Advances in the Understanding and Use of the Genomic Base of Microbial Secondary Metabolite Biosynthesis for the Discovery of New Natural Products[⊥]

James B. McAlpine*

Thallion Pharmaceuticals Inc., 7150 Alexander-Fleming, Montréal, Canada, H4S 2C8

Received November 19, 2008

Over the past decade major changes have occurred in the access to genome sequences that encode the enzymes responsible for the biosynthesis of secondary metabolites, knowledge of how those sequences translate into the final structure of the metabolite, and the ability to alter the sequence to obtain predicted products via both homologous and heterologous expression. Novel genera have been discovered leading to new chemotypes, but more surprisingly several instances have been uncovered where the apparently general rules of modular translation have not applied. Several new biosynthetic pathways have been unearthed, and our general knowledge grows rapidly. This review aims to highlight some of the more striking discoveries and advances of the decade.

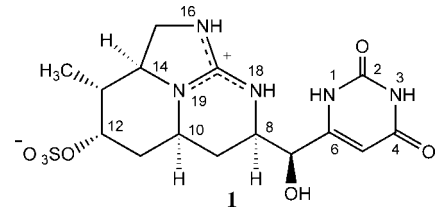
Introduction

One of the most basic and striking changes has been in sequencing itself. Modern sequencers have made whole genome sequencing commonplace, rapid, and relatively inexpensive. Recently, the author pointed out that the number of whole bacterial genomes entering the public literature was obeying Moore's law with a doubling time of 20 months.[1] In the last two years the semilog plot has continued in a straight line, maintaining this doubling time (Figure 1).

The two major pathways to microbial secondary metabolites, polyketide synthases (PKSs), and nonribosomal peptide synthetases (NRPSs) have been the subject of substantive recent reviews,[2−4] which cover extensively the structure and function of these enzyme systems. These two systems, and combinations of them, provide the basis for a plethora of natural products covering a wide variety of structural types. As the genomic basis of each was unearthed, the initial impression was that it would be relatively simplistic; however both the generation time for bacteria and the period of evolution have worked together to build a much more complex picture.

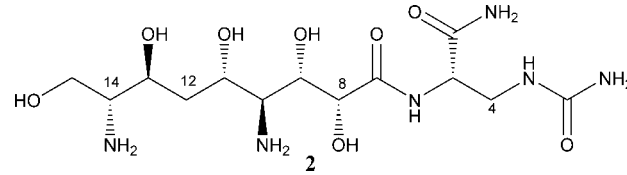
Type I Polyketides

Non-iterative type I polyketide synthases have always seemed to provide for straightforward gene-to-secondary metabolite analyses. The distinct modules and the colinearity rule[5] together with an understanding of active sites of the various reducing enzymes and the critical amino acids in the acyl transferases (ATs) that determine the extender units[6] have all conspired to give the chemist a false confidence that, for any type I PKS, one could predict the structure from the genes and/or the arrangement of genes that would lead to the structure. Indeed, in most cases, this was true and some were bold enough to claim that the genetic data should be relied on as much as conventional spectral data in structure determination.[7] Recently, several unusual type I PKSs have surfaced that might give one pause in jumping to conclusions. When Richard Moore's group elucidated the structure of cylindrospermopsin (**1**), the hepatotoxin from the blue-green alga *Cylindrospermopsis raciborskii*, they modestly made no predictions on its biosynthesis, as one might not immediately have confidently predicted that this was a type I polyketide. [8] Their subsequent feeding experiments established this as the probable biosynthesis, suggesting a novel starter unit, guanidinoacetic acid, that had been derived from glycine.[9] All of the other carbons of the backbone were labeled from acetate, except the branched methyl. The hybrid NRPS/type I PKS nature was confirmed at the genetic level, and the amidotransferase-encoding gene was characterized.[10] The complete biosynthetic locus was sequenced and shown to comprise 15 ORFs and cover 43 kb.[11] This analysis provides data on a new bacterial gene encoding an amidinotransferase (homologous to the human arginine:glycine amidinotransferase), the unique specificity code selecting for NRPS adenylation of guanidinoacetic acid (DFHFITHD), and outlines the genes for a novel biosynthetic pathway leading to the uracil moiety.



1

Another NRPS/PKS that might fool the novice is zwittermicin A (**2**). Unlike the case of cylindrospermopsin, where the classical PKS structure is masked by elaborative reactions on the polyketide chain (the Michael additions of the guanidine group to the olefins of the chain and the unusual synthesis of the uracil group), the polyketide nature of zwittermicin is masked by the use of uncommon extender units in normal polyketide growth.



2

As previously noted,[12] the backbone of zwittermicin A might be thought to arise from an amino sugar. However, it is a staightforward type I polyketide, fused to an NRPS starter unit, but with aminomalonate and hydroxymalonate as the last two extender units. Thomas and colleagues have used MS and biochemical studies[13] to confirm their previous assignment[14] of these two ACP conjugates as the sources of the last two units. Such functionalized extender units hold the key to greater chemical diversity in combinatorial biosynthesis. Methoxymalonate, another functionalized extender unit, occurs in the biosynthesis of soraphen A,[15] ansamitocin P-3,[16] ascomycin,[17] spiramycin,[18] and geldana-

[⊥] Dedicated to Dr. David G. I. Kingston of Virginia Polytechnic Institute and State University for his pioneering work on bioactive natural products.

* To whom correspondence should be addressed. Tel: 514-923-6800. Fax: 514-940-3620. E-mail: jmcalpine@thallion.com.

