anthranilate, a common intermediate from the shikimate pathway (Fig. 2E). The colocalization of SgcD and SgcD1 along with the other C-1027 production genes assures the availability of anthranilate for secondary metabolite biosynthesis. Although it remains unclear what the origin of the C_3 unit is and how it is fused to the anthranilate intermediate to form the morpholinone moiety of **7**, the latter is apparently activated as an acyl-*S*-CoA for attachment to **4** by SgcD6 (Fig. 2B). To support this hypothesis, we inactivated *sgcD6* (see SOM). The resultant *S. globisporus* SB1004 mutant strain completely lost its ability to produce **1** (Fig. 3F). The fact that the biosynthetic building blocks are activated as aminoacy-*S*-PCP, acyl-*S*-CoA, and nucleotide diphosphosugar; and are attached to the enediyne core by an NRPS condensation enzyme, an acyltransferase, and a glycosyl transferase, respectively, highlights nature's efficiency and versatility in synthesizing complex molecules.

Finally, we inactivated the *sgcC* hydroxylase gene to demonstrate the production of enediyne metabolites by manipulating genes governing C-1027 biosynthesis (see SOM). The resulting *S. globisporus* SB1006 mutant strain still produces a chromoprotein complex that is biologically active, as judged by bioassay against *M. luteus*, but is distinct from **1** upon HPLC analysis (Fig. 3G). The resulting enediyne chromophores were isolated (see SOM) and subjected to ESI-MS analysis. **2** exhibited an $(M + H)^+$ ion at $m/z = 828$, consistent with the molecular formula $C_{43}H_{42}N_3O_{12}Cl$, and **8** showed an $(M + H)^+$ ion at $m/z = 830$, consistent with the molecular formula $C_{43}H_{44}N_3O_{12}Cl$. By comparison with **1** and **3**, **2** and **8** were deduced to be deshydroxy-C-1027 (**2**) and its aromatized product (**8**) (Fig. 1), as would be predicted according to Fig. 2D. Intriguingly, **2** is at least fivefold more stable than is **1** at 25°C, in respect to undergoing the Bergman cyclization (see SOM), a property that could be potentially explored in developing C-1027 into a clinically useful drug. We envisage applying methods of combinatorial biosynthesis to the enediyne system for the production of polyketides.

References and Notes

1. T. Otani, Y. Minami, T. Marunaka, R. Zhang, M.-Y. Xie, *J. Antibiot.* **41**, 1580 (1988).
2. K. C. Nicolaou, W.-M. Dai, *Angew. Chem. Int. Ed. Engl.* **30**, 1387 (1991).
3. Z. Xi, I. H. Goldberg, in *Comprehensive Natural Products Chemistry*, D. Barton, K. Nakanish, O. Meth-Cohn, Eds. (Elesvier, New York, 1999), vol. 7, pp. 553–592.
4. I. Brukner, *Curr. Opin. Oncol. Endocr. Met. Invest. Drugs* **2**, 344 (2000).
5. J. S. Thorson *et al.*, *Curr. Pharm. Design* **6**, 1841 (2000).
6. H. Maeda, T. Konno, in *Neocarzinostatin: The Past, Present, and Future of an Anticancer Drug*, H. Maeda, K. Edo, N. Ishida, Eds. (Springer-Verlag, New York, 1997), pp. 227–267.
7. E. L. Sievers *et al.*, *Blood* **93**, 3678 (1999).
8. D. E. Cane, C. T. Walsh, C. Khosla, *Science* **282**, 63 (1998).
9. L. Du, B. Shen, *Curr. Opin. Drug Discov. Dev.* **4**, 215 (2001).
10. J. Staunton, K. J. Weissman, *Nat. Prod. Rep.* **18**, 380 (2001).
11. B. Shen, *Top. Curr. Chem.* **209**, 1 (2000).
12. W. Liu, B. Shen, *Antimicrobiol. Agents Chemother.* **44**, 382 (2000).
13. K.-I. Yoshida, Y. Minami, R. Azuma, M. Saeki, T. Otani, *Tetrahedron Lett.* **34**, 2637 (1993).
14. K.-I. Iida *et al.*, *Tetrahedron Lett.* **37**, 4997 (1996).
15. N. Funa *et al.*, *Nature* **400**, 897 (1999).
16. B. S. Moore, J. N. Hopke, *ChemBioChem* **2**, 35 (2001).
17. O. D. Hensens, J.-L. Giner, I. H. Goldberg, *J. Am. Chem. Soc.* **111**, 3295 (1989).
18. Y. Tokiwa *et al.*, *J. Am. Chem. Soc.* **114**, 4107 (1992).
19. K. S. Lam *et al.*, *J. Am. Chem. Soc.* **115**, 12340 (1993).
20. J. Ahlert *et al.*, *Science* **297**, 1173 (2002).
21. The neocarzintostatin cluster was cloned, sequenced, and characterized from *Streptomyces carzinostaticus* ATCC15944 by W. Liu *et al.*, unpublished data.
22. M. Lee *et al.*, *Science* **280**, 915 (1998).
23. J. Kennedy *et al.*, *Science* **284**, 1368 (1999).
24. J. G. Metz *et al.*, *Science* **293**, 290 (2001).
25. T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem. Biol.* **6**, 493 (1999).
26. G. L. Challis, J. Ravel, C. A. Townsend, *Chem. Biol.* **7**, 211 (2000).
27. H. Chen, B. K. Hubbard, S. E. O'Connor, C. T. Walsh, *Chem. Biol.* **9**, 103 (2002).

