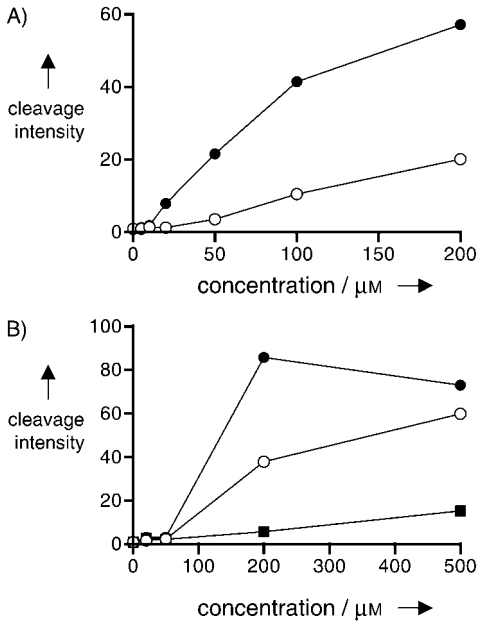


Figure 3. *Extent of RNA cleavage at the U40 reactive site within TAR RNA, (A) with* ***3*** *in the absence (○) and presence (●) of an equimolar concentration of MPA, and (B) with the* para *isomer* ***3*** *(●) or the related* ortho *(○) and* meta *(■) isomers. Data are compiled from quantitative analysis of three series of phosphor images such as those shown in Figure 2 and must be considered as a set of averaged values.*

required for the cleavage of TAR RNA by **3**. The metal complex must fit into the cavity created by the trinucleotide bulge to produce RNA-reactive species directed at the opposite U40 reactive site. The specificity of **3** clearly differs from that of a Ni^{II}–Gly-Gly-His metallopeptide[7] and a nonapeptide–cyclen conjugate,[8] both of which cleave the TAR RNA within the C29–G36 apical loop. In contrast, the specificity of our salen–Cu complex is reminiscent of that of the rhodium complex $Rh(phen)_2phi^{3+}$, which induces specific cleavage at residues C39 and U40.[10]

In conclusion, the present study conveys two important pieces of information. First, the combined results of XANES and EXAFS spectroscopy confirm that the hydroxysalen ligand forms a stable complex with Cu^{III} in solution. Second, we show that this unusual Cu^{III} complex functions as a site-specific ribonuclease capable of inducing precise cutting of the TAR RNA at the uridine bulge. These findings open a new avenue to the design of ribonucleases to probe specific structures of RNA.

Experimental Section

The syntheses of the salen–Cu complex **1** and its *para* hydroxy derivative **2** have previously been described[3, 4] and the same procedure was used to obtain the morpholino derivative **3**.

The procedure used for the XANES and EXAFS measurements is given in the Supporting Information.

In vitro transcription of TAR RNA was performed as previously described.[10] The cleavage conditions are described in the Supporting Information.

[1] L. Canali, D. C. Sherrington, *Chem. Soc. Rev.* **1999**, *28*, 85.
[2] a) D. J. Gravert, J. H. Griffin, *Inorg. Chem.* **1996**, *35*, 4837; b) K. Sato, M. Chikira, Y. Fujii, A. Komatsu, *J. Chem. Soc. Chem. Commun*. **1994**, 625; c) J. G. Muller, S. J. Paikoff, S. E. Rokita, C. J. Burrows, *J. Inorg. Biochem*. **1994**, *54*, 199; d) S. Bhattacharya, S. S. Mandal, *J. Chem. Soc. Chem. Commun*. **1995**, 2489; e) G. Pratviel, J. Bernadou, B. Meunier, *Adv. Org. Chem*. **1998**, *45*, 251.
[3] S. Routier, J.-L. Bernier, M. J. Waring, P. Colson, C. Houssier, C. Bailly, *J. Org. Chem*. **1996**, *61*, 2326.
[4] E. Lamour, S. Routier, J.-L. Bernier, J.-P. Catteau, C. Bailly, H. Vezin, *J. Am. Chem. Soc.* **1999**, *121*, 1862.
[5] J. L. DuBois, P. Mukherjee, T. D. P. Stack, B. Hedman, E. I. Solomon, K. O. Hodgson, *J. Am. Chem. Soc*. **2000**, *122*, 5775.
[6] M. M. Bhadbhade, D. Srivinas, *Inorg. Chem*. **1993**, *32*, 5458.
[7] I. J. Brittain, X. Huang, E. C. Long, *Biochemistry* **1998**, *37*, 12113.
[8] K. Michaelis, M. Kalesse, *Angew. Chem.* **1999**, *111*, 2382; *Angew. Chem. Int. Ed.* **1999**, *38*, 2243.
[9] H. R. Neenhold, T. M. Rana, *Biochemistry* **1995**, *34*, 6303.
[10] C. Bailly, P. Colson, C. Houssier, F. Hamy, *Nucleic Acids Res*. **1996**, *24*, 1460–1464.

