*J. Afr. Ass. Physiol. Sci. 2 (1): 117-129 (2014)*

Journal of African Association of Physiological Sciences

*Official Publication of the African Association of Physiological Sciences*

*http://www.jaaps.aapsnet.org*

*Research Article*

# Pentocin KCA1: a novel circular bacteriocin gene encoded in the genome of *Lactobacillus pentosus* KCA1 with putative basic property.

**Kingsley C Anukam**

*TWAS Genomics Research Unit, Department of Medical Laboratory Sciences, School of Basic Medical Sciences, University of Benin, Benin City, Nigeria.*

**Keywords:**
*Circular bacteriocin, Pentocin KCA1, Lactobacilli, Antimicrobial peptide*

**ABSTRACT**

**Background**: The use of bacteriocin and bacterial strains that produces the antimicrobial peptide has shown to possess potential applications in the conferment of health benefits on the host. We isolated and carried out comprehensive genome sequence analysis of the first *Lactobacillus pentosus* KCA1 of human origin encoding genes for the biosynthesis of antimicrobial bacteriocin peptide. Due to the growing number of antimicrobial resistance, the need for developing alternatives to traditional antibiotics is now more germane. **Aims:** To describe the first circular bacteriocin predicted in the genome sequence of *Lactobacillus pentosus* KCA1 isolated from the vagina of a healthy Nigerian Ibo woman using *in silico* bioinformatic tools. **Methods** The translated open reading frame (ORF) coding pentocin KCA1 was compared with the non-redundant database (nrdb) using BLASTp for protein similarity search. Clustalw algorithm was used for alignment with other published circular bacteriocins. **Results**: The genome of *L. pentosus* KCA1 contains a 7-gene cluster, chromosomally encoded for biosynthesis of a predicted circular bacteriocin. The bacteriocin designated as ñpentocin KCA1ò is synthesized as a precursor gene consisting of 273 nucleotide base sequence encoding the translated product of pentocin KCA1 with 91 amino acid residues in length. The peptide is cleaved off between asparagine (Asn33) and isoleucine (Ile34) to produce the 58 amino acid pentocin KCA1 as an outer membrane peptide. The mature pentocin KCA1 has a high proportion of basic (positively charged-Lysine, Histidine and Asparagine) to acidic (negatively charged-Glutamate and Aspartate) amino acids in the ratio of 8:0. **Conclusions:** Off the 11 circular bacteriocins known to date, amino acid residue asparagine (8.62%) is utilized more in the biosynthesis of pentocin KCA1. The mature putative circular pentocin KCA1 consists of four alpha-helical structures and has a high proportion of basic amino acid residues when compared with other circular bacteriocins, thereby suggesting that pentocin KCA1 is a circular bacteriocin peptide with strong basic property. The relevance of this basic property lends credence for investigation in subsequent functional studies.



.

## INTRODUCTION

Bacteriocins are ribosomally synthesized antimicrobial peptides produced by prokaryotes to facilitate ecological competition (Majeed et al., 2011). The inhibitory spectrum of these agents is predominantly directed toward bacteria that are closely related to the producer strain (Jack et al., 1995). The modes of action exhibited by many bacteriocins makes them more active than conventional antibiotics against pathogenic and drug-resistant Gram-positive bacteria, yet display no toxicity toward eukaryotic cells (Nissen-Meyer & Nes, 1997). Bacteriocins have potential applications in human and veterinary medicine in the treatment of local and systemic bacterial infections (Sang & Blecha, 2008; Sit & Vederas, 2008).

The functions of bacteriocins as colonizing peptides, includes antimicrobial activities, signaling either through quorum sensing or cross talk with the host immune system from some bacterial species colonizing the genitourinary and gastro-intestinal tracts have been documented (Dobson et al., 2012). Traditionally, bacteriocins have been classified into two major groups: class I lantibiotics with post-translationally

modified amino acids and class II nonlantibiotics with nonmodified amino acids (Nes et al., 2007). Over a couple of years ago, a new class of bacteriocins has been proposed, circular bacteriocins, which are characterized by a covalent linkage between the N and C termini (Maqueda, et al., 2008; Heng & Tagg, 2006). The circular structure appears to enhance the thermodynamic stability and structural integrity of the peptide, which is required to improve its biological stability and activity (Gong et al., 2009).

To date, about 11 circular bacteriocins have been reported from diverse bacterial species. These includes gassericin A (Kawai et al., 1998), enterocin AS-48 ( Sanchez-Barrena et al., 2003), reutericin 6 (Toba et al., 1991), circularin A (Kemperman et al., 2003), butyrivibriocin AR10 (Kalmokoff et al., 2003), subtilosin A (Zheng et al., 1999), uberolysin (Wirawan et al., 2007), lactocyclicin Q (Sawa et al., 2009), acidocin B (Leer et al., 1995), carnocyclin A (Martin-Visscher et al., 2008) and recently Garvicin ML (Borrero et al., 2011). One of the cell-killing mechanisms of circular bacteriocins depends on the ability to form membrane pores as observed in enterocin AS-48 (Sanchez-Barrena et al., 2003). Interestingly, production of bacteriocin is one of the characteristics required for the selection of potential probiotic strains (Kawai et al., 1998).

It should be noted that all the published circular bacteriocin peptides were recovered from bacterial species colonizing sites other than the human vagina. However, these peptides and their nucleotide base sequences, obtained from the non-redundant database of the NCBI have provided leeway to characterize a circular bacteriocin here designated pentocin KCA1. The first chromosomally encoded circular bacteriocin predicted in the genome sequence of *Lactobacillus pentosus* KCA1 (Anukam et al., 2013) isolated from the vagina of a healthy woman is hereby described.

**Methods:**

***Bacterial strain, DNA sequencing and gene predictions***

*Lactobacillus pentosus* KCA1 was originally isolated from the vagina of a healthy Ibo- Nigerian woman. DNA sequencing, gene predictions and annotations have been described previously (Anukam et al., 2013). Briefly, genomic DNA from *Lactobacillus pentosus* KCA1 was used to prepare a genomic library using the Illumina paired-end sample preparation protocol at the Centre for Applied Genomics, Toronto, Canada (www.tcag.ca). The sequencing was done with the Next-Generation Illumina GA II facility. The reads were filtered to include only those with a Q score greater than 10 for all nucleotides, leaving 16,920,226 reads, about 8.45 million from each side with approximate mate-pair distances of 450 base-pairs. The paired-end reads were assembled into contigs with a maximum *kmer* length of 57 using the VELVET assembler tool. The final assembly has 602 nodes and n50 of 108429, a maximum of 217,505, and a total of 3,418,159 bp. All the reads were used for the assembly initially giving 281 contigs > 200bp in length, which were trimmed down to 83 contigs (1 scaffold) for gene predictions. Open-reading frames (ORFs) greater than 100 nt were predicted using GeneMark (Isono et al., 1994) and Glimmer software (Salzberg et al., 1998). The translated ORF predictions were compared to the NCBI non-redundant database (nrdb) using BLASTp to evaluate gene predictions. All predicted ORFs were manually checked with the Artemis software (Carver et al., 2008) and corrected when necessary (e.g. start codons, frameshifts).

***Location of the gene clusters coding pentocin KCA1***

DNA plotter and Arthemis ACT tool (Carver et al., 2008) were used to locate the position of the bacteriocin gene clusters and bacteriocin immunity proteins within the *L. pentosus* KCA1 chromosome.

***Determination of the predicted optimal aligned spacing between the Shine-Dalgarno (SD) sequence and the initiation codon ATG for the pentocin KCA1.***

The sequence of SD and SD sequence complementary to the 3ô-end of the 16 S rRNA was adapted from Chen et al. (1994) and aligned to the region of *L. pentosus* KCA1 genome sequence coding for pentocin KCA1 precursor, in order to determine the spacing between the SD sequence and the initiation codon.

