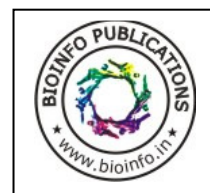


Phylogenetic signatures of functional conservedness in lantibiotics- an *in-silico* regulomics study

Protip Basu, Hirak Jyoti Chakraborty, Sayak Ganguli*, Paushali Roy and Abhijit Datta

DBT-Centre for Bioinformatics, Presidency College, Kolkata, sayakbif@yahoo.com, Ph: +919433081029



Abstract- The name Lantibiotics was introduced in 1988 as an abbreviation for "Lanthionine-containing peptide antibiotics". In spite of this naming, Lantibiotics are not classed as antibiotics. The first structures of these antimicrobial agents were produced by pioneering work by Gross and Morell in the late sixties and early seventies, thus marking the formal introduction of Lantibiotics. Since then Lantibiotics such as Nisin have been for food preservation and have yet to encounter significant bacterial resistance. These attributes of lantibiotics have led to more detailed research into them.

Introduction

A comprehensive structural analysis of the lantibiotics was performed. Phosphorylation sites were detected for the sequences. Conserved domains were also analyzed for phylogenetic signatures (PHYLIP) and molecular modeling of representative members of the major groups of the lantibiotics were performed using Modeler 9.2. Several plant homologues were detected sharing sequence identity of over 85% with the lantibiotics mostly belonging to the bacteriocin superfamily. Most of the plant homologues identified were transcription factors induced by phytohormones. Phosphorylation sites detected in the sequences show a bias for serine phosphorylation in most of the conserved regions. Phylogenetic analysis revealed that blocks of lantibiotics are phylogenetically related to certain members of the mammalian immune response pathways as they share a stretch of conserved region in them. Apart from them certain sequences had metallothionein sequences as their sister groups indicative of a stress related functional conservedness.

Background

Lantibiotics are highly modified peptides from Gram-positive bacteria. They contain alpha, beta-unsaturated amino acids (dehydroalanine and dehydrobutyrine) and lanthionine or 3-methylanthionine rings. There are 2 types of lantibiotics:

Type A are strongly cationic and bactericidal - nisin, subtilin and Pep5 inhibit the growth of Gram-positive bacteria, probably by voltage-dependent pore formation in the cytoplasmic membrane, resulting in cellular efflux of electrolytes, amino acids and ATP.

Type B lantibiotics possess at most one positive charge and are not bactericidal.

Ribosomally synthesized peptide bacteriocins from Gram-positive bacteria can be subdivided into two major classes. Bacteriocins of class I are characterized by having modified amino acid residues (e.g., lantibiotics) and bacteriocins of

class II are characterized by not possessing modified amino acid residues (e.g., small heat stable non-lantibiotics). These two classes are the most studied due to their abundance and their potential use for industrial applications.

They are similar in size (approx 20-60 amino acids), mostly cationic, and possess a hydrophobic domain and/or amphiphilic region, which may relate to their action on membranes. The defining characteristic of lantibiotics is that they contain the unusual amino acid lanthionine (or β -methyllanthionine). Generally, type A lantibiotics are characterized by being strongly cationic (with 2 to 7 net positive charge), having molecular masses more than 2 kDa, and having rigid ring conformations separated by areas of flexibility [2, 3]. It is currently believed that their primary bactericidal activity is mediated through the formation of voltage-dependent membrane channels.

Contrastingly, type B lantibiotics are characterized by being neutral or slightly anionic (with a 0 to -1 net charge), having molecular masses of less than 2 kDa, and having a more compact globular structure. Their primary function is mediated through the inhibition of essential enzymes. There is a unique group of lantibiotics called the two-peptide bacteriocins. In this system, two genes encoding each of the peptides are situated next to one another and both of the peptides are needed for bactericidal activity. Each peptide is either inactive or only slightly active when tested individually. The mode of bactericidal activity is believed to be the same as for type A lantibiotics. In some cases, they have been grouped along with type A lantibiotics [7, 9]. Lantibiotics are produced by a large number of Gram positive bacteria such as *Streptococcus* and *Streptomyces* to attack other gram positive bacteria and as such they are considered a member of the bacteriocins. Lantibiotics are well studied because of the commercial use of these bacteria in the food industry for making dairy products such as cheese [12, 14].

