ВІСНИК ЛЬВІВ. УН-ТУ
Серія біологічна. 2003. Вип. 32. С. 84-91

VISNYK OF L'VIV UNIV.
Biology Series. 2003. Is. 32. P. 84-91

УДК.238.218.60

SEQUENCING AND ANALYSIS OF PUTATIVE 3^{D}-4^{H} RING CYCLASE GENE *lndF* OF *STREPTOMYCES GLOBISPORUS* 1912 LANDOMYCIN E BIOSYNTHETIC GENE CLUSTER

B. Ostash*, Yu. Rebets*, M. Samborskyy*, J. A. Salas, V. Fedorenko***

**Ivan Franko National University of L'viv*
Hrushevskyy st.4, L'viv 79005, Ukraine
e-mail: v_fedorenko@franko.lviv.ua
*** Universidad de Oviedo, Departamento de Biologia Funcional*
Julian Claveria s/n 33006, Oviedo, Spain
e-mail: jasf@sauron.quimica.uniovi.es

DNA fragment of landomycin E biosynthesis gene cluster 700 bp in size has been completely sequenced. 3´-end of *lndE* (oxygenase) was identified, 5´-end of *lndA* (ketosynthase) and entire ORF for previously not sequenced *lanF* homologue, *lndF*. Analysis of *lndF* putative translation product revealed that it is highly conservative in comparison with other known cyclases from antibiotic biosynthesis gene clusters. Unlike urdamycin, jadomycin biosynthetic clusters, *lndF* and *lndA* are uncoupled, as well as genes *lanF* and *lanA*. Genes *lndF* and *lndA* are not preceded by direct or inverted repeats, putative sites for binding of transcriptional activator LndI. In contrast, *lanF* is flanked at it's 5´-end by three direct repeats, possible target for regulatory protein LanI.

Key words: *Streptomyces*, angucyclines, cyclases, landomycin E

Last decade was marked by tremendous progress in understanding genetics and enzymology of polyketide framework synthesis. Efficient systems for heterologous gene expression and cell-free synthesis of polyketides were developed, many clusters for polyketide compounds production were cloned and studied via sequence analysis, gene knockouts, heterologous expression experiments, various biochemical studies on purified proteins [3,11,16]. Recent advances approached us to the stage, when obtaining of novel polyketide compounds can be realized through simple "plug'n'play" manipulations, where artificially designed gene sets direct the production of compounds with predicted structure [11]. On other hand, identification of novel unusual polyketide antibiotics rase new questions about basic steps leading to aglycon formation. One of such challenges was represented by angucycline group of antibiotics [5,10]. An unique trait of angucyclines is specific character of 3rd ring cyclization (Fig.1). Genetic control of this biosynthetic step is obscure [6,7], thus identification of plausible genes governing this cyclization step is of great interest in view of "angu"-cyclases potential use in novel bioactive polyketides combinatorial synthesis.