28. We thank Y. Li, Institute of Medicinal Biotechnology, Chinese Academy of Medical Sciences, Beijing, China, for the wild-type *S. globisporus* strain; and Ecopia BioSciences Inc., Montreal, Canada, for assistance in sequencing the neocarzinostatin gene cluster and for sharing unpublished data. Supported in part by NIH grant CA78747. B.S. is a recipient of an NSF CAREER Award (MCB9733938) and an NIH Independent Scientist Award (AI51689). S.D.C. was supported in part by NIH grant T32 GM07377.

Supporting Online Material

www.sciencemag.org/cgi/content/full/297/5584/1170/DC1
Materials and Methods
Fig. S1

21 March 2002; accepted 10 June 2002

The Calicheamicin Gene Cluster and Its Iterative Type I Enediyne PKS

Joachim Ahlert,[1] Erica Shepard,[1] Natalia Lomovskaya,[1] Emmanuel Zazopoulos,[2] Alfredo Staffa,[2] Brian O. Bachmann,[2] Kexue Huang,[2] Leonid Fonstein,[1] Anne Czisny,[1] Ross E. Whitwam,[1] Chris M. Farnet,[2]* Jon S. Thorson[1]*

The enediynes exemplify nature's ingenuity. We have cloned and characterized the biosynthetic locus coding for perhaps the most notorious member of the nonchromoprotein enediyne family, calicheamicin. This gene cluster contains an unusual polyketide synthase (PKS) that is demonstrated to be essential for enediyne biosynthesis. Comparison of the calicheamicin locus with the locus encoding the chromoprotein enediyne C-1027 reveals that the enediyne PKS is highly conserved among these distinct enediyne families. Contrary to previous hypotheses, this suggests that the chromoprotein and nonchromoprotein enediynes are generated by similar biosynthetic pathways.

Calicheamicin γ_1^I (Fig. 1) from *Micromonospora echinospora* ssp. *calichensis* is the most prominent member of the nonchromoprotein enediyne family and, as an antitumor agent, is > 5000 times as potent as adriamycin (*1*). Calicheamicin has two distinct structural regions: The aryltetrasaccharide is composed of a set of carbohydrate and aromatic units, which delivers the metabolite specifically into the minor groove of DNA (*2*); the aglycone, or "warhead," consists of a highly functionalized bicyclo[7.3.1]tridecadiynene core structure with an allylic trisulfide serving as the initial trigger for warhead cycloaromatization (*3*). Once the aryltetrasaccharide is docked to DNA, aromatization of the bicyclo[7.3.1]tridecadiynene core structure, by means of a 1,4-dehydrobenzene-diradical, results in the site-specific oxidative double-strand scission of the targeted DNA (*4*). In vitro and in vivo studies confirm the role of calicheamicin as a DNA-damaging agent and even suggest that calicheamicin may favor cleavage at certain chromosomal sites (*5*). Although this extraordinary reactivity has sparked considerable interest in the pharmaceutical industry, calicheamicin lacks tumor specificity. To circumvent this problem, conjugation of calicheamicin to tumor-specific monoclonal antibodies has been used as a targeting strategy. The recent success of calicheamicin-CD33 antibody conjugates (Mylotarg) to treat acute myelogenous leukemia

[1]Laboratory for Biosynthetic Chemistry, Pharmaceutical Sciences Division, School of Pharmacy, University of Wisconsin-Madison, 777 Highland Avenue, Madison, WI 53705, USA. [2]Ecopia Biosciences Inc., 7290 Frederick-Banting, Montreal, Quebec, Canada H4S 2A1.