Received: August 28, 2002 [Z481]

The Biosynthesis of Indolocarbazoles in a Heterologous *E. coli* Host

Chang-Gu Hyun,[a] Tsion Bililign,[a, b] Jianchun Liao,[a] and Jon S. Thorson*[a]

KEYWORDS:

biosynthesis · gene locus · heterologous expression · natural products · rebeccamycin

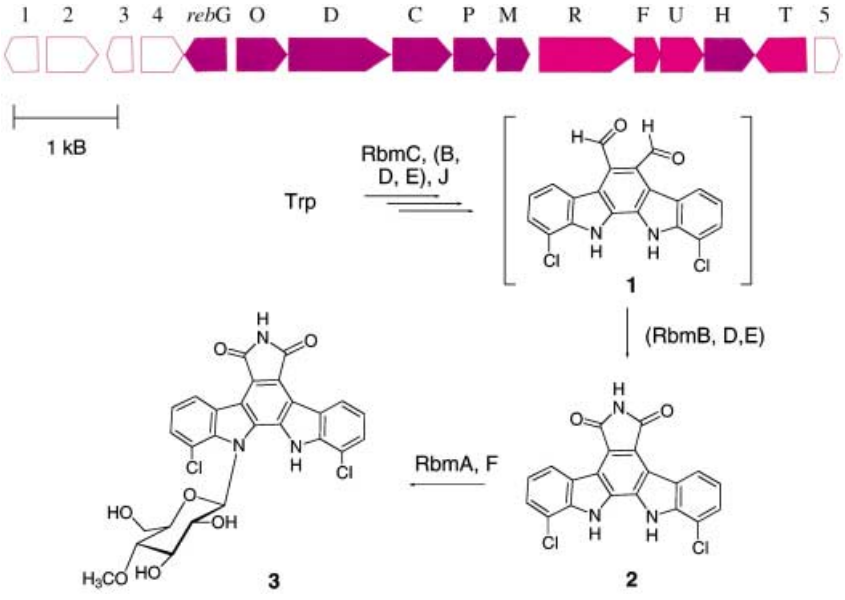
Rebeccamycin (**3**, Scheme 1), an antitumor antibiotic produced by the bacterium *Saccharothrix aerocolonigenes*, is a prototype of a class of complex natural products called indolocarbazoles. The indolocarbazole alkaloids are an emerging class of natural products that are now typically divided into two major classes according to their structure and mechanism of action.[1] The first class, exemplified by staurosporine and K252a, possess an indolocarbazole subunit wherein the indole nitrogens are bridged by glycosyl linkages; these compounds are potent protein kinase C inhibitors.[2] The second group, exemplified by rebeccamycin,[3] typically carry a single β-glycoside unit, which is

[a] *Dr. C.-G. Hyun, T. Bililign, Dr. J. Liao, Prof. J. S. Thorson*
Laboratory for Biosynthetic Chemistry
Pharmaceutical Sciences Division, School of Pharmacy
University of Wisconsin–Madison
777 Highland Avenue, Madison, WI 53705 (USA)
Fax: (+1) 608-262-5345
E-mail: jsthorson@pharmacy.wisc.edu
[b] *T. Bililign*
Joan and Sanford I. Weill
Graduate School of Medical Sciences
Cornell University
New York, NY 10021 (USA)

critical to their interaction with DNA topoisomerase.[4] Members of both classes have entered clinical trials in the search for novel antitumor therapies and, more recently, have been noted for their antiviral properties.[5]

