BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **247,** 674–680 (1998)
ARTICLE NO. RC988848

Determinants of Recombinant Production of Antimicrobial Cationic Peptides and Creation of Peptide Variants in Bacteria

L. Zhang,* T. Falla,*,[1] M. Wu,* S. Fidai,*,[2] J. Burian,†,[2] W. Kay,† and R. E. W. Hancock*,[3]
**Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3; and †Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia, Canada V8W 3P5*

Received May 14, 1998

Cationic peptides possessing antibacterial activity are virtually ubiquitous in nature, and offer exciting prospects as new therapeutic agents. We had previously demonstrated that such peptides could be produced by fusion protein technology in bacteria and several carrier proteins had been tested as fusion partners including glutathione-S-transferase, *S. aureus* protein A, IgG binding protein and *P. aeruginosa* outer membrane protein OprF. However these fusion partners, while successfully employed in peptide expression, were not optimized for high level production of cationic peptides (Piers, K., Brow, M. L., and Hancock, R. E. W. 1993, *Gene* 137, 7–13). In this paper we took advantage of a small replication protein RepA from *E. coli* and used its truncated version to construct fusion partners. The minimal elements required for high level expression of cationic peptide were defined as a DNA sequence encoding a fusion protein comprising, from the N-terminus, a 68 amino acid carrier region, an anionic prepro domain, a single methionine and the peptide of interest. The 68 amino acid carrier region was a block of three polypeptides consisting of a truncated RepA, a synthetic cellulose binding domain and a hexa histidine domain. The improved system showed high level expression and simplified downstream purification. The active peptide could be yielded by CNBr cleavage of the fusion protein. This novel vector was used to express three classes of cationic peptides including the α-helical peptide CEMA, the looped peptide bactenecin and the extended peptide indolicidin. In addition, mutagenesis of the peptide gene to produce peptide variants of CEMA and indolicidin using the improved vector system was shown to be successful. © 1998 Academic Press

The recent discovery of natural peptides possessing a broad range of antimicrobial activities has led to a rapidly increasing number of novel peptides being identified. These peptides, all of which are polycationic, are produced by various species of animals and insects, often in response to bacterial infection (1). Structurally, these peptides are quite diverse, allowing them to be classified into different families based on their primary and secondary structure, including the β-stranded (eg. defensins) (2), α-helical (eg., cecropins and magainins) (3,4,5), extended helix (eg., indolicidin) (6) and loop (eg., bactenecin) (7) classes. These peptides tend to be relatively short (13-33 amino acids), contain multiple positive charges (eg, on lysine or arginine residues), and function to kill bacterial cells by inserting into their membranes and aggregating to form channels (8,9). They have several properties that have led them to be considered as potential therapeutics including a broad spectrum of antibacterial activity (2), anti-endotoxin activity (10,11), synergy with conventional antibiotics (12), and rapid killing action (13). For these reason they are being considered for commercial exploitation.

Due to the relatively short length of these peptides, solid phase chemical synthesis has generally been a useful method for producing moderate amounts of the desired peptides for carrying out structure-function analysis (14). However, there are at least two major drawbacks to producing peptides in this manner, including the significant synthesis costs, particularly for large scale purposes, and to a lesser extent, limitations in synthesizing peptides that have an abundance of tryptophan or even arginine residues. Recombinant DNA technology offers the ability to overcome both problems, provided the host is insensitive to the effects

[1] Current address: Department of Microbiology, Old Medical School, University of Leeds, Leeds, UK LS2 9JT.
[2] Current address: Micrologix Biotech. Inc., Vancouver, BC, Canada V6S 2L2.
[3] Corresponding author. Fax: 604-822-6041. E-mail: bob@cmdr.ubc.ca.

0006-291X/98 $25.00

of the expressed peptide. For expression in bacteria, this necessitates expressing the cationic peptide as a fusion protein to negate the antimicrobial activity of the peptide (15).

Perhaps not surprisingly, in view of the potent antimicrobial activity of many of these peptides, there are few examples of successful recombinant expression in bacteria. Callaway et al (16) demonstrated expression of cecropin A in inclusion bodies in *E. coli* using the *araBAD* promoter from *S. typhimurium* and an amino terminal fusion with the first 490 amino acids of the *araB* gene product L-ribulokinase. However the proportion of the fusion protein that was cecropin was less than 8% of the isolated protein. At the same time, we demonstrated the need for an anionic stabilizing fragment in the fusion protein to counteract the action of the cationic peptide portion and to prevent proteolytic cleavage of the fusion protein (15). This anionic fragment could be a carrier protein (eg., protein A) if expressed as a secreted protein in *S. aureus*, but if the fusion protein was expressed in inclusion bodies in *E. coli*, an extra anionic sequence equivalent to the prepro sequence from the gene for human definsin (HNP-1) was required (15). In agreement with this, when fused to the amino-terminal of subtilisin inhibitor, the cationic peptide apidacein IB was reasonably well expressed in an insensitive host *Streptomyces* sp, but expression was poor in susceptible *E. coli* (17). The reason is likely due to loss of cell viability during the late logarithmic phase of growth (17). To further investigate the elements that are required for stable expression of the fusion protein and to optimize maximal expression, a systematic approach was taken here to define the minimal sequence necessary for expression of broad cationic peptides. In addition creation of peptide variants via random or semi-random DNA mutagenesis was also carried out.