10.1021/np800742z CCC: $40.75
Published on Web 02/06/2009

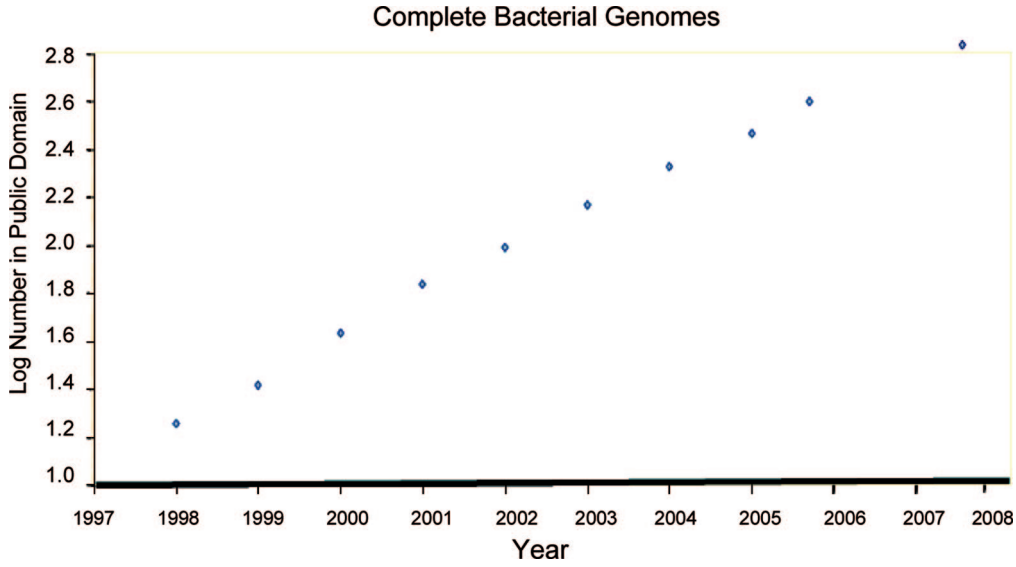
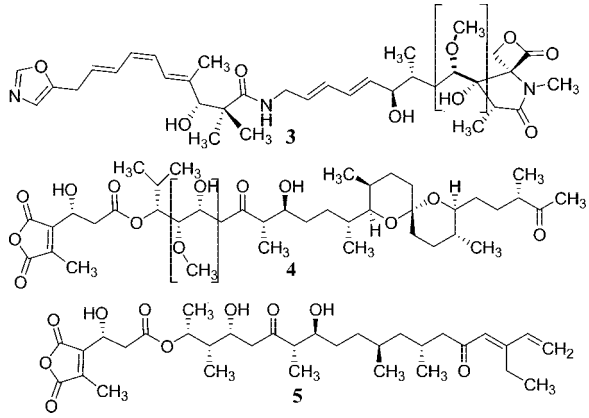
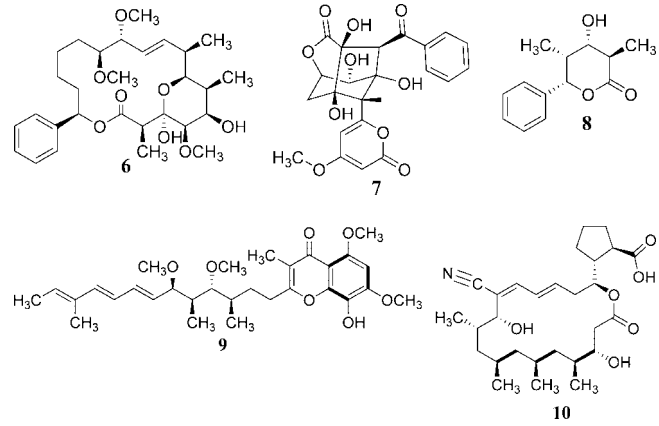


Figure 1. Semilog plot of the number of bacteria for which the complete genome was available in the public domain over the past decade is a straight line, indicating that the number of bacterial genomes fully sequenced doubles every 20 months.

mycin[19] and has been implicated as originating from the glycolytic pathway. Although it provides a less facile entry into branching points in combinatorial biosynthesis than hydroxyl- or amino-malonate, it provides a search tool for other loci. The genes encoding methoxymalonyl-ACP have been used by the Shen group to locate the biosynthetic loci of oxazolomycin (**3**) from *Streptomyces albus* JA3453[20] and of tautomycin (**4**) from *Streptomyces spiroverticillatus,*[21] whereas the Sherman group had to take the somewhat less certain route to the related metabolite tautomycetin (**5**) via a methylmalonyl AT gene probe.[22] Glyceric acid has been shown to be a chain extension unit particularly as the terminal group in the tetronate antibiotics.[23,24] Shen and Kelleher have identified the enzyme OzmB from the oxazolomycin biosynthetic pathway that is responsible for carrying 1,3-diphosphoglycerol from the primary glycolytic pathway to the thiolated glyceric acid, from which it can be used directly, or presumably shuttled to hydroxyl- or methoxy-malonate.[25]



A wide variety of starter units have been characterized in polyketides, particularly as NRPS/PKS fusions are relatively common. However, a somewhat surprising one is benzoic acid, which one does not think of as a common natural product. This occurs in the myxobacterial (*Sorangium cellulosum*)-produced soraphen A (**6**) and in the very atypical type II polyketide enterocin (**7**). The Leadlay group have expressed heterologously the acyl transferase (AT) of the PKS loading domain from *S. cellulosum* in a *Saccharopolyspora erythrea* construct and shown that under appropriate conditions they could produce the 5-phenyltriketide lactone (**8**).[26] Bradley Moore and Xiang have characterized the genes encoding the biosynthetic enzymes for benzoyl-Co-A in the enterocin producer "*Streptomyces maritimus*"[27] and the enzyme EncM, which carries the polyoxo compound through a Favorskii rearrangement, two aldol condensations, and the closure of both heterocycles, to arrive at the penultimate product ready for the *O*-methyl transferase.[28]



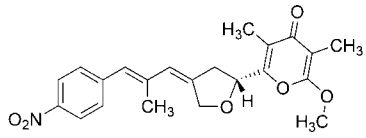
Iteration in Non-iterative Polyketide Synthases

Iterative operations in presumably non-iterative PKSs present a minor level of insecurity to the predictability of type I polyketides. A mutant strain of *Saccharopolyspora erythrea* was shown to produce minor quantities of two 16-membered macrolides, which appear to have arisen from a "stuttering" of module 4 of the PKS.[29] (The two differed only in the starter unit, acetyl and propionyl, a variability seen in the parent strain).