Given the importance of these metabolites, significant effort has been put forth to generate indolocarbazole derivatives with enhanced biological properties.[6] Yet, challenges in developing scalable fermentation processes for natural indolocarbazoles still represent a bottleneck to preclinical and clinical development. The ability to induce and/or manipulate the biosynthesis of complex indolocarbazoles in an established heterologous host might significantly benefit indolocarbazole-based drug development. Moreover, since precursors anticipated for rebeccamycin production are standard to the well-developed host *Escherichia coli*, host engineering, like that required for producing other complex natural products in *E. coli*,[7] should not be necessary. Thus, an *E. coli* strain transformed with a plasmid harboring the rebeccamycin gene locus might be capable producing rebeccamycin and providing a system for generating potentially useful indolocarbazole derivatives through standard bioengineering.

To test this hypothesis, a *S. aerocolonigenes* genomic cosmid library in *E. coli* was generated. The *ngt* gene, previously demonstrated to encode an indolocarbazole *N*-glycosyltransferase in *S. aerocolonigenes*,[8] served as the basis for our rebeccamycin locus screen. Four hundred colonies from the genomic library were screened from which a single positive clone carrying the *ngt* gene was found. This clone (pJST2301), shown to carry an insert of roughly 35 Kb in size by restriction mapping, was submitted to shotgun sequencing. Sequence data obtained from forward and reverse readings of 700 single random subclones assembled into three large contigs. The largest of these three, contig 1, (28 658 bp) contained the putative rebeccamycin locus (Scheme 1) based upon DNA sequence analysis.[9]



Scheme 1. *The rebeccamycin locus and the postulated late biosynthetic intermediates. Dark shades represent proposed structural genes, light shades represent possible genes for transport and resistance, and white shapes represent nearby genes outside the locus.*

During this time an independent report of cloning the identical locus was reported.[10] Thus, our gene nomenclature was reassigned to be consistent with this previously published work.

Based upon database comparison and the recently published work by Sanchez et al.,[10] it is likely five gene products are essential for forming the indolocarbazole core—RebD is a homologue of VioB,[11] suggested to catalyze the decarboxylative fusion of two tryptophans en route to violacein; RebO belongs to the flavin-containing amine oxidases and could be involved in deamination or amination; RebH is a putative tryptophan halogenase; and RebC falls among the flavin-dependent monooxygenases while RebP is a cytochrome P450 homologue, both of which may play a role in ring oxidation or oxidative ring closure. In addition, this analysis revealed two gene products consistent with the final modification steps—RebG is the *N*-glycosyltransferase (Ngt) homologue while RebM is an *S*-adenosylmethionine dependent sugar methyltransferase. Of the remaining gene products, RebR belongs to the family of adenosine triphosphate dependent transcriptional activators; RebF is a putative flavin mononucleotide:reduced nicotinamide adenine dinucleotide oxidoreductase and may thereby be essential for providing reducing equivalents for one or more Reb biosynthetic enzymes; and RebU/T are integral membrane transporters probably involved in resistance and/or secretion.

To test production in a heterologous host, strain pJST2301/*E. coli* XL-1 Blue was analyzed for its ability to produce rebeccamycin and/or relevant biosynthetic intermediates with HPLC-MS. As a control, an *E. coli* strain harboring the control plasmid pBluescript (pBS/*E. coli* XL-1 Blue) was analyzed for metabolites in a similar manner. In comparison to the control, three new species were identified in the pJST2301 strain (Figure 1). The major component among these three was indistinguishable from **3** by coelution with a commercially available standard. Of the two minor new species identified, the first was consistent with rebeccamycin aglycon **2**, based both on HRMS and on coelution with hydrolytically generated rebeccamycin aglycon.[12] The second minor constituent was most consistent with the proposed biosynthetic intermediate **1**. Another signature in this analysis was the distinguishing isotopic enrichment found in all three species, consistent with dichlorinated structures for **1** – **3**.

