg 180
catctgggga acgtggaagt ggtcgggttt agtgatttaa tggcgaactt cgcccgtaat 240
caacacgcta cggtgctgat tcgtggcctg cgtgcggtgg cagattttga atatgaaatg 300
cagctggcgc atatgaatcg ccacttaatg ccggaactgg aaagtgtgtt tctgatgccg 360
tcgaaagagt ggtcgtttat ctcttcacgc ttggtgaaag aggtggcgcg ccatcagggc 420
gatgtcacc c atttcctgcc ggagaatgtc catcaggcgc tgatggcgaa gttagcgtag 480

```

<210> 5
 <211> 138
 <212> PRT
 <213> Escherichia coli

<220>
 <223> YQGF

<400> 5
 Met Ser Gly Thr Leu Leu Ala Phe Asp Phe Gly Thr Lys Ser Ile Gly
 1 5 10 15
 Val Ala Val Gly Gln Arg Ile Thr Gly Thr Ala Arg Pro Leu Pro Ala
 20 25 30
 Ile Lys Ala Gln Asp Gly Thr Pro Asp Trp Asn Ile Ile Glu Arg Leu
 35 40 45
 Leu Lys Glu Trp Gln Pro Asp Glu Ile Ile Val Gly Leu Pro Leu Asn
 50 55 60
 Met Asp Gly Thr Glu Gln Pro Leu Thr Ala Arg Ala Arg Lys Phe Ala
 65 70 75 80
 Asn Arg Ile His Gly Arg Phe Gly Val Glu Val Lys Leu His Asp Glu
 85 90 95
 Arg Leu Ser Thr Val Glu Ala Arg Ser Gly Leu Phe Glu Gln Gly Gly
 100 105 110
 Tyr Arg Ala Leu Asn Lys Gly Lys Val Asp Ser Ala Ser Ala Val Ile
 115 120 125
 Ile Leu Glu Ser Tyr Phe Glu Gln Gly Tyr
 130 135

<210> 6
 <211> 417
 <212> DNA
 <213> Escherichia coli

<220>
 <223> yggF Position 11818...12234 of accession number
 AE000377; U00096

<400> 6
 atgagtggaa ccttactcgc cttcgacttc ggcaccaaaa gcattggcgt agcggtcggc 60
 caacgcatta ccggcaccgc tcgccctttg cctgcaatta aagcacagga cggtagcggc 120
 gactggaaca ttatcgagcg tttactgaaa gactggcagc cggacgaaat catcgtcggc 180
 ttgccgctga atatggacgg caccgagcaa ccattgactg ccagagcgcg taaatttgcc 240
 aaccgtattc atggccggtt cggtgttgaa gtaaagctcc atgacgagcg tcttagcact 300
 gtggaagccc gttccggtct gtttgaacag ggcggctatc gggcgctcaa caaaggcaaa 360
 gttgactctg cctctgcggc tattattctc gaaagctatt tcgagcaggg atattaa 417

Le A 33 616-Foreign Countries

- 37 -

<210> 7
 <211> 252
 <212> PRT
 <213> Escherichia coli

<220>
 <223> YGGJ

<400> 7
 Met Gly Arg Arg Arg Gly Phe Leu Thr Met Arg Ile Pro Arg Ile Tyr
 1 5 10 15
 His Pro Glu Pro Leu Thr Ser His Ser His Ile Ala Leu Cys Glu Asp
 20 25 30
 Ala Ala Asn His Ile Gly Arg Val Leu Arg Met Gly Pro Gly Gln Ala
 35 40 45
 Leu Gln Leu Phe Asp Gly Ser Asn Gln Val Phe Asp Ala Glu Ile Thr
 50 55 60
 Ser Ala Ser Lys Lys Ser Val Glu Val Lys Val Leu Glu Gly Gln Ile
 65 70 75 80
 Asp Asp Arg Glu Ser Pro Leu His Ile His Leu Gly Gln Val Met Ser
 85 90 95
 Arg Gly Glu Lys Met Glu Phe Thr Ile Gln Lys Ser Ile Glu Leu Gly
 100 105 110
 Val Ser Leu Ile Thr Pro Leu Phe Ser Glu Arg Cys Gly Val Lys Leu
 115 120 125
 Asp Ser Glu Arg Leu Asn Lys Lys Leu Gln Gln Trp Gln Lys Ile Ala
 130 135 140
 Ile Ala Ala Cys Glu Gln Cys Gly Arg Asn Arg Val Pro Glu Ile Arg
 145 150 155 160
 Pro Ala Met Asp Leu Glu Ala Trp Cys Ala Glu Gln Asp Glu Gly Leu
 165 170 175
 Lys Leu Asn Leu His Pro Arg Ala Ser Asn Ser Ile Asn Thr Leu Pro
 180 185 190
 Leu Pro Val Glu Arg Val Arg Leu Leu Ile Gly Pro Glu Gly Gly Leu
 195 200 205
 Ser Ala Asp Glu Ile Ala Met Thr Ala Arg Tyr Gln Phe Thr Asp Ile
 210 215 220
 Leu Leu Gly Pro Arg Val Leu Arg Thr Glu Thr Thr Ala Leu Thr Ala
 225 230 235 240
 Ile Thr Ala Leu Gln Val Arg Phe Gly Asp Leu Gly
 245 250

<210> 8
 <211> 759
 <212> DNA
 <213> Escherichia coli

Le A 33 616-Foreign Countries

- 38 -

<220>

<223> yggJ Position 9425..10813 of accession number
AE000377; U00096

<400> 8