MATERIALS AND METHODS

Bacterial strains and plasmids. Bacterial strains and plasmids used in this study are listed in Table 1. Generally *E. coli* strains were grown in Luria broth (LB) at 37°C unless otherwise indicated. For induction with isopropyl β-o-thiogalactopyranoside (IPTG), *E. coli* BL21 (DE3) harboring the various expression constructs was grown in LB medium containing 100 μg/ml ampicillin until the cell density reached an OD_{600} of about 0.4. The cultures were then induced with 0.2-1 mM IPTG and grown for an additional 2–5 h. For heat shock induction, the *E. coli* XL1 blue strain, in which the T7 polymerase gene was carried on plasmid pGP1-2, was allowed to grow in LB medium containing 100 μg/ml ampicillin and 50 μg/ml kanamycin to the same OD_{600} value. Then heat induction was carried out by shifting the incubation temperature to 42°C for 15 min. Cells were than grown for an additional 2–3 h at 37°C.

SDS-PAGE and AU-PAGE. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed as previously described (18). Acid-urea polyacrylamide gel electrophoresis (AU-PAGE) was done according to the method of Panyim and Chalkey with minor modification (19). Briefly, 15% polyacrylamide gels containing 5M urea and 5% acetic acid were pre-electrophoresed at 100V for 30 min. Samples were solubilized in 5% acetic acid and diluted 2:1 with sample buffer (9 M urea in 5% acetic acid containing methyl green as a tracking dye). Gels were run in the reverse polarity (towards the cathode) at 100V until the methyl green migrated off the bottom of the gel.

TABLE 1

Bacterial Strains and Plasmids Used in This Study

Name	Description	Reference
Bacteria		
DH5α	*supE*44Δ*lacU*169 (Φ80 *lacZ*Δ*M*1) *hsdR*17 *recA*1 *endA*1*gyrA*96 *thi*-1*rel*A1	(29)
HMS174	$r^- m^+$ *rif*^{r}*thi recA*1 *sup*	(30)
BL21(DE3)	*hsdS gal* λcIts857 *ind*1 *sam*7 *nin*5 *lac*UV5-T7 gene 1	(30)
Plasmids		
pKL1	Small criptic plasmid carrying replication protein RepA	(20)
pTZE07	Plasmid encoding a cellulose binding domain from *C. fimi*	(21)
pSP72	Cloning vector	
pT7-7	Expression vector	(22)
pIRT5	*S. aureu* proetin A-derived expression vector	(15)
pEP12	Expression vector containing Rep21-CBD_{syn}-his-prepro fusion protein cloned into pT7-7	This study
pECP1	The peptide gene encoding for CP2600 was cloned to pEP12	This study

DNA techniques. Plasmid isolations, restriction enzyme digestions, DNA ligations and transformations were performed as described by Ausubel et al (20). To confirm all the final constructs, plasmids were sequenced using the Taq Dyedeoxy Terminator Cycle Sequencing Kit (Applied Biosystems, Inc.) and the Applied Biosystems Inc., Model 373 automated DNA sequencer following the manufacture's protocols.

Oligonucleotide synthesis. The amino acid sequence of CEMA variants, bactenecin and indolicidin were used to design nucleotide sequences for these cationic peptides. The DNA sequence encoding a portion of a cellulose binding domain (designated CBD_{syn}) from the fungus *Trichoderma ressei* (21) was deduced from the corresponding amino acid sequence. The hexa-histidine tag was also synthesized as an oligonucleotide. All the oligonucleotides and primers were synthesized on an Applied Biosystems Inc., Model 380B DNA synthesizer according to manufacturer's instruction.

