Commentary

Organic synthesis and cell biology: Partners in controlling gene expression

Randall L. Halcomb

Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309-0215

A major direction of the field of organic synthesis is the integration of its powerful techniques into the solution of problems that are directly relevant to medicine and biology. The major advantage of incorporating organic synthesis into these approaches is to provide access to compounds that are designed to study or probe a particular biological phenomenon and that cannot be obtained from natural or commercial sources. Organic chemists in recent years have increasingly realized the power of their field, and they have entered into productive collaborations with biologists and biochemists which have afforded unique insights into a number of biological processes. The fruits of one particularly notable collaboration are reported by Danishefsky, Schreiber, Crabtree, and their colleagues in this issue of the *Proceedings* (1). This report elegantly demonstrates how basic chemical research into a group of natural products, the calicheamicins, when combined with some biochemical intuition, can open new avenues to investigate cellular processes.

Because of its unique biological activity, unusual chemical structure (Fig. 1), and unprecedented mechanism of action, calicheamicin, arguably more than any other natural product, has spurred synthetic chemists to consider applying their knowledge to biological problems (2, 3). Calicheamicin (4–7) is a member of the enediyne class of antibiotics, which also includes esperamicin, dynemicin, kedarcidin chromophore, C-1027 chromophore, and necarzinostatin chromophore (8). The compounds of this class are characterized by their ability to bind to and cleave nucleic acids. Each undergoes an activation event which serves to trigger in some manner a Bergman rearrangement (9, 10), thus producing a highly reactive diradical intermediate. This diradical then abstracts hydrogen atoms from the sugar backbone of DNA, ultimately resulting in oxidative DNA damage.

Calicheamicin exhibits particularly good selectivity in cleaving duplex DNA (11–13) and has a preference for tracts containing the sequence TCCT·AGGA (14), although other binding sequences have been identified (15). Additionally, experiments were performed which
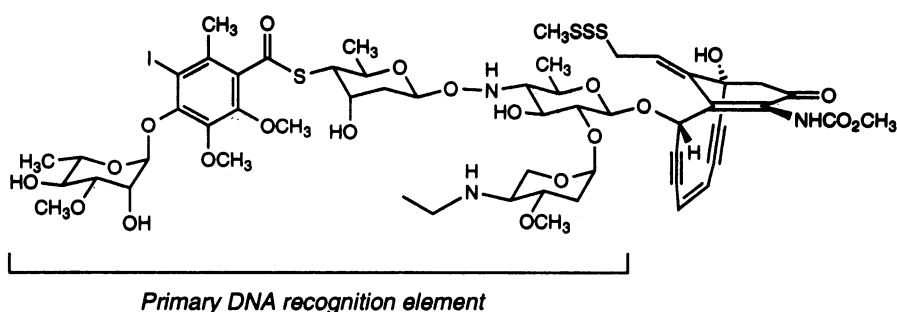

Primary DNA recognition element

FIG. 1. Structure of calicheamicin γ_1^I. The carbohydrate domain is the primary DNA recognition element.

showed that calicheamicin binds in the minor groove (14). The presence of known minor groove binders, such as netropsin, disrupts the cleavage pattern of calicheamicin. The specific hydrogen atoms that are abstracted from the deoxyribose moieties of DNA have also been identified (16, 17).

Early hypotheses concerning the origins of the sequence-selective cleavage by calicheamicin were the subject of much debate. It quickly became clear, however, that most of the DNA binding and recognition was due to contacts made by the aryl-tetrasaccharide domain. Kahne and coworkers (18, 19) showed, through NMR studies, that calicheamicin adopts an extended, highly organized conformation in solution, making it well suited to function as a minor groove binder. They also demonstrated that the hydroxylamine glycosidic linkage plays a key role in maintaining this extended structure (20).