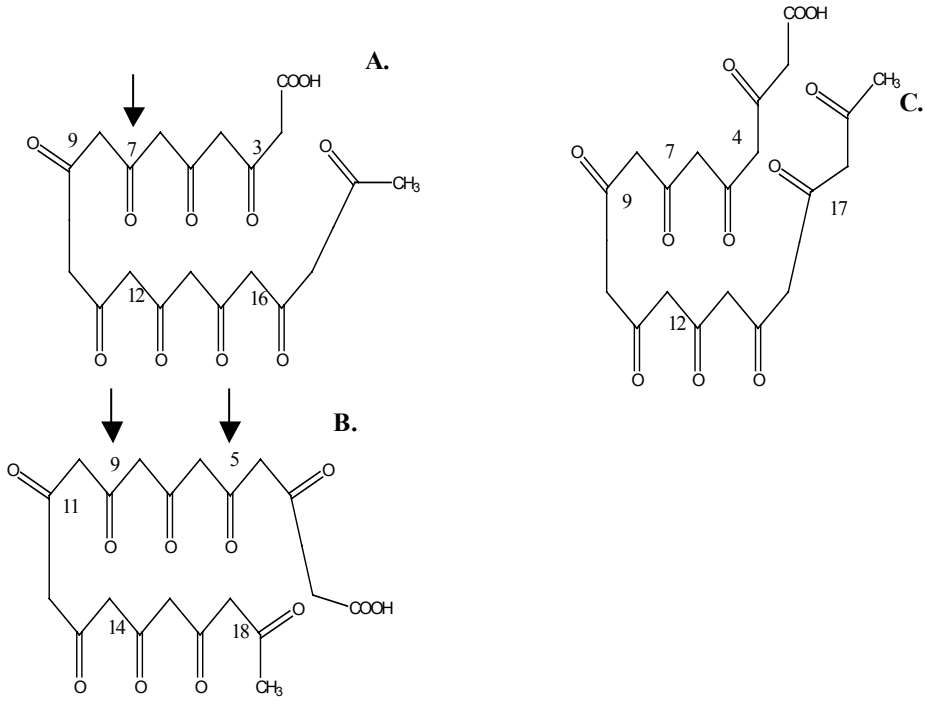

Fig.1. Folding pattern of tetracyclines (A), tetracenomycins (B) and angucyclines (C). Arrows indicate ketogroups taking part in intramolecular aldol condensation [10].

S. globisporus 1912 gene cluster for antitumor angucycline landomycin E [2] biosynthesis (*lnd*-cluster) has been cloned, but gene encoding 3rd ring cyclase has not been identified [1]. In this work we set out to localize and sequence plausible cyclase gene *lndF* for landomycin E third ring formation.

1.8kb KpnI-fragment of *lnd*-cluster spanning from *lndE* (oxygenase) to *lndA* (α-subunit of ketosynthase) gene was used for sequencing (Fig.2). Nucleotide sequence was determined by dideoxynucleotide chain termination method on ALF Express sequencer (Pharmacia). The sequence was analyzed with the GCG sequence analysis software package (version 8; Genetics Computer Group, Madison, Wis.). BLAST X searches were performed to find *lndF* homologues. Phylogenetic tree was built with help of CLUSTAL W program.

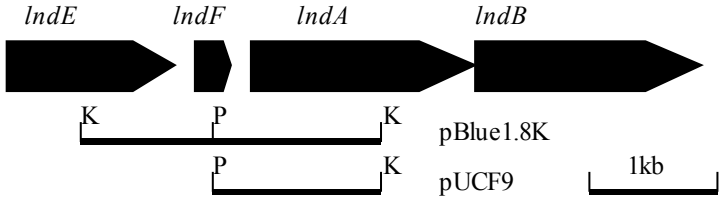

Fig.2. Fragment of *S. globisporus* 1912 *lnd*-cluster. *lndE* – oxygenase, *lndF* – 3rd-4th ring cyclase, *lndA* – α-subunit od ketosynthase, *lndB* – β-subunit of ketosynthase [1]. At the bottom two subclones (pBlue1.8K and pUCF9) used for sequencing and scale are shown. Abbreviations: K - *Kpn*I, P – *Pst*I.