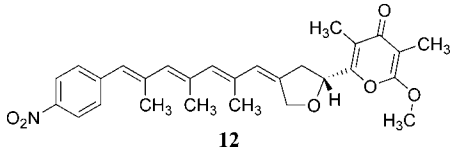
This "stuttering" occurs occasionally in natural strains. In the case of stigmatellin (**9**), the structure requires that either module StiH or StiJ be used iteratively to insert the three terminal units (accented).[30] In the case of borrelidin (**10**), the structure requires that one module in the PKS be used three times successively to insert the three last methylmalonates (accented).[31]

An analysis of the aureothin (**11**) and neoaureothin (**12**) pathways indicates that in each case the initial PKS module is used iteratively.[32,33]

The congener of neoaureothin, orniocin (**13**), arises from a UV-induced elimination of mesitylene from neoaureothin and is not

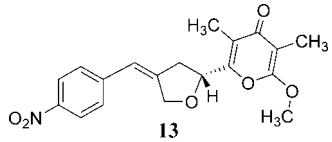


11

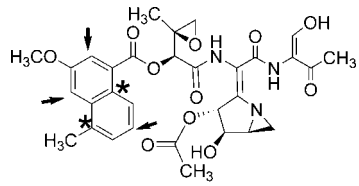


12

the direct product of a biosynthetic locus;[34] hence it does not represent the product of the aureothin locus without the iteration.



13



14

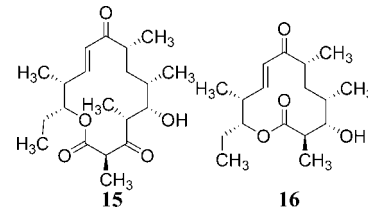
The naphthoate biosynthesis in azinomycin B (**14**) is via a less common iterative type I PKS and with an interesting distinction from the other actinomycete iterative type I PKSs, in that the single ketoreductase presumably acts specifically at exactly three of the carbonyls (arrowed). The keto groups at C7 and C11 (starred) of the poly oxo chain are needed, presumably, for formation of the naphthalene nucleus. The methoxyl group is the result of two separate genes (in the locus but outside the PKS) encoding a P-450 hydroxylase and a SAM methyltransferase.[35]

Trans-AT Polyketides

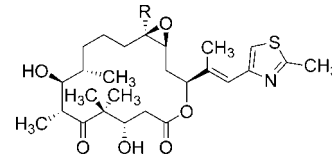
An increasingly large subgroup of type I polyketides, which diverge from the canonical rules of colinearity, are those with trans-AT domains. In these loci, one or more AT is located outside of the string of PKS ORFs, which are devoid of such modules. This configuration was first noticed in the dorrigocin/migrastatin biosynthetic loci[36] and was described in detail with respect to the leinamycin biosynthetic locus.[37] Normally, the AT domain specifies the subunit incorporated into the polyketide; however for trans-AT polyketides, one or two ATs are responsible for many subunits, sometimes with a variety of subunits in these polyketides, which commonly exceeds the number of ATs. This problem has recently been solved by Piel and Hertweck,[38] who showed that the ketosynthases, in trans-AT PKSs, hold the specificity for the extending unit. However, in many cases, trans-AT polyketides acquire their branching by incorporating methyltransferases in the modules.

Domain Skipping

In contrast to the "stuttering" disconnect between genes and secondary metabolite structure is a "skipping" of an apparently functional module during processing. This was observed with a hybrid PKS construct in which a module from the rapamycin locus was paired with the loading module, the second module, and the thioesterase from the 6-deoxyerythronolide B (DEBS) PKS.[39] Although the expected tetraketide was generated, it was by far the minor product (<5%), and major products were the two triketides, which were generated by the construct missing the module from the rapamycin locus. The more studied case is the intriguing natural one in which *Streptomyces venezuelae*, under different fermentation conditions, can produce either the 12-membered macrolide, methymycin, or the 14-membered macrolide, pikromycin, via the same PKS, which contains functional modules for the production of the hexaketide precursor to pikromycin, narbonolide (**15**). Recent studies by the Sherman group suggest that the conformation of the docking domain between the ultimate and penultimate PKS modules determines whether the growing ketide chain is passed from the penultimate thiolation (ACP) domain to the ultimate thiolation domain or to the thioesterase, with the former leading to narbonolide, and the latter to 10-deoxymethynolide (**16**).[40]



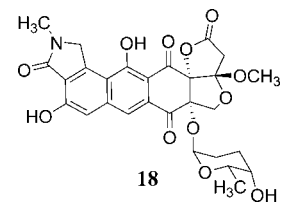
15 **16**



17a R = H
17b R = CH3

Undefined Motifs

Another variation in PKSs, which can lead to less exact reading of the genetic information, is the coding for nondefinitive motifs. This is noticed in the locus for the biosynthesis of the epothilones A and B (**17a** and **17b**, respectively) and their precursors C and D, in which the epoxide is replaced by an olefin. The variation between methyl and hydrogen arises because the specifying motif in the AT introducing the fourth unit in the ketide chain is neither HAFH (specifying malonyl) nor YASH (specifying methylmalonyl) but rather an indeterminate hybrid HASH.[41] However, domain swapping experiments with ATs from the epothilone locus and DEBS module 1 indicated that the situation may be more complex than it appears. If the HASH containing AT was included in DEBS, the distribution of products was as would be predicted from that seen in the native *S. cellulosum.* However, if site-directed mutation changed the motif to YASH in this AT, the DEBS construct products did not change significantly, and if mutated to HAFH, the construct did not produce any triketides. Clearly myxobacterial genes and actinomycete genes are not, haphazardly, interchangeable. Genes encoding hybrid motifs are not likely to lead to incorrect predictions of secondary metabolite structure but rather to an ambiguity as to the distribution of products or perhaps, even the functionality of the module.