That the oligosaccharide domain was crucial for selective and tight binding prior to cleavage was demonstrated by a series of experiments by Danishefsky, Crothers, and coworkers (21, 22). These experiments investigated the cleavage of DNA by the aglycone of calicheamicin (calicheamicinone), which itself was prepared in racemic (23) or in either antipodal (22, 24) form by total synthesis, since it could not be obtained from degrading the natural product. The aglycone was also synthesized in an enantioselective fashion by the Nicolaou group (25). Danishefsky and Crothers found that calicheamicinone cleaves DNA nonselectively and primarily in a single-

strand fashion. Interestingly, the unnatural enantiomer of the aglycone produced a slightly higher ratio of double-strand to single-strand cleavage than that of the natural enantiomer (22). From these studies it was concluded that the "sensing" interactions leading to sequence selectivity are due to the aryl-tetrasaccharide region.

The hypothesis of Danishefsky was supported by subsequent work of Kahne *et al.* (15) which involved the examination of the cleavage selectivity of a calicheamicin derivative obtained by semisynthesis. This derivative, termed calicheamicin T, contained only the A and E sugar residues. Calicheamicin T exhibited minimal selectivity, implying that the distal residues of the aryl-tetrasaccharide are important in recognition.

Several models have been advanced to rationalize the binding of calicheamicin to DNA (26–33). One of the earliest was proposed by Schreiber and coworkers (26), who sought to accurately explain the relative binding orientation of calicheamicin. This model suggests the importance of an interaction between the iodine of calicheamicin and an amino group of a guanine. Binding studies carried out by Nicolaou (see below) showed that the iodine atom was indeed important for binding and selectivity (29). However, the specific nature of the contacts made by the iodine is still unclear, since binding sites which do not have a guanine have been identified by the Kahne group (15). A current view is that a conformational change of DNA is induced upon binding to calicheamicin to accommodate the interactions necessary

for selectivity and high affinity (30–33). This theory is supported by NMR (27, 30) and CD (31) studies, as well as by molecular modeling (32), of calicheamicin–DNA complexes. On the basis of NMR analysis, Kahne and coworkers (30) have suggested that the conformational flexibility of oligopyrimidine sequences allows these regions of DNA to distort to accommodate binding of the drug. Chazin, Nicolaou, and coworkers (27) have identified two modes of binding by NMR. They proposed that the two modes are due to a shift of the aglycone within the minor groove. Their NMR data are also consistent with the occurrence of an interaction between the iodine and a guanine amino group. Hydrophobic interactions have also been implicated as making a major energetic contribution to binding (34).

It was not until the carbohydrate domain of calicheamicin was chemically synthesized that its role in selective DNA binding could be directly evaluated. Early degradation studies of the related antibiotic esperamicin indicated that the intact carbohydrate could not be recovered from degradations of the natural product (35, 36). Therefore, the only way to obtain this compound for study was by synthesis. The first synthesis of the methyl glycoside of this carbohydrate domain was accomplished by Nicolaou and coworkers (37). To date, two other syntheses of the calicheamicin aryl-tetrasaccharide have been reported (38, 39). Also, the total synthesis of calicheamicin has been accomplished (40, 41). The first was a landmark effort by Nicolaou and coworkers (40) that set the standard for future efforts. The total synthesis was also achieved by Danishefsky and coworkers (41), using a highly convergent strategy.

The first evaluations of the synthetic aryl-tetrasaccharide, in complementary investigations by Nicolaou and Danishefsky, revealed that it binds to the same sites and in a similar manner as the parent natural product (29, 42, 43). Competition experiments by Danishefsky and coworkers (42) revealed that the oligosaccharide will block the cleavage of DNA by calicheamicin at its preferred binding site. Therefore, the major DNA recognition machinery of the drug must reside within the oligosaccharide domain. Additionally, footprinting studies by Nicolaou and coworkers (29, 43) led to the similar conclusion that the oligosaccharide is primarily responsible for the sequence selectivity, although the aglycone most likely confers minor additional selectivity and binding energy. Nicolaou and Joyce (29, 43) have also probed the nature of the iodine–guanine interaction by replacing the iodine with hydrogen, methyl, or other halogens. Each of these compounds has lower affinity for DNA

than does the iodide. They conclude that indeed the iodine is involved in binding and selectivity, although the nature of this effect still remains to be unambiguously determined.