*To whom correspondence should be addressed. E-mail: farnet@ecopiabio.com; jthorson@pharmacy.wisc.edu

(AML) (*6*) points to therapeutic promise for calicheamicin.

Although innovative synthetic schemes have been devised for calicheamicin (*7*), surprisingly little is known about how the warhead is synthesized in nature. Our understanding of calicheamicin biosynthesis is limited to a set of metabolic labeling experiments performed on the related enediyne esperamicin (Fig. S1, compound **5**) (*8*). We report the elucidation of the calicheamicin gene locus from *Micromonospora echinospora* ssp. *calichensis*, which reveals an unusual iterative polyketide synthase (PKS) gene (*calE8*). Disruption of this PKS gene in the calicheamicin-producer *Micromonospora* provides a strain no longer capable of producing calicheamicin and suggests that this PKS is critical for enediyne formation. A homolog of this PKS has also been found in the C-1027 gene cluster encoding a member of the structurally distinct chromoprotein enediynes [see companion paper (*9*)]. In contrast to previous biosynthetic hypotheses (*10*), this finding suggests that all enediynes may share a common biosynthetic origin. This work also lays the foundation for future metabolic engineering strategies to generate enediyne analogs.

Using a combination of polymerase chain reaction (PCR)–based screening and screens for clones capable of conferring calicheamicin resistance (*11*), nine overlapping cosmid clones were isolated from a genomic library of *M. echinospora* ssp. *calichensis* (NRRL 15839). Although sequence analysis of these clones was consistent with genes encoding aryltetrasaccharide biosynthesis and calicheamicin resistance, initial attempts to complete the biosynthetic locus by chromosomal walking were unsuccessful (*12*). To circumvent this problem, we used a shotgun-based approach to finish the calicheamicin locus. DNA sequence analysis revealed 74 open reading frames (ORFs) that span more than 90 kb (Fig. 2). They include 16 genes expected to participate in polyketide construction or modification (*O1* to *O6* and *E1* to *E10*), nine putative regulatory elements (*R1* to *R9*), seven genes associated with membrane transport (*T1* to *T7*), 14 genes consistent with the expected production of four unusual activated nucleotide sugars (*S1* to *S14*), four glycosyltransferase genes (*G1* to *G4*), an insertional element (IS), and 22 ORFs (*U1* to *U22*) of unknown function (*13*). The known resistance gene *calC* (*11*) was also found to reside near the middle of this locus.

Given the long-standing controversy surrounding the biosynthetic origin of the enediyne moiety, the polyketide-associated genes within the calicheamicin gene cluster are most relevant to our hypothesis. The two distinct calicheamicin structural elements—the calicheamicin warhead (Fig. 1, PKS E) and the orsellinic acid derivative found within the aryltetrasaccharide (Fig. 1, PKS O)—are each expected to derive from a separate polyketide precursor. Consistent with this, the calicheamicin locus (Fig. 2, genes *calE8* and *calO5*) encodes two separate iterative type I PKSs (CalE8 and CalO5). CalO5 shows striking similarity (65% identity, 93% similarity) to the PKS (AviM) responsible for orsellinic acid biosynthesis in *S. viridochromogenes* en route to the biosynthesis of the orthosomycin antibiotic avilamycin A (*14*). Thus, CalO5 likely plays an identical role in the biosynthesis of the calicheamicin aryltetrasaccharide. In contrast, CalE8 demonstrates the greatest sequence homology to two iterative PKSs involved in the biosynthesis of the polyunsaturated fatty acids (PUFAs) eicosapentaenoic acid (EPA) in *Shewanella* (*15, 17*) and docosahexaenoic acid (DHA) in *Moritella marina* (*16, 17*).

