UDC 577.21:579.222:7:577.182
ISSN 0352-9193

conference paper

The Impact of Molecular Genetics on Understanding Antibiotic Biosynthesis: Design of New Drugs

Utjecaj molekularne genetike na razumijevanje biosinteze antibiotika: konstrukcija novih lijekova

K. J. Linton, Arinthip Thamchaipenet and I. S. Hunter

Institute of Biomedical and Life Sciences, Robertson Laboratory of Biotechnology, University of Glasgow, Glasgow G12 8QQ, U.K.

Received: June 15, 1994
Accepted: September 26, 1994

Summary

A number of gene clusters for antibiotic production have now been cloned and analysed in considerable detail. Attention has focussed on the genes for β-lactams and for the polyketides, such as erythromycin and oxytetracycline. It is now clear that the polyketides fall into two classes: simple (aromatic) structures are made by complexes of enzymes, each subunit of which has a single enzyme activity, whereas more complex structures (such as the macrolides) are made by very large multi-functional polypeptides. Paradoxically, it is easier to rationalise the biology behind the biosynthesis of the more complex structures. Our understanding of the systems is now at a stage where alteration of the genes has resulted in different compounds being made. Many of the 'rules' which determine that a particular structure is made have now been deduced. It is also now possible to engineer strains to make 'biological synthons', which can be modified subsequently by chemical means or by further fermentation. This approach will provide a novel route to discovery of new antibiotics. In recent years, it has been realised that many of the metabolites made by microorganisms are useful as therapeutic agents for treating many diseases. The same methodology that is being used for antibiotics can also be applied to them. Molecular genetics is now an integral part of drug-discovery programs, and presents many exciting new prospects.

Sažetak

Do sada je kloniran i iscrpno analiziran niz genskih nakupina za proizvodnju antibiotika. Pozornost je usmjerena prema genima za β-laktame i poliketide, kao što su eritromicin i oksitetraciklin. Danas je jasno da poliketidi pripadaju dvjema skupinama: jednostavne (aromatske) strukture sintetiziraju se s pomoću enzimskih kompleksa, od kojih svaka podjedinica ima zasebnu enzimsku aktivnost, dok se složenije strukture (poput makrolida) sintetiziraju s pomoću vrlo velikih višefunkcionalnih polipeptida. Paradoksalno je što je lakše objasniti biologiju biosinteze složenijih struktura. Naše današnje spoznaje tih sustava takve su da izmjena gena omogućava proizvodnju različitih spojeva pa su iz toga izvedena mnoga »pravila« prema kojima se može odrediti određena struktura. Ujedno je moguće konstruirati sojeve koji proizvode »biološke sintone«. Oni se naknadno mogu izmijeniti kemijskim putem ili daljnjom fermentacijom. To će omogućiti nov pristup otkriću novih antibiotika. Posljednjih nekoliko godina otkriveno je da su mnogi mikrobni metaboliti korisna terapeutska sredstva za liječenje različitih bolesti. Prema tome, na njih se može primijeniti ista metodologija kao i za antibiotike. Zbog toga je molekularna genetika danas sastavni dio programa za otkriće novih lijekova pa puno obećava.

Introduction

For over 50 years, human health care has benefited from the discovery and use of antibiotics. Nowadays, it is difficult to imagine life without antibiotics – we take them so much for granted. However, the widespread use of antibiotics throughout the last half-century has resulted in some clinical infections now becoming resistant to many of the antibiotics in common use. Natural systems for genetic exchange (by transposition or transduction) ensure that, once a single strain has become resistant to an antibiotic, a resistance determinant may transmit quickly to others within the species and subsequently across the species barriers. For example, many hospital operating theatres are infested with methicillin-resistant *Staphylococcus aureus* (MRSA), which is resistant to all clinical antibiotics except vancomycin. It is probably only a matter of time before vancomycin-resistant strains of MRSA evolve. Similarly, infectious diseases such as tuberculosis, which were almost eradicated as a result of

antibiotic therapy in the early days, are now reappearing with resistance determinants. It is clear that we must continue our search for novel antibiotics to supply the clinicians with new drugs to combat this inevitable 'antibiotic-resistance' problem *(1)*.