With a basic understanding of the mode of binding of calicheamicin in hand, the consortium of Danishefsky, Schreiber, and Crabtree realized the relevance of this information to cell biology (1). That realization was that calicheamicin binding sites are present in the binding sequences of many transcription factors (1). This, along with the knowledge that the methyl glycoside of the calicheamicin oligosaccharide (CLM-MG) binds to sequences similar to those bound by calicheamicin, suggested that the oligosaccharide would block the binding of transcription factors to DNA. This was indeed found to be the case. CLM-MG inhibited the formation of the transcription factor–DNA complexes when the recognition sequence of the protein contained a calicheamicin binding site. That the inhibition of the transcription factor NFAT (44, 45) was due to the binding of CLM-MG to the calicheamicin binding site was demonstrated by changing a base within the CLM-MG binding site. Minimal inhibition of protein–DNA complex formation was observed when the transcription factor recognition sequence did not contain a CLM-MG binding site. It is interesting to note that CLM-MG bound in the minor groove blocks the association of transcription factors in the major groove. This finding lends further support to the theory that calicheamicin binding induces a distortion and change of conformation of DNA.

An *in vivo* study showed that CLM-MG effectively blocked the expression of a reporter gene under the control of NFAT, which contains the CLM-MG binding site within its binding sequence. CLM-MG was also found to inhibit the proliferation of T cells in a dose-dependent manner (1). This is apparently a direct result of blocking the formation of transcription factor–DNA complexes.

It was kindly brought to the author's attention that a similar phenomenon was observed by K. C. Nicolaou, P. Vogt, and coworkers (personal communication). This team observed that, *in vitro*, the calicheamicin carbohydrate domain interferes with the binding of an interleukin-6-dependent transcription factor to its target DNA sequence, within which is the TCCT tract.

The work presented by Danishefsky, Schreiber, Crabtree, and coworkers (1) suggests a new direction for designing DNA-binding agents. Perhaps glycoconjugates such as CLM-MG will be the first step toward new classes of synthetic antiproliferative and immunosuppressive compounds. The chemical synthesis of CLM-MG is also worthy of comment.

The entire oligosaccharide was rapidly constructed from simple glycal building blocks in a modular fashion (46–48). The synthetic methodology should also allow the rapid assembly of many other congeners for studying the DNA-binding properties of oligosaccharides and glycoconjugates.

This investigation illustrates the power of combining organic synthesis and cell biology to arrive at solutions to scientific problems. The study was dependent on both organic synthesis, since the intact CLM-MG cannot be obtained in any other way, and cell biology. The work also demonstrates how information obtained from basic research into the synthesis and mechanism of action of natural products can ultimately lead to insights into how to control cellular processes.