```

gtgggggagac gacgcggatt tttaactatg cgtatcccc gcatttatca tcctgaacca 60
ctgaccagcc attctcacat cgcgctttgc gaagatgccg ccaaccatat cgggcgcgta 120
ctgcgcatgg ggccgggggca ggcgttgcaa ttgtttgacg gtagcaacca ggtctttgac 180
gccgaaatta ccagcgccag caaaaaaagc gtggaagtga aggtgctgga aggccagatc 240
gacgatcgcg aatctccgct gcatattcac ctcggtcagg tgatgtcgcg tggtgaaaaa 300
atggaattta ctatccagaa atcgatcgaa ctcgggtgtaa gcctcattac gccacttttt 360
tctgagcgct gcggcggttaa actggatagt gaacgtctga acaagaagct tcagcagtgg 420
cagaagattg caattgctgc ctgtgagcag tgtggtcgta accgggtgcc ggaaatccgt 480
ccagcgatgg atctggaagc ctgggtgtgca gagcaggatg aaggactgaa actgaatctt 540
caccgcgcgc ccagtaacag catcaatacg ttgccgttac cggttgaacg cgtccgcctg 600
ctgattggcc cggaaggcgg tttatcgga gatgaaattg ccatgactgc ccgctatcaa 660
tttactgata tcctgttggg acctcgcgtt ttgcgtacag agacaactgc gctcaccgcc 720
attaccgcgc tacaagtacg atttggcgat ttgggctaa 759

```

<210> 9

<211> 152

<212> PRT

<213> Escherichia coli

<220>

<223> YHBC

<400> 9

```

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

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

Gly Arg Thr Ser Thr Leu Arg Ile Tyr Ile Asp Ser Glu Asp Gly Ile
          35                      40                      45

Asn Val Asp Asp Cys Ala Asp Val Ser His Gln Val Ser Ala Val Leu
          50                      55                      60

Asp Val Glu Asp Pro Ile Thr Val Ala Tyr Asn Leu Glu Val Ser Ser
          65                      70                      75                      80

Pro Gly Leu Asp Arg Pro Leu Phe Thr Ala Glu His Tyr Ala Arg Phe
          85                      90                      95

Val Gly Glu Glu Val Thr Leu Val Leu Arg Met Ala Val Gln Asn Arg
          100                      105                      110

Arg Lys Trp Gln Gly Val Ile Lys Ala Val Asp Gly Glu Met Ile Thr
          115                      120                      125

Val Thr Val Glu Gly Lys Asp Glu Val Phe Ala Leu Ser Asn Ile Gln
          130                      135                      140

Lys Ala Asn Leu Val Pro His Phe
          145                      150

```

<210> 10
 <211> 459
 <212> DNA
 <213> Escherichia coli

<220>
 <223> yhbC Position complement 14177..14635 of accession
 number AE000397; U00096

<400> 10
 gtgggcttgt ccacattaga gcaaaaatta acagagatga ttactgcgcc agttgaggcc 60
 ctgggttttg aactgggttg catcgaattt attcgcggtc gcacatccac actgcgcac 120
 tatattgata gtgaagatgg catcaatgtt gatgattgtg ctgatgtgag ccaccaggta 180
 agtgctgtgc tggatgttga agatcccatc accgttgctt ataacctgga agtctcctca 240
 ccgggtctcg atcgcccact gttcacggct gaacactacg cccgttttgt cggagaagag 300
 gtgactctgg ttctccgtat ggcggtacaa aaccgtcgta aatggcaggg cgttatcaaa 360
 gcggtagacg gtgaaatgat cacagttacc gtcgaaggta aagatgaagt gttegcgctg 420
 agtaatatcc agaaggcgaa cctgggtccc cacttttaa 459

<210> 11
 <211> 236
 <212> PRT
 <213> Escherichia coli

<220>
 <223> YGBP

<400> 11
 Met Ala Thr Thr His Leu Asp Val Cys Ala Val Val Pro Ala Ala Gly
 1 5 10 15
 Phe Gly Arg Arg Met Gln Thr Glu Cys Pro Lys Gln Tyr Leu Ser Ile
 20 25 30
 Gly Asn Gln Thr Ile Leu Glu His Ser Val His Ala Leu Leu Ala His
 35 40 45
 Pro Arg Val Lys Arg Val Val Ile Ala Ile Ser Pro Gly Asp Ser Arg
 50 55 60
 Phe Ala Gln Leu Pro Leu Ala Asn His Pro Gln Ile Thr Val Val Asp
 65 70 75 80
 Gly Gly Asp Glu Arg Ala Asp Ser Val Leu Ala Gly Leu Lys Ala Ala
 85 90 95
 Gly Asp Ala Gln Trp Val Leu Val His Asp Ala Ala Arg Pro Cys Leu
 100 105 110
 His Gln Asp Asp Leu Ala Arg Leu Leu Ala Leu Ser Glu Thr Ser Arg
 115 120 125
 Thr Gly Gly Ile Leu Ala Ala Pro Val Arg Asp Thr Met Lys Arg Ala
 130 135 140
 Glu Pro Gly Lys Asn Ala Ile Ala His Thr Val Asp Arg Asn Gly Leu
 145 150 155 160
 Trp His Ala Leu Thr Pro Gln Phe Phe Pro Arg Glu Leu Leu His Asp
 165 170 175

Le A 33 616-Foreign Countries

- 40 -

Cys Leu Thr Arg Ala Leu Asn Glu Gly Ala Thr Ile Thr Asp Glu Ala
180 185 190

Ser Ala Leu Glu Tyr Cys Gly Phe His Pro Gln Leu Val Glu Gly Arg
195 200 205

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

Phe Tyr Leu Thr Arg Thr Ile His Gln Glu Asn Thr
225 230 235

<210> 12

<211> 711

<212> DNA

<213> Escherichia coli

<220>

<223> ygbP Position complement 6754..7464 of accession
number AE000358; U00096

<400> 12