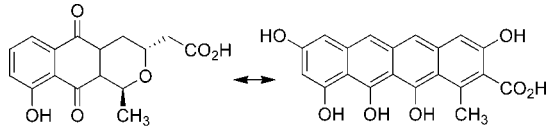


18

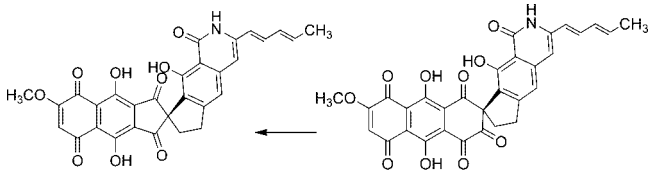
Type II Polyketides

The biosynthetic locus for lactonamycin (**18**) has been cloned and sequenced and together with the results of labeled precursor feeding experiments indicate that the type II PKS is unusual in both its starter unit, glycine or sarcosine, and in that that the cluster includes a thioesterase, not commonly used in type II PKSs.[42]

Classical type II PKSs present a puzzle with respect to chain length, region-specific chain bending, and aromatization, as the determining parameters of these functions have not been fully elucidated. The heterodimeric ketosythase, KS_α/KS_β, is believed to control chain length, and Khosla's group has shown that site-directed mutation of amino acids at the interface between the subunits can interchange the predominant chain lengths obtained by the actinorhodin and tetracenomycin PKS products (**19**).[43]

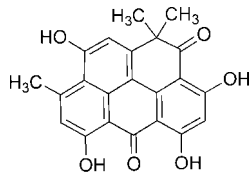


19



20 **21**

The Shen group has cloned and characterized the biosynthetic locus for fredericamycin (**20**) and noted that, atypically, it has two adjacent cysteine residues in the active site of the KS_α subunit.[44] They subsequently identified fredericamycin E (**21**), a key intermediate in the presumed path to forming the bisfurano-carbospiro cyclic structural unit.[45]



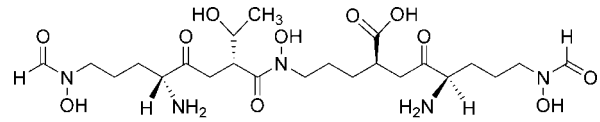
22

Resistomycin (**22**) represents a particular challenge among type II PKSs in that the discoid structure requires an "S"-type folding of the polyketide chain. The biosynthetic locus has been cloned and initial analysis reveals some unusual features.[46] The ACP domain more closely resembles those of fatty acid synthases than other type II PKSs and is separated from the KS_α/KS_β genes by a phosphopantetheinyl transferase gene. Atypically for type II PKSs, there is a malonyl-CoA:ACP transacylase gene in the cluster. Moreover, the AT gene more closely resembles those of modular PKSs than type IIs. Hertweck's and Piel's groups have gone a long way to uncover the features controlling the folding pattern for large type II polyketides. They observed that a phylogenetic analysis of the ketoreductases in these systems grouped them according to the carbon at which the reductase operated and that this corresponded to a folding point in the chain.[47]

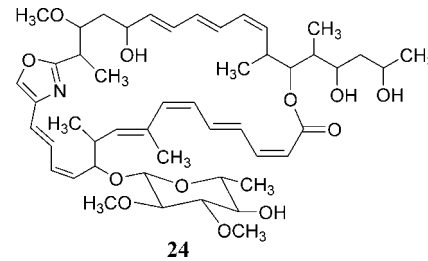
Nonribosomal Peptide Synthetases

Also, NRPSs are not perfectly behaved. The sequence of the first actinomycete to be fully read, *Streptomyces coelicolor*, when annotated, revealed the presence of almost an order of magnitude more biosynthetic loci than were already implicated by secondary metabolite discovery. One new NRPS had *N*-hydroxyamino acids and hence was implicated as an iron chelator, which would presumably be overexpressed in iron-deficient medium. Challis and co-workers[48] deduced the structure from the gene sequence as a tripeptide and proceeded to grow *S. coelicolor* and isolate the metabolite, which, on structural studies turned out to be the tetrapeptide, they named coelichelin (**23**).

This can only be rationalized with the locus if the enzyme complex, having formed the tripeptide, reuses the first module to



23



24

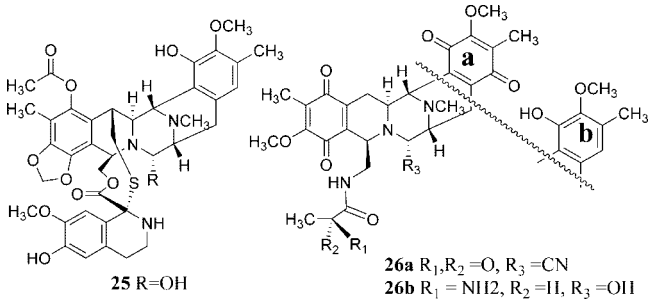
insert another epimerized *ω*-*N*-hydroxyornithine. Note that only the terminal ornithines are epimerized. Not only does this represent iterative use of a module, it is a little more complex than the situations outlined for the PKSs, in that the iterations are separated by passage through other modules.

An analysis of the biosynthetic locus (92 kb) for chivosazol A (**24**) from *S. cellulosum* indicates that this megasynthetase mixture of PKS and PKS/NRPS ORFs is very noncanonical.[49] It contains a single cis AT domain in the first module, with all other modules lacking AT domains, and this function must be supplied in trans and must specify malonate, as feeding experiments indicate that all of the branched methyl groups are derived from methionine. There are methyl transferase-encoding genes in each of the modules required by the structure, with an extra one in the penultimate module, which does not appear to function. There is an apparently nonfunctional dehydratase in module 3, but none in modules 4, 10, and 14 as required by the structure. Both module 4 and module 5 contain adjacent ACP domains. In the former, one would appear to be inactive, lacking a serine in the active site; however, this is not the case in the latter. The ketoreductase domains resulting in the diene in the southwest corner of **24** have the normally diagnostic Asp indicative of the opposite stereochemistry for each of these olefins. Clearly, the biosynthetic sequence of this myxobacteral metabolite does not follow the first-order rules of analysis that have been deduced mainly from actinomycetes. The tools for genetic manipulation of myxobacteria are less well developed than are those for actinomycetes, and the myxobacteria are slower growing, making progress here more arduous. However, their larger genome (up to 13 Mb) and proven capability to produce a wide variety of novel bioactive metabolites ensure that they will be a major source of future discoveries.

The Marine−Microbial Interface

The use of microbial fermentations to access complex structures, which are otherwise difficult to come by, is appealing in that the large-scale production of microbial products has usually been very economic. Natural products isolated from marine macro-organisms often bear a striking resemblance to microbial natural products, and the question arises as to the exact source of the compound: the macro-organism or a comensal microbe. Even if it is the macro-organism, a structural similarity raises the possibility of microbial production of an analogue with subsequent conversion into the original natural product, or tailored mutasynthesis of the microbial product to produce the natural product derived from the marine organism.