Assembly of fusion partners. For fusion protein design, four blocks of genes were created by either PCR or oligonucleotide synthesis as shown in Figure 1. The first block was taken from the *repA* gene of a small cryptic plasmid pKL1 of *E. coli* (22). A fragment encoding 78 amino acids and a fragment encoding 21 amino acids of RepA were amplified. Both fragments retained the very N-terminal portion and contain an endogenous promoter and a putative ribosomal binding site (rbs). These fragments were constructed to carry an *Eco*RI site at the 5′-end and *Bam*HI and *Hin*dIII sites at the 3′-end of the gene to facilitate cloning. The second gene block was an oligonucleotide encoding for a synthetic CBD domain designated as CBD_{syn} which carried *Bgl*II and *Bam*HI at the 5′-end and 3′-end, respectively. As an alternative, a 0.35 kb fragment encoding for another CBD (designed CBD_{nat}) from a family II CBD, an exogluconase/xylanase of bacterium *Cellulomonas fimi* (23), was taken for in comparison to CBD_{syn}. The third addition was the hexa-histidine tag inserted next to the pre pro region. The fourth block of gene was the

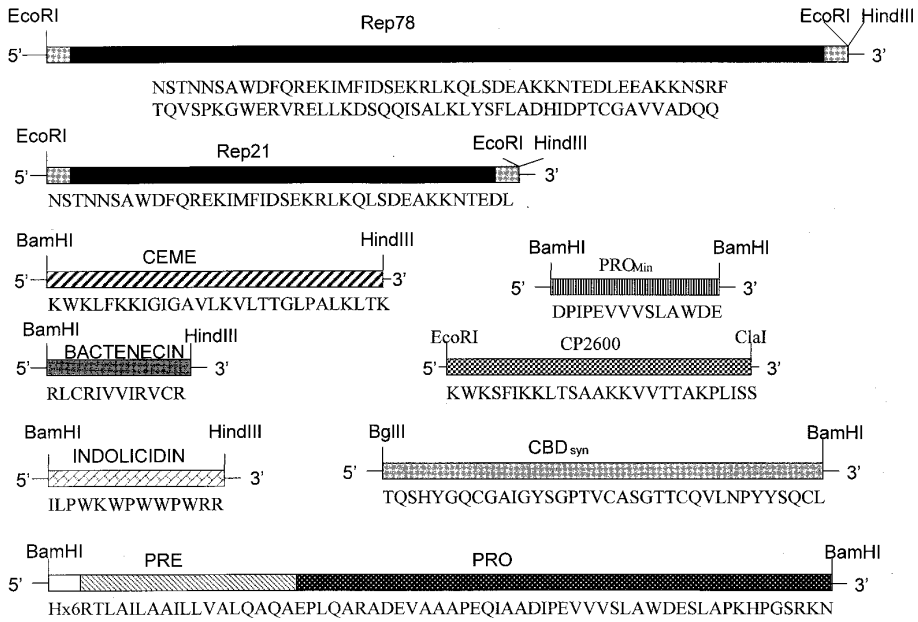

FIG. 1. Schematic diagram of PCR amplified fragments and their corresponding amino acid sequence. Hx6: the hexa-histidine tag. The rep78 and rep21 fusion proteins started at the first M (methionine) of the indicated sequences.

prepro piece of human defensin 1. A methionine residue was designed immediately downstream adjacent to the prepro sequence to permit removal of the cationic peptide from the resultant fusion protein by CNBr cleavage. All four blocks of genes were assembled together by appropriate restriction endonuclease digestions and ligations to appropriately cleaved vector pSP72, which resulted in a series of expression vectors of different combinations (Figure 1 and Table 2). All of the final constructs were verified by DNA sequencing.

Construction of expressing vector pEP12. To transfer the fusion construct from vector pSP72 to pT7-7, a primer was designed to carry a *Nde*I site at the 5′-end and 18 bp complementary to the first 18 bp of *repA* gene. After PCR amplification with help of a second primer downstream of the fusion construct, the 45 bp upstream sequence containing the endogenous *repA* promoter and rbs was deleted and a *Nde*I site was incorporated right before the start coding ATG of *repA*. Digestion of the PCR fragment with *Nde*I and ligation to the

TABLE 2

Dependence of Expression Level on the Specific Nature of the Construction

Construct No.	Fusion partner	Level[a]	Size (kDa)
1.	Rep78-CBD_{syn}-his-pre-pro-CEMA	+++	24
2.	Rep78-his-pre-pro-CEMA	++	20
3.	Rep78-his-pre-pro-CBD_{syn}-CEMA	−	24
4.	Rep78-his-pro_{min}-CBD_{syn}-CEMA	−	23
5.	Rep78-CEMA	−	12
6.	Rep78-CBD_{nat}-his-pre-pro-CEMA	−	32
7.	Rep78-his-pre-pro-CBD_{nat}-CEMA	−	32
8.	Rep78-CBD_{syn}-pro_{min}-CEMA	−	17
9.	Rep21-CBD_{syn}-his-pre-pro-CEMA	+++	18
10.	Rep21-CBD_{syn}-his-pre-pro-bac[b]	+++	17
11.	Rep21-CBD_{syn}-pre-pro-bac	+++	16
12.	Rep21-CBD_{syn}-pre-pro-ind[c]	++	16
13.	Rep21-CBD_{syn}-pro-CEMA	++	16
14.	Rep21-his-pre-pro-CEMA	+	13
15.	Rep21-CBD_{syn}-CEMA	−	10
16.	Rep21-CBD_{syn}-his-promin-CEMA	−	12
17.	Rep21-his-pre-pro-his-pre-pro-CEMA	+++	22
18.	Rep21-his-pre-pro-CBD_{syn}-CEMA	−	18

[a] Expression levels are based on Coomasie-stained SDS-PAGE. +++, major band; ++, band easily observable; +, minor band; −, no band observed.