```
atggcaacca ctcatttggga tgttttgcgcc gtggttccgg cggccggatt tggccgtcga 60
atgcaaacgg aatgtcctaa gcaatatctc tcaatcggta atcaaaccat tcttgaacac 120
tcggtgcatg cgctgctggc gcatccccgg gtgaaacgtg tcgtcattgc cataagtcct 180
ggcgatagcc gttttgcaca acttcctctg gcgaatcatc cgcaaatacac cgttgtagat 240
ggcggatgat agcgtgccga ttccgtgctg gcaggctctga aagccgctgg cgacgcgcag 300
tgggtattgg tgcattgacgc cgctcgtcct tgtttgcatc aggatgacct cgcgcgattg 360
ttggcggtga gcgaaaccag ccgcacgggg gggatcctcg ccgcaccagt gcgcgatact 420
atgaaacgtg ccgaaccggg caaaaatgcc attgctcata ccggtgatcg caacggctta 480
tggcacgcgc tgacgccgca atttttccct cgtgagctgt tacatgactg tctgacgcgc 540
gctctaaatg aaggcgcgac tattaccgac gaagcctcgg cgctggaata ttgcggattc 600
catcctcagt tggtcgaagg ccgtgcggat aacattaaag tcacgcgccc ggaagatttg 660
gcactggccg agttttacct caccggaacc atccatcagg agaatacata a 711
```

<210> 13

<211> 159

<212> PRT

<213> Escherichia coli

<220>

<223> YGBB

<400> 13

Met Arg Ile Gly His Gly Phe Asp Val His Ala Phe Gly Gly Glu Gly
1 5 10 15

Pro Ile Ile Ile Gly Gly Val Arg Ile Pro Tyr Glu Lys Gly Leu Leu
20 25 30

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

Leu Gly Ala Ala Ala Leu Gly Asp Ile Gly Lys Leu Phe Pro Asp Thr
50 55 60

Asp Pro Ala Phe Lys Gly Ala Asp Ser Arg Glu Leu Leu Arg Glu Ala
65 70 75 80

Trp Arg Arg Ile Gln Ala Lys Gly Tyr Thr Leu Gly Asn Val Asp Val
85 90 95

Le A 33 616-Foreign Countries

- 41 -

Thr Ile Ile Ala Gln Ala Pro Lys Met Leu Pro His Ile Pro Gln Met
100 105 110

Arg Val Phe Ile Ala Glu Asp Leu Gly Cys His Met Asp Asp Val Asn
115 120 125

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

Gly Ile Ala Cys Glu Ala Val Ala Leu Leu Ile Lys Ala Thr Lys
145 150 155

<210> 14

<211> 480

<212> DNA

<213> Escherichia coli

<220>

<223> ygbB Position complement 6275..6754 of accession
number AE000358; U00096

<400> 14

atgcgaattg gacacggttt tgacgtacat gcctttggcg gtgaaggccc aattatcatt 60
ggtggcgtag gcattcctta cgaaaaagga ttgctggcgc attctgatgg cgacgtggcg 120
ctccatgcgt tgaccgatgc attgcttggc gcggcgggcg tgggggatat cggcaagctg 180
ttcccgata ccgatccggc atttaaaggt gccgatagcc gcgagctgct acgcgaagcc 240
tggcgtcgta ttcaggcgaa gggttatacc cttggcaacg tcgatgtcac tatcatcgct 300
caggcaccga agatgttgcc gcacattcca caaatgcgcg tgttttattgc cgaagatctc 360
ggctgccata tggatgatgt taacgtgaaa gccactacta cggaaaaact gggatttacc 420
ggacgtgggg aagggtattgc ctgtgaagcg gtggcgctac tcattaaggc aacaaaatga 480

<210> 15

<211> 283

<212> PRT

<213> Escherichia coli

<220>

<223> YCHB

<400> 15

Met Arg Thr Gln Trp Pro Ser Pro Ala Lys Leu Asn Leu Phe Leu Tyr
1 5 10 15

Ile Thr Gly Gln Arg Ala Asp Gly Tyr His Thr Leu Gln Thr Leu Phe
20 25 30

Gln Phe Leu Asp Tyr Gly Asp Thr Ile Ser Ile Glu Leu Arg Asp Asp
35 40 45

Gly Asp Ile Arg Leu Leu Thr Pro Val Glu Gly Val Glu His Glu Asp
50 55 60

Asn Leu Ile Val Arg Ala Ala Arg Leu Leu Met Lys Thr Ala Ala Asp
65 70 75 80

Ser Gly Arg Leu Pro Thr Gly Ser Gly Ala Asn Ile Ser Ile Asp Lys
85 90 95

Arg Leu Pro Met Gly Gly Gly Leu Gly Gly Gly Ser Ser Asn Ala Ala
100 105 110

Le A 33 616-Foreign Countries

- 42 -

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

<210> 16
 <211> 852
 <212> DNA
 <213> Escherichia coli

<220>
 <223> ychB Position complement 5720..6571 of accession
 number AE000219; U00096

<400> 16
 atgcggacac agtggccctc tccggcaaaa cttaatctgt ttttatacat taccggtcag 60
 cgtgcggatg gttaccacac gctgcaaacg ctgtttcagt ttcttgatta cggcgacacc 120
 atcagcattg agcttcgtga cgatggggat attcgtctgt taacgcccgt tgaaggcgtg 180
 gaacatgaag ataacctgat cgttcgcgca gcgcgattgt tgatgaaaac tgcggcagac 240