The anticancer drug Yondelis (ecteinascidin-743) (**25**), from the ascidian *Ecteinascidia turbinata*, is clearly related structurally to saframycin A (**26a**) and safracin B (**26b**). No microbial origin of ecteinascidin-743 has been identified, but the possibility of producing it, or other potent analogues, from a genetic construct of the saframycin or safracin producers *Streptomyces lavendulae* and

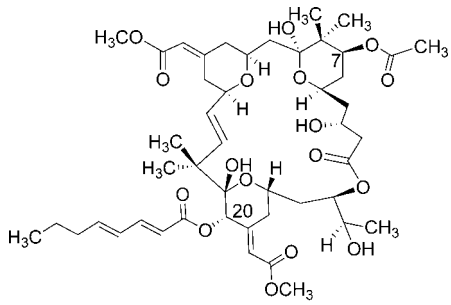


25 R=OH

26a R_1,R_2 =O, R_3 =CN
26b R_1 = NH2, R_2 =H, R_3 =OH

Pseudomonas fluorescens, respectively, either directly or by a combination of biosynthesis and chemical modification, has provided the incentive to characterize these biosynthetic pathways. The biosynthetic locus for saframycin Mx1 was originally described from *Myxococcus xanthus* in 1995.[50] More recently, Lui and colleagues[51] have characterized the saframycin A biosynthetic locus (62 kb) from *S. lavendulae* and shown that the molecule is assembled by a NRPS in which the terminal unit is used iteratively, to attach two molecules of 3-hydroxy-4-methoxy-5-methylphenylalanine to the two initial units, alanine and glycine.

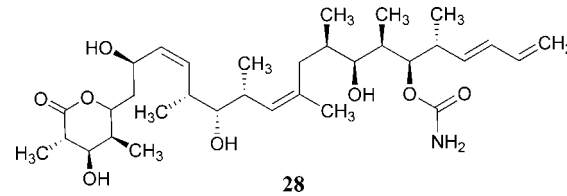
The safracin biosynthetic pathway was discovered earlier by the Pharmamar group[52] and consists of only 17.5 kb. They identified the NRPS nature of the locus but did not definitively define the modules. Although the goal of using this to provide analogues of ecteinascidin-743 is expressed, none are described. The Schmidt group characterized the catechol *O*-methyl transferase, Saf C, from *M. xanthus* and, on the basis of substrate specificity, showed that the biosynthetic pathway for the aromatic amino acid probably proceeded through 4-*O*-methyldopamine before introduction of the 5-methyl group.[53]



27

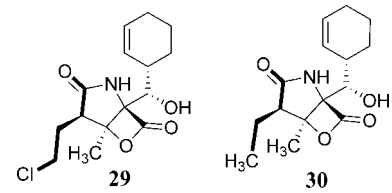
Bryostatin 1 (**27**) is currently still in clinical trials, more than 25 years after it was initially discovered. Much of the delay has been associated with supply problems. It was originally isolated from the bryazoan *Bugula neritina* in very low yield. Both its structure and the variation in yield seen from different harvests of *Bugula* suggested that it was the product of a comensal microbe, and this has proved to be the case, although, to date, the microbe "*Candidatus Endobugula sertula*" has not been cultured. The biosynthetic locus has been isolated from different specimens of *B. neritina*.[54] In two, from shallow collections, the complete locus was contiguous, but in another, from deep water, the PKS genes were separated from the elaborative genes. In earlier work,[55] part of the PKS had been sequenced, and this was used to locate the entire locus (~80 kb) encoding the biosynthesis of the core (i.e., without the appended oxyester groups at 7 and 20) comprising 13 PKS modules, with two trans AT domains in a single ORF. This has been annotated, and several modules appear inactive. The heterologous expression of this very large locus in a high-producing host would be a major step forward in the drug development of bryostatins.

The marine sponge *Discodermia dissoluta* was the original source of the potent anticancer agent discodermolide (**28**),[56] the structure of which strongly suggests that it is the product of a bacterial type I PKS. In an extensive search for the genes encoding the



28

biosynthesis of discodermolide, Hutchinson and collaborators[57] were able to clone, into cosmids, a large number of PKS genes from both the bacterial-enriched and the total sponge DNA from a sample of *D. dissoluta* tissue, but were unable to find the genes for discodermolide biosynthesis.

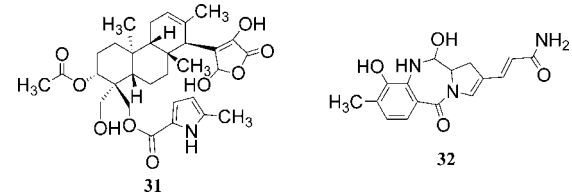


29 **30**

Marine-derived microbes have provided several exciting advances in recent years, and the discovery of novel genera[58] has, not surprisingly, led to secondary metabolites of new structural types and with novel bioactivities. Interesting from a biosynthetic point of view are the 20S proteasome inhibitors salinosporamides A (**29**) and B (**30**). The four-carbon accented portion on the left of the molecules arises from butyrate in the case of salinosporamide B, whereas the analogous carbons definitely did not arise from butyrate but came from glucose, with the carbonyl carbon having a disjointed pathway to the other three carbons, which came from contiguous glucose carbons.[59]

Recent New Methodologies

The sheer size of secondary metabolite biosynthetic loci in actinomycetes, together with the high GC content of their DNA, makes the more simple genetic engineering techniques using *E. coli*−Actinomycete shuttle-vector cosmids and standard restriction-ligation methods less useful as tools for heterologous expression. This has been overcome by the use of recombinogenic cassettes in which two (in theory presumably more) cosmids with overlapping sequences and containing different selection markers are able to, between them, encompass the entire (or near entire) locus. Müller's group used this technique to move the biosynthetic genes for phenalinolactone from *Streptomyces* sp. Tü6071 into *S. lividans* and *S. coelicolor* and obtain expression of the total biosynthtetic pathway. They also obtained the unglycosylated phenalinolactone E (**31**) from an *S. coelicolor* construct in which the glycosyl transferase was deleted.[60] Bachmann's group used this technique to transfer the biosynthetic locus for anthramycin (**32**) from *S. refuineus* into *S. lividans* and to show that, as in the parent organism, this biosynthesis was expressed only at elevated temperatures.[61] Interestingly, the yield of phenalinolactones in Müller's *S. lividans* construct was 100-fold higher at 37 °C than at 28 °C.