Historically, the pharmaceutical industry has screened large numbers of microbial isolates in the hope that novel structures will be produced which can be isolated directly from the microorganisms. The search has prompted microbiologists to investigate exotic environments (such as thermal springs, arid deserts and marine ecosystems) to isolate unusual microbes which may make novel isoteric structures. A second essential element of discovery programs for new antibiotics has been contributed by chemists, who have taken the antibiotic structures isolated from these microorganisms and modified them chemically – to derive semi-synthetic products with more potent or other useful pharmacological activity.

The advent of molecular biology has added a new dimension to the discovery of antibiotics. Although these techniques were devised initially to understand better how biosynthesis of antibiotics occurred, it is clear that the technology can also be applied to generate diversity in antibiotic structures *(2)* – with the hope that novel activities effective against drug-resistant infections might result.

The filamentous microorganisms produce the vast majority of antibiotics. Within them, there are two major groupings. The eukaryotic filamentous fungi (*Acremonium, Cephalosporium*) produce β-lactams, such as the penicillins, cephalosporins and the starting materials for the semi-synthetic β-lactam derivatives. The prokaryotic filamentous bacteria, typified by the *Actinomycetes*, produce a diversity of metabolites such as the tetracyclines, erythromycin, streptomycin as well as a number of unusual β-lactams such as cephamycin.

This short review will limit its scope to discuss the *Streptomyces* – the actinomycetes on which most progress has been made. It will focus on work undertaken on a particular class of antibiotics made by the *Streptomyces* – the so-called 'polyketides' – derived from intermediates of central metabolism activated by Coenzyme A such as acetyl-CoA, propionyl-CoA, methylmalonyl-CoA and butyryl-CoA.

Results and Discussion

Cloning of antibiotic pathways

The discovery that genes for antibiotic biosynthesis are invariably clustered (reviewed *3)* on the single chromosome of *Streptomyces* [now shown to be linear *(4)*, in marked contrast to most other bacteria] has simplified the cloning of entire antibiotic pathways. When any gene within an antibiotic cluster is cloned, it can be reasonably assumed that the other biosynthetic genes will lie alongside it. Hence, a limited amount of 'chromosome walking' using lambda or cosmid libraries will enable the entire gene cluster to be cloned. A number of strategies have been used to tackle the problem of cloning for the first time a gene within a particular pathway (reviewed *3)*.

The discovery that an antibiotic-resistance gene most often resides within the cluster of genes encoding production of that antibiotic has provided a facile means by which to clone initially a gene of a pathway. A microorganism which produces an antibiotic must, itself, be resistant to that antibiotic – otherwise it will commit suicide. Construction of a gene bank of the producing organism in a streptomycete host which is sensitive to the antibiotic (most often *S. lividans*) has been the route by which many resistance genes have been cloned. The initial step in cloning of the gene cluster for biosynthesis of oxytetracycline (OTC) used this strategy, except that an OTC-sensitive mutant of the producing strain (*S. rimosus*) was used as the primary host *(5)*.

In the case of OTC, two resistance genes (*otrA* and *otrB*) were cloned *(6)*. It is now commonplace to discover more than one gene for resistance to an antibiotic within a producer organism. For OTC, DNA sequencing of *otrA* revealed that it encodes a ribosomal resistance factor which protects the ribosomes of the producing organism from translational arrest by the endogenous OTC that it is synthesising *(7)*. A similar analysis of *otrB* revealed that it encodes an exporter for OTC – which may be legitimately considered to be the last step of the biosynthetic pathway, in which intracellular OTC is excreted from the cell *(8)*. Many of the resistance genes isolated from production gene clusters have turned out to encode exporters. This is not surprising, as antibiotics must get out of the cell to be effective.