 agcgggcgtc ttccgacggg aagcgggtgcg aatatcagca ttgacaagcg ttgcccgatg 300
 ggcggcggtc tcggcggtgg ttcatccaat gccgcgacgg tcctggtggc attaatcat 360
 ctctggcaat gcgggctaag catggatgag ctggcggaag tggggctgac gctgggcgca 420
 gatgttcctg tctttgttcg ggggcatgcc gcgtttgccg aaggcgttgg tgaaatacta 480
 acgccggtgg atccgccaga gaagtggat ctggtggcgc accctggtgt aagtattccg 540
 actccggtga tttttaaga tcctgaactc ccgcgcaata cgccaaaag gtcaatagaa 600
 acgttgctaa aatgtgaatt cagcaatgat tgcgagggtta tcgcaagaaa acgttttcgc 660
 gaggttgatg cgggtgctttc ctggctgtta gaatacgccc cgtcgcgcct gactgggaca 720
 ggggcctgtg tctttgctga atttgataga gagtctgaag cccgccaggt gctagagcaa 780
 gccccggaat ggctcaatgg ctttgtggcg aaaggcgcta atctttcccc attgcacaga 840
 gccatgcttt aa 852

<210> 17
<211> 21
<212> DNA
<213> Escherichia coli

<220>
<223> 21 bp linker sequence from the Plasmid pAL761

<400> 17
aggctagctt tagctcgagg g 21

<210> 18
<211> 23
<212> DNA
<213> Escherichia coli

<220>
<223> Primer A

<400> 18
cctgctggca atcaatcccg ata 23

<210> 19
<211> 24
<212> DNA
<213> Escherichia coli

<220>
<223> Primer B

<400> 19
aggcgggtggc ggcacatcgg cgtt 24

<210> 20
<211> 28
<212> DNA
<213> Escherichia coli

<220>
<223> Primer YGBP1A

<400> 20
cgcggatccc cacatgggtca ctgcctgg 28

<210> 21
<211> 42
<212> DNA
<213> Escherichia coli

<220>
<223> Primer YGBP1B

<400> 21
cccatccact aaactgcagc tcaaatagagt ggttgccatg tt 42

<210> 22
<211> 40
<212> DNA
<213> Escherichia coli

<220>
<223> Primer YGBP2A

<400> 22
agctgcagtt tagtggatgg gtacctcacc cgaaccatcc 40

<210> 23
<211> 29
<212> DNA
<213> Escherichia coli

<220>
<223> Primer YGBP2B

<400> 23
cgcggatcca cctggcagcc ttccagttg 29

<210> 24
<211> 30
<212> DNA
<213> Escherichia coli

<220>
<223> Primer YGBPR

<400> 24
atcgctagca tcagcccggg aattaacatg 30

<210> 25
<211> 32
<212> DNA
<213> Escherichia coli

<220>
<223> Primer YGBPT

<400> 25
atcctcgaga attcgcatTA tgtattctcc tg 32

<210> 26
<211> 27
<212> DNA
<213> Escherichia coli

<220>
<223> Primer ARA2B

<400> 26
atcgcggccg caaagccgtg ctgcgcg 27

<210> 27
<211> 33
<212> DNA
<213> Escherichia coli

<220>
<223> Primer YJEE-START

<400> 27
tgaagatcta atcgagtaat tccgctccct gat 33

<210> 28
<211> 38
<212> DNA
<213> Escherichia coli

Le A 33 616-Foreign Countries

- 45 -

<220>

<223> Primer YJEE-STOP

<400> 28

tcagaattct taaccggcta aacgcgccag caacaatc

38

SEQUENZPROTOKOLL

<110> Bayer AG

<120> Neue essentielle Gene und ihre Proteine

<130> Essential Genes of EColi

<140>

<141>

<160> 28

<170> PatentIn Vers. 2.0

<210> 1

<211> 153

<212> PRT

<213> Escherichia coli

<220>

<223> YJEE

<400> 1

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

Leu	Gly	Glu	Arg	Val	Ala	Lys	Ala	Cys	Asp	Gly	Ala	Thr	Val	Ile	Tyr
			20					25					30		

Leu	Tyr	Gly	Asp	Leu	Gly	Ala	Gly	Lys	Thr	Thr	Phe	Ser	Arg	Gly	Phe
		35					40					45			

Leu	Gln	Ala	Leu	Gly	His	Gln	Gly	Asn	Val	Lys	Ser	Pro	Thr	Tyr	Thr
	50					55					60				

Leu	Val	Glu	Pro	Tyr	Thr	Leu	Asp	Asn	Leu	Met	Val	Tyr	His	Phe	Asp
65					70					75					80

Leu	Tyr	Arg	Leu	Ala	Asp	Pro	Glu	Glu	Leu	Glu	Phe	Met	Gly	Ile	Arg
			85						90					95	

Asp	Tyr	Phe	Ala	Asn	Asp	Ala	Ile	Cys	Leu	Val	Glu	Trp	Pro	Gln	Gln
			100					105					110		

Gly	Thr	Gly	Val	Leu	Pro	Asp	Pro	Asp	Val	Glu	Ile	His	Ile	Asp	Tyr
		115					120					125			

Gln	Ala	Gln	Gly	Arg	Glu	Ala	Arg	Val	Ser	Ala	Val	Ser	Ser	Ala	Gly
	130					135					140				

Glu	Leu	Leu	Leu	Ala	Arg	Leu	Ala	Gly
145					150			

<210> 2

<211> 462

<212> DNA

<213> Escherichia coli

<220>

<223> yjeE Position 3299..3760 in accession number
AE000489; U00096

<400> 2