Preliminary analysis of gene clusters for angucycline biosynthesis revealed presence of one conservative gene which invariantly is localized upstream of ketosynthase gene [5,6,10,17,18]. Functional characterization of such a gene from *S. venezulae* ISP5230 jadomycin B biosynthesis gene cluster has proved it's involvement in angucyclinone formation [6,20]. Thus it was decided to sequence the region of *lnd*-cluster corresponding to *lanE-lanF-lanA* region of *S. cyanogenus* S136 landomycin A gene biosynthetic cluster, very similar in general cluster organization to that of *lnd*- cluster [18]. Sequencing revealed an ORF (named *lndF*, by analogy to *lanF*) 330nt in size. *lndF* begins with ATG codon which is preceeded by putative RBS (GAGG) 7nt upstream of *lndF*. First in-frame stop codon (TAA) in *lndF* is recognized 327nt downstream of start codon, thus specifying 109 aa protein (Fig.3).

```
     TACGCCAAGC TCGAAATTAA CCCTCACTAA AGGGAACAAA AGCTGGTACC
51   TCCGCTCTGC TGATCCGCCC CGACGGTCAT GTCGCCTGGG CCGCTCCCGG
101  CAGCCACCAC GACCTGCCCA TGGCTCTGAC CCGCTGGTTC GGCCGCCCGC
151  CGGTCGGACG CCGGGTCTGA TCCCGCAGCG CCCCTACGGA ACGGAAGAGG
                       lndF ⟶
201  GAGAACCATG CACAGCACAC TGATCGTCGC CCGGATGGAC CCCGCGTCGA
251  GCATCGACGT GGCGGAACTC TTCGGCGAGT TCGACCGCAC CGAGATGCCC
301  CACCGCATGG GCACCAGGCG TCGGCAGCTC TTCTCGTATC GCGGACTGTA
351  CTTCCACCTG CAGGACTTCG ACTCCGACAA CGGCGGAGAG CTGATCGAGG
401  AGGCCAAGAG CGACCCGCGC TTCGCGGCGA TCAGCCAGGA CCTGAAGCCC
451  TTCATCGAAG CGTACGACCC GGCCACCTGG CGTTCCCCGG CCGATGCGAT
501  GGCCACCCGC TTCTACAACT GGACGACGTC GTCATAAGGC CGTCCACCGA
                                   lndA ⟶
551  CACGAGGGAG GCCCCAGTGG GGCGCCGGGT AGTAATCACT GGAATCGGGG
601  TGCTGGCGCC GGGCGGTGTC GGCACCAAGA ACTTCTGGGA AGCTGCTGAG
651  CGAAGGGCCG TACGGCGACG CCGGGGGATC ACCTTCTTCG ATCCGTCGCC
```

Fig. 3. Nucleotide sequence of 700 bp fragment containing *lndF* gene. The gene possess high GC bias typical for *Streptomyces*. Putative RBS sites for *lndF* and *lndA* are underlined, translation start and stop codons are typed in bold.

Probable product of *lndF* translation shows 89% of identity and 95% of similarity to putative 3rd ring cyclase LanF from *S. cyanogenus* S136 *lan*-cluster [18], 83% and 90% to angucyclinone-forming cyclase JadI from *S. venezulae* ISP5230 jadomycin B biosynthetic gene cluster [6], 85% and 89% to PgaF, presumed 3rd ring cyclase from silent angucycline-type gene biosynthetic cluster in rubromycin-producing strain strain *S.* sp. PGA64, 80% and 87% to possible 3rd ring cyclase UrdF from *S. fradiae* Tu2717 *urd*-cluster [10], 78% and 82% to probable 3rd ring cyclase SimA4 from *S. antibioticus* Tu6040 simocyclinone biosynthesis gene cluster [17], 69% and 83% to Aur1C, putative cyclase involved in biosynthesis of a proposed polyketide auricin in *S. aureofaciens* CCM3239 , 48% and 62% to WhiE-ORFVII protein, cyclase involved in *S. coelicolor* spore pigment biosynthesis [15], 38% and 56% to TcmI, involved in D-ring formation during tetracenomycin C biosynthesis in *S. glaucescens* [12]. Multiple sequence alignment of LndF with the best BLAST hits is presented on fig.4.

```
LanF  MHSTLIVAKM DPASSIDVAK LFGDFDRTEM PHRMGTRRRQ LFSYRGLYFH
LndF  MHSTLIVARM DPASSIDVAE LFGEFDRTEM PHRMGTRRRQ LFSYRGLYFH
PgaF  MHSTLIVARM DVESSAQVAE LFGDFDRTEM PHRMGTRRRQ LFSYNGLYFH
JadI  MHSTLIVARM EPGSSTDVAK LFAEFDASEM PHLMGTRRRQ LFSYRGLYFH
UrdF  MHSTLIVARM EPRSAEDVAR LFSEFDGTDM PHRMGTRRRQ LFSYRGLYFH
      **********    *  ***  **  **   * ** ******* **** *****

LanF  LQDFDENNGG ELIEEAKTDP RFVAISQDLK PFIQAYDPET WRSPADAMAT
LndF  LQDFDSDNGG ELIEEAKSDP RFAAISQDLK PFIEAYDPAT WRSPADAMAT
PgaF  LQDFDGDNGG ELIEEAKSDS RFIRISEDLK PYIEAYDPAT WRSPADAMAK
JadI  LQDFDADNGG ELIERAKTDP RFVGISEDLK PFIEAYDPAT WRSPADAMAT
UrdF  LQDFDSEDGG ERIEAAKTDQ RFIGISEDLK PFIAAYDPDT WRSPADAMAQ
      *****   ** * ** ** *  **  ** *** *** **** * *********
LanF  RFYDWTASS
LndF  RFYNWTTSS
PgaF  RFYDWTATQ
JadI  RFYNWEANA
UrdF  RFYHWTAL.
      *** **
```