Bacteriocins are classified according to their extent of posttranslational modification. The lantibiotics are a class of more extensively modified bacteriocins, also called Class I. Bacteriocins for which disulfide bonds are the only modification to the peptide are Class II bacteriocins. Most bacteriocins are biologically active single-chain peptides. Some are only active as partners with a second peptide. Nisin and epidermin are members of a family of lantibiotics that bind to a cell wall precursor lipid component of target bacteria and disrupt cell wall production. The duramycin families of lantibiotics bind phosphoethanolamine in the membranes of its target cells and seem to disrupt several physiological functions [13, 14, 15].

Type A Lantibiotics are long flexible molecules – e.g. nisin, subtilin, epidermin, *etc.* Subgroup AI includes Mutacin II; subgroup AII includes Mutacin I & III. Type B Lantibiotics are globular – e.g. mersacidin, actagardine, cinnamycin, *etc.*

They are synthesized with a leader polypeptide sequence which is only removed during the transport of the molecule out of the synthesizing cell. They are synthesized by ribosomes, which distinguishes them from antibiotics which are synthesized by enzymes. Lantibiotics show substantial specificity for some components (e.g. lipid II) of bacterial cell membranes especially of Gram positive bacteria [6, 8]. Type A Lantibiotics kill rapidly by pore formation, type B inhibit peptidoglycan biosynthesis. They are active in very low concentrations.

Lantibiotics are produced by Gram-positive bacteria and show strong antimicrobial action towards a wide range of other Gram-positive bacteria. As such they have become attractive candidates for use in food preservation (by inhibiting pathogens that cause food spoilage) and the pharmaceutical industry (to prevent or fight infections in humans or animals) [10, 14].

Methodology

Lantibiotic sequences were obtained from public repositories – NCBI and Swiss – Prot and were analyzed primarily using BLASTn and BLASTp to generate the nearest homologues. Out of the total of about 500 sequences that were obtained using BLAST, 350 sequences were selected based on *e* - values.

These sequences were then grouped together based on the conserved domain signatures that were identified using CDART.

Each representative members were then subjected to BLOCKS which revealed a further 600 sequences having representative signature motifs.

All these sequences were aligned to identify the conserved residues. The PHYLIP 3.68 package was used for the construction of the phylogenetic trees and bootstrapping was performed.

Discussion

Homology search using BLAST shows that there is indeed homology of the lantibiotic sequences with the plant sequences and most of the plant homologues of Lantibiotics correspond to the stress response proteins of the plants as revealed by the maximum homology of the lantibiotic sequences with Arabidopsis chromosome 5 which has been reported to contain the genes for stress response and plant immune system.

Screening of proteins for Phosphorylation sites

The maximum frequency of phosphorylation was found at sites on proteins containing serine residues. This proves that regulation of lantibiotics is a highly specialized process, due to the fact that, serine residues are an important part, of most active sites of functional proteins, and phosphorylation at those sites is an evidence of highly advanced evolutionary constraint. (Fig 1)

Analysis of Conserved Domains

Analysis of conserved domains revealed the presence of the Gallidermin superfamily of proteins as the predominant class of conserved domain. Another group that had only a single instance was the Bacteriocin-II superfamily.

Multiple Alignment and Phylogenetic analysis

There are a number of lantibiotic peptides that show high similarity with the lantibiotic lactacin 481. These include lactacin J46, mutacin II, butyrivibriocin, ruminococcin, streptococcin SA-F22, nukacin, variacin and the salivaricins [11]. Such is the number of peptides/variants within this group and the variety of residues that differ upon alignment that it becomes obvious which amino acids are highly conserved i.e. Gly5, Thr9, Ser/Thr11, Glu/Asp13, Cys14, Asn17, Ser/Thr18, Phe23, Cys25 and Cys26 (as located in lactacin 481) while the others differ with high frequency.

The existence of two salivaricin variants, A and A1, which differ by only two amino acids, is particularly noteworthy, and there are also instances where two or more genes that are identical with respect to the region encoding mature structural peptides are present in a single strain. This is the case with streptococcin SA-F22 which possesses 2 such genes located adjacently while ruminococcin producers have three such structural genes i.e. *rumA1*, *rumA2* and *rumA3* [1, 5].