```

atgatgaatc gagtaattcc gctccctgat gagcaggcaa cattagacct gggcgagcgg 60
gtagcgaaag cctgcgatgg cgcaaccgta atctatctgt atggcgattt aggcgcaggt 120
aaaaccacct ttagccgggg ctttttacag gctctgggtc atcagggtaa tgtcaaaagc 180
cccacttata cgctggtcga accctatacg ctcgacaact taatgggtcta tcactttgat 240
ttgtaccgcc ttgccgatcc cgaggagctg gagtttatgg ggatccgcga ttattttgcc 300
aacgatgcca tctgcctggg ggagtggcca caacaaggta cagggtgttct tcctgacccg 360
gatgtcgaaa tacacattga ttatcaggca caaggccgtg aggcgcgcgt gagtgcggtt 420
tcctctgcgg gtgaattggt gctggcgcgt ttagccggtt aa 462

```

<210> 3

<211> 159

<212> PRT

<213> Escherichia coli

<220>

<223> KDTB

<400> 3

```

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

Gly His Ile Asp Ile Val Thr Arg Ala Thr Gln Met Phe Asp His Val
          20              25              30

Ile Leu Ala Ile Ala Ala Ser Pro Ser Lys Lys Pro Met Phe Thr Leu
          35              40              45

Glu Glu Arg Val Ala Leu Ala Gln Gln Ala Thr Ala His Leu Gly Asn
          50              55              60

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

Gln His Ala Thr Val Leu Ile Arg Gly Leu Arg Ala Val Ala Asp Phe
          85              90              95

Glu Tyr Glu Met Gln Leu Ala His Met Asn Arg His Leu Met Pro Glu
          100              105              110

Leu Glu Ser Val Phe Leu Met Pro Ser Lys Glu Trp Ser Phe Ile Ser
          115              120              125

Ser Ser Leu Val Lys Glu Val Ala Arg His Gln Gly Asp Val Thr His
          130              135              140

Phe Leu Pro Glu Asn Val His Gln Ala Leu Met Ala Lys Leu Ala
          145              150              155

```

<210> 4

<211> 480

<212> DNA

<213> Escherichia coli

<220>

<223> kdtB Position 1720..2199 of accession number
AE000441; U00096

<400> 4

```

atgcaaaaac gggcgattta tccgggtact ttcatccca ttaccaatgg tcatatcgat 60
atcgtgacgc gcgccacgca gatgttcgat cacgttattc tggcgattgc cgccagcccc 120

```

agtaaaaaac cgatgtttac cctggaagag cgtgtggcac tggcacagca ggcaaccgcg 180
 catctgggga acgtggaagt ggtcgggttt agtgatttaa tggcgaactt cgcccgtaat 240
 caacacgcta cgggtgctgat tcgtggcctg cgtgcggtgg cagattttga atatgaaatg 300
 cagctggcgc atatgaatcg ccacttaatg ccggaactgg aaagtgtgtt tctgatgccg 360
 tcgaaagagt ggtcgtttat ctcttcatcg ttggtgaaag aggtggcgcg ccatcagggc 420
 gatgtcacc c atttctgccc ggagaatgtc catcaggcgc tgatggcgaa gttagcgtag 480

<210> 5

<211> 138

<212> PRT

<213> Escherichia coli

<220>

<223> YQGF

<400> 5

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

Val Ala Val Gly Gln Arg Ile Thr Gly Thr Ala Arg Pro Leu Pro Ala
 20 25 30

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

Leu Lys Glu Trp Gln Pro Asp Glu Ile Ile Val Gly Leu Pro Leu Asn
 50 55 60

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

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

Arg Leu Ser Thr Val Glu Ala Arg Ser Gly Leu Phe Glu Gln Gly Gly
 100 105 110

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

Ile Leu Glu Ser Tyr Phe Glu Gln Gly Tyr
 130 135

<210> 6

<211> 417

<212> DNA

<213> Escherichia coli

<220>

<223> yggF Position 11818...12234 of accession number
 AE000377; U00096

<400> 6

atgagtggaa ccttactcgc cttcgacttc ggcaccaaaa gcattggcgt agcggtcggc 60
 caacgcatta ccggcaccgc tcgccctttg cctgcaatta aagcacagga cggtagccg 120
 gactggaaca ttatcgagcg tttactgaaa gactggcagc cggacgaaat catcgtcggt 180
 ttgccgctga atatggacgg caccgagcaa ccattgactg ccagagcgcg taaatttgcc 240
 aaccgtattc atggccgttt cgggtgttgaa gtaaagctcc atgacgagcg tcttagcact 300
 gtggaagccc gttccggtct gtttgaacag ggcggctatc gggcgctcaa caaaggcaaa 360
 gttgactctg cctctgcggt tattattctc gaaagctatt tcgagcaggg atattaa 417

<210> 7

<211> 252

<212> PRT

<213> Escherichia coli

<220>

<223> YGGJ

<400> 7

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

<210> 8

<211> 759

<212> DNA

<213> Escherichia coli

<220>

<223> yggJ Position 9425..10813 of accession number
 AE000377; U00096

<400> 8