Fig.4. Multiple sequence alignment of LndF with putative polyketide cyclases from different clusters for angucycline biosynthesis (see the text above). Aminoacid residues identical in all proteins are marked with asterisks below the sequences.

So, all proteins putatively involved in 3rd-4th ring cyclization of angucyclinones are shown to form very conservative group and their genetic determinants are localized upstream of ketosynthase gene. Notably, "angu"-cyclases share rather high level of homology (more than 50% of similarity) with cyclases known to catalyze last rings formation during biosynthesis of non-angucyclic polyketides (TcmI, WhiE-ORFVII, ElmI). This hints at existence of some conserved domains in all cyclases.

Another interesting trait concerning genes for 3rd-4th ring cyclases is about their genetic organization. Although genes by themselves are very similar, they fall into two groups: those overlapped with downstream ketosynthase gene (*urdF*, *jadI*, *ovdC*, *simA4*) and those that are not (*pgaF*, *lanF*, *lndF*). Particularly, *lanF* and *lanA* genes from *S. cyanogenus lan*-cluster are separated by 35nt region [18], *lndF* and *lndA* are separated by 29nt (fig.3). In other cases stop-codon of cyclase gene is overlapped with start codon of ketosynthase gene. Since first cloned and sequenced clusters for angucycline biosynthesis were those for urdamycin and jadomycin biosynthesis [5,6], it was hypothesized, that overlapping of genes for cyclase and α-subunit of ketosynthase has special biological meaning. Transcriptional and translational coupling of genes for α and β PKS subunits facilitate production of respective proteins in equimolar quantities. KS_α and KS_β form heterodimer with ratio 1:1, that is why gene coupling conceivably is kind of regulatory mechanism preventing from undesired excessive production of synthase proteins (potential competitive inhibitors of PKS activity in monomer forms) [3,14]. In several cases where genes for α and β PKS subunits are not coupled (for example, gene cluster for frenolicin biosynthesis), respective PKS complex is capable of producing polyketides with differ-

ent chain length [10]. Minimal PKS for angucyclinone biosynthesis represent new stage of ketosynthase complex organization. Lack of functional *jadI* gene in *jad*-cluster could not be complemented with *tcmI* or *dpsY* cyclase genes, catalyzing 3rd ring formation in anthracycline-like antibiotics [20]. This could be accounted for existence of specific protein-protein interactions between minimal *jad*-PKS and JadI cyclase. As it was shown in case of 1st ring cyclases involved in tetracenomycin C and actinorhodin biosynthesis, their presence highly increase the efficiency of polyketide synthesis and could affect polyketide chain length [3,7,9,13]. Coupling of cyclase and ketosynthase genes could be an indirect evidence for specific fine interactions of respective proteins. Finding that in *lan*/*lnd* cluster genes for 3rd ring cyclases are not physically coupled shows the diversity of regulatory ways controlling formation of PKS complex for angucycline biosynthesis.