1. Ho, S. N., Boyer, S. H., Schreiber, S. L., Danishefsky, S. J. & Crabtree, G. R. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 9203–9207.
2. Nicolaou, K. C. & Smith, A. L. (1992) *Acc. Chem. Res.* **25,** 497–503.
3. Nicolaou, K. C., Smith, A. L. & Yue, E. W. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 5881–5888.
4. Lee, M. D., Dunne, T. S., Siegel, M. M., Chang, C. C., Morton, G. O. & Borders, D. B. (1987) *J. Am. Chem. Soc.* **109,** 3464–3466.
5. Lee, M. D., Dunne, T. S., Chang, C. C., Ellestad, G. A., Siegel, M. M., Morton, G. O., McGahren, W. J. & Borders, D. B. (1987) *J. Am. Chem. Soc.* **109,** 3466–3468.
6. Lee, M. D., Ellestad, G. A. & Borders, D. B. (1991) *Acc. Chem. Res.* **24,** 235–243.
7. Lee, M. D., Dunne, T. S., Chang, C. C., Siegel, M. M., Morton, G. O., Ellestad, G. A., McGahren, W. J. & Borders, D. B. (1992) *J. Am. Chem. Soc.* **114,** 985–997.
8. Nicolaou, K. C. & Dai, W.-M. (1991) *Angew. Chem. Int. Ed. Engl.* **30,** 1387–1416.
9. Bergman, R. G. (1973) *Acc. Chem. Res.* **6,** 25–31.
10. Darby, N., Kim, C. U., Salaün, J. A., Shelton, K. W., Takada, S. & Masamune, S. (1971) *J. Chem. Soc. Chem. Commun.*, 1516–1517.
11. Zein, N., Poncin, M., Nilakantan, R. & Ellestad, G. A. (1989) *Science* **244,** 697–699.
12. Sugiura, Y., Uewawa, Y., Takahashi, Y., Kuwahara, J., Golik, J. & Doyle, T. W. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7672–7676.
13. Dedon, P. C., Salzberg, A. A. & Xu, J. (1993) *Biochemistry* **32,** 3617–3622.
14. Zein, N., Sinha, A. M., McGahren, W. J. & Ellestad, G. A. (1988) *Science* **240,** 1198–1201.
15. Walker, S., Landovitz, R., Ding, W. D., Ellestad, G. A. & Kahne, D. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 4608–4612.
16. De Voss, J. J., Townsend, C. A., Ding, W.-D., Morton, G. O., Ellestad, G. A., Zein, N., Tabor, A. B. & Schreiber,

S. L. (1990) *J. Am. Chem. Soc.* **112**, 9669–9670.

17. Hangeland, J. J., De Voss, J. J., Heath, J. A., Townsend, C. A., Ding, W.-d., Ashcroft, J. S. & Ellestad, G. A. (1992) *J. Am. Chem. Soc.* **114**, 9200–9202.

18. Walker, S., Valentine, K. G. & Kahne, D. (1990) *J. Am. Chem. Soc.* **112**, 6428–6429.

19. Walker, S., Gange, D., Gupta, V. & Kahne, D. (1994) *J. Am. Chem. Soc.* **116**, 3197–3206.

20. Walker, S., Yang, D., Kahne, D. & Gange, D. (1991) *J. Am. Chem. Soc.* **113**, 4716–4717.

21. Drak, J., Iwasawa, N., Crothers, D. M. & Danishefsky, S. J. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 7464–7468.

22. Aiyar, J., Hitchcock, S. A., Denhart, D., Liu, K. K.-C., Danishefsky, S. J. & Crothers, D. M. (1994) *Angew. Chem. Int. Ed. Engl.* **33**, 855–858.

23. Haseltine, J. N., Cabal, M. P., Mantlo, N. B., Iwasawa, N., Yamashita, D. S., Coleman, R. S., Danishefsky, S. J. & Shulte, G. K. (1991) *J. Am. Chem. Soc.* **113**, 3850–3866.

24. Rocco, V. P., Danishefsky, S. J. & Shulte, G. (1991) *Tetrahedron Lett.* **32**, 6671–6674.

25. Smith, A. L., Pitsinos, E. N., Hwang, C.-K., Mizuno, Y., Saimoto, H., Scarlato, G. R., Suzuki, T. & Nicolaou, K. C. (1993) *J. Am. Chem. Soc.* **115**, 7612–7624.

26. Hawley, R. C., Kiessling, L. L. &

Schreiber, S. L. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1105–1109.