Comparisons of the deduced peptide sequence of OtrA with those on the database revealed that OtrA has a high similarity to TetM, the resistance factor of some Gram-positive tetracycline-resistant transposons. OtrB has significant similarity to the tetracycline exporter of Gram-negative transposons. Indeed, genes with extremely high similarities to *otrA* and *otrB* have now been identified in tetracycline-resistant clinical infections of *Streptomyces* and Mycobacteria *(9)*. This leads further credence to the hypothesis that one source of antibiotic resistance in clinical infections is, in fact, by horizontal gene transfer from the microorganism which produces that antibiotic. Thus, horizontal transfer contributes to the overall 'antibiotic resistance problem'.

When the *otrA* and *otrB* genes were mapped onto cosmids constructed from the genomic DNA of *S. rimosus*, they were found to lie approximately 30 kb apart, with the genes for production of OTC between them *(6)*. Subcloning into *S. lividans* of a 34 kb *EcoRI* fragment including both resistance genes enabled *S. lividans* to make OTC, confirming that this DNA segment contained all the information necessary for OTC biosynthesis *(10)*. This example demonstrates how an entire pathway for biosynthesis of an antibiotic can be cloned starting with a single, easily-selectable gene.

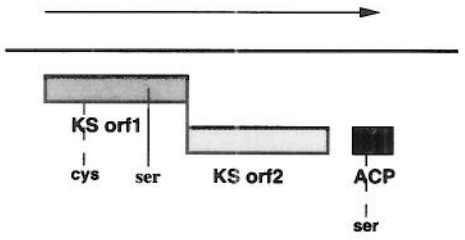
The OTC pathway is only one of an increasing number of gene clusters for antibiotic production which have been cloned, particularly those for polyketides (reviewed *11, 12)*.

Analysis of genes for the biosynthesis of the polyketide backbone

DNA sequencing of genes for biosynthesis of polyketide antibiotics was, logically, the simplest strategy to gain information on the biosynthetic enzymes re-

sponsible – the so-called 'polyketide synthase' (PKS). For tetracenomycin *(13)* and granatacin *(14)*, which are antibiotics composed solely of acetate units polymerised together 'head-to-tail' ('poly-acetate' antibiotics), DNA sequencing of the PKS's revealed that the gene products were somewhat similar to those responsible for biosynthesis of fatty acids in *Escherichia coli*. By analogy with the fatty acid synthases, these PKS proteins – which were simple, easily-dissociable, monofunctional units, each capable of catalysing a single type of biochemical reaction – were termed 'Type II' PKS's.

A similar architecture for the PKS genes for OTC biosynthesis was revealed by DNA sequencing *(15)*. The PKS genes for each 'polyacetate' polyketide are very similar to each other in architecture (Fig. 1.), and consist of three open reading frames (orfs).



orf 1 &2

- **separate proteins**
- **very similar**
- **translationally-coupled (usually)**
- **orf1 has 'catalytic' cys**
- **probably heterodimers**

ACP

- **requires phosphopantotheine gp**

Fig. 1. Architecture of 'Type II' polyketide synthase
Slika 1. Arhitektura poliketid sintaza tipa II.

Two orfs (orf1 and orf2) are very similar. They are translationally-coupled, implying that they are synthesised in strict stoichiometry, and may exist normally as heterodimers. Orf1 contains cysteine and serine residues in positions within the protein which are consistent with it being the catalytic subunit of the 'keto synthase' – responsible for sequential additions of acetate units to the growing polyketide chain. By contrast, orf 2 does not contain the 'catalytic' cysteine or serine residues, and this observation has generated the hypothesis that orf2 may be implicated, somehow in the heterodimer *(16)*, in determining the specificity of the length of polyketide chain made. The third orf (Fig. 1.) encodes an acyl carrier protein (ACP), a small (≈ 8 kDa) protein which is modified post-translationally by addition of 5' phosphopantetheine. Conceptual models of fatty acid biosynthesis assign the *holo*-ACP to be a protein with a 'swinging arm' which 'docks' the growing polyketide chain into the active sites of the various enzymes of fatty acid biosynthesis. By analogy, a similar role for *holo*-ACP is proposed in polyketide biosynthesis. In principle, these three orfs, termed the 'minimal PKS', are likely to provide the catalytic potential and the information necessary to determine the length of the polyketide chain assembled by the type II PKS's.