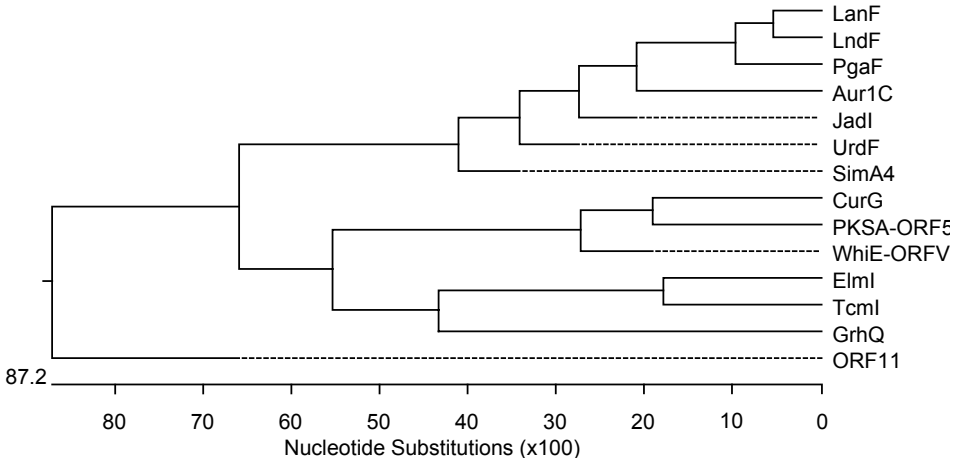

Fig.5. Phylogenetic tree showing relatedness of 3rd -4th ring cyclases from different polyketide biosynthesis gene clusters. Cyclases LanF, LndF, PgaF, Aur1C, JadI, UrdF, SimA4, WhiE-ORFVII and TcmI have been mentioned in the text. CurG – cyclase from *S. curacoi* curamycin gene cluster, ElmI – D-ring cyclase from *S. olivaceus* elloramycin gene cluster, GrhQ – probable cyclase for aromatic spiroketal polyketide griseorhodin biosynthesis in marine streptomycete *S. sp* JP95, ORF11 – probable cyclase from pradimycin-producing *Actinomadura hibisca* [10].

Genes for minimal PKS from *lnd*-cluster, as key determinants of polyketide chain synthesis, should be the primary target for transcriptional activation by regulatory protein LndI, whose gene is identified within the cluster and studied [8]. As it has been shown in case of gene clusters for actinorhodin and daunorubicin production, specific short DNA sequences upstream of promoters are recognized by activating proteins (tandemly arrayed heptameric repeats with consensus sequence 5'-$T_1C_2G_3A_4(G/C)_5C_6G_7$ are separated from each another by 11-22 bp, 3rd repeat is localized 4 bp downstream of 2nd) [19]. Preliminary search for possible direct/inverted repeats upstream of sequenced *lnd*-genes allowed to identify 3 DNA boxes with consensus sequence 5'-$T_1C_2G_3C_4(G/C)_5(G/A)_6(C/A)_7$ upstream of *lndI* gene; first two are separated by 12 bp, third is localized 4 bp downstream of second. No inverted or direct repeats are observed between *lndE* and *lndF*, *lndF* and *lndA*. On other hand, in *S. cyanogenus* S136 *lan*-cluster 5 bp

downstream of *lanE* stop codon three direct repeats are revealed with consensus sequence 5'-$T_1G_2(C/G)_3C_4(C/G)_5(C/G)_6G_7$, and localized in the same manner as those upstream of *lndI*. This obvious discrepancy in nucleotide sequence and localization should be supported by more extensive screening for repeats upstream of structural *lnd*/*lan*-genes in order to get clues into general organization of regulatory DNA sequences.