31 **32**

The abundance of cryptic (orphan, silent) biosynthetic loci (i.e., loci encoding the biosynthetic machinery for secondary metabolites that had not been detected in fermentations of the parent organism) became apparent early on with the complete genome sequencing of *Streptomyces avermitilis*[62] and *S. coelicolor*,[63] both well-studied

organisms, with almost an order of magnitude more secondary metabolic capability than was known at the time. The use of genome mining to discover novel metabolites from these cryptic pathways has blossomed, and a number of approaches have emerged. The development of bioinformatics systems to mine the genome and predict structures was the basis of the discovery platform at Thallion Pharmaceuticals (formerly Ecopia BioSciences),[64,65] which has led to several novel potential drugs,[7,66−68] of which one is in phase II clinical trials.

Gross and colleagues have pioneered the "genome isotopic approach"[69,70] with an example in which, on the basis of genome analysis, a precursor amino acid labeled with ^{15}N or $^{15}N-^{13}C$ is fed to the culture of an organism containing a cryptic NRPS locus for a cyclic lipopeptide. The low natural abundance of ^{15}N allows H−N HMBC to be used confidently to follow the metabolic fate of the labeled amino acid and isolate the predicted product of the cryptic locus.

Another approach that has received some considerable attention is the search for biosynthetic loci from environmental DNA (eDNA). Banik and Brady point out that with a gram of soil, estimated to contain >10 000 different bacteria, the number of undiscovered secondary metabolites must be considerable. They surveyed eDNA from both Americas and Africa looking for homologues of the OxyC genes, which encodes the very distinctive oxidizing enzyme responsible for the C−C bond linking the hydroxyphenylglycine and the dihydroxyphenylglycine in glycopeptides. From all 11 eDNA samples examined, they were able isolate homologues of OxyC, suggesting that glycopeptides are much more widespread than might have been assumed from the relatively small number that have been discovered, given that in the 1980s a specific screening test and isolation system dependent on Lysyl-D-Ala-D-Ala binding of this class were in broad use throughout the pharmaceutical industry. They built a 10^7-membered cosmid-based megalibrary from a Utah soil sample and sought specific novel glycopeptides by stitching together overlapping cosmids, starting from an OxyC analogue. Two loci were detected from a subset of this libray, one encoding a vancomycin homologue and another, a teichoplanin analogue. The latter contained genes encoding three sulfotranferases. These were cloned and used to generate 6-His-tagged enzymes, which were purified and used to prepare all of the possible mono-, di-, and trisulfated teichoplanin aglycones specified by these three enzymes.[71]

Conclusion

Clearly there is a wealth of new natural products to be discovered from the bacterial kingdom alone. Genetic approaches are helping to uncover these and to use the biosynthetic machinery to build combinatorial libraries, furthering our understanding of SAR within known classes and generating NCEs for broader discoveries.

Although the general unity of biosynthetic processes remains intact and allows knowledge to advance by standing on the shoulders of prior discoveries, the minor variations that Mother Nature begrudgingly reveals continue to fascinate the researcher studying the laws governing secondary metabolite biosynthesis at the genome level. The goal of a definitive understanding of these laws seems to move a little further away as they are approached, but exceptions to early formulated rules lead only to more thorough understanding in the long run.

Note Added after ASAP Publication: The structure of **2** was incorrect in the version posted on February 6, 2009. The correct structure appears after March 27, 2009.