To confirm that CalE8 is critical to the biosynthesis of calicheamicin in *Micromonospora*, a *calE8* gene disruption was constructed by introduction of an apramycin selectable marker (Fig. 3A). The calicheamicin producer *M. echinospora* is resistant to most antibiotics, and is inherently indisposed to genetic manipulation, with a best reported transformation efficiency of 10^{-7} transformants per microgram DNA (*18*). In spite of these difficulties, a combination of many gene disruption attempts in *M. echinospora* LL6000 (*19*) led to nine independent apramycin-resistant clones. All nine isolates mapped consistently with the expected *calE8* gene disruption by both PCR fragment amplification and Southern hybridization. The representative hybridization results from two of these clones (pJAW117/LL6000-1 and pJAW117/LL6000-3) are illustrated in Fig. 3B. All nine *calE8* disruption mutants and two representative LL6000 isolates were sub-

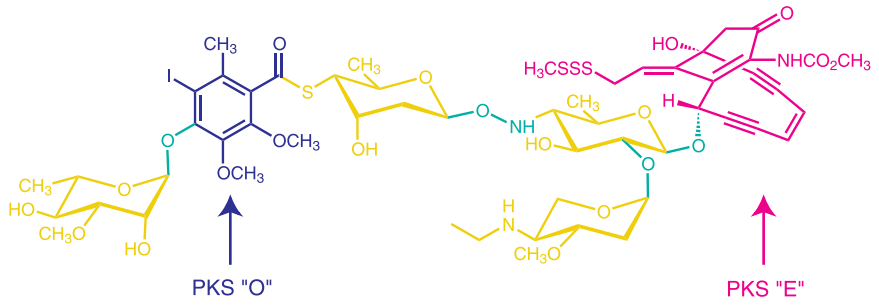


Fig. 1. The structure of the nonchromoprotein enediyne calicheamicin. The enediyne families are typically distinguished by either the number of carbons in the enediyne ring (9-membered versus 10-membered) or by their association, or lack thereof, with an apoprotein (chromoprotein versus nonchromoprotein). With one exception (N1999A), the 9-membered enediynes are chromoprotein enediynes, whereas all 10-membered enediynes to date are nonchromoprotein. The colors correlate the calicheamicin final structure to the functions of genes within the calicheamicin locus (Fig. 2) with the specific PKS structural components designated as PKS E and PKS O.

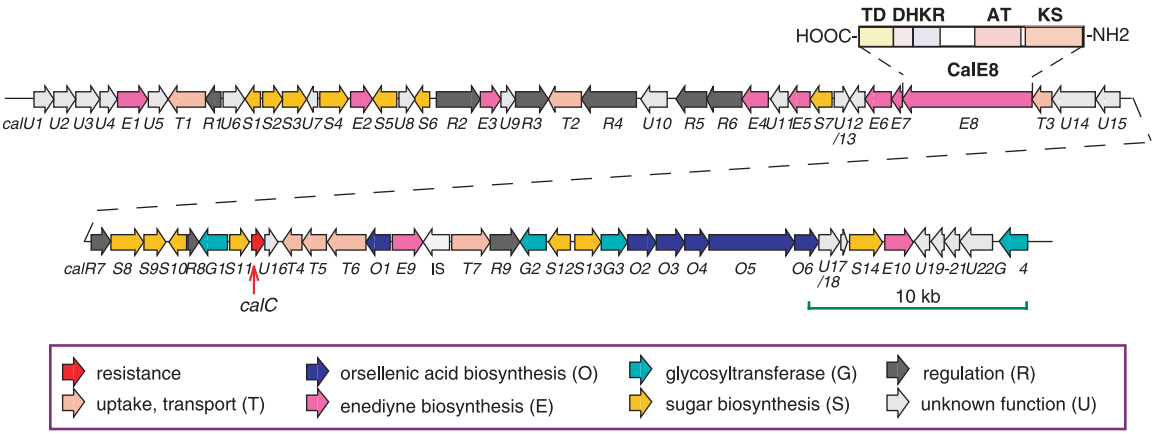
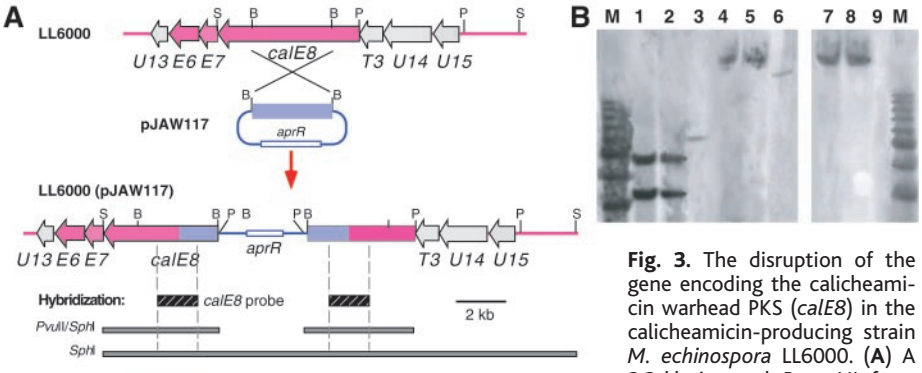