Sequencing of 700 bp DNA fragment from *S. globisporus* 1912 *lnd*-cluster has revealed new gene *lndF* for landomycin E biosynthesis, possibly encoding 3rd-4th ring cyclase. It's function could be deduced from sequence comparison with other cyclase genes involved in angucycline antibiotics biosynthesis. Phylogenetic analysis of conservative proteins is valuable source of information about process of evolution of given protein family. Close relatedness of "angu"-cyclases with those involved in spore pigment production could point that gene clusters for angucyclinone synthesis take origin from spore pigment biosynthesis gene clusters. Several differences in *lndF*/*lanF* genes organization with respect to the cyclase genes from other clusters for angucycline biosynthesis could reflect differences in mode of their expression regulation. Finally, initial identification and comparative analysis of possible target sequences for regulatory proteins within respective clusters revealed novel stage of dissimilarities between the clusters, suggesting that further detailed studying of landomycins production genetic control will be promising source of information on processes of antibiotic gene clusters evolution and enrich us with new insights about basic steps of polyketide biosynthesis.

Acknowledgements

L. Castro is acknowledged for help in sequencing. This work was supported by INTAS grant YSF 00-208 to B.O.

---

1. *Fedorenko V.O., Basiliya L.I., Pankevych K.O,* et al. Genetic control of polyketide biosynthesis by actinomycetes-producers of antitumor antibiotics // Bull. Ins. Agr.Microbiol. 2000. Vol.2.P.27-31.
2. *Henkel T., Rohr J., Beale J., Schwenen L.* Landomycins, new angucycline antibiotics from Streptomyces sp. Structural studies on landomycinsA-D// J.Antibiot.1990.Vol.43.P. 492-503.
3. *Khosla C., Gokhale R.S., Jacobsen J.R. and Cane D.E.* Tolerance and specificity of polyketide synthases // Annu. Rev. Biochem. 1999.Vol. 68.P. 219-253.
4. *Kramer P.J., Zawada .X., McDaniel R., Hopwood D.A., Khosla C.* Rational design and engineered biosynthesis of a novel 18-carbon aromatic polyketide // J.Am.Chem. Soc.1997.Vol.119. P. 636-639.
5. *Krohn K., Rohr J.* Angucyclines: total syntheses, new structures and biosynthetic studies of an emerging new class of antibiotics // Topics In Curr. Chem. 1997.Vol.188.P.127-195.
6. *Kulowski K., Wendt-Pienkowski E., Han L., Yang K., Vining L. and Hutchinson C. R.* Functional characterization of the *jadI* gene as a cyclase forming angucyclinones // J. Am. Chem. Soc.1999.Vol.121.P.1786-1794.
7. *Meurer G., Gerlitz M., Wendt-Pienkowski E.,*et al. Iterative type II polyketide synthases, cyclases and ketoreductases exhibit context-dependent behavior in the biosynthesis of linear and angular decapolyketides //Chem.Biol.1997.Vol. 4.P.433-443.
8. *Ostash B.O., Pankevych K.O., Luzhetskiy A.M.,*et al. Studying of genes involved in landomycin E biosynthesis by *Streptomyces globisporus* 1912 // 12th international symposium on the biology of actinomycetes; Vancouver, Canada, August 5-9, 2001. Theses.P.133.