[b] Bactenecin.

[c] Indolicidin.

*Nde*I-digested vector pT7-7 resulted in another expression vector pEP12. Expression of the fusion partner in pEP12 was purely under the control of T7 promoter and the rbs site provided by the pT7-7 vector (24).

Construction of peptide variants via DNA mutagenesis. We choose the DNA sequence encoding for CP2600 for semi-random mutagenesis. Since multiple bases can be added to the support-bound oligonucleotide by simultaneous delivery of the specified phosphoramidites, we designed a 60 mer oligonucleotide in which, within a 36 nt central region, a mixture of two bases were allowed to be inserted at 9 separate positions during oligonucleotide synthesis. The purpose of the semi-random mutagenesis was to ensure only relatively conservative alterations could take place. The exact positions where substitutions were to take place are shown below:

```
 A         C          C    T   G
AGG ACG TTA AAG TCT GCA GCT AAG ACA

                       A   T   T                    G
                       GTG CTC TAC ACC GCT TTA AAA CCA
```

After oligonucleotide synthesis, a pool of 2^9 different DNA molecules was generated. This permitted amino acid substitutions exclusively between K and R; A and V; Y and H; T and F; and, A and P. Cloning of these DNA molecules into the expression vector pEP12 gave rise to a representative library and individual clones were characterized for novel peptide production. The cloning strategy involves construction of three sections. The first section, the 3′ end containing restriction endonuclease sites and stop codon, was synthesized using standard protocols. The second section was the fragments generated by semi-random mutagenesis, and the third section, the 5′ end, also was synthesized according to standard protocols. The resulting oligonucleotides contained a pool of semi-randomly mutated sequences flanked by conserved 3′ and 5′ sequences. The three sections were joined together by PCR.

For creation of indolicidin variants, a library was created via a random-mutagenesis approach in *S. aueus* using the expression vector pIRT5 (15). The DNA sequence encoding for indolicidin was synthesized using nucleotide stocks that were doped with 3% of each of the other three nucleotides. The relationship between the doping concentration and the mutation rate was defined by an equation described by NcNeil and Smith (25). The resulting oligonucleotides contained a pool of randomly mutated indolicidin sequences flanked by conserved 3′ and 5′ sequences.

Isolation of inclusion bodies. Inclusion bodies were purified from the insoluble cell lysate by extraction with 8M urea under reducing conditions. A modified approach was designed to extract the fusion protein from the insoluble cell lysate by using a mixture of organic solvents originally described for the purification of the bacteriophage T4 RegA protein (26). The solvent system was composed of 40% (v/v) 1-propanol, and 0.63% (v/v) triethylamine. Following probe sonication of the inclusion body pellet resuspended in the hydrophobic solvent system, a majority of the soluble fusion protein was identified in the first extract. After centrifugation, the clarified extract was concentrated and the protein content analyzed.

CNBr cleavage and minimal inhibitory concentration (MIC). CNBr cleavage was performed in 70% formic acid and 1 M CNBr at 23°C overnight. AU-PAGE was used to confirm the presence of a band corresponding to the migration of chemically synthesized peptide. The MIC of peptides was determined by the broth microdilution method. Serial dilutions of each peptide were made in Luria-Bertani (1% w/v tryptone and 0.5% w/v yeast extract) medium in 96-well microtiter plates. Each well was inoculated with 10 μl of 10^4–10^5 colony forming unit per ml of the test organism. The MIC, determined after overnight incubation of the plates at 37°C, was taken as the lowest peptide concentration at which growth was inhibited.

RESULTS AND DISCUSSIONS

Vector improvement. Cationic peptides have good activity against bacteria and having a high content of basic amino acids are susceptible to proteolysis by trypsin/chymotrypsin like proteases. Thus when recombinantly expressed in a susceptible host such as *E. coli*, they must be produced as a protein fusion (15). The primary function of the fusion partner in this case is to inhibit the activity of antimicrobial peptide and to prevent the proteolytic cleavage of the peptide after its synthesis. We therefore set out to test a series of amino terminal fusions for their ability to stably express various cationic peptides.