```

gtgggggagac gacgcgggatt tttaactatg cgtatccccc gcatttatca tcctgaacca 60
ctgaccagcc attctcacat cgcgctttgc gaagatgccg ccaaccatat cgggcgcgta 120
ctgcgcatgg ggccgggggca ggcgttgcaa ttgtttgacg gtagcaacca ggtctttgac 180
gccgaaatta ccagcgccag caaaaaaagc gtggaagtga aggtgctgga aggccagatc 240
gacgatcgcg aatctccgct gcatattcac ctcggtcagg tgatgtcgcg tggtgaaaaa 300
atggaattta ctatccagaa atcgatcgaa ctcgggtgtaa gcctcattac gccacttttt 360
tctgagcgct gcggcggttaa actggatagt gaacgtctga acaagaagct tcagcagtgg 420
cagaagattg caattgctgc ctgtgagcag tgtgggtcgta accgggtgcc ggaaatccgt 480
ccagcgatgg atctggaagc ctgggtgtgca gagcaggatg aaggactgaa actgaatctt 540
caccgcgcgc ccagtaacag catcaatacg ttgccgttac cggttgaacg cgtccgcctg 600
ctgattggcc cggaaggcgg tttatcgcca gatgaaattg ccatgactgc ccgctatcaa 660
tttactgata tcctgttggg acctcgcgtt ttgcgtacag agacaactgc gctcaccgcc 720
attaccgcgc tacaagtacg atttggcgat ttgggctaa 759

```

<210> 9

<211> 152

<212> PRT

<213> Escherichia coli

<220>

<223> YHBC

<400> 9

```

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

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

Gly Arg Thr Ser Thr Leu Arg Ile Tyr Ile Asp Ser Glu Asp Gly Ile
          35              40              45

Asn Val Asp Asp Cys Ala Asp Val Ser His Gln Val Ser Ala Val Leu
          50              55              60

Asp Val Glu Asp Pro Ile Thr Val Ala Tyr Asn Leu Glu Val Ser Ser
          65              70              75              80

Pro Gly Leu Asp Arg Pro Leu Phe Thr Ala Glu His Tyr Ala Arg Phe
          85              90              95

Val Gly Glu Glu Val Thr Leu Val Leu Arg Met Ala Val Gln Asn Arg
          100             105             110

Arg Lys Trp Gln Gly Val Ile Lys Ala Val Asp Gly Glu Met Ile Thr
          115             120             125

Val Thr Val Glu Gly Lys Asp Glu Val Phe Ala Leu Ser Asn Ile Gln
          130             135             140