***Comparisons of nucleotide sequences of the 12 circular bacteriocin precursors including pentocin KCA1***

Nucleotide sequences coding for the circular bacteriocin precursors including pentocin KCA1 were downloaded from the non-redundant database of European Molecular Biology Laboratory (EMBL). The sequences were imported into Jalview program (Waterhouse et al., 2009) for ClustalW and

Multiple Alignment using Fast Fourier Transform (MAFFT) alignments.

***Prediction of the transmembrane region of the circular bacteriocins including pentocin KCA1 precursor***

Transmembrane regions in peptides were deduced using the SOSUI program (Hirokawa et al., 1998) (http://bp.nuap.nagoya-u.ac.jp/sosui/) and PRED-TMBB: a web server for predicting the topology of beta-barrel outer membrane proteins (Bagos et al., 2004).

***Secondary structure prediction***

The secondary structure of pentocin KCA1 was predicted with online position specific iterated (PSIPRED) (McGuffin et al., 2000) and Jnet algorithm prediction (Jpred3) (Cole et al., 2008) [31] protein structure prediction server, which compares with non-redundant set of protein database (PDB) structures.

***Comparative physicochemical properties of the cyclic bacteriocins predicted from bacteriocin-BACTIBASE database.***

Individual mature circular bacteriocins were extracted from the online repository of bacteriocin natural antimicrobial peptides database-BACTIBASE (http://bactibase.pfba-lab-tun.org/main.php).

Physicochemical properties such as total amino acids present and absent, mass (Da), isoelectric point, net charges, basic and acidic residues, hydropathy index, instability index and half life were automatically predicted.

**Results:**

The genome of *L. pentosus* KCA1 encodes seven bacteriocin-related genes including a new class of bacteriocin identified by *in silico* analysis as circular bacteriocin, belonging to class V, here designated as pentocin KCA1 (*PenA*) (locus tag KCA1_0433).

The protein coding sequences for the pentocin KCA1 appeared in contig 12 (NZ_AKAO01000012.1, NCBI DB XREF; GI:392947746) in the region of 489167 to 489442 along the *L. pentosus* KCA1 chromosome (3, 418, 159bp) as shown in **Figure 1**.

The pentocin KCA1 shows a limited 49% amino acid residue identity to the class IIc bacteriocin, gassericin A from *L. gasseri* LA39, 50% id to acidocin B from *L. acidophilus*, 34% id to butyrivibriocin AR10 from Butyrrivibrio, and 38% id to an unknown bacteriocin of *Streptococcus sp*. 2_1_36FAA. The gene penT, annotated as ABC transporter ATP-binding protein is 237 amino acid long, having 46% identity to gaaT with 226 amino acids (**Table 1**).

The pentocin KCA1 precursor gene consists of 273 nucleotide base sequence open reading frame encoding the primary translated product of pentocin KCA1 with 91 amino acid residues in length as shown in **Figure 2-A**. A putative ribosomal binding site precedes the precursor gene, with the Shine-Dalgarno sequence (AAAGGAGGA) (Shine & Dalgarno 1974) upstream of the methionine translation initiation codon (ATG). The *L. pentosus* KCA1 region coding for pentocin A precursor aligned 100% to the Shine-Dalgarno (SD) sequence 5ô-UAAGGAGGU and the initiation codon AUG (**Figure 2-B**). The SD sequence is complementary to the 3ô-end of the 16S rRNA (anti-codon to the SD region) as underlined. The SD is separated by 7 nucleotides before the initiation codon ATG. **Figure 3** shows the alignment of the 12 circular bacteriocins, including pentocin KCA1 (penA). The bacteriocins with common amino acid residue motifs are indicated in red colour.

The putative 33 leader peptides of the pentocin KCA1 precursor cleaves off between asparagine (Asn) and Isoleucine and cyclization taking place between the N-terminal isoleucine (Ile34) and the C-terminal Alanine (Ala91) by a peptide bond formation.

Using the MAFFT tool to align and compare the nucleotide base sequences of both group i and ii circular bacteriocins, it can be shown that termination codon ñTAGò is utilized for the precursors of pentocin KCA1, gassericin A, reutericin 6, acidocin B. In contrast, codon ñTAAò terminated the precursors of butyrivibriocin AR10, subtilosin A, enterocin AS-48, circularin A, uberolysin, lactocyclin Q, carnocyclin A, and garvicin ML (**Figure 4**).

**Figure 5** shows the ClustalW amino acid alignment tree of the 12 published circular bacteriocins including pentocin KCA1 displaying the relationships of the peptides.

The mature circular pentocin KCA1 has a high proportion of basic (positively charged-Lysine K [1], Histidine H [2], and Asparagine [5]) to acidic (negatively charged-Glutamate [0], and Aspartate [0]) amino acids in the ratio of 8:0. In comparison, gassericin A and acidocin B have positively charged amino acids-Lysine [1] and Histidine [1] and

Negatively charged amino acids-Glutamate [0] and Aspartate [2].

The utilization of amino acids in the biosynthesis of the pentocin KCA1 varies in comparison with other group ii circular bacteriocins (**Figure 6**). For example, alanine (Ala) is the most predominant amino acid in the mature circular pentocin KCA1, making up to 20.69% of the total peptide, while it is 31.0% for gassericin A and acidocin B respectively. Seven (12.07%) leucine amino acid residues are utilized in the biosynthesis of pentocin KCA1, while eight (13.8%) leucine are used for gassericin A, and acidocin B. Comparatively, among all the published circular bacteriocins, asparagine (8.62%) is utilized more in the biosynthesis of pentocin KCA1. Also, amino acid residue valine (12.07%), histidine