In a previous study, we constructed four fusion partners including fusions to GST, *P. aeruginosa* outer membrane protein OprF, IgG binding protein and to the *S. aureus* protein A (15). We could show expression of the above fusion proteins, but we also observed disadvantages of these systems. The existence of twelve Met residues in GST made subsequent purification complex and poorly efficient, while fusion to the IgG binding protein A without the prepro region was not sufficient to prevent proteolysis, and fusion protein was degraded during synthesis. Fusion to OprF led to fusion protein that was localized to the outer membrane of the producing bacterium which made the final purification difficult. Thus, although the above systems could be used to prepare cationic peptides by recombinant DNA technology, they were rather inefficient. The *S. aureus* protein A system allowed the fusion protein to be exported to the external medium which provided an advantage in preventing proteolytic degradation, but the yield was poor and the use of the host strain *S. aureus* provided a disadvantage since it was not a standard industry fermentation vector.

The reason we choose RepA as a fusion partner was that previous studies had shown that it was possible to generate very high levels of this protein by over expression in *E. coli* (20,27). In addition there was only a single N-terminal Met residue in the first 100 amino acid residues of this protein, which is another advantage when CNBr cleavage is utilized. Thus we generated two fragments, one encoding for 78 aa and the other 21 aa of RepA for fusion partner construction. The pattern of expression between the Rep78 and Rep21 analogs was observed to be similar and both gave rise to reasonable production of target proteins (Table 2, construct 1 and 9; Figure 2, lane 2 and 3). Since both variants retained the amino terminal sequence, we reasoned that the N-terminal part of RepA was helpful for optimal expression. The importance of this part was also evident by deletion analysis. If the rep21 portion was totally removed from the fusion partner, expression of the remaining part became inconsistent (data not shown).

A summary of the fusion constructs and their levels

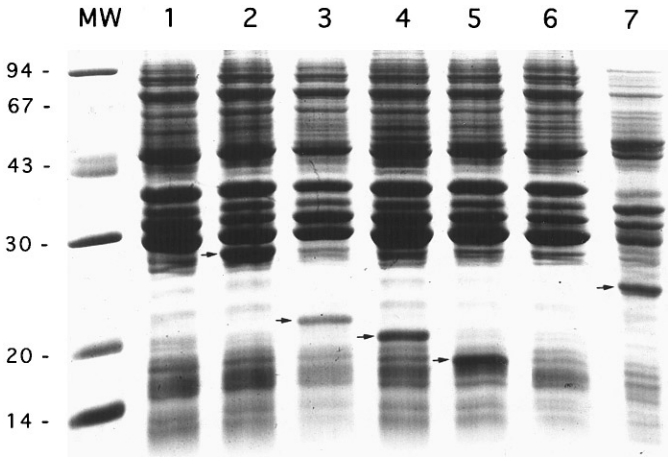

FIG. 2. Coomasie-strained SDS-PAGE analysis of crude cell lysates induced for recombinant expression and containing the following plasmids (also see Table 2): Lane 1, construct 1 uninduced control; lane 2, construct 1; lane 3, construct 9; lane 4, construct 10; lane 5, construct 11; lane 6, construct 14; lane 7, construct 17. Lane 2 to lane 7 (except lane 6) were induced by 0.2 mM IPTG. MW: molecular weight marker (Pharmacia Biotech.)

of expression is presented in Table 2. An indispensable element in recombinant peptide expression is the prepro piece. When Rep78 or Rep21 was fused directly to CEMA, there was no observable expression detected on SDS-PAGE (Table 2, constructs 5 and 15). But this could be corrected if the prepro sequence of HNP-1 was inserted between the Rep and CEMA genes. This is consistent with our earlier finding that inclusion of prepro between GST and either human HNP-1 or a CEMA lead to dramatic increase in expression level. We assume that formation of a folded structure between the cationic peptide and the anionic prepro piece decreased proteolysis. Liu and Ganz demonstrated that the minimal domain of the HNP-1 prepro sequence essential for expression and subcellular trafficking in human neutrophils could be localized to a central stretch of 12 aa (28). We therefore replaced the entire prepro region of clone (Table 2, construct 9) with the minimal pro domain designated pro_{min}. The expression level of this construct was below the detection limits of SDS-PAGE (Table 2, construct 8). We also attempted to delete the pre region; however there was only a trace amount of fusion protein present in the crude cell lysate, indicating that the pre region is advantageous.

In an attempt to simplify the downstream purification procedure, three different affinity tags were tested; including a hexa-histidine tag, a CBD_{nat} from *C. fimi* (23) and an CBD_{syn} from the fungus *T. reesei* (21). Sites chosen for insertion of these tags were between Rep and prepro or between prepro and CEMA. To our surprise, introduction of CBD_{nat} did not seem to be useful since no discernable levels of expression were detected in any constructs containing this tag. However constructs with inserted CBD_{syn} were well expressed and in several cases, especially when locating CBD_{syn} upstream of the prepro domain, yielded expression levels higher than the analogous constructs that did not contain CBD_{syn} (Table 2, construct 1 verses construct 2). We also noticed that CBD_{syn} could be replaced by a his-prepro duplication. Thus the requirement of CBD_{syn} is not indispensable but assists in stabilizing the fusion construct.