Fig. 2. The calicheamicin locus from *Micromonospora echinospora* spp. *calichensis*. The colors of genes delineate putative roles in calicheamicin biosynthesis, based on BLAST analysis and correlate by color to the calicheamicin structure presented in Fig. 1. The corresponding *calE8* translation, CalE8, is also shown to highlight the location of the main domains KS, AT, KR, DH, and TD.

sequently tested in parallel for calicheamicin production. Extracts from these strains were analyzed by (i) the biochemical induction assay, a modified prophage induction assay used in the original discovery of the calicheamicins (*20*); (ii) the molecular break light assay, a DNA-cleavage assay based on intramolecular fluorescence quenching optimized for DNA-cleavage by enediynes (in which femtomolar calicheamicin concentrations are detectable) (*21*); and (iii) high-performance liquid chromatography (HPLC). Paternal LL6000 isolates produced 0.5 to 0.8 mg/liter calicheamicin, as expected. In contrast, the mutant strains with the *calE8* gene disrupted were both devoid of any calicheamicin, known calicheamicin derivatives, and/or enediyne activity by all three methods of detection (Fig. 3C). Because disruption of PKS O machinery is anticipated to provide a structure lacking PKS O and the terminal rhamnose (similar to that of namenamicin, Fig. S1, compound **6**) (*22*), the result of our *calE8* disruption is consistent with elimination of the PKS machinery involved in warhead construction (PKS E) and supports our assignment of CalE8 as this essential entity.



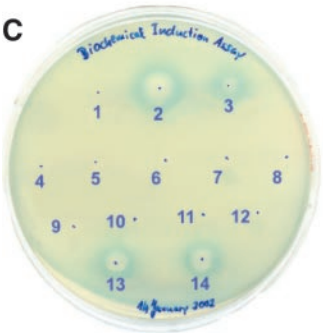


Fig. 3. The disruption of the gene encoding the calicheamicin warhead PKS (*calE8*) in the calicheamicin-producing strain *M. echinospora* LL6000. (**A**) A 3.3-kb internal Bam HI fragment of *calE8* (purple) was ligated into pOJ260 (*24*) to give plasmid pJAW117. This disruption construct was then transformed into LL6000 protoplasts by a slight modification of established methods. Proper integration of pJAW117, and *calE8* disruption in LL6000, was confirmed by both PCR with primers adjacent to the cloned 3.3-kb fragment (*25*) and Southern hybridization with a 1.8 = kb internal *calE8* DNA fragment to give the anticipated Pvu II–Sph I and Sph I fragments shown (highlighted as dark grey bars). (**B**) For Southern hybridization, chromosomal DNA of mutants (pJAW117/LL6000-1 and pJAW117/LL6000-3 as the two representative samples shown) and paternal LL6000 DNA was digested with Pvu II–Sph I (lanes 1 to 3) and Sph I (lanes 4 to 9), respectively. Lanes 1 to 6 were hybridized with the 1.8 = kb internal fragment of *calE8* (lanes 1 to 6). In a separate experiment, lanes 7 to 9 were hybridized with an 0.8-kb internal Sac I fragment of the apramycin resistance gene *apr*^{R}. Hybridization with the *calE8* probe revealed the integration of plasmid pJAW117 at the correct location in the genomes of all nine knockout mutants as represented by pJAW117/LL6000-1 and pJAW117/LL6000-3 (lanes 1 and 2 and 4 and 5, respectively) in comparison with the paternal genome (lanes 3 and 6), whereas hybridization with the *apr*^{R} probe showed integration only in the mutants (lanes 7 and 8), but not in the paternal LL6000 genome (lane 9). (**C**) One (the BIA) of three methods used to detect calicheamicin production in mutant and paternal strains. In this assay (*20*), calicheamicin provides both a zone of growth inhibition and the blue rings are indicative of DNA damage in the modified prophage indicator strain *E. coli* BR513. Zones 1 to 3 are controls in which zone 1 contained only 1 μl of CH_3CN (the solvent used to deliver calicheamicin), whereas zones 2 and 3 contained 1 μl of different calicheamicin/CH_3CN standards (100 ng and 10 ng, respectively). As revealed in zones 4 to 12, calicheamicin production was lacking in all nine pJAW117/LL6000 *calE8* gene disruption mutants. As a separate control, a 1-μl sample of extracts from two representative isolates from the paternal LL6000 strains clearly revealed calicheamicin production (samples 13 and 14). These results were also confirmed by the MBL assay (*21*) and HPLC analysis of fermentation extracts.