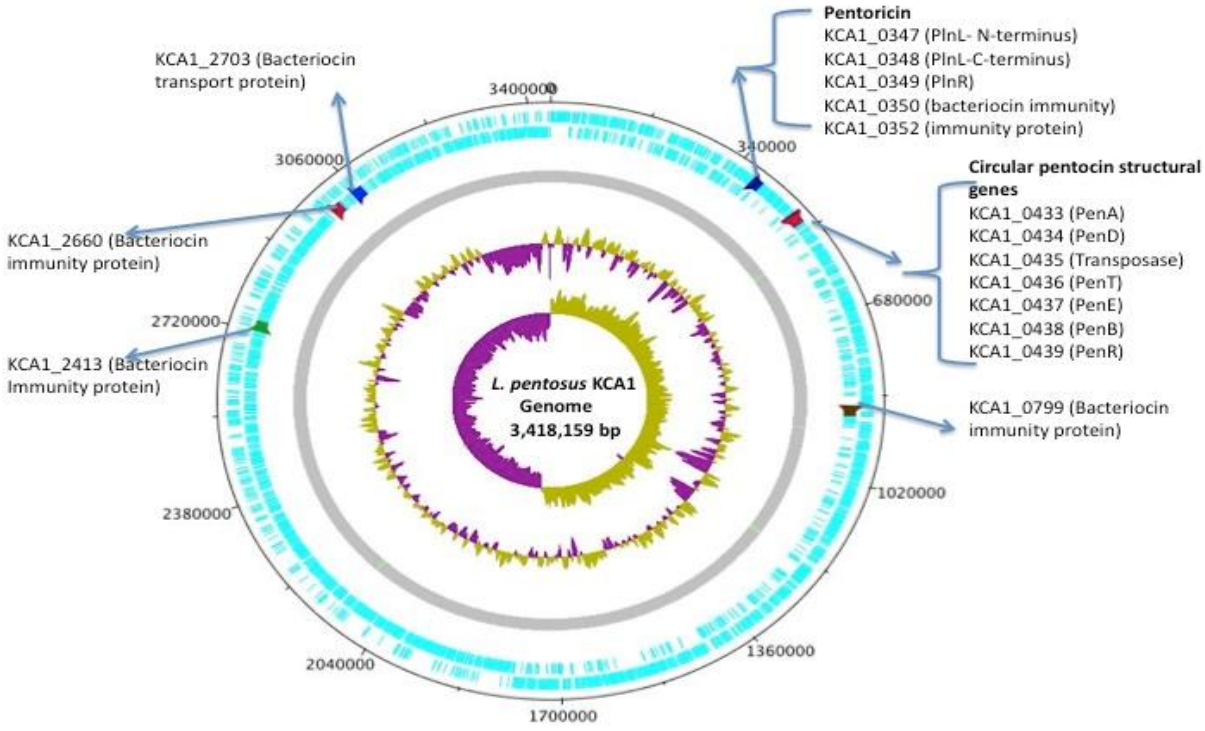


**Fig. 1:** Genome atlas view of the *L. pentosus* KCA1 chromosome showing the locations of the circular pentocin KCA1 structural genes and associated bacteriocin immunity proteins. Location of the pentocin KCA1 was done using DNA plotter and Artemis tools.

**Table 1**: Pentocin KCA1 gene clusters showing the locus tags, number of amino acid (AA), product annotations, protein family identity number (Pfam ID), transmembrane helical prediction (#TMHMM), and % species identity

| **KCA1 gene** | Gene | AA | Product | #Pfam ID | #TMHMM | % Species Identity |
|---|---|---|---|---|---|---|
| KCA1_0433 | PenA | 91 | Circular bacteriocin (**PentocinA KCA1**) | PF12173.1 | PredHel=2 | 49% ID to gaaA (91 AA) |
| KCA1_0434 | PenD | 137 | Hypothetical protein | PF01944.10 | PredHel=4 | 31% ID to gaaD (162 AA) |
| KCA1_0435 | IS30 (Ts) | 340 | Transposase (IS30 Family) | PF13936 | PredHel=1 | 92% to is15 LPST_C0978 |
| KCA1_0436 | PenT | 237 | ABC transporter ATP-binding protein | PF00005.20 | PredHel=0 | 46% ID to gaaT (226 AA) |
| KCA1_0437 | PenE | 213 | ABC transporter, permease protein | No hit | PredHel=6 | 40% ID to gaaE (212 AA) |
| KCA1_0438 | PenB | 177 | Hypothetical protein | No hit | PredHel=5 | 27% ID to gaaB (174 AA) |
| KCA1_0439 | PenR | 67 | Transcriptional regulator PBSX family | PF01381.15 | PredHel=0 | 59% ID to B. cereus (63 AA) |

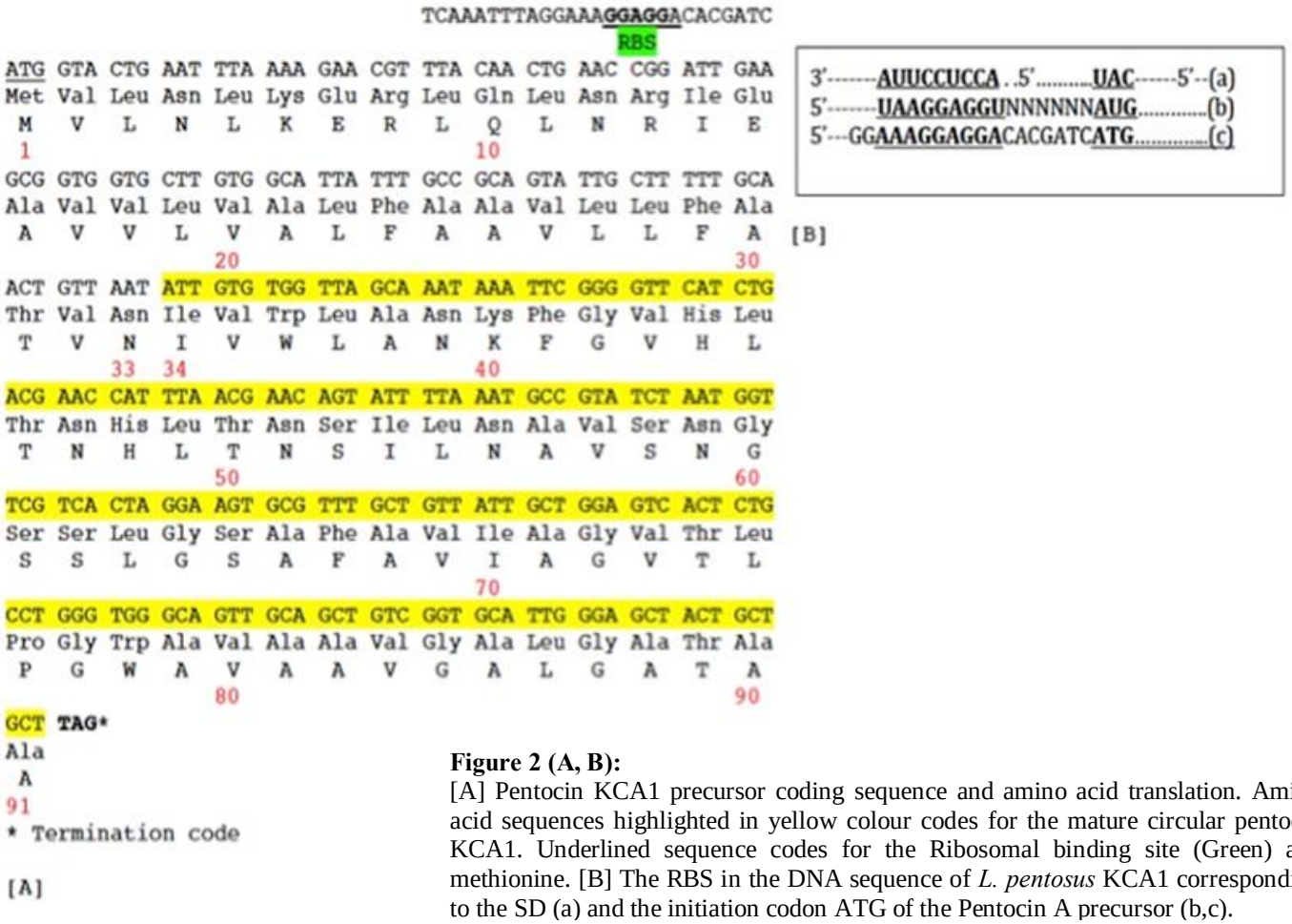


**Figure 2 (A, B):**
[A] Pentocin KCA1 precursor coding sequence and amino acid translation. Amino acid sequences highlighted in yellow colour codes for the mature circular pentocin KCA1. Underlined sequence codes for the Ribosomal binding site (Green) and methionine. [B] The RBS in the DNA sequence of *L. pentosus* KCA1 corresponding to the SD (a) and the initiation codon ATG of the Pentocin A precursor (b,c).

(3.45%), and serine (8.62%) appear to be higher in pentocin KCA1 than gassericin A and acidocin B.