The more complex polyketides, typified by the macrolides (such as erythromycin) and by the polyethers (such as tetronasin) have different features in their biosynthetic schemes. Instead of being composed solely of acetate 'extender' units, these complex polyketides are derived from a number of different extender units. Moreover, at each extension step the chemistry of the newly-added keto group may be modified – which results eventually in one of the following: formation of a keto group (if no further modification takes place), of a hydroxyl group (if the keto group is further modified by a keto reductase), of an unsaturated double bond (if the hydroxyl group is further modified by a dehydrase) and of a fully reduced bond (if that unsaturated band is further modified by an enoyl reductase). Thus, the diversity of backbone structures of complex polyketides is immense. It is dictated by the choice of the various extenders which can be used, and this diversity is enhanced further by the different chemistries which may be performed on the keto group which is added at each extension step.

Recognising the vast structural diversity of the complex polyketides suggested that unravelling the mechanism by which the proteins might interact to 'program' the biosynthesis of that polyketide would be a monumental undertaking. The solution to the problem, which resulted from the DNA sequencing of the genes responsible for the biosynthesis of the backbone of erythromycin *(17,18)*, provides a delightful example of the simplicity of nature in resolving how to undertake such a complex operation.

The deduced peptide sequence of the erythromycin 'PKS' showed that, in marked contrast to the polyaromatic (type II) PKS's described above, this synthase was composed of three large (> 300 kDa) multifunctional polypeptides. By analogy with the mammalian fatty acid synthases which also contain multifunctional polypeptides, these synthases were termed 'Type I' PKS's. Each polypeptide contained two 'modules', each module of which was uniquely responsible for the addition of one extender unit to the growing polyketide which eventually ended up as the backbone of erythromycin. The biosynthesis of erythromycin requires six extension steps in total, each of which is catalysed by a separate module (3 large proteins each containing 2 modules). Thus, unlike the type II PKS's (in which each monofunctional protein is used iteratively to build up the chain), the Type I PKS's for complex polyketides are composed of modules each of which is responsible for a unique step in the biosynthetic scheme. The choice of extender unit which is added at a particular step is governed by the identity of the acyltransferase domain embedded in the module for that step. If the acyltransferase catalyses transfer of a methylmalonyl-CoA to the 'PKS', then this is inserted: likewise for propionyl-CoA and so on. For each extension step, the module responsible for that step will also contain domains which determine the final chemical functionality of the keto group added. For example a module, responsible for an extension step which

eventually results in the keto group which is added becoming a hydroxyl group, will contain domains for a keto synthase, acyl transferase, keto reductase and ACP. When the functionality is extended to become a fully saturated bond, the module will contain, in addition, domains for a dehydrase and enoyl reductase (reviewed *18).*

In summary, polyketide synthases fall into two main classes – those for biosynthesis of polyaromatic structures derived solely from acetate (Type II PKS's) and those which are derived from a variety of extender units (Type I PKS's). The programming of the complex 'Type I' PKS's turns out to be a triumph in the simplicity of nature – the chemistry for addition of each extender unit is determined by the domain structure of each large multifunctional protein in which each domain is used only once during the biosynthetic scheme. For the 'Type II' PKS's an understanding of the strategy for programming the length of polyketide (i.e. the number of acetate groups added) has come mainly from attempts to make novel polyketide structures (see below).