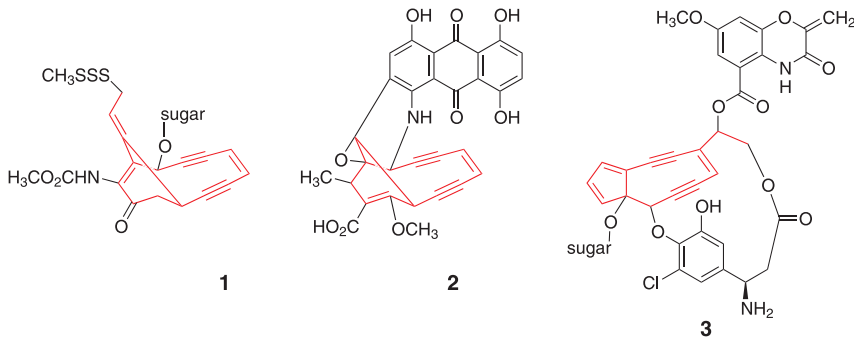


Fig. 4. A comparison of the warheads from calicheamicin (compound **1**), dynemicin (compound **2**) and C-1027 (compound **3**). The common cyclododecylpolyene skeleton is highlighted in red.

Full-length CalE8 is 1919 amino acids in length (Fig. 2). Comparing CalE8 with the known domains of both type I PKS protypes (*23*) and the PUFA PKSs (which include domains for ketosynthase, KS; acyltransferase, AT; acyl carrier protein, ACP; ketoreductase, KR; and dehydratase, DH), we find domains as follows: A CalE8 KS domain spans positions 3 to 467; an AT domain spans positions 482 to 905; a KR domain spans positions 1153 to 1414 and a DH domain spans positions 1422 to 1563. Given the similarity in organization between CalE8 and the PUFA PKSs, the region bridging the AT and KR domains might harbor a putative ACP domain. Yet, in comparison with either the PUFA PKSs or other iterative type I PKSs, there are few obvious ACP markers within this CalE8 domain 1, and thus, this assignment is speculative. The COOH-terminal domain of CalE8 (positions 1582 to 1919, designated TD) also lacks homology to any known PKS or protein.

Interestingly, CalE8 shows remarkable similarity (56% identity and 67% similarity, overall) in both sequence and organization to an enediyne PKS (SgcE) from the C-1027 biosynthetic gene locus (*9*). C-1027 (Fig. S1, compound **4**) is a member of structurally distinct chromoprotein enediyne family, previously postulated to derive from a fatty acid precursor (*10*). In both CalE8 and SgcE, all six domains are evident and four of these six (KS, AT, KR, and DH) are highly conserved. Significant homology divergence is observed in the region bridging the AT and KR domains, a phenomenon common to type I PKSs (*23*), and the COOH-terminal region (TD) (Fig. 4) (*9*). In contrast to previous hypotheses, the observed CalE8/SgcE relationship clearly implicates a common PKS progenitor for both the chromoprotein (C-1027) and nonchromoprotein (calicheamicin) enediynes. In further support of this hypothesis, inspection of representative chromoprotein (C-1027) and nonchromoprotein (dynemicin and calicheamicin) warheads reveals a common cyclododecylpolyene skeleton (Fig. 4). This similarity in structure suggests that the biosynthesis of all enediyne warheads may proceed via similar polyunsaturated

polyketide intermediates constructed by the highly conserved warhead PKS. This remarkable CalE8/SgcE relationship establishes a new paradigm for the enediyne PKS, distinct from other known bacterial type I interative PKSs and opens the door for bioengineering novel enediyne variants.