The secondary structure prediction as determined by JPRED and PSI servers, indicates that pentocin KCA1 has four putative alpha helices (**Figure 7**) similar to gasserin A, carnocyclin A and lactocyclicin Q, unlike AS-48, Circularin A and Uberolysin that haven been proven to have five alpha helices (Martin-Visscher et al., 2009). The four putative α-helices of pentocin KCA1 comprised of residues $Trp^{3}$-$Phe^{8}$, $His^{11}$-$Val^{24}$, $Leu^{30}$-$Ala^{38}$, and $Gly^{44}$-$Ala^{55}$. For pentocin KCA1, there are presence of a high number of helix breaking residues glycine and proline, similar to what has been observed in the structure of carnocyclin A, in which a long extended loop separates α2 from α3 and a short tight loop demarcates α3 and α4. The mature circular pentocin KCA1 has GXXG, GXXXG, and GXXXXG motifs (Figure 3), which are used potentially to strengthen the helical structure. From the hidden Markov model (http://bioinformatics.biol.uoa.gr//PRED-TMBB/), pentocin KCA1 amino acid precursor sequence scored a value of **2.814**, which is lower than the threshold value of 2.965. The difference between the value and the threshold indicates the possibility of the protein being an outer membrane protein.

However, the mature cleaved cyclic pentocin KCA1 had a sequence score value of **2.833**, which is still lower than the threshold value of 2.965.

The transmembrane region of the mature circular pentocin KCA1 was predicted to occur at position 33 for the N terminal and 55 for C-terminal with a total of 23 amino acids in length.

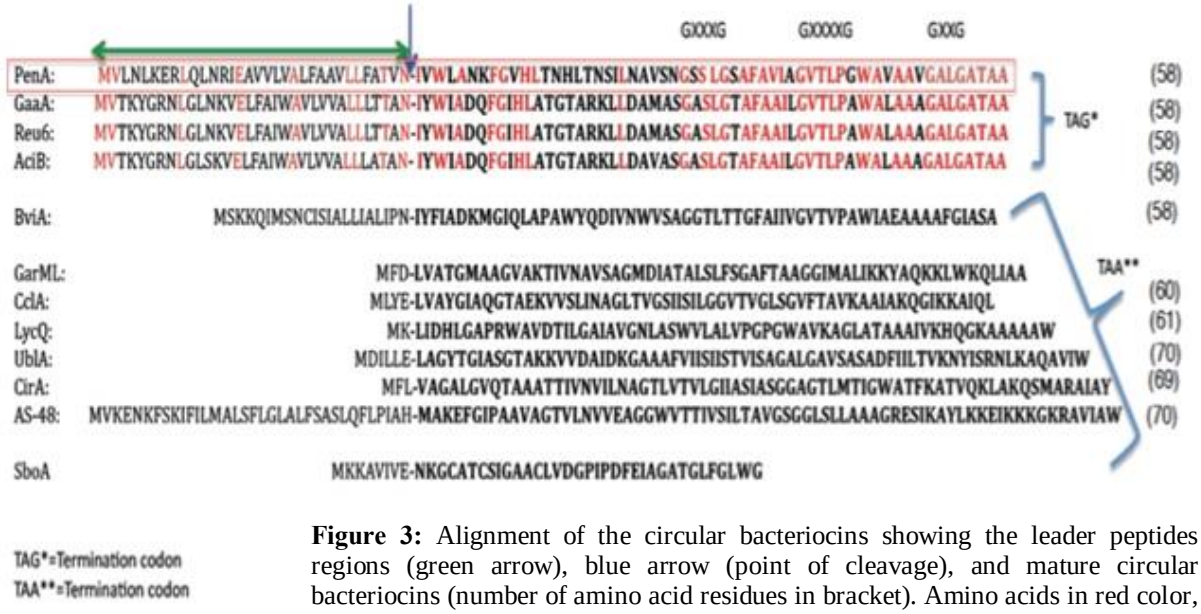


**Figure 3:** Alignment of the circular bacteriocins showing the leader peptides regions (green arrow), blue arrow (point of cleavage), and mature circular bacteriocins (number of amino acid residues in bracket). Amino acids in red color, shows common motifs.

Table 2: Physicochemical properties of group ii cyclic bacteriocins including pentocin KCA1, predicted from bacteriocin-BACTIBASE database **(**http://bactibase.pfba-lab-tun.org/main.php).

| Cyclic Bacteriocins | Pentocin KCA1 | Gassericin A | Acidocin B | Reutericin 6 | Butyrivibriocin AR10 | Subtilocin-A |
|---|---|---|---|---|---|---|
| Total amino acids | 58 | 58 | 59 | 58 | 58 | 35 |
| Absent amino acids | CDEMRY | CEN | CEM | CEN | CHR | HMQRY |
| Common AA | A | A | A | A | A | G |
| Mass (DA) | 5672..85 | 5672.18 | 5772.73 | 5672.18 | 6019 | 3444.56 |
| Isoelectric point | 9.7 | 7.54 | 7.54 | 7.54 | 3.88 | 3.88 |
| Net charge | 3 | 1 | 1 | 1 | -2 | -2 |
| Basic residues | 3 | 3 | 3 | 3 | 1 | 1 |
| Acidic residues | 0 | 2 | 2 | 2 | 3 | 3 |
| Hydrophobic residues | 33 | 35 | 36 | 35 | 34 | 15 |
| Polar residues | 21 | 15 | 16 | 15 | 15 | 14 |
| Aliphatic residues | 17 | 13 | 14 | 13 | 15 | 7 |
| Tiny residues | 24 | 27 | 27 | 27 | 21 | 13 |
| Boman Index | 41.05 | 47.31 | 42.36 | 47.31 | 51.97 | 16.34 |
| Hydropathy Index | 0.964 | 0.997 | 0.96 | 0.997 | 1 | 0.69 |
| Aliphatic Index | 122.93 | 31.62 | 119.66 | 31.62 | 114.66 | 89.43 |
| Instability Index | 10.44 | 3.73 | 9.19 | 3.73 | 16.29 | 24.71 |
| Extinction Coefficient | 11000 | 19490 | 12490 | 19490 | 19480 | 5625 |
| Absorbance 280nm | 192.98 | 219.12 | 215.34 | 219.12 | 341.75 | 165.44 |
| Half life | | | | | | |
| Mammalian | 20 hours | 20 hour | 1.4 hour | 20 hour | 20 hour | 1.4 hour |
| Yeast:30 minutes | 30 min | 30 min | 3 min | 30 min | 30 min | 3 min |
| E. coli:>10 hour | >10 hours | >10 hour | >10 hour | >10 hour | >10 hour | >10 hour |