Lys Ala Asn Leu Val Pro His Phe
          145             150

```

<210> 10
 <211> 459
 <212> DNA
 <213> Escherichia coli

<220>
 <223> yhbC Position complement 14177..14635 of accession
 number AE000397; U00096

<400> 10
 gtgggcttgt ccacattaga gcaaaaatta acagagatga ttactgcgcc agttgaggcc 60
 ctgggttttg aactgggttg catcgaattt attcgcggtc gcacatccac actgcgcac 120
 tatattgata gtgaagatgg catcaatgtt gatgattgtg ctgatgtgag ccaccaggta 180
 agtgctgtgc tggatgttga agatcccatc accgttgctt ataacctgga agtctcctca 240
 ccgggtctcg atcgcccact gttcacggct gaacactacg cccgttttgt cggagaagag 300
 gtgactctgg ttctccgtat ggcggtacaa aaccgtcgta aatggcaggg cgttatcaaa 360
 gcggtagacg gtgaaatgat cacagttacc gtcgaaggta aagatgaagt gttcgcgctg 420
 agtaatatcc agaaggcgaa cctggttccc cacttttaa 459

<210> 11
 <211> 236
 <212> PRT
 <213> Escherichia coli

<220>
 <223> YGBP

<400> 11
 Met Ala Thr Thr His Leu Asp Val Cys Ala Val Val Pro Ala Ala Gly
 1 5 10 15
 Phe Gly Arg Arg Met Gln Thr Glu Cys Pro Lys Gln Tyr Leu Ser Ile
 20 25 30
 Gly Asn Gln Thr Ile Leu Glu His Ser Val His Ala Leu Leu Ala His
 35 40 45
 Pro Arg Val Lys Arg Val Val Ile Ala Ile Ser Pro Gly Asp Ser Arg
 50 55 60
 Phe Ala Gln Leu Pro Leu Ala Asn His Pro Gln Ile Thr Val Val Asp
 65 70 75 80
 Gly Gly Asp Glu Arg Ala Asp Ser Val Leu Ala Gly Leu Lys Ala Ala
 85 90 95
 Gly Asp Ala Gln Trp Val Leu Val His Asp Ala Ala Arg Pro Cys Leu
 100 105 110
 His Gln Asp Asp Leu Ala Arg Leu Leu Ala Leu Ser Glu Thr Ser Arg
 115 120 125
 Thr Gly Gly Ile Leu Ala Ala Pro Val Arg Asp Thr Met Lys Arg Ala
 130 135 140
 Glu Pro Gly Lys Asn Ala Ile Ala His Thr Val Asp Arg Asn Gly Leu
 145 150 155 160
 Trp His Ala Leu Thr Pro Gln Phe Phe Pro Arg Glu Leu Leu His Asp
 165 170 175
 Cys Leu Thr Arg Ala Leu Asn Glu Gly Ala Thr Ile Thr Asp Glu Ala

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

<210> 12
 <211> 711
 <212> DNA
 <213> Escherichia coli

<220>
 <223> ygbP Position complement 6754..7464 of accession
 number AE000358; U00096

<400> 12
 atggcaacca ctcatttgga tgtttgcgcc gtgggttcgg cggccggatt tggccgtcga 60
 atgcaaacgg aatgtcctaa gcaatatctc tcaatcggtat atcaaaccat tcttgaacac 120
 tcgggtgcatg cgctgctggc gcatccccgg gtgaaacgtg tcgtcattgc cataagtcct 180
 ggcgatagcc gttttgcaca acttcctctg gcgaatcatc cgcaaatacac cgttgtagat 240
 ggcggtgatg agcgtgccga ttccgtgctg gcaggtctga aagccgctgg cgacgcgcag 300
 tgggtattgg tgcattgacgc cgctcgctcct tgtttgcatc aggatgacct cgcgcgattg 360
 ttggcggtga gcgaaaccag ccgcacgggg gggatcctcg ccgcaccagt gcgcgatact 420
 atgaaacgtg ccgaaccggg caaaaatgcc attgctcata ccgttgatcg caacggctta 480
 tggcacgcgc tgacgcgcga atttttccct cgtgagctgt tacatgactg tctgacgcgc 540
 gctctaaatg aaggcgcgac tattaccgac gaagcctcgg cgctggaata ttgcggattc 600
 catcctcagt tggtcgaagg ccgtgcggat aacattaaag tcacgcgccc ggaagatttg 660
 gcaactggccg agttttacct caccggaacc atccatcagg agaatacata a 711

<210> 13
 <211> 159
 <212> PRT
 <213> Escherichia coli

<220>
 <223> YGBB

<400> 13
 Met Arg Ile Gly His Gly Phe Asp Val His Ala Phe Gly Gly Glu Gly
 1 5 10 15
 Pro Ile Ile Ile Gly Gly Val Arg Ile Pro Tyr Glu Lys Gly Leu Leu
 20 25 30
 Ala His Ser Asp Gly Asp Val Ala Leu His Ala Leu Thr Asp Ala Leu
 35 40 45
 Leu Gly Ala Ala Ala Leu Gly Asp Ile Gly Lys Leu Phe Pro Asp Thr
 50 55 60
 Asp Pro Ala Phe Lys Gly Ala Asp Ser Arg Glu Leu Leu Arg Glu Ala
 65 70 75 80
 Trp Arg Arg Ile Gln Ala Lys Gly Tyr Thr Leu Gly Asn Val Asp Val
 85 90 95
 Thr Ile Ile Ala Gln Ala Pro Lys Met Leu Pro His Ile Pro Gln Met

	100		105		110
Arg	Val Phe Ile Ala Glu Asp	Leu Gly Cys His Met	Asp Asp Val Asn		
	115	120	125		
Val	Lys Ala Thr Thr Thr Glu	Lys Leu Gly Phe Thr	Gly Arg Gly Glu		
	130	135	140		
Gly	Ile Ala Cys Glu Ala Val	Ala Leu Leu Ile	Lys Ala Thr Lys		
145	150	155			

<210> 14
 <211> 480
 <212> DNA
 <213> Escherichia coli

<220>
 <223> ygbB Position complement 6275..6754 of accession
 number AE000358; U00096

<400> 14
 atgcgaattg gacacgggtt tgacgtacat gcctttggcg gtgaaggccc aattatcatt 60
 ggtggcgctac gcattcctta cgaaaaagga ttgctggcgc attctgatgg cgacgtggcg 120
 ctccatgcgt tgaccgatgc attgcttggc gcggcggcgc tgggggatat cggcaagctg 180
 ttcccggata ccgatccggc atttaaaggt gccgatagcc gcgagctgct acgcgaagcc 240
 tggcgctcgta ttcaggcgaa gggttatacc cttggcaacg tcgatgtcac tatcatcgct 300
 caggcaccga agatgttgcc gcacattcca caaatgcgcg tgtttattgc cgaagatctc 360
 ggctgccata tggatgatgt taacgtgaaa gccactacta cggaaaaact gggatttacc 420
 ggacgtgggg aagggattgc ctgtgaagcg gtggcgctac tcattaaggc aacaaaatga 480

<210> 15
 <211> 283
 <212> PRT
 <213> Escherichia coli

<220>
 <223> YCHB

<400> 15
 Met Arg Thr Gln Trp Pro Ser Pro Ala Lys Leu Asn Leu Phe Leu Tyr
 1 5 10 15
 Ile Thr Gly Gln Arg Ala Asp Gly Tyr His Thr Leu Gln Thr Leu Phe
 20 25 30
 Gln Phe Leu Asp Tyr Gly Asp Thr Ile Ser Ile Glu Leu Arg Asp Asp
 35 40 45