The predominant groups of BLOCKS that were found were Gallidermin signature family (17 query sequences), Lantibiotic domain family (19 query sequences), Nisin signature family (16 query sequences), Echinoidea (sea urchin) metallothionein family (6 query sequences),

CD59 antigen family (1 query sequence), Bacteriocin class II family (1 query sequence), tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase family (1 query sequence), and Squalene Epoxidase family (1 query sequence). [Fig 2 – 4].

Homology modeling of Lantibiotics

Several homology models were prepared using Modeler 9.2 [16]; these models were then validated using the Molprobit analyzer. The hinge region in the mature peptides of most classes was found to be highly conserved. The hinge region has been imparted importance as trypsin cleavage has been found to occur mostly in this region [4] (Fig 5).

Conclusion

Most of the plant homologues identified were transcription factors induced by phytohormones. Phosphorylation sites detected in the sequences show a bias for serine phosphorylation in most of the conserved regions. Tyrosine phosphorylation is mostly absent.

Phylogenetic analysis revealed that blocks of lantibiotics are phylogenetically related to certain members of the mammalian immune response pathways as they share a stretch of conserved region in them. Apart from them certain sequences had metallothionein sequences as their sister groups indicative of a stress related functional conservedness.

The analysis reveals that lantibiotics are probably orthologues to plant phytohormones activated transcription factors and also some mammalian immune response proteins. The overall functional similarity between all these members indicates a general response to stress. Evolutionary implication of such a finding is that probably all these groups of proteins are orthologues and have evolved from the same group of a divergent family.

References

- [1] Hynes W.L., Friend V.L., Ferretti J.J., (1994) *Appl. Environ. Microbiol.*, 60: 4207-4209.
- [2] Brotz H., Sahl H.G., (2000), *Journal of Antimicrobial Chemotherapy*, 46: 1-6.
- [3] Van Kraaij C., de Vos W.M., Siezen R.J. and Kuipers O.P. (1999) *Nat. Prod. Rep.*, 16: 575–587.
- [4] Cotter P.D., Hill C., Ross R.P. (2005) *Current Protein and Peptide Science*, 6: 61-75.
- [5] Netz D.J.A., Sahl H.G., Marcolino R., Nascimento J.D., de Oliveira S.S., Soares M.B. and Bastos M.D.D. (2001) *Journal of Molecular Biology*, 311: 939-949.
- [6] Nes I.F., Tagg J.R. (1996) *International Journal of General and Molecular Microbiology*, 69: 89-97.
- [7] Martin N.I., Breukink E. (2007) *Future Microbiol.*, 2(5): 513-525.
- [8] Hyink O., Wescombe P.A., Upton M., Ragland N., Burton J.P. and Tagg J.R. (2007) *Appl. Environ. Microbiol.* 73(4): 1107–1113.
- [9] Sahl H.G., Bierbaum G. (1998), *Annual Review of Microbiology*, 52: 41-79.
- [10] Sahl H.G., Bierbaum G. (2008) *Microbe*, 3(10):467-473.
- [11] Savadogo A., Ouattara C.A.T., Bassole I.H., Traore S.A. (2006) *African Journal of Biotechnology*, 5(9): 678-683,
- [12] Dimov S., Ivanova P., Harizanova N. (2005) *Biotechnol. & Biotechnol. Eq. General & Applied Genetics*, 19(3): 4-10.
- [13] Hasper H.E., Kramer N.E., Smith J.L., Hillman J.D., Zachariah C., Kuipers O.P., de Kruijff B., Breukink E. (2006) *Science*, 313: 1636-1637.
- [14] Boakes S., Wadman S. (2008) *IPT*, 27: 22-24.
- [15] Oscariz J.C., Pisabarro A.G. (2001) *Int. Microbiology*, 4: 13-19.
- [16] Sali A., Blundell T.L., (1993) *J. Mol. Biol.* 234(3):779-815.