Folding of the polyketide chain

Considerable DNA sequencing has been undertaken for a number of gene clusters for biosynthesis of polyketides catalysed by a 'Type II' PKS. Invariably these clusters contain one or more 'cyclase/dehydrase' genes thought to be responsible for the cyclisation (at least in part) of the *de novo* polyketide backbone structure to the fully-folded final structure of the molecule *(19).* Many of the clusters also contain a specific ketoreductase gene (for which the paridigm is the *act*III gene of *S. coelicolor (20)*) which reduces the keto group present on the ninth carbon atom from the end of the growing chain. The result is that a double bond is introduced at this point within the backbone structure of the polyketide. Unsaturated bonds within the backbone confer structural rigidity, whereas fully-saturated bonds are flexible, and it has been hypothesised that this may predicate (in part) the pattern of folding of the polyketide. However, tetracenomycin, which does not have such a ketoreduction step in its biosynthetic scheme, folds perfectly well. Thus the ketoreductase step is not a necessary prerequisite for folding in every case.

For the polyketides synthesised by 'Type I' PKS's (most common of which are the macrolides which often contain rings of 16 or 14 carbon atoms), it is still not clear how the folding of these large rings takes place. Certainly, there are no genes similar to the cyclase/dehydrases within the known clusters for 'Type I' polyketides.

Novel structures derived by recombinant DNA technology

The first report of a novel structure derived by genetic engineering of a polyketide pathway was published in 1985 *(21).* In this case, a hydroxylase gene from the actinorhodin pathway was cloned into the producer of medermycin – an antibiotic which has a structure similar to actinorhodin but lacks a particular hydroxyl group. The recombinant produced a hydroxylated version of medermycin, mederrhodin, which was easily recognised by its different colour. This report demonstrated that it was possible to use molecular genetics to derive structures which had not been seen before in nature. Much of the work on such so-called 'hybrid' antibiotics has been reviewed recently *(12).*

The example of mederrhodin showed that post-polyketide-assembly modifications of antibiotic structures was possible. A particularly striking example came from cloning the *car*E gene from *S. thermotolerans* into *S. ambofaciens (22).* This gene is responsible for addition of isovaleryl or butyryl groups to a sugar on carbomycin, which is produced by *S. thermotolerans.* Whereas wild-type *S. ambofaciens* produces spiramycin, the recombinant produced isovaleryl-spiramycin, a novel product.

For the 'Type I' PKS's, for which erythromycin biosynthesis is the universally-accepted model, attempts have been made to reprogramme some of the modules. These experiments have been designed to be predictive – if modular programming of biosynthesis of complex polyketides is real, then it should be possible to deduce the structure of a polyketide which would result from genetic intervention of a module. When an in-frame deletion of the ketoreductase domain of the 5th module for erythromycin biosynthesis was created, the resulting recombinant produced the predicted structure which retained the keto group at the postion in the backbone structure which is added at the fifth step *(23).* Similarly, the enoyl reductase domain of module 4 was mutated to result in the novel structure which now contained a double bond at that position rather than the fully-saturated bond (cited *12).* Obviously, only the successful attempts to reprogram the PKS for erythromycin have been reported – many other attempts may have failed due to the PKS being intolerant to these mutations. However, the work does demonstrate the potential utility of this type of approach.

For the 'Type II' PKS's, matrix studies have been undertaken to 'mix and match' polyketide synthase genes from various sources. Substitution of the ACP with one from a heterologous source certainly changed the level of productivity of antibiotic made, but not its identity *(24).* However, the number of additions of acetate units to the growing chain does appear to be specified by the catalytically-inert orf2 which is the partner of the catalytic ketosynthase subunit in the heterodimer of the PKS complex *(25).* When a hybrid PKS is created, the length of the polyketide chain reflects the identity of this subunit.

Conclusion and perspective

The study of polyketide biosynthesis using molecular genetics has not only enhanced our understanding of the fundamental processes of biosynthesis, but has also provided the opportunity to use the technology to create novel structures not seen previously in nature. These structures have the potential to act as novel agents useful in human health care.