Footnote: AA=Amino acid

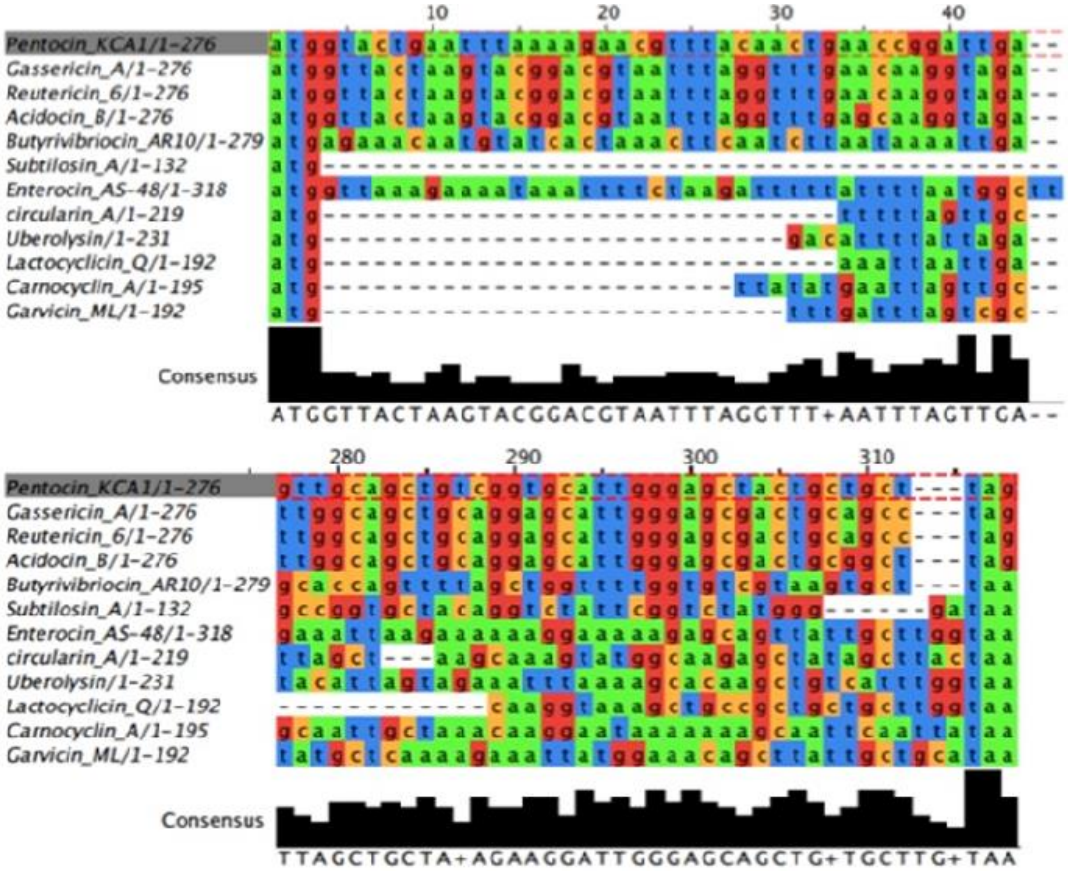


**Figure 4**: Alignment of nucleotide base sequences of circular bacteriocin precursors showing the termination codons . The alignment was done using MAFFT tool.

The Transmembrane regions in the peptides were deduced using the SOSUI program (http://bp.nuap.nagoya-u.ac.jp/sosui/) indicating the possibility of the pentocin KCA1 being an outer membrane protein as shown in **Figure 8.** The leader pentocin KCA1 precursor peptide has two transmembranes, which appear to be providing protection against the putative circular bacteriocin. This assumption is based on the fact that Lysine ($K^+$)-Glutamic acid ($E^-$)-Arginine ($R^+$) residues are located in the transmembrane region, with the possibility of countering the net positive charge of the mature pentocin in re-entering the cytoplasm of the bacterial cell. Interestingly this is fortified by the amino acids at position 12-25 that are in the inner cytoplasm having a negatively charged glutamic acid and a positively charged arginine.

**Discussion:**

The differences observed in the organization of the gene clusters and varied sequence identity of pentocin KCA1 when compared with other circular bacteriocins, indicates that the peptide may be novel. The region encoding the pentocin KCA1 gene clusters with locus tags (KCA1_0433-KCA1_0439) represents both the biosynthetic and secretory systems involving various transporters, including bacteriocin immunity proteins, used to protect against self-destruction. It is interesting to note that these gene clusters for pentocin KCA1 (penA-penD-IS30-penT-penE-penB-penR) are chromosomally encoded (Anukam et al., 2013), while gassericin A/ reutericin 6 (Kawai et al., 1998; Toba et al., 1991), acidocin B (Leer et al., 1995), AS-48 (Sanchez-Barrena et al., 2003) and

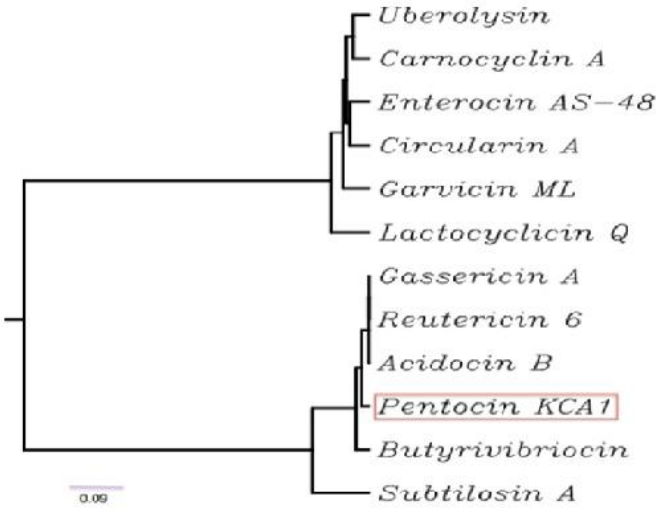


**Figure 5**: ClustalW alignment tree of the 12 published circular bacteriocins including pentocin KCA1 displaying the relationships of the peptides.

garvicin ML (Borrero et al., 2011) are located on plasmids. Heterologous expression of putative *penADTsTEBR* may result in the synthesis of pentocin KCA1. A transposase -*Ts* (KCA1_0435) of the IS30 family is found between penD and penT and it may be involved in the selective non-covalent integration of the gene product within the chromosome (Nagy et al., 2004). The order of gene clusters appears to be re-arranged, when compared with the clusters of gassericin A (*gaaBCADITE*), Enterocin AS-48 (*as-48ABCC$_1$DD$_1$*), circularin A (*cirABCDE*), and carnocyclin A (*cclBITCDA*). It appears that pentocin KCA1 is the only circular bacteriocin with a nearby putative transcriptional regulator (*penR*), which belongs to the PBSX family. This regulatory gene may be involved in the transcription of the pentocin structural gene.

The process of prokaryotic translation initiation involves binding of the 16S rRNA and the initiator tRNA to the mRNA ribosome binding site (RBS) on the mRNAs. The RBS generally extends 20 nucleotides (nt) on either side of the translation initiation codon (usually AUG) and contains, upstream from the AUG, part or all of a polypurine sequence (UAAGGAGGU) known as the Shine-Dalgarno(SD) sequence (Shine & Dalgarno, 1974). Excessive long or short spacing between the SD and initiation codon have been shown to be detrimental to efficient translation initiation (Roberts et al., 1979). The SD is separated by 7 nucleotides before the initiation codon ATG in pentocin KCA1, thus suggesting that translation initiation may take place efficiently.

The pentocin KCA1 is synthesized as 91 amino acid precursor peptide, which is processed between Asparagine (Asn33) and isoleucine (Ile34) to produce the mature 58 amino acid pentocin KCA1 peptide. The alignment of pentocin KCA1 with other circular bacteriocins justifies the hypothesis that pentocin KCA1 is cleaved between Asn33 and Ile34. The same position of cleavage was observed for both Gassericin A and Acidocin B with the same corresponding amino acid residues. The putative 33 leader peptides of the pentocin KCA1 precursor is predicted to cleave off and cyclization taking place between the N-terminal isoleucine (Ile34) and the C-terminal Alanine (Ala91) by a peptide bond formation, similar to Gassericin A.

The MAFFT tool aligned and compared the nucleotide base sequences of both group i and ii circular bacteriocins, showing that termination codon ñTAGò is utilized for the precursors of pentocin KCA1, gassericin A, reutericin 6, and acidocin B. In contrast, codon ñTAAò terminated the precursors of butyrivibriocin AR10, subtilosin A, enterocin AS-48, circularin A, uberolysin, lactocyclin Q, carnocyclin A, and garvicin ML (**Figure 4**). However, the functional relevance of this distinction is yet to be determined.