From this analysis, two refinements of the postulated rebeccamycin biosynthetic pathway can be introduced (Scheme 1). First, halogenation apparently occurs prior to completion of aglycon closure which could occur either at the monomer level (for example, upon tryptophan or tryptophan-derived monomers) or during the early stages of the first ring fusion. Second, the first ring closure and aromatization in the biosynthesis of **3** apparently occurs prior to the formation of the final pyrrolocarbazole ring. While the present study is consistent with the incorporation of tryptophan, indolepyruvic acid, and indoleacetaldehyde observed in previous

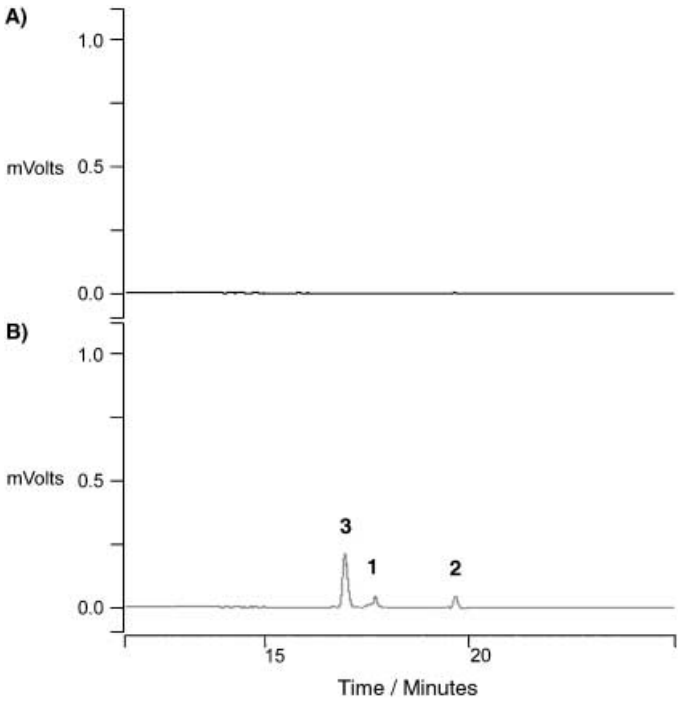


Figure 1. *Analysis of production strain extracts by HPLC. (A) Control strain pBS/*E. coli. *(B) Producing strain pJST2301/*E. coli.

indolocarbazole metabolic labeling studies,[13] it contrasts earlier speculation of *N*,*N*-dimerization as an early biosynthetic step.[13d]

More importantly, this is the first production of **3** and the corresponding analogues in the heterologous host *E. coli*. The availability of a heterologous expression system indicates rapid advancement in several important areas. First, the ease of protein and metabolic engineering of *E. coli* expression systems, in comparison to the poorly understood and slow-growing wild-type producer *S. aerocolonigenes* or the *Streptomyces* heterologous systems, clearly lends itself to enhancing the availability of novel and potentially superior indolocarbazoles. On the same basis, it should now also be possible to rapidly advance the mechanistic dissection of this unique pathway in a 'user-friendly' host, thereby setting the stage for indolocarbazole glycorandomization.[14]

Experimental Section

Materials: All culture media and biochemical reagents were purchased from Difco (Detroit, MI), Fisher Scientific (Pittsburgh, PA), or Sigma (St. Louis, MO). HPLC was performed on a Varian Dynamax SD-200 instrument controlled with Dynamax HPLC software. Routine mass spectra were recorded on a PE SCIEX API 100 LCMS mass spectrometer and HRMS was accomplished by the Riverside Mass Spectrometry Facility of the University of California.

Library production: Chromosomal DNA isolated from *S. aerocolonigenes* (American type culture collection strain 39243) was partially digested with *Sau*3AI for the optimal production of 30 – 40 kb fragments. These fragments were size-selected, ligated with calf intestinal phosphatase-treated BamHI-digested SuperCos (Stratagene, La Jolla, CA) and packaged in vitro by using Gigapack II-XL λ extracts (Stratagene) and transduction of *E. coli* XL1-Blue MR performed as recommended by the manufacturer.

Library screening: Each of 400 ampicillin resistant clones from the genomic library was separately inoculated in 1.5-mL eppendorf tubes that contained Luria-Bertani broth (500 μL) containing ampicillin (100 $\mu g\,\mu L^{-1}$). After overnight growth, liquid (30 μL) from 50 distinct tubes was combined in a new 1.5-mL tube, to give a total of 8 samples. The purified plasmid DNA from each of these 8 samples was suspended in distilled water (100 μL) and submitted to PCR screening. The primers for this screen were NGT-F (5′-CTACCGTCGGACCATATGGGGGCACGAGTG-3′) and NGT-R (5′-GATCAGTCGACGCCTCGAGGACCGCGTCG-3′), designed to amplify a 1.2-Kb cassette harboring the entire *ngt* gene. Sample 3 revealed a positive PCR fragment of the appropriate size. Fifty plasmid samples were subsequently prepared from plating a portion of sample 3 culture broth following colony selection growth. A second PCR screen of these 50 samples revealed a single positive (sample 23).