References and Notes

1. J. S. Thorson *et al*., *Curr. Pharm. Des.* **6**, 1841 (2000).
2. N. Zein, A. M. Sinha, W. J. McGaharen, G. A. Ellestad, *Science* **240**, 1198 (1988).
3. A. G. Myers, S. B. Cohen, B. M. Kwon, *J. Am. Chem. Soc.* **116**, 1255 (1994).
4. J. J. DeVoss, J. J. Hangeland, C. A. Townsend, *J. Am. Chem. Soc.* **112**, 4554 (1990).
5. C. M. H. Watanabe, L. Supekova, P. G. Schultz, *Chem. Biol.* **9**, 245 (2002).
6. E. L. Sievers, M. Linenberger, *Curr. Opin. Oncol.* **13**, 522 (2001).
7. S. J. Danishefsky, M. D. Shair, *J. Org. Chem.* **61**, 16 (1996).
8. K. S. Lam *et al*., *J. Am. Chem. Soc.* **115**, 12340 (1993).
9. W. Liu, S. D. Christenson, S. Standage, B. Shen, *Science* **247**, 1170 (2002).
10. O. D. Hensens, J.-L. Giner, I. H. Goldberg, *J. Am. Chem. Soc.* **111**, 3295 (1989).
11. R. W. Whitwam, J. Ahlert, T. R. Holman, M. Ruppen, J. S. Thorson, *J. Am. Chem. Soc.* **122**, 1556 (2000).
12. Most likely owing to the toxic nature of some of the gene products when cloned into *Escherichia coli* (J. Ahlert *et al*., unpublished observations).
13. The sequence is submitted under GenBank Accession numbers AF505622 and AF497482.
14. G. Weitnauer *et al*., *Chem. Biol.* 8, 569 (2001).
15. H. Takeyama, D. Takeda, K. Yazawa, A. Yamada, T. Matsunaga, *Microbiology* **143**, 2725 (1997).
16. N. Morita, M. Tanaka, H. Okuyama, *Biochem. Soc. Trans.* **28**, 943 (2000).
17. J. G. Metz *et al*., *Science* **293**, 290 (2001).
18. S. F. Love, W. M. Maiese, D. M. Rothstein, *Appl. Environ. Microbiol.* **58**, 1376 (1992).
19. Strain LL6000 is an optimized production strain provided by Wyeth Research which, under typical growth conditions, provides ~0.8 mg/liter of the calicheamicins.
20. M. Greenstein, T. Monji, R. Yeung, W. M. Maiese, R. J. White, *Antimicrob. Agents Chemother.* **29**, 861 (1986).
21. J. B. Biggins, J. R. Prudent, D. J. Marshall, M. Ruppen, J. S. Thorson, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13537 (2000).
22. Using the HPLC conditions described in the supporting material on *Science* Online, the namenamicin standard has a retention time of 18.3 min. This compound also displays strong activity in both the BIA (biochemical lambda prophage induction) and the MBL ("molecular break lights") assays (J. B. Biggins, unpublished data).
23. J. Staunton, K. J. Weissman, *Nat. Prod. Rep.* **18**, 380 (2001).
24. M. Bierman *et al*., *Gene* **116**, 43 (1992).
25. J. Ahlert *et al*., unpublished observations.

26. We gratefully acknowledge Wyeth Research for graciously providing materials and P. R. Hamann for helpful discussion. J.S.T. is a Rita Allen Foundation Scholar and Alfred P. Sloan Fellow. This work was supported in part by grants from the National Institutes of Health (CA84374 and GM58196), the Mizutani Foundation for Glycoscience, and a Cancer Center Support Grant (CA08748). We also thank the Charles A. Dana and Norman and Rosita Winston Foundations for postdoctoral fellowship support (J.A. and R.E.W.).