In conclusion, as summarized in Table 2, it was apparent that both the specific sequences of the fusion partner and their positioning relative to each other and the peptide were crucial. Specifically in the optimal constructions described above, the fusion must consist of a highly expressed amino terminal at least 21 aa of the RepA protein, and a stabilizing prepro domain located immediately upstream of peptide of interest. In addition to these two requirements, a linker sequence encoding for CBD_{syn} and hexa-histidine tag was also advantageous to separate the Rep and prepro domains, and to significantly boost expression levels.

To test the versatility of the fusion partner created in this study for expression of other polycationic antibacterial peptides, oligonucleotides encoding for two additional peptides bactenecin (Figure 2, lane 5) and indolicidin (data not shown) were fused to the fusion partner carried on pSP72. Bactenecin is a smaller cyclic peptide of 12 amino acids with two cysteines that form a disulphide bridge. Indolicidin is a 13 amino acid tryptophan- and proline-rich peptide that forms an extended helix. Good yields of fusion proteins containing these peptides were observed using our expression system (Table 2, constructs 10-12), indicating that this improved system could be used to express a wide variety of antimicrobial cationic peptide.

All of the above constructs utilized the vector pSP72 (Table 1). Since pSP72 is not as efficient as the high expression vector pT7-7, which carries a strong endogenous T7 promoter and an optimized rbs sequence, we transferred the Rep21 construct to the pT7-7 vector giving rise to vector pEP12. The difference between construct 9 in Table 2 and pEP12 is that the 45 nt sequence carrying the endogenous *repA* promoter and rbs upstream of the region encoding 21 amino acids of *repA* was totally removed. Vector pEP12 showed dramatically higher expression of fusion protein. The peptide gene encoding for CP2600 was cloned into pEP12 to generate pECP1. Expression of pECP1 is shown in Figure 3A.

In all cases, the fusion proteins were expressed as insoluble inclusion bodies under either IPTG induction or heat induction. We observed that IPTG induction was more reliable and consistent than the heat induction. Inclusion bodies were not seen as a disadvantage, since these peptides are basically unstructured in solution and only achieve their folded structure after interaction with the appropriate (membrane-like) environ-

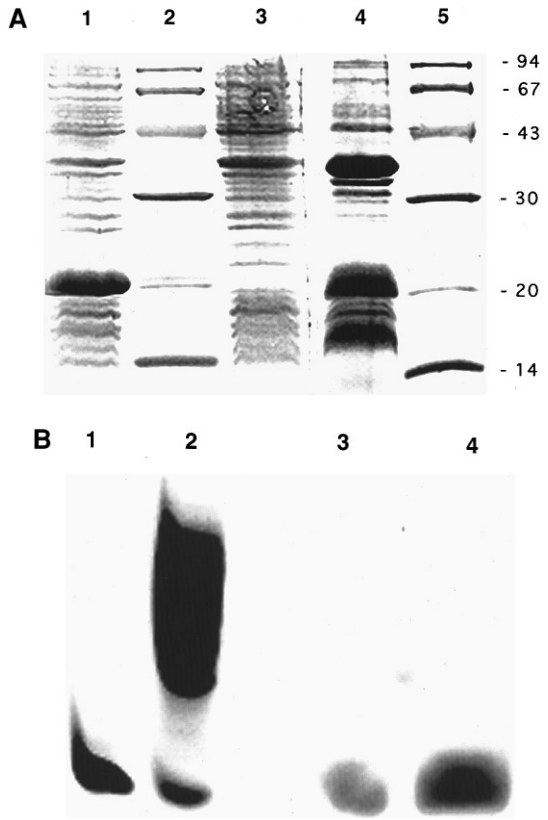

FIG. 3. (A) SDS-PAGE analysis of protein expression and inclusion bodies of BL21 (DE3) harboring pECP1 which carry peptide gene encoding for cationic peptide CP2600 in expression vector pEP12. Lane 1, pECP1 induced by 0.2 mM IPTG; lane 2, molecular weight marker; lane 3, pECP1 uninduced control; lane 4, isolated inclusion bodies; lane 5, molecular weight marker. (B). AU-PAGE analysis of CNBr cleaved inclusion bodies and recombinant peptide. Lane 1, synthetic peptide control; lane 2, CNBr cleaved inclusion bodies; lane 3, synthetic peptide control; lane 4, recombinant peptide purified by reverse-phase FPLC.