27. Poloma, L. G., Smith, J. A., Chazin, W. J. & Nicolaou, K. C. (1994) *J. Am. Chem. Soc.* **116**, 3697–3708.

28. Mah, S. C., Price, M. A., Townsend, C. A. & Tillius, T. D. (1994) *Tetrahedron* **50**, 1361–1378.

29. Nicolaou, K. C., Tsay, S.-C., Suzuki, T. & Joyce, G. F. (1992) *J. Am. Chem. Soc.* **114**, 7555–7557.

30. Walker, S., Murnick, J. & Kahne, D. (1993) *J. Am. Chem. Soc.* **115**, 7954–7961.

31. Krisnamurthy, G., Ding, W.-d., O'Brien, L. & Ellestad, G. A. (1994) *Tetrahedron* **50**, 1341–1349.

32. Langley, D. R., Golik, J., Krishnan, B., Doyle, T. W. & Beveridge, D. L. (1994) *J. Am. Chem. Soc.* **116**, 15–29.

33. Uesugi, M. & Sugiura, Y. (1993) *Biochemistry* **32**, 4622–4627.

34. Ding, W.-d. & Ellestad, G. A. (1991) *J. Am. Chem. Soc.* **113**, 6617–6620.

35. Golik, J., Wong, H., Krishnan, B., Vyas, D. & Doyle, T. W. (1991) *Tetrahedron Lett.* **32**, 1851–1854.

36. Halcomb, R. L., Wittman, M. D., Olson, S. H., Danishefsky, S. J., Golik, J., Wong, H. & Vyas, D. (1991) *J. Am. Chem. Soc.* **113**, 5080–5082.

37. Groneberg, R. D., Miyazaki, T., Stylianides, N. A., Schulze, T. J., Stahl, W., Schreiner, E. P., Suzuki, T., Iwabuchi, Y., Smith, A. L. & Nicolaou, K. C. (1993) *J. Am. Chem. Soc.* **115**, 7593–7611.

38. Halcomb, R. L., Boyer, S. H. & Danishefsky, S. J. (1992) *Angew. Chem. Int. Ed. Engl.* **31**, 338–340.

39. Kim, S.-H., Augeri, D., Yang, D. & Kahne, D. (1994) *J. Am. Chem. Soc.* **116**, 1766–1775.

40. Nicolaou, K. C., Hummel, C. W., Nakada, M., Shibayama, K., Pitsinos, E. N., Saimoto, H., Mizuno, Y., Baldenius, K.-U. & Smith, A. L. (1993) *J. Am. Chem. Soc.* **115**, 7625–7635.

41. Hitchcock, S. A., Boyer, S. H., Chu-Moyer, M. Y., Olson, S. H. & Danishefsky, S. J. (1994) *Angew. Chem. Int. Ed. Engl.* **33**, 858–862.

42. Aiyar, J., Danishefsky, S. J. & Crothers, D. M. (1992) *J. Am. Chem. Soc.* **114**, 7552–7554.

43. Li, T., Zeng, Z., Estevez, V. A., Baldenius, K. U., Nicolaou, K. C. & Joyce, G. F. (1994) *J. Am. Chem. Soc.* **116**, 3709–3715.

44. Durand, D. B., Bush, M. R., Morgan, J. G., Weiss, A. & Crabtree, G. R. (1987) *J. Exp. Med.* **165**, 395–407.

45. Shaw, J. P., Utz, P. J., Durand, D. B., Toole, J. J., Emmel, E. A. & Crabtree, G. R. (1988) *Science* **241**, 202–205.

46. Danishefsky, S. J., Koseki, K., Griffith, D. A., Gervay, J., Peterson, J. M., McDonald, F. E. & Oriyama, T. (1992) *J. Am. Chem. Soc.* **114**, 8331–8333.

47. Randolph, J. T. & Danishefsky, S. J. (1993) *J. Am. Chem. Soc.* **115**, 8473–8474.

48. Danishefsky, S. J., McClure, K. F., Randolph, J. T. & Ruggeri, R. B. (1993) *Science* **260**, 1307–1309.