 Gly Asp Ile Arg Leu Leu Thr Pro Val Glu Gly Val Glu His Glu Asp
 50 55 60
 Asn Leu Ile Val Arg Ala Ala Arg Leu Leu Met Lys Thr Ala Ala Asp
 65 70 75 80
 Ser Gly Arg Leu Pro Thr Gly Ser Gly Ala Asn Ile Ser Ile Asp Lys
 85 90 95
 Arg Leu Pro Met Gly Gly Gly Leu Gly Gly Gly Ser Ser Asn Ala Ala
 100 105 110
 Thr Val Leu Val Ala Leu Asn His Leu Trp Gln Cys Gly Leu Ser Met
 115 120 125

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

<210> 16
 <211> 852
 <212> DNA
 <213> Escherichia coli

<220>
 <223> ychB Position complement 5720..6571 of accession
 number AE000219; U00096

<400> 16
 atgcggacac agtggccctc tccggcaaaa cttaatctgt ttttatacat taccgggtcag 60
 cgtgcggatg gttaccacac gctgcaaacg ctgtttcagt ttcttgatta cggcgacacc 120
 atcagcattg agcttcgtga cgatggggat attcgtctgt taacgcccgt tgaaggcgtg 180
 gaacatgaag ataacctgat cgttcgcgca gcgcgattgt tgatgaaaac tgcggcagac 240
 agcgggcgctc ttccgacggg aagcgggtgcg aatatcagca ttgacaagcg tttgccgatg 300
 ggcggcggtc tcggcggtgg ttcattccaat gccgcgacgg tcctgggtggc attaaatcat 360
 ctctggcaat gcgggctaag catggatgag ctggcggaag tggggctgac gctgggcgca 420
 gatgttcctg tctttgttcg ggggcatgcc gcgtttgccg aaggcggttg tgaaatacta 480
 acgccggtgg atccgccaga gaagtggat ctggtggcgc accctggtgt aagtattccg 540
 actccggtga tttttaaaga tcctgaactc ccgcgcaata cgccaaaag gtcaatagaa 600
 acgttgctaa aatgtgaatt cagcaatgat tgcgagggtta tcgcaagaaa acgttttcgc 660
 gaggttgatg cgggtgcttc ctggctgtta gaatacgccc cgtcgcgcct gactgggaca 720
 ggggcctgtg tctttgctga atttgataca gagtctgaag cccgccaggt gctagagcaa 780
 gccccggaat ggctcaatgg ctttgtggcg aaaggcgcta atctttcccc attgcacaga 840
 gccatgcttt aa 852

<210> 17
 <211> 21
 <212> DNA
 <213> Escherichia coli

<220>

<223> 21-bp-Linker-Sequenz aus dem Plasmid pAL761

<400> 17

aggctagctt tagctcgagg g

21

<210> 18

<211> 23

<212> DNA

<213> Escherichia coli

<220>

<223> Primer A

<400> 18

cctgctggca atcaatcccg ata

23

<210> 19

<211> 24

<212> DNA

<213> Escherichia coli

<220>

<223> Primer B

<400> 19

aggcgggtggc ggcacatcgg cggt

24

<210> 20

<211> 28

<212> DNA

<213> Escherichia coli

<220>

<223> Primer YGBP1A

<400> 20

cgcggatccc cacatgggtca ctgcctgg

28

<210> 21

<211> 42

<212> DNA

<213> Escherichia coli

<220>

<223> Primer YGBP1B

<400> 21

cccatccact aaactgcagc tcaaagtgt ggttgccatg tt

42

<210> 22

<211> 40

<212> DNA

<213> Escherichia coli

<220>

<223> Primer YGBP2A

<400> 22

agctgcagtt tagtggatgg gtacctcacc cgaaccatcc

40

<210> 23

<211> 29

<212> DNA
<213> Escherichia coli

<220>
<223> Primer YGBP2B

<400> 23
cgcggatcca cctggcagcc ttccagttg 29

<210> 24
<211> 30
<212> DNA
<213> Escherichia coli

<220>
<223> Primer YGBPR

<400> 24
atcgctagca tcagcccggg aattaacatg 30

<210> 25
<211> 32
<212> DNA
<213> Escherichia coli

<220>
<223> Primer YGBPT

<400> 25
atcctcgaga attcgcatta tgtattctcc tg 32

<210> 26
<211> 27
<212> DNA
<213> Escherichia coli

<220>
<223> Primer ARA2B

<400> 26
atcgcgccg caaagccgtg ctcgcgc 27

<210> 27
<211> 33
<212> DNA
<213> Escherichia coli

<220>
<223> Primer YJEE-START

<400> 27
tgaagatcta atcgagtaat tccgctccct gat 33

<210> 28
<211> 38
<212> DNA
<213> Escherichia coli

<220>
<223> Primer YJEE-STOP

<400> 28
tcagaattct taaccggcta aacgcgccag caacaatc 38

Patent Claims:

1. Use of the proteins from the group of YQGF, YHBC, YGGJ, YGBP, YCHB, YGBB, YJEE, KDTB and their orthologous gene products in an assay for finding substances which bind to these proteins.
5
2. Use according to claim 1, where the proteins are 95-100% identical to those listed.
- 10 3. Vectors comprising one or more of the genes which code for the proteins according to claim 1 and 2.
4. Transformed microorganisms comprising one or more of the genes which code for the proteins according to claim 1 and 2.
15
5. Compounds which bind to the proteins according to claim 1 and 2.
6. Compounds which bind to the proteins according to claim 1 and 2 and modulate their activity.
20
7. Pharmaceutical comprising one or more compounds according to claim 5 or 6.
8. Use of compounds according to claims 5 and 6 for producing pharmaceuticals.
25
9. Method for finding compounds which modulate proteins from the group of YQGF, YHBC, YGGJ, YGBP, YCHB, YGBB, YJEE, KDTB and their orthologous gene products, characterized in that the compounds are brought into contact with the proteins, and binding thereof is checked.
30

Le A 33 616-Foreign Countries

- 47 -

10. Method for finding compounds which have antibacterial activity and which modulate proteins from the group of YQGF, YHBC, YGGJ, YGBP, YCHB, YGBB, YJEE, KDTB and their orthologous gene products, characterized in that the compounds are brought into contact with the proteins, and binding thereof is checked, and they are subsequently subjected to an MIC test.
- 5

Fetherstonhaugh & Co.
Ottawa, Canada
Patent Agents