ment. We used a modified approach to purify inclusion bodies from *E. coli* cells harboring different constructs. The isolated inclusion bodies were shown in Figure 3A, lane 4. Approximately 140 mg of fusion protein was recovered from 1 liter of cells grown to an OD_{600} of 1.0. The yield was significantly higher than that we had previously described using the *S. aureus* protein A system, in which affinity purification on an IgG column yielded only less than 10 mg of fusion protein per liter of cells at the same OD_{600} value. An added advantage in optimizing the yield of cationic peptide was that our present fusion construct was considerably smaller (15 kDa) than the much larger protein A fusion (30 kDa after processing of the signal sequence) or the AraB fusion system (52 kDa) according to Callaway (16). When the size of fusion protein is small, it would result in higher stoichiometric amounts of cationic peptide after cleavage. Furthermore, our fusion protein has only a single internal Met residue placed prior to the peptide sequence, which made downstream purification considerably easier after CNBr cleavage. The CNBr cleaved inclusion bodies are shown in Figure 3B, lane 2; and the recombinant peptide CP2600 was shown in Figure 3B, lane 4.

We also attempted to purify the fusion protein before or after renaturation of the fusion protein from urea, using a prepackaged chelating sepharose column (Hi-Trap column, Pharmacia Biotech.) to affinity bind the hexa-histidine tag, or Avicil (phosphoric acid-treated cellulose) to bind CBD_{syn} tag. However neither procedure was successful. We reasoned that the failure of affinity purification is likely due to the position of the hexa-hitidine tag in the middle position of the fusion protein which would make it poorly exposed to nickel metal, and similar considerations may also have accounted for the failure to bind to cellulose. We are currently trying to improve affinity binding by relocating the hexa-histidine tag to the N-terminal of the fusion protein.

Three classes of cationic peptides, CEMA variants, bactenecin and indolicidin, were recombinantly produced in this study (Table 2). To compare the activity of purified peptides with their chemically synthesized counterparts, minimal inhibitory concentration (MIC) assays were performed. Results showed no difference in killing activity against bacteria (data not shown).

Construction of peptide variants via DNA mutagenesis. Recombinant DNA technology provides an efficient way to generate a pool of peptide variants via DNA mutagenesis. We first applied a random mutagenesis technique to alter the DNA sequence encoding for indolicidin as described in Materials and methods. The mutations were introduced into the DNA fragment by utilizing doped nucleotide bottles containing 3% of each of the other nucleotides (ie. 3% of each A, T and G in C bottle, etc) during oligonucleotide synthesis, and these fragments were subsequently cloned downstream of the *S. aureus* protein A-derived expression vector. Sequencing of a random selection of three clones revealed 4-8 base pair changes per clone that resulted in 4-6 amino acid substitutions. As shown in Table 3, variant 1 where K was substituted by I at position 5, and at position 6, W was substituted by C etc. This method was shown to be useful for producing a recombinant peptide library expressing variants with a predetermined mutation rate. However we found this method also often introduced nonsense mutations to the DNA sequence, or very great changes in the character of amino acids at particular sites, and in most cases, the mutations were unpredictable.

To overcome this difficulty, we developed a semi-random mutagenesis approach. Unlike the random mutagenesis which allowed us to mix four nucleotides within a stretch of DNA, the targetted-mutagenesis approach involved the mixing of two nucleotides at carefully de-

TABLE 3

Peptide Variants Produced by DNA Mutagenesis

Peptide	Amino acid sequence
Indolicidin	ILPWKWPWWPWRR
Variant 1	ILPWICPWRPSKAN
Variant 2	IVPWKWTLWPWRR
Variant 3	TLPCLWPWWPWSI
CP2600 variants	
V1	KWKSFLKTLKSPVKTVFYTALKPISS
V8	KWKSFLRTFKSPVRTVFHTALKPISS
V10	KWKSFLKTLKSAVRTEFHTALKPISS
V11	KWKSFLRTLKSAVRTEFHTALKPISS
V14	KWKSFLKTFKSPVKTVFYTALKPISS
V25	KWKSFLRTLKSPAKTVFHTALKAISS
V31	KWKSFLKTFKSPARTVLHTALKPISS
V68	KWKSFLKTFKSPARTVLYTALKPISS
V681	KWKSFLKTFKSAVKTVLHTALKAISS
CP2600	KWKSFLKKLTSAAKKVLTTALKPISS

signed positions to change specific codons, leading to specific amino acid substitution at those positions. For example, with codon TTC, if we mixed both A and C at the third wobble position, it would give rise to two possible codons, TTC, encoding F, or TTA, encoding L. This method in particular, allowed us to switch between any two similar amino acids, such as between K and R, A and V, F and L etc. We utilized three strategies in changing amino acids, one was to substitute hydrophobic amino acids with different hydrophobic amino acid, e.g., L with I; the second was to substitute hydrophobic amino acids with hydrophilic amino acids, like V and E, the third was to substitute charged amino acids with a similarly charged amino acid, like R with K. We randomly selected 9 clones and all of them carried mutations as expected (Table 3).