Remarkably, the mature circular pentocin KCA1 is devoid of Cystine (Cys-C), Glutamic acid (Glu-E), Methionine (Met-M), Glutamine (Gln-Q), Arginine (Arg-R), Tyrosine (Tyr-Y) and Aspartic acid (Asp-D), while only Cys-C, Glu-E, and Asn-N amino acid residues are absent in gassericin A. In comparison, the isoelectic point of pentocin KCA1 appear to be higher (9.7) than all the known circular bacteriocins in group ii with lower isoelectric point, thereby indicating the basic nature of the peptide.

The amino acid composition analysis clearly show the absence of Cystine (Cys-C) in all the circular bacteriocins known to date including pentocin KCA1, with the exception of subtilosin-A having three Cystine amino acid residues. The subtilosin A is a lantibiotic produced by *Bacillus subtilis* subsp. *subtilis* strain 168 (Kawulka et al., 2004).

Circular bacteriocins can be subdivided into two groups (i and ii) as suggested by van Belkum et al. (2011), based on sequence similarity and biochemical properties (Martin-Visscher et al., 2009). Comparatively, pentocin KCA1 based on sequence identity, appears to belong to group ii, which includes gassericin A, reutericn 6, acidocin B, and butyrivibriocin AR10. Acidocin B is considered a

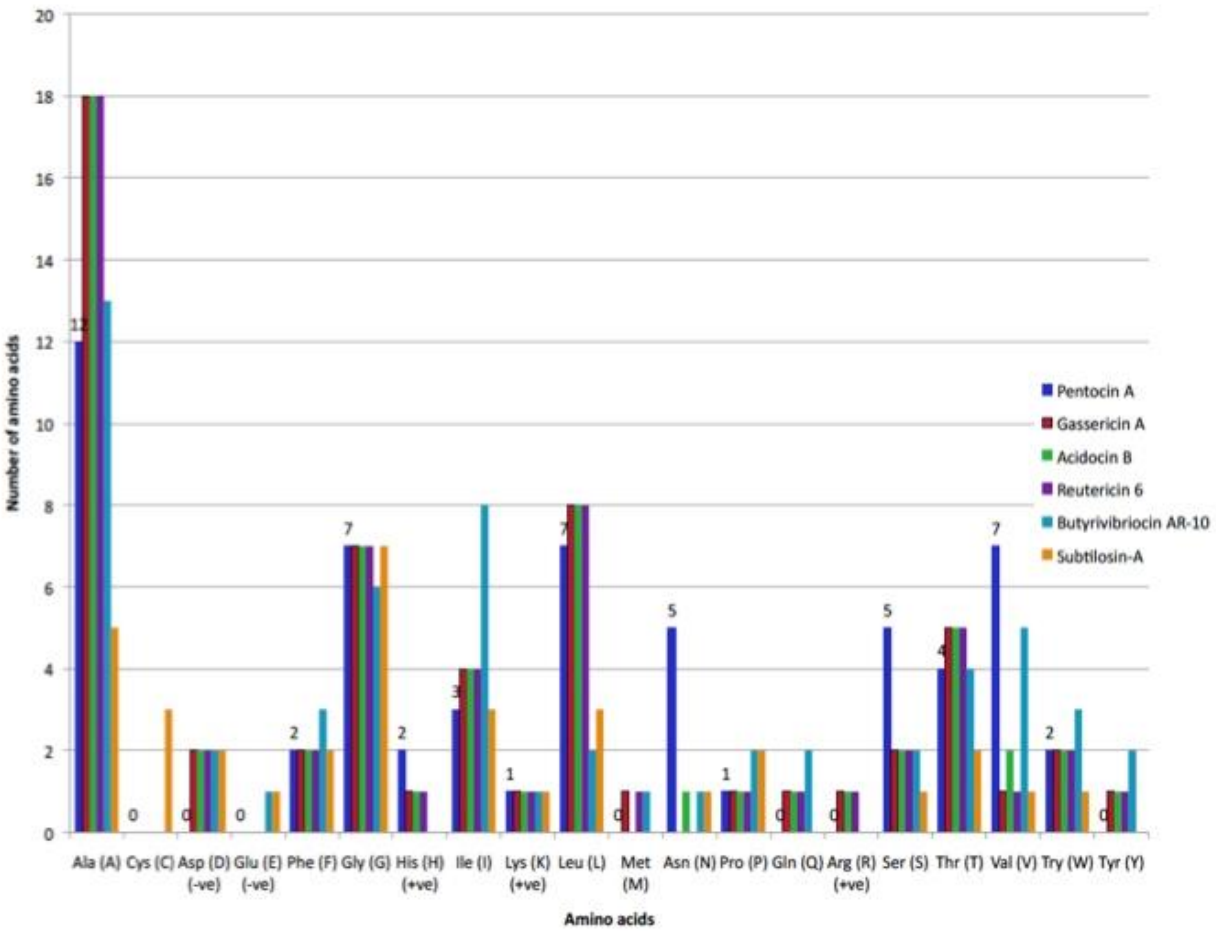


**Figure 6**: Frequency of amino acid utilization in the biosynthesis of group ii circular bacteriocins. The number displayed on top of the bar chart indicates the number of amino acid for pentocin KCA1.

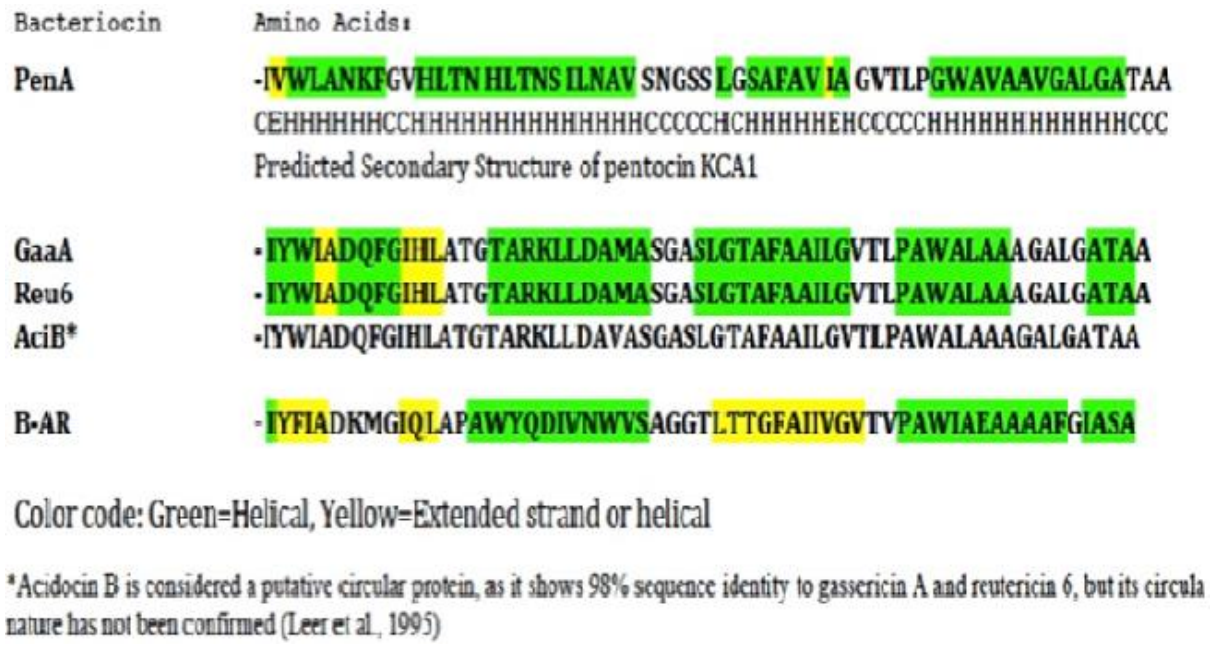


**Figure 7**: Secondary structure prediction as determined using JPRED and PSI servers, indicating four putative alpha helices present in pentocin KCA1.