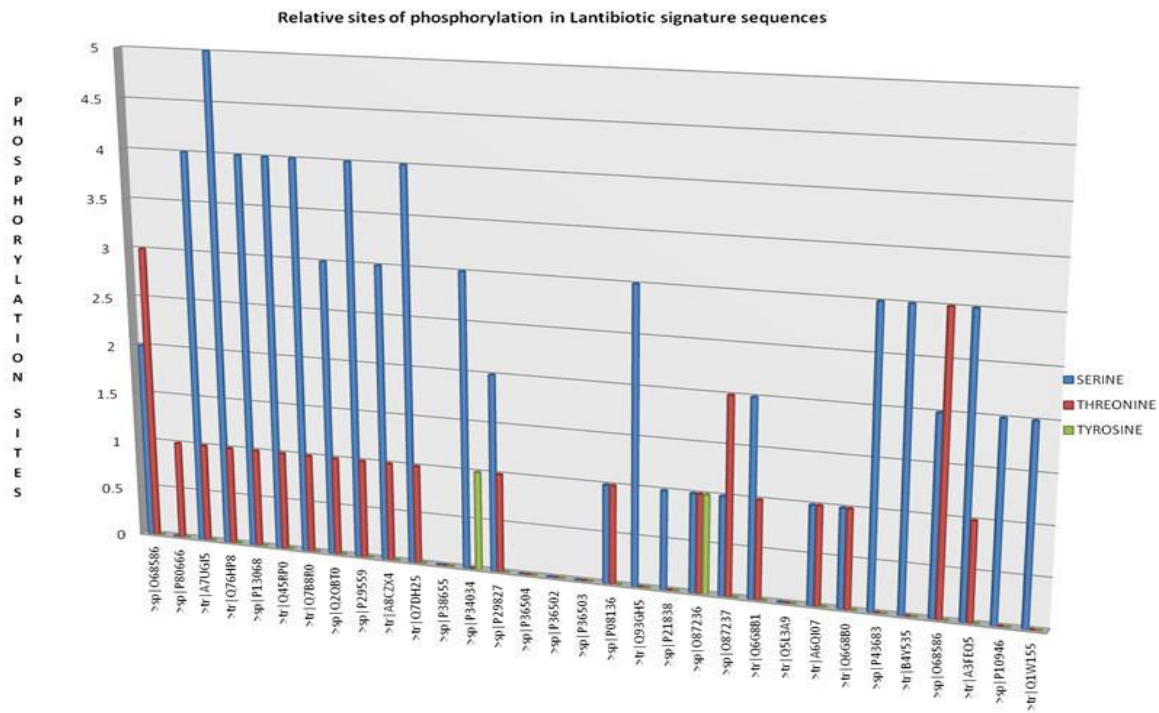


Fig. 1-
Fig 2: Gallidermin Signature domain phylogeny



Fig 3: Nisin signature domain phylogeny