9. *Petkovic H., Thamchaipenet A., Zhou L.-H.*,et al. Disruption of an aromatase/cyclase from the oxytetracycline gene cluster of *Streptomyces rimosus* results in production of novel polyketides with shorter chain lengths // J. Biol. Chem.1999.Vol.274.P. 32829-32834.
10. *Rawlings B*. Biosynthesis of polyketides (other than actinomycete macrolides) // Nat. Prod.Rep.1999.Vol.2.P.425-484.
11. *Salas J.A. and Mendez C.* Genetic manipulation of antitumor-agent biosynthesis to produce novel drugs // TIBTECH.1998.Vol. 161.P. 475-482.
12. *Shen B. and C.R. Hutchinson*. Tetracenomycin F2 cyclase: intramolecular aldol condensation in the biosynthesis of tetracenomycinC in *Streptomyces glaucescens* // Biochem.1993.Vol.32.P. 11149-11154.
13. *Shen B., Summers R.G., Wendt-Pienkowski E., C.R. Hutchinson.* The *Streptomyces glaucescens tcmKL* polyketide synthase and *tcmN* polyketide cyclase genes govern the size and shape of the aromatic antibiotics// J.Am.Chem.Soc.1995.Vol.117.P.6811-6821.
14. *Shen B and C.R.Hutchinson*. Deciphering the mechanism for the assembly of aromatic polyketides by a bacterial polyketide synthase// Proc.Natl.Acad.Sci.USA.1996.Vol.63.P.6600-6604.
15. *Shen Y., Yoon P., Yu T.-W.*,et al. Ectopic expression of the minimal *whiE* polyketide synthase generates a library of aromatic polyketides of diverse sizes and shapes // PNAS.1999.Vol.96.P.3622-3627.
16. *Staunton J., Weissman K.J.* Polyketide biosynthesis: a millenium review // Nat. Prod. Rep. 2001. Vol. 4. P. 380-416.
17. *Trefzer, A., Pelzer, S., Schimana, J* et al. Biosynthtetic gene cluster of simocyclinone, a natural multihybrid antibiotic. Antimicrob. Agents Chemother.2002. Vol. 46.P. 1174-1182.
18. *Westrich L., Domann S., Faust B.*,et al. Cloning and characterization of a gene cluster from Streptomyces cyanogenus S136 probably involved in landomycinA biosynthesis// FEMS Microbiol. Lett.1999.Vol.170.P.381-387.
19. *Wietzorrek A., Bibb M.* A novel family of proteins that regulates antibiotic production in streptomycetes appear to contain an OmpR-like DNA-binding fold.// Mol. Microbiol.1997.Vol.39.p.9447-9450.
20. *Wohlert S.-E., Wendt-Pienkowsk E.,Bao W., and C. R. Hutchinson.* Production of aromatic minimal polyketides by the daunorubicin polyketide synthase genes reveals the incompatibility of the heterologous DpsY and JadI ycclases, *J. Nat. Prod.* 2001.Vol.64.P. 1077-1080.

СЕКВЕНУВАННЯ ТА АНАЛІЗ НУКЛЕОТИДНОЇ ПОСЛІДОВНОСТІ ГЕНУ *LNDF* , ЩО КОДУЄ ІМОВІРНУ ЦИКЛАЗУ 3-ГО-4-ГО КІЛЕЦЬ В КЛАСТЕРІ ГЕНІВ БІОСИНТЕЗУ ЛАНДОМІЦИНУ Е *STREPTOMYCES GLOBISPORUS* 1912

Б.Осташ*, Ю.Ребець*, М.Самборський*, Х.А.Салас, В.Федоренко***

**Львівський національний університет імені Івана Франка*
вул. Грушевського 4, Львів 79005, Україна
e-mail: v_fedorenko@franko.lviv.ua
*** Університет м. Ов'єдо, відділ функціональної біології*
вул. Х.Клаверії 2, Ов'єдо 33006, Іспанія
e-mail: jasf@sauron.quimica.uniovi.es

Була визначена нуклеотидна послідовність фрагменту *lnd*-кластеру S.globisporus 1912 розміром 700 п.н. 3′-кінець гену оксигенази *lndE*, 5′-кінець гену кетосинтази *lndA* та повна ORF для гену циклази 3-го кільця *lndF* були ідентифіковані. Аналіз імовірного продукту трансляції гену *lndF* виявив його високий ступінь гомології з ідентифікованими циклазами 3-го та 4-го кілець ангуциклінів. На відміну від генів *jadI*, *urdF*, стоп-кодони яких перекриваються із старт-кодонами генів-кетосинтаз, ген *lndF* (як і *lanF*) не перекривається із *lndA* (*lanA*). Філогенетичний аналіз *lndF* дозволяє припустити, що гени біосинтезу ангуциклінів беруть свій початок із кластерів генів біосинтезу спорових пігментів актиноміцетів. У просеквенованій ділянці *lnd*-кластеру не виявлено прямих або інвертованих тандемних повторів, які є потенційними мішенями для регуляторного білка LndI.

Ключові слова: Streptomyces, ангуцикліни, циклази, ландоміцин Е.

Стаття надійшла до редколегії 14.12.2002
Прийнята до друку17.02.2003