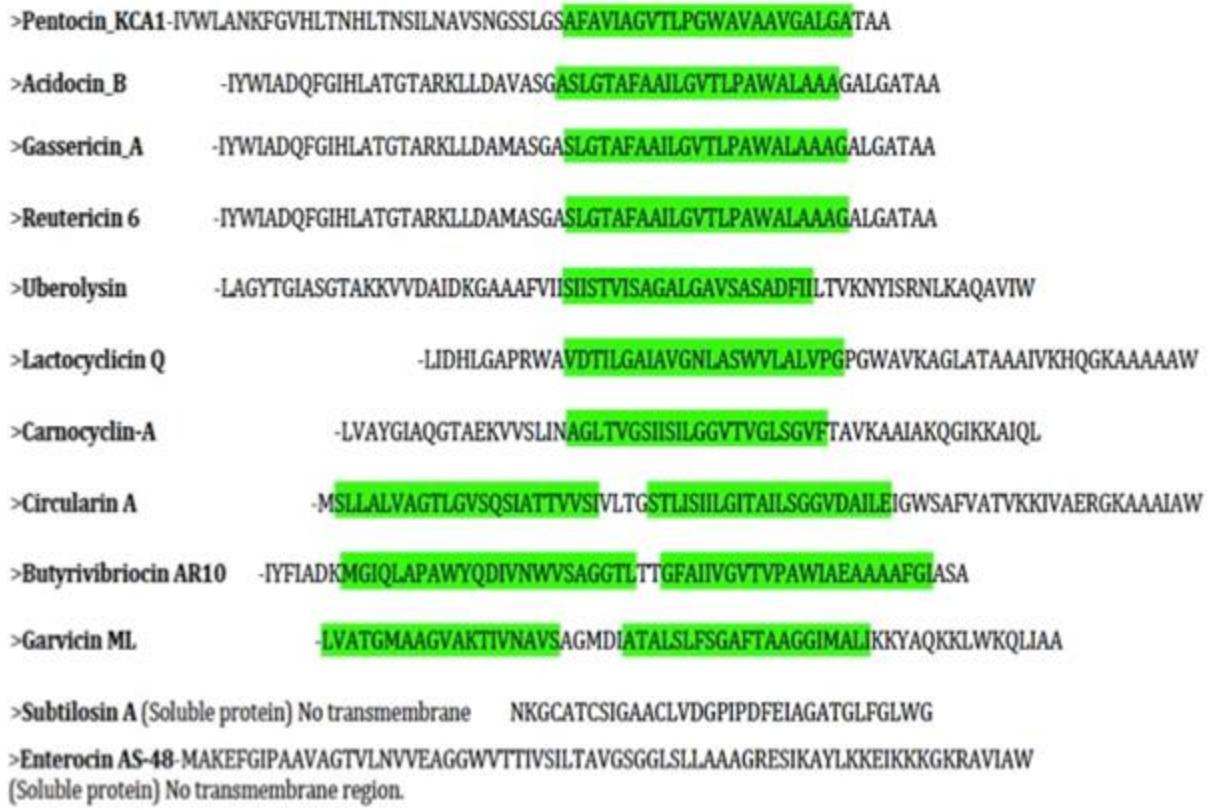


**Figure 8**: Mature circular bacteriocins showing transmembrane helices highlighted in green colour. Transmembrane regions in peptides were deduced using the SOSUI program (http://bp.nuap.nagoya-u.ac.jp/sosui/)

putative circular protein, as it shows 98% sequence identity to gassericin A and reutericin 6, but its circular nature has not been confirmed (Leer et al., 1995). However, while circular bacteriocins in group ii share significant sequence identity and more acidic residues, pentocin KCA1 has a high proportion of basic (positively charged) to acidic (negatively charged) amino acids in the ratio of 8:0, compared with gassericin A and acidocin B having a proportion of 3:2, thereby suggesting that pentocin KCA1 is a circular bacteriocin peptide with a strong basic property. The extent to which this basic property influences the microbiota composition is yet to be addressed. Future functional studies will provide such information in order to confirm the nature of the predicted pentocin KCA1 sequences and how the bacteriocin influences the biology of surrounding Gram-positive microbes. It will be interesting also to know how the differences in amino acid composition will impact on the functionality of the pentocin KCA1.

Both Nuclear Magnetic Resonance (NMR) and x-ray crystallography have solved the structure of AS-48 (González et al., 2000). However, the 3-D structure of the putative pentocin KCA1 will possibly be determined in subsequent studies.

**Conclusion**

*In silico* characterization of pentocin KCA1 based on bioinformatics tools shows differences in the organization of the gene clusters and varied sequence identity when compared with other bacteriocins. In addition, a high proportion of basic amino acid residues was observed in pentocin KCA1 relative to other circular bacteriocins in the same group but demonstrated a common structural motif similar to known circular bacteriocins consisting of four alpha-helical structures. The biological activities of the predicted circular pentocin KCA1 is yet to be determined, however, it is anticipated that the unique structural characteristics may impact on its putative medical and biotechnological applications. As the gene clusters of pentocin KCA1 are chromosomally encoded, its synthesis requires a coordinated expression of the genetic determinants leading to maturation and cleavage via the alanine C-terminus involving different transporters in the cascade. With the unique extra polar residues, it has the potential to attract the polypeptide to the negatively charged bacterial membranes, which may be one of the mechanisms for antimicrobial properties observed in *L. pentosus* KCA1 strain (Anukam & Reid, 2007; Lebeer et al., 2008). This postulation is based on the fact that cell-killing mechanism of circular bacteriocins depends on the possible ability to form membrane pores as observed in AS-48 (Sanchez-Barrena et al., 2003). Future directions will require pure isolation and purification of the pentocin KCA1 peptide, structural determination using Nuclear Magnetic Resonance (NMR) spectroscopy, Matrix-Assisted Laser Desorption Ionization- Time of Flight Mass Spectrometry (MALDI-TOF MS) and functional studies.

**Acknowledgement:**

The initial support from the Third World Academy of Sciences (TWAS) is highly appreciated

**REFERENCES**

Anukam KC, Macklaim JM, Gloor GB, Reid G, Boekhorst J, Renckens B, *et al* (2013). Genome sequence of *Lactobacillus pentosus* KCA1: vaginal isolate from a healthy premenopausal woman. *Plos ONE*. **8** (3): e59239.

Anukam KC, Reid G (2007). *Lactobacillus plantarum* and *Lactobacillus fermentum* with probiotic potentials isolated from the vagina of healthy Nigerian women. *Res. J. Microbiol.* **2** (1): 81ï 87

Bagos PG, Liakopoulos TD, Spyropoulos IC, Hamodrakas SJ (2004). PRED-TMBB: a web server for predicting the topology of beta-barrel outer membrane proteins. *Nucleic Acids Res*. **32**: W400-W404.

Borrero J, Brede DA, Skaugen M, Diep DB, Herranz C, Nes IF, *et al*. (2011). Characterization of Garvicin ML, a novel circular bacteriocin produced by Lactococcus garvieae DCC43, isolated from Mallard Ducks (*Anas platyrhynchos*). *Appl. Environ. Microbiol*. **77** (1):369-373

Carver T, Berriman M, Tivey A, Patel C, Bohme U, Barrell BG, et al. (2008). Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics*. **24**(23):2672-2676.

Chen H, Bjerknes M, Kumar R, Jay E (1994). Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of Escherichia coli mRNAs. *Nucleic Acid Res*. **22**(23): 4953-4957.