Whereas it is generally assumed that the *Streptomyces* only produce antibiotics, it has become clear over the last decade that the genus also provides a rich source of metabolites, such as immunosuppressants, vasoconstrictors, muscle relaxants and nemodectocides, which themselves are not capable of killing infectious agents – but which have enormous potential in health care.

The application of recombinant DNA technology to generate novel structures promises to contribute both to

our search for new products to combat the 'antibiotic resistance problem' and to enhance the quality of life in general by providing new therapeutic agents of use to mankind.

References

1. I. Chopra, *J. Antimicrob. Chemother. 30* (1992) 737.
2. C. R. Hutchinson, *Biotechnology, 12* (1994) 375.
3. I. S. Hunter, S. Baumberg, *Symp. Soc. Gen. Microbiol. 44* (1989) 121.
4. Y.-S. Lin, H. M. Kieser, D. A. Hopwood, C. W. Chen, *Mol. Microbiol. 10* (1993) 923.
5. I. S. Hunter, *Biochem. Soc. Trans. 12* (1984) 586.
6. M. J. Butler, E. J. Friend, I. S. Hunter, F. S. Kaczmarek, D. A. Sugden, M. Warren, *Mol. Gen. Genet. 215* (1988) 231.
7. D. Doyle, K. J. McDowall, M. J. Butler, I. S. Hunter, *Mol. Microbiol. 5* (1991) 2923.
8. J. P. Reynes, T. Calmeb, D. Drocourt, G. Tiraby, *J. Gen. Microbiol. 134* (1988) 585.
9. Y. Pang, B. A. Brown, V. A. Steingrube, R. J. Wallace, M. C. Roberts, *Antimicrob. Agent. Chemother. 38* (1994) 1408.
10. C. Binnie, M. Warren, M. J. Butler, *J. Bacteriol. 171* (1989) 887.
11. D. A. Hopwood, D. H. Sherman, *Annu. Rev. Genet. 24* (1990) 37.
12. L. Katz, S. Donadio, *Annu. Rev. Microbiol. 47* (1993) 875.
13. M. J. Bibb, S. Biro, H. Motamedi, J. F. Collins, C. R. Hutchinson, *EMBO J. 8* (1989) 2727.
14. D. H. Sherman, F. Malpartida, M. J. Bibb, H. M. Kieser, D. A. Hopwood, *EMBO J. 8* (1989) 2717.
15. E. S. Kim, M. J. Bibb, M. J. Butler, D. A. Hopwood, D. Sherman, *Gene, 141* (1994) 141.
16. R. McDaniel, S. Ebert-Khosla, D. A. Hopwood, C. Khosla, *Science, 262* (1993) 1546.
17. J. Cortes, S. F. Haydock, G. A. Roberts, D. J. Bevitt, P. F. Leadlay, *Nature, 348* (1990) 176.
18. S. Donadio, M. J. Stauer, J. B. McAlpine, S. J. Swanson, L. Katz, *Science, 252* (1991) 675.
19. D. H. Sherman, M. J. Bibb, T. J. Simpson, D. Johnson, F. Malpartida, M. Fernandez-Moreno, E. Martinez, C. R. Hutchinson, D. A. Hopwood, *Tetrahedron, 47* (1991) 6023.
20. S. E. Hallam, F. Malpartida, D.A. Hopwood, *Gene, 74* (1988) 305.
21. D. A. Hopwood, F. Malpartida, H. M. Kieser, H. Ikeda, J. Duncan, I. Fujii, B. A. M. Budd, H. G. Ross, S. Onura, *Nature, 314* (1985) 642.
22. J. K. Epp, M. L. B. Huber, J. R. Turner, T. Goodson, B. Schoner, *Gene, 85* (1989) 293.
23. S. Donadio, L. Katz, *Gene, 111* (1992) 51.
24. C. Khosla, S. Ebert-Khosla, D. A. Hopwood, *Mol. Microbiol. 6* (1992) 3237.
25. R. McDaniel, S. Ebert-Khosla, D. A. Hopwood, C. Khosla, *Science, 262* (1993) 1546.