ACKNOWLEDGMENTS

This work was financially supported by Micrologix Biotech Inc. and the Canadian Cystic Fibrosis Foundation. SF was an MRC Industrial Postdoctoral Fellow. REWH was an MRC Distinguished Scientist.

REFERENCES

1. Hancock, R. E. W., Falla, T., and Brown, M. M. (1995) *Adv. Microbial Physio.* **37,** 135–137.
2. Ganz, T., Selsted, M. E., and Lehrer, R. I. (1990) *Eur. J. Haemat.* **44,** 1–8.
3. Hultmark, D., Engstrom, A., Bennich, H., Kapur, R., and Boman, H. G. (1982) *Eur. J. Biochem.* **127,** 207–217.
4. Hultmark, D., Steiner, H., Rasmuson, T., and Boman, H. G. (1980) *Eur. J. Biochem.* **106,** 7–16.
5. Zasloff, M. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 5449–5453.
6. Selsted, M. E., Novotny, M. J., Morries, W. L., Tang, Y. Q., Smith, W., and Cullor, J. S. (1992) *J. Biol. Chem.* **267,** 4292–4295.
7. Skerlavaj, B., Romeo, D., and Gennaro, R. (1990) *Infect. Immun.* **58,** 3724–3730.
8. Agawa, Y., Lee, S., Ono, S., Aoyagi, H., Ohno, M., Tamguchi, T., Anzai, K., and Kirino, Y. (1991) *J. Biol. Chem.* **266,** 20218–20222.
9. Christensen, B., Fink, J., Merrifield, R. B., and Mauzerall, D. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 5072–5076.
10. Sawyer, J. G., Martin, N. L., and Hancock, R. E. W. (1988) *Infect. Immun.* **56,** 693–698.
11. Gough, M., Hancock, R. E. W., and Kelly, N. M. (1996) *Infect. Immun.* **64,** 4922–4927.
12. Darveau, R. P., Cunningham, M. D., Seachord, C. L., Cassiano-Clough, L., Cosand, W. L., Blake, J., and Watkins, C. S. (1991) *Antimicrob. Agents. Chemother.* **35,** 1153–1159.
13. Selsted, M. E., Szklarek, D., Ganz, T., and Lehrer, R. I. (1985) *Infect. Immun.* **49,** 206–206.
14. Merrifield, R. B. (1963) *J. Am. Chem. Soc.* **85,** 2149–2154.
15. Piers, K. L., Brown, M. M., and Hancock, R. E. W. (1993) *Gene* **134,** 7–13.
16. Callaway, J. E., Lai, J., Haselbeck, B., Baltaian, M., Bonnesen, S. P., Weickmann, J., Wilcox, G., and Lei, S. P. (1993) *Antimicrob. Agents. Chemother.* **37,** 1614–1619.
17. Taguchi, S., Nakagawa, K., Maeno, M., and Momose, H. (1994) *Appl. Environ. Microbiol.* **60,** 3566–3572.
18. Hancock, R. E. W., and Carey, A. M. (1979) *J. Bact.* **140,** 902–910.
19. Panyim, S., and Chalkey, R. (1969) *Arch. Biochem. Biophys.* **130,** 337–346.
20. Ausubel, E. M., Brent, R., Kingston, R. E., Moore, O. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York.
21. Tomme, P., Van Tilbeurgh, H., Petterson, G., Van Damme, J., Vandekerckhove, J., Knowles, J., Teeri, T., and Claeyssens, M. (1988) *Eur. J. Biochem.* **170,** 575–581.
22. Burian, J., Stuchlik, S., and Kay, W. W. Submitted.
23. Ong, E., Kilburn, D. G., Miller, R. C. Jr., and Warren, R. A. (1994) *J. Bact.* **176,** 999–1008.
24. Tabor, S., and Richardson, C. C. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1074–1078.
25. McNeil, J. B., and Smith, M. (1985) *Mol. Cell. Biol.* **5,** 3545–3551.
26. Adari, H. Y., Rose, K., Williams, K. R., Konigsberg, W. H., Lin, T. C., and Speicer, E. K. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1901–1905.
27. Burian, J., Guller, L., Macor, M., and Key, W. W. (1997) *Plasmid* **37,** 2–14.
28. Liu, L., and Ganz, T. (1995) *Blood* **85,** 1095–1103.
29. Hanahan, D. (1983) *J. Mol. Biol.* **166,** 557–580.
30. Studier, F. W., and Moffatt, B. A. (1986) *J. Mol. Biol.* **189,** 113–130.