Supporting Online Material
www.sciencemag.org/cgi/content/full/297/5584/1173/DC1
Materials and Methods
Fig. S1

21 March 2002; accepted 10 June 2002

Structures of Glycoprotein Ibα and Its Complex with von Willebrand Factor A1 Domain

Eric G. Huizinga,[1]*† Shizuko Tsuji,[2]* Roland A. P. Romijn,[2] Marion E. Schiphorst,[2] Philip G. de Groot,[2] Jan J. Sixma,[2] Piet Gros[1]†

Transient interactions of platelet-receptor glycoprotein Ibα (GpIbα) and the plasma protein von Willebrand factor (VWF) reduce platelet velocity at sites of vascular damage and play a role in haemostasis and thrombosis. Here we present structures of the GpIbα amino-terminal domain and its complex with the VWF domain A1. In the complex, GpIbα wraps around one side of A1, providing two contact areas bridged by an area of solvated charge interaction. The structures explain the effects of gain-of-function mutations related to bleeding disorders and provide a model for shear-induced activation. These detailed insights into the initial interactions in platelet adhesion are relevant to the development of antithrombotic drugs.

Transient interactions of platelet-receptor glycoprotein Ibα (GpIbα) and immobilized von Willebrand factor (VWF) mediate the rolling of platelets at sites of vascular damage. Rolling reduces platelet velocity and prolongs the contact time with reactive components of the cell matrix. This facilitates platelet activation and subsequent integrin-mediated firm attachment (*1*). Platelet GpIbα and VWF coexist in the circulation but do not interact at a detectable level unless shear stress is applied or exogenous modulators like the snake venom botrocetin are added (*2*). Four types of congenital bleeding disorders have been defined that are caused by mutations in GpIbα or VWF, either enhancing or reducing complex formation. Shear-induced GpIbα-VWF interaction in occluded atherosclerotic arteries or at the surface of ruptured atherosclerotic plaques contributes critically to the onset of arterial thrombosis (*3*).

GpIbα is the central component of a receptor complex consisting of glycoproteins Ibα and Ibβ, IX, and V. It anchors the complex to the cytoskeleton and harbors the VWF-binding function in its ~290 NH_2-terminal residues. The VWF-binding site is exposed well above the platelet surface, being connected to a ~45-nm-long highly O-glycosylated stalk (*4*). The ~250-kD VWF protein forms large disulfide-bonded multimers with molecular sizes of up to 10 MD. It is found in plasma and the subendothelial cell matrix and is released from storage granules when platelets and endothelial cells are activated. A VWF multimer acts as the bridging ligand between platelets and the cell matrix through collagen binding by its A3 domain and through GpIbα binding by its A1 domain (*5*).

Although the crystal structure of VWF-A1 is known (*6*, *7*) and there is a large body of mutagenesis data (*8*–*15*), the precise interactions between GpIbα and A1, the mechanism of shear-induced activation, and the molecular basis of related bleeding disorders are poorly understood. We present crystal structures of the NH_2-terminal domain of GpIbα (residues 1 to 290) and its complex with the VWF-A1 domain (residues 498 to 705; VWF residue numbering used here starts at the first residue of the mature subunit, and the addition of 763 converts the numbering to that of preproVWF) (Table 1). We expressed recombinant GpIbα with the mutations N21Q and N159Q to remove N-glycosylation sites (*16*). Crystallization of the complex required the use of gain-of-function mutants GpIbα-M239V and A1-R543Q, which are associated with platelet-type and type 2B von Willebrand diseases and cause an enhanced affinity for complex formation [dissociation constant (K_d) = 5.8 nM (*16*)].

The crystal structure of the VWF-binding domain of GpIbα displays an elongated, curved shape (Fig. 1A) that is typical for proteins containing leucine-rich repeats (*17*). Eight short leucine-rich repeats, seven of which were predicted on the basis of the amino acid sequence, make up the central region of the molecule. Flanking sequences, which are conserved among numerous extracellular proteins including the other members of the GpIb-IX-V com-

[1]Department of Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, Netherlands. [2]Thrombosis and Haemostasis Laboratory, Department of Haematology, Institute of Biomembranes, University Medical Center Utrecht, Netherlands.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: e.g.huizinga@chem.uu.nl (E.G.H.), p.gros@chem.uu.nl (P.G.)