Cole C, Barber JD, Barton GJ (2008). The Jpred 3 secondary structure prediction server. *Nucleic Acids Res*. **36**: W197-W201.

Dobson A, Cotter PD, Ross RP, Hill C (2012). Bacteriocin production: a probiotic trait? *Appl. Environ. Microbiol*. **78** (1):1-6.

Gong X, Martin-Visscher LA, Nahirney D, Vederas JC, Duszyk M (2009). The circular bacteriocin, carnocyclin A, forms anion-selective channels in lipid bilayers. *Biochim. Biophys. Acta.* **1788**:1797ï 1803.

González C, Langdon GM, Bruix M, Gálvez A, Valdivia E, Maqueda M, *et al*. (2000)

Bacteriocin AS-48, a microbial cyclic polypeptide structurally and functionally related to mammalian NK-lysin. *Proc. Natl. Acad. Sci. USA.* **97** (21): 11221-11226.

Heng NCK, Tagg JR (2006). Whatôs in a name? Class distinction for bacteriocins. *Nat. Rev. Microbiol.* 4 (2): doi:10.1038/nrmicro1273-c1.

Hirokawa T, Boon-Chieng S, Mitaku S (1998). SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics*. **14**: 378ï 379.

Isono K, JMcIninch JD, Borodovsky M (1994). Characteristic features of the nucleotide sequences of yeast mitochondrial ribosomal protein genes as analyzed by computer program GeneMark. *DNA Res*. **1**: 263-269.

Jack RW, Tagg JR, Ray B (1995). Bacteriocins of Gram-positive bacteria. *Microbiol. Rev* . **59**:171ï 200.

Kalmokoff ML, Cyr TD, Hefford MA, Whitford MF, Teather RM (2003). Butyrivibriocin AR10, a new cyclic bacteriocin produced by the ruminal anaerobe Butyrivibrio fibrisolvens AR10: characterization of the gene and peptide. *Can. J. Microbiol.* **49**:763ï 773.

Kawai Y, Saito T, Kitazawa H, Itoh T (1998). Gassericin A; an uncommon cyclic bacteriocin produced by Lactobacillus gasseri LA39 linked at N- and C-terminal ends. *Biosci. Biotechnol. Biochem*. **62**:2438ï 2440.

Kawulka K, Sprules T, Diaper CM, Whittal RM, McKay RT, Mercier P, *et al*. (2004). Structure of subtilosin A, a cyclic antimicrobial peptide from Bacillus subtilis with unusual sulfur to R-carbon cross-links: formation and reduction of R-thio-R-amino acid derivatives. *Biochem.* **43**: 3385ï 3395.

Kemperman R, Kuipers A, Karsens H, Nauta A, Kuipers O, Kok J (2003). Identification and characterization of two novel clostridial bacteriocins, circularin A and closticin 574. *Appl. Environ. Microbiol.* **69**:1589ï 1597.

Lebeer S, Vanderleyden J, De K, Sigrid CJ (2008). Genes and Molecules of Lactobacilli Supporting Probiotic Action. *Microbiol. Mol. Biol. Reviews*. **72** (4): 728-764.

Leer RJ, van der Vossen JM, van Giezen M, van Noort JM, Pouwels PH (1995). Genetic analysis of acidocin B, a novel bacteriocin produced by Lactobacillus acidophilus. *Microbiol*. **141**:1629ï 1635.

Majeed H, Gillor O, Kerr B, Riley MA (2011). Competitive interactions in Escherichia coli populations: the role of bacteriocins. *ISME J*. **5**:71ï 81.

Maqueda M, Sa Ñchez-Hidalgo M, Fernа Ñndez M, Montalba Ñ-Lo Ñpez M, Valdivia E, Martа Ñnez-Bueno M (2008). Genetic features of circular bacteriocins produced by gram-positive bacteria. *FEMS. Microbiol. Rev*. **32**:2ï 22.

Martin-Visscher LA, van Belkum MJ, Garneau-Tsodikova S, Whittal RM, Zheng J, McMullen LM, *et al*. (2008). Isolation and characterization of carnocyclin A, a novel circular bacteriocin produced by Carnobacterium maltaromaticum UAL307. *Appl. Environ. Microbiol.* **74**: 4756ï 4763.

Martin-Visscher LA, Gong X, Duszyk M, Vederas JC (2009). The three-dimensional structure of carnocyclin A reveals that many circular bacteriocins share a common structural motif. *J. Biol. Chem*. **284**: 28674ï 28681.

McGuffin LJ, Bryson K, Jones DT (2000). The PSIPRED protein structure prediction server. *Bioinformatics*. **16**: 404-405.

Nagy Z, Szabo M, Chandler M, Olasz F (2004). Analysis of the N-terminal DNA binding domain of the IS30 transposase. *Molecular Microbiol*. **54** (2): 478-488.

Nes IF, Diep DB, Holo H (2007). Bacteriocin diversity in Streptococcus and Enterococcus. *J. Bacteriol*. **189**:1189ï 1198.

Nissen-Meyer J, Nes IF (1997). Ribosomally synthesized antimicrobial peptides: their function, structure, biogenesis, and mechanism of action. *Arch. Microbiol*. **167**: 67ï 77.

Roberts TM, Bikel I, Yocum RR, Livinstone D, Ptashne M (1979). Synthesis of simian virus 40 t antigen in Escherichia coli. *Proc. Natl. Acad. Sci. USA*. **76**: 5596-5600.

Salzberg S, Delcher A, Kasif S, White O (1998). Microbial gene identification using interpolated markov models. *Nucleic Acids Res*. **26**: 544-548.

Sanchez-Barrena MJ, Martïnez-Ripoll M, Gálvez A, Valdivia E, Maqueda M, Cruz V, *et al*. (2003). Structure of bacteriocin AS-48: from soluble state to membrane bound state. *J. Mol. Biol.* **334**:541ï 549.

Sang Y, Blecha F (2008). Antimicrobial peptides and bacteriocins: alternatives to traditional antibiotics. *Anim. Health. Res Rev*. **9**:227ï 235.

Sawa N, Zendo T, Kiyofuji J, Himeno K, Nakayama J, Sonomoto K (2009). Identification and characterization of lactocyclicin Q, a novel cyclic bacteriocin produced by Lactococcus sp. strain QU12. *Appl. Environ. Microbiol*. **75**:1552ï 1558.

Shine J, Dalgarno L (1974). The 3ô-Terminal sequence of Escherichia coli 16S ribosomal RNA: complementary to nonsense triplets and ribosomal binding sites. *Proc. Natl. Acad. Sci. USA*. **71**:1342-1346.

Sit CS, Vederas JC (2008). Approaches to the discovery of new antibacterial agents based on bacteriocins. *Biochem. Cell. Biol*. **86**: 116ï 123.

Toba T, Samant SK, Yoshioka E, Itoh T (1991). Reutericin 6, a new bacteriocin produced by *Lactobacillus reuteri* LA6. *Lett. Appl. Microbiol*. **13**: 281ï 286.

Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ (2009). Jalview version 2: a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. **25**: 1189-1191.

van Belkum MJ, Martin-Visscher LA, Vederas JC (2011). Structure and genetics of circular bacteriocins. *Trends Microbiol*. **19** (8): 411-418.

Wirawan RE, Swanson KM, Kleffmann T, Jack RW, Tagg JR (2007). Uberolysin: a novel cyclic bacteriocin produced by Streptococcus uberis. Microbiol. **153**: 1619ï 1630.

Zheng G, Yan LZ, Vederas JC, Zuber P (1999). Genes of the sbo-alb locus of *Bacillus subtilis* are required for production of the antilisterial bacteriocin subtilosin. *J. Bacteriol.* **181**:7346ï 7355.