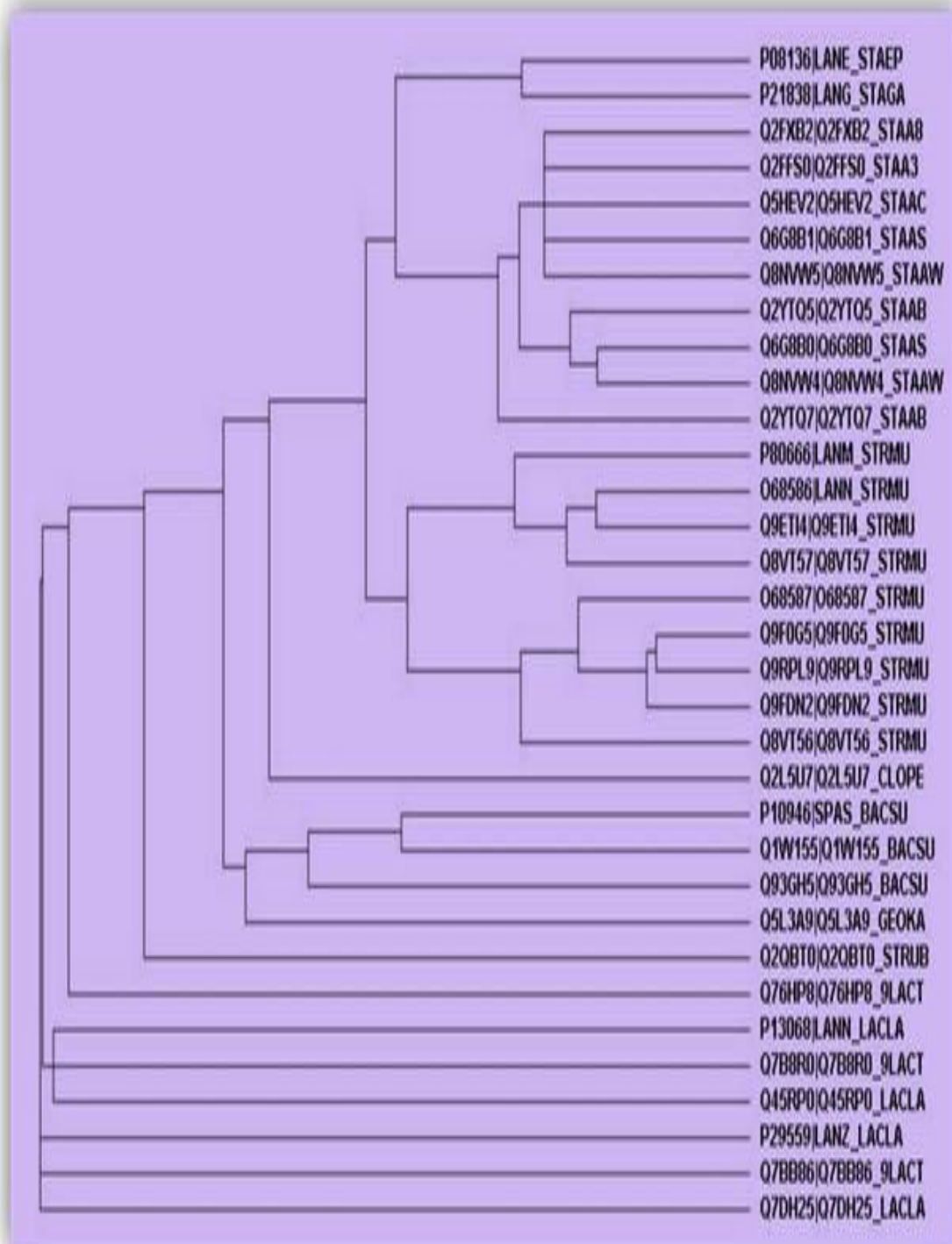


Fig 4: General Lantibiotic associated peptides, tRNA methyltransferase signature member's phylogeny