References and Notes

(1) McAlpine, J. B. *Ann. Missouri Bot. Gard.*, in press.
(2) Fischbach, M. A.; Walsh, C. T. *Chem. Rev.* **2006**, *106*, 3468–3496.
(3) Felnagle, E. A.; Jackson, E. E.; Chan, Y. A.; Podevels, A. M.; Berti, A. D.; McMahon, M. D.; Thomas, M. G. *Mol. Pharmaceutics* **2008**, *5*, 191–211.
(4) Van Lanen, S. G.; Shen, B. *Curr. Opin. Drug Discovery Dev.* **2008**, *11*, 186–195.
(5) Donadio, S.; Staver, M. J.; McAlpine, J. B.; Swanson, S. J.; Katz, L. *Science* **1991**, *252*, 675–679.
(6) Haydock, S. F.; Aparicio, J. F.; Molnár, I.; Schwecke, T.; Khaw, L. E.; König, A.; Marsden, A. F. A.; Galloway, I. S.; Staunton, J.; Leadlay, P. F. *FEBS Lett.* **1995**, *374*, 246–248.
(7) McAlpine, J. B.; Bachmann, B. O.; Piraee, M.; Tremblay, S.; Alarco, A.-M.; Zazopoulos, E.; Farnet, C. M. *J. Nat. Prod.* **2005**, *68*, 493–496.
(8) Ohtani, I.; Moore, R. E.; Runnegar, M. T. C. *J. Am. Chem. Soc.* **1992**, *114*, 7941–7942.
(9) Burgoyne, D. L.; Hemscheidt, T. K.; Moore, R. E.; Runnegar, M. T. C. *J. Org. Chem.* **2000**, *65*, 152–156.
(10) Shalev-Alon, G.; Sukenik, A.; Livnah, O.; Schwarz, R.; Kaplan, A. *FEMS Lett.* **2002**, *209*, 87–91.
(11) Mihali, T. K.; Kellmann, R.; Muenchhoff, J.; Barrow, K. D.; Neilan, B. *Appl. Environ. Microbiol.* **2008**, *74*, 716–729.
(12) Rogers, E. W.; Dalisay, D. S.; Molinski, T. F. *Angew. Chem. Int. Ed.* **2008**, *47*, 8086–8089.
(13) Chan, Y. A.; Boyne II, M. T.; Podevels, A. M.; Klimowicz, A. K.; Handelsman, J.; Kelleher, N. L.; Thomas, M. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 14349–14354.
(14) Emmert, E. A. B.; Klimowicz, A. K.; Thomas, M. G.; Handelsman, J. *Appl. Environ. Microbiol.* **2004**, *70*, 104–113.
(15) Ligon, J.; Hill, S.; Beck, J.; Zirkle, R.; Molnár, I.; Zawodny, J.; Money, S.; Schupp, T. *Gene* **2002**, *285*, 257–267.
(16) Carroll, B. J.; Moss, S. J.; Bai, L.; Kato, Y.; Toelzer, S.; Yu, T. W.; Floss, H. G. *J. Am. Chem. Soc.* **2002**, *124*, 4176–4177.
(17) Wu, K.; Chung, L.; Revill, W. P.; Katz, L.; Reeves, C. D. *Gene* **2000**, *251*, 81–90.
(18) Karray, F.; Darbon, E.; Oestreicher, N.; Dominguez, H.; Tuphile, K.; Gagnat, J.; Blondelet-Rouault, M.-H.; Gerbaud, C.; Pernodet, J.-L. *Microbiology* **2007**, *153*, 4111–4122.
(19) Rascher, A.; Hu, Z.; Viswanathan, N.; Schirmer, A.; Reid, R.; Nierman, W. C.; Lewis, M.; Hutchinson, C. R. *FEMS Microbiol. Lett.* **2003**, *218*, 223–230.
(20) Zhao, C.; Ju, J.; Christenson, S. D.; Smith, W. C.; Song, D.; Zhou, X.; Shen, B.; Deng, Z. *J. Bacteriol.* **2006**, *188*, 4142–4147.
(21) Li, W.; Ju, J.; Osada, H.; Shen, B. *J. Bacteriol.* **2006**, *188*, 4148–4152.
(22) Choi, S.-S.; Hur, Y.-A.; Sherman, D. H.; Kim, E.-S. *Microbiology* **2007**, *153*, 1095–1102.
(23) Zhang, H.; White-Phillip, J. A.; Melançon, C. E.; Kwon, H.-j.; Yu, W.-l.; Liu, H.-w. *J. Am. Chem. Soc.* **2007**, *129*, 14670–14683.
(24) Sun, Y.; Hong, H.; Gillies, F.; Spencer, J. B.; Leadlay, P. F. *ChemBioChem* **2008**, *9*, 150–156.
(25) Dorrestein, P. C.; Van Lanen, S. G.; Li, W.; Zhao, C.; Deng, Z.; Shen, B.; Kelleher, N. L. *J. Am. Chem. Soc.* **2006**, *128*, 10386–10387.
(26) Wilkinson, C. J.; Frost, E. J.; Stauton, J.; Leadlay, P. F. *Chem. Biol.* **2001**, *8*, 1197–1208.
(27) Xiang, L.; Moore, B. S. *J. Bacteriol.* **2003**, *185*, 399–404.
(28) Xiang, L.; Kalaitzis, J. A.; Moore, B. S. *Proc. Nat. Acad. Sci. U.S.A.* **2004**, *101*, 15609–15614.
(29) Wilkinson, B.; Foster, G.; Rudd, B. A. M.; Taylor, N. L.; Blackaby, A. P.; Sidebottom, P. J.; Cooper, D. J.; Dawson, M. J.; Buss, A. D.; Gaisser, S.; Böhm, I. U.; Rowe, C. J.; Cortés, J.; Leadlay, P. F.; Stauton, J. *Chem. Biol.* **2000**, *7*, 111–117.
(30) Gaitatzis, N.; Silakowski, B.; Kunze, B.; Nordsiek, G.; Blöker, H.; Hölfe, G.; Müller, R. *J. Biol. Chem.* **2002**, *277*, 13082–13090.
(31) Olano, C.; Wilkinson, B.; Moss, S. J.; Brana, A. F.; Mendez, C.; Leadlay, P. F.; Salas, J. A. *Chem. Commun.* **2003**, 2780–2782.
(32) He, J.; Hertweck, C. *ChemBioChem* **2005**, *6*, 908–912.
(33) Traitcheva, N.; Jenke-Kodama, H.; He, J.; Dittmann, E.; Hertweck, C. *ChemBioChem.* **2007**, *8*, 1841–1849.
(34) Müller, M.; Kusebauch, B.; Liang, G.; Beaudry, C. M.; Trauner, D.; Hertweck, C. *Angew. Chem., Int. Ed.* **2006**, *45*, 7835–7838.
(35) Zhao, Q.; He, Q.; Ding, W.; Tang, M.; Kang, Q.; Yu, Y.; Deng, W.; Zhang, Q.; Fang, J.; Tang, Q.; Liu, W. *Chem. Biol.* **2008**, *15*, 693–705.
(36) Farnet, C. M.; Zazopoulos, E.; Staffa, A.; Yang, X. Eur. Pat. EP1381685, 2004.
(37) Cheng, Y.-Q.; Tang, G.-L.; Shen, B. *Proc. Nat. Acad. Sci.* **2003**, *100*, 3149–3154.
(38) Nguyen, T.; Ishida, K.; Jenke-Kodama, H.; Dittmann, E.; Gurgui, C.; Hochmuth, T.; Taudien, S.; Platzer, M.; Hertweck, C.; Piel, J. *Nat. Biotechnol.* **2008**, *26*, 225–233.
(39) Thomas, I.; Martin (née Rowe), C. J.; Wilkinson, C. J.; Stauton, J.; Leadlay, P. F. *Chem. Biol.* **2002**, *9*, 781–787.
(40) Kittendorf, J. D.; Beck, B. J.; Buchholz, T. J.; Seufert, W.; Sherman, D. *Chem. Biol.* **2007**, *14*, 944–954.
(41) Petković, H.; Aandmann, A.; Challis, I. R.; Hecht, H.-J.; Silakowski, B.; Low, L.; Beeston, N.; Kuščer, E.; Garcia-Bernardo, J.; Leadlay, P. F.; Kendrew, S. G.; Wilkinson, B.; Müller, R. *Org. Biomol. Chem.* **2008**, *6*, 500–506.