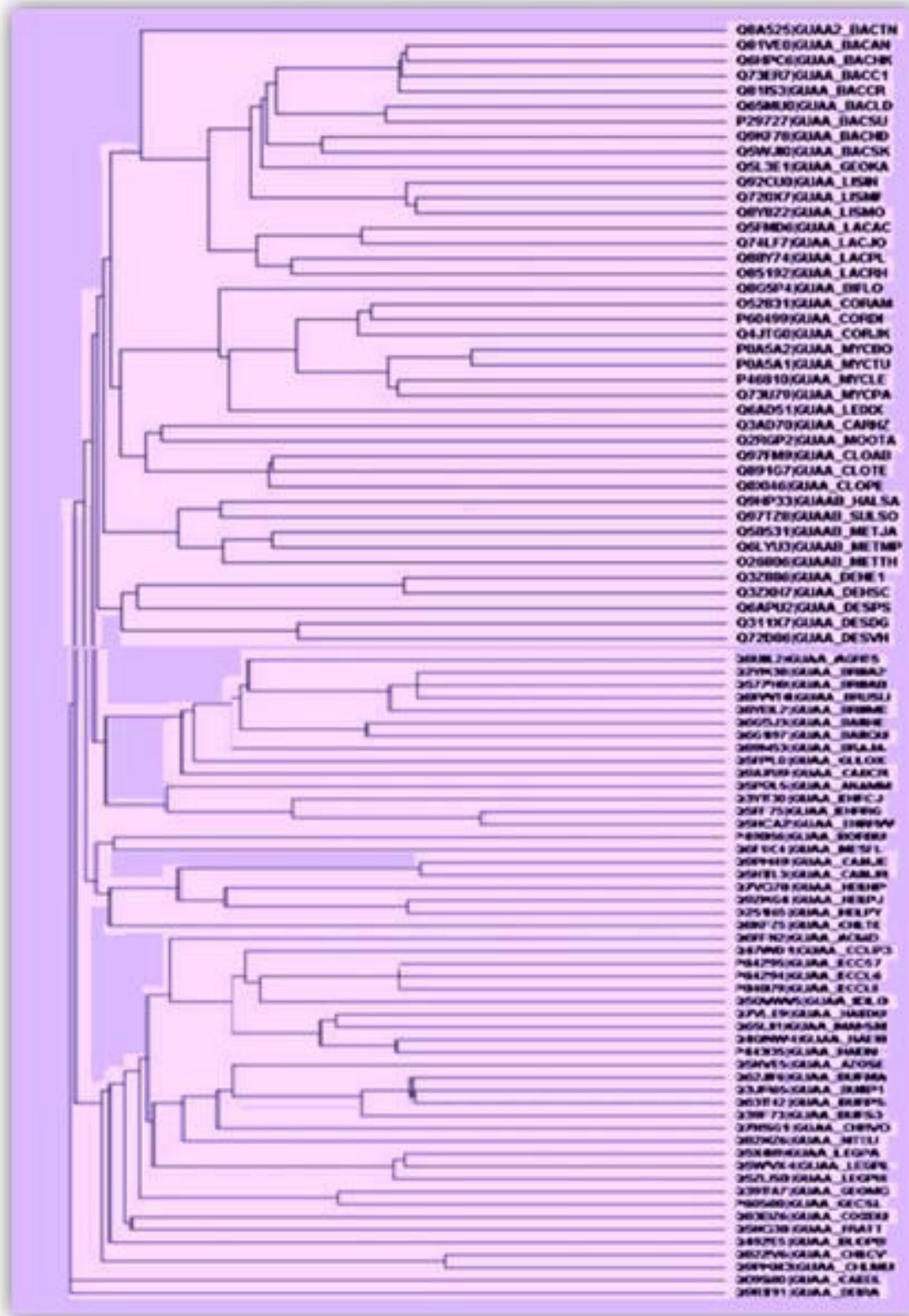
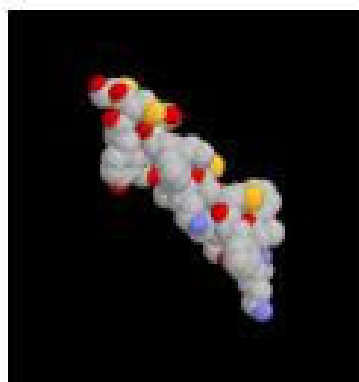
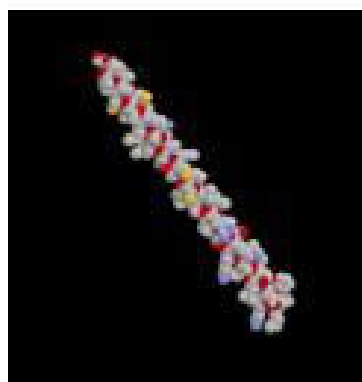


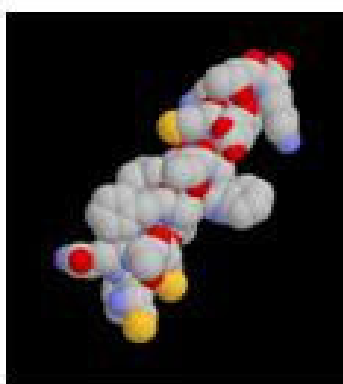
Fig 5: Homology models of major lantibiotics



MUTACIN B



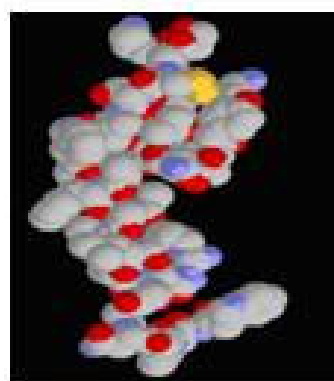
NISIN F



ANCOVENIN



DURAMICIN



LEUCOCIN