(42) Zhang, X.; Alemany, L. B.; Fiedler, H.-P.; Goodfellow, M.; Parry, R. J. *Antimicrob. Agents Chemother.* **2008**, *52*, 574–585.
(43) Tang, Y.; Tsai, S. C.; Khosla, C. *J. Am. Chem. Soc.* **2003**, *125*, 12708–12709.
(44) Wendt-Pienkowski, E.; Huang, Y.; Zhang, J.; Li, B.; Jiang, H.; Kwon, H.; Hutchinson, C. R.; Shen, B. *J. Am. Chem. Soc.* **2005**, *127*, 16442–16452.
(45) Chen, Y.; Luo, Y.; Ju, J.; Wendt-Pienkowski, E.; Rajski, S. R.; Shen, B. *J. Nat. Prod.* **2008**, *71*, 431–437.
(46) Jakobi, K.; Hertweck, C. *J. Am. Chem. Soc.* **2004**, *126*, 2298–2299.
(47) Lackner, G.; Schenk, A.; Xu, Z.; Reinhardt, K.; Yunt, Z. S.; Piel, J.; Hertweck, C. *J. Am. Chem. Soc.* **2007**, *129*, 9306–9312.
(48) Lautru, S.; Deeth, R. J.; Bailey, L. M.; Challis, G. L. *Nat. Chem. Biol.* **2005**, *1*, 265–269.
(49) Perlova, O.; Gerth, K.; Kaiser, O.; Hans, A.; Müller, R. *J. Biotechnol.* **2006**, *121*, 174–191.
(50) Pospiech, A.; Cluzel, B.; Bietenhader, J.; Schupp, T. *Microbiology* **1995**, *141*, 1793–1803.
(51) Li, L.; Deng, W.; Song, L.; Ding, W.; Zhao, Q.-F.; Peng, C.; Song, W.-W.; Tang, G.-L.; Liu, W. *J. Bacteriol.* **2008**, *190*, 251–263.
(52) Velasco, A.; Acebo, P.; Gomez, A.; Schleissner, C.; Rodriguez, P.; Aparicio, T.; Conde, S.; Muñoz, R.; de la Calle, F.; Garcia, J. L.; Sánchez-Puelles, J. M. *Mol. Microbiol.* **2005**, *56*, 144–154.
(53) Nelson, J. T.; Lee, J.; Sims, J. W.; Schmidt, E. W. *Appl. Environ. Microbiol.* **2007**, *73*, 3575–3580.
(54) Sudek, S.; Lopanik, N. B.; Waggoner, L. E.; Hildebrand, M.; Anderson, C.; Liu, H.; Patel, A.; Sherman, D. H.; Haygood, M. G. *J. Nat. Prod.* **2007**, *70*, 67–74.
(55) Hildebrand, M.; Waggoner, L. E.; Liu, H.; Sudek, S.; Allen, S.; Anderson, C.; Sherman, D. H.; Haygood, M. *Chem. Biol.* **2004**, *11*, 1543–1552.
(56) Gunsakera, S. P.; Gunsakera, M.; Longley, R. E.; Schulte, G. K. *J. Org. Chem.* **1990**, *55*, 4912–4916.
(57) Schirmer, A.; Gadkari, R.; Reeves, C. D.; Ibrahim, F.; DeLong, E. F.; Hutchinson, C. R. *Appl. Environ. Microbiol.* **2005**, *71*, 4840–4849.
(58) Maldonado, L. A.; Fenical, W.; Jensen, P. R.; Kaufmann, C. A.; Mincer, T. J.; Ward, A. C.; Bull, A. T.; Goodfellow, M. *Int. J. Syst. Evol. Microbiol.* **2005**, *55*, 1759–1766.
(59) Beer, L. L.; Moore, B. S. *Org. Lett.* **2007**, *9*, 845–848.
(60) Binz, T. M.; Wenzel, S. C.; Schnell, H.-J.; Bechthold, A.; Müller, R. *ChemBioChem* **2008**, *9*, 447–454.
(61) Hu, Y.; Phelan, V. V.; Farnet, C. M.; Zazopoulos, E.; Bachmann, B. O. *ChemBioChem* **2008**, *9*, 1603–1608.
(62) Ōmura, S.; Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Takahashi, C.; Shinose, M.; Takahashi, Y.; Horikawa, H.; Nakazawa, H.; Osonoe, T.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M. *Proc. Nat. Acad. Sci. U.S.A.* **2001**, *98*, 12215–12220.
(63) Bentley, S. D.; Chater, K. F.; Cerdeñeno-Tárraga, A.-M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C.-H.; Kieser, T.; Larke, L.; Murphy, L.; Oliver, K.; O'Neil, S.; Rabbinowitsch, E.; Rajadream, M.-A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrel, B. G.; Parkhill, J.; Hopwood, D. A. *Nature* **2002**, *417*, 141–147.
(64) Zazopoulos, E.; Huang, K.; Staffa, A.; Lui, W.; Bachmann, B. O.; Nonaka, K.; Ahlert, J.; Thorson, J. S.; Shen, B.; Farnet, C. M. *Nat. Biotechnol.* **2003**, *2*, 187–190.
(65) Farnet, C. M.; Zazopoulos, E. In *Natural Products: Drug Discovery and Therapeutic Medicine*; Zhang, L., Demain, A. L., Eds.; Humana Press Inc.: Totowa, NJ, 2005; pp 95−106.
(66) Banskota, A. H.; McAlpine, J. B.; Sørensen, D.; Aouidate, M.; Piraee, M.; Alarco, A.-M.; Ômura, S.; Shiomi, K.; Farnet, C. M.; Zazolpoulos, E. *J. Antibiot.* **2006**, *59*, 168–176.
(67) Banskota, A. H.; McAlpine, J. B.; Sørensen, D.; Ibrahim, A.; Aouidate, M.; Piraee, M.; Alarco, A.-M.; Farnet, C. M.; Zazopoulos, E. *J. Antibiot.* **2006**, *59*, 533–542.
(68) McAlpine, J. B.; Banskota, A. H.; Charan, R. D.; Schlingmann, G.; Zazopoulos, E.; Piraee, M.; Janso, J.; Bernan, V. S.; Aouidate, M.; Farnet, C. M.; Feng, X.; Zhao, Z.; Carter, G. T. *J. Nat. Prod.* **2008**, *71*, 1585–1590.
(69) Gross, H.; Stockwell, V. O.; Henkels, M. D.; Nowak-Thompson, B.; Loper, J. E.; Gerwick, W. H. *Chem. Biol.* **2007**, *14*, 53–63.
(70) Gross, H. *Appl. Microbiol. Biotechnol.* **2007**, *75*, 267–277.
(71) Banik, J. J.; Brady, S. F. *Proc. Nat. Acad. Sci. U.S.A.* **2008**, *105*, 17273–17277.

NP800742Z