DNA sequencing and gene functional assignments: Intact cosmids were mechanically sheared, and fragments in the size range of 1 – 3 Kb were purified by preparative agarose gel electrophoresis and subcloned into SmaI-digested plasmid Bluescript II SK$^-$. Purified plasmid subclones were submitted to double-stranded DNA automated sequencing by the University of Wisconsin – Madison Genome Center. Assembly and analysis was accomplished with the DNASTAR DNA analysis package. Putative functions of open reading frames were determined by comparison to protein sequences present in the GenBank nonredundant database by using the BLAST protein database search software available at http://www.ncbi.nlm.nih.gov. The sequence was submitted under GenBank accession number AF534707.

Production in *E. coli*: For each strain (in triplicate), a single colony (strain pJST2301/*E. coli* XL-1 Blue) from a plate of Luria-Bertani broth containing ampicillin (LB/amp) was used to inoculate a culture (100 mL) of LB/amp (100 $\mu g\,mL^{-1}$) which was grown overnight, with shaking, at 37 °C. This culture was then used to inoculate a fresh LB/amp culture (1 L) enriched with 1 mM tryptophan which was grown for 48 hr, with shaking, at 26 °C. The culture was extracted with an equal volume of EtOAc, and the organic layer was isolated, dried ($NaSO_4$), filtered, and evaporated to dryness. The remaining residue was resuspended in EtOAc (50 μL) and submitted for HPLC analysis.

Product isolation and characterization: Samples (20 μL) were analyzed by reversed-phase chromatography (Microsorb-MV C-18; 4.6 × 250 mm; 80 % 0.1 % trifluoroacetic acid (TFA)/H_2O, 20 % CH_3CN to 25 % 0.1 % TFA/H_2O, 75 % CH_3CN; 20 min gradient, 1 $mL\,min^{-1}$; λ = 316 nm) on an instrument coupled to a photodiode array detector. Eluted peaks were collected and submitted for HRMS. HRMS: **1**: calcd for $C_{20}H_{10}Cl_2N_2O_2$ 380.0119, found *m/z* 381.0201 [*M* + H]; **2**: calcd for $C_{20}H_9Cl_2N_3O_2$ 393.0072, found *m/z* 394.0150 [*M* + H]; **3**: HRMS calcd for $C_{27}H_{21}Cl_2N_3O_7$ 569.0757, found *m/z* 570.0833 [*M* + H]. Production of **3** (based upon peak integration) was 0.3 – 0.5 $mg\,L^{-1}$.

This work was supported in part by the National Institutes of Health (Grant nos.: GM58196, CA84374, and AI52218). J.S.T. is an Alfred P. Sloan Fellow.

[1] a) S. Omura, Y. Sasaki, Y. Iwai, H. Takeshima, *J. Antibiot.* **1995**, *48*, 535; b) G. W. Gribble, S. J. Berthel in *Studies in Natural Products Chemistry, Vol. 12*, Elsevier Science, Amsterdam, **1993**, p. 365.

[2] a) H. Kase, K. Iwahashi, Y. Matsuda, *J. Antibiot.* **1986**, *39*, 1059; b) S. Nakanishi, Y. Matsuda, K. Iwahashi, H. Kase, *J. Antibiot.* **1986**, *39*, 1066; c) T. Yasuzawa, T. Iida, M. Yoshida, N. Hirayama, M. Takahashi, K. Shirahata, H.

Sano, *J. Antibiot*. **1986**, *39*, 1072; d) T. Tamaoki, H. Nomoto, I. Takahashi, Y. Kato, M. Morimoto, F. Tomita, *Biochem. Biophys. Res. Commun*. **1986**, *135*, 397.

[3] a) D. E. Nettleton, T. W. Doyle, B. Krishnan, G. K. Matsumoto, J. Clardy, *Tetrahedron Lett*. **1985**, *26*, 4011; b) J. A. Bush, B. H. Long, J. J. Catino, W. T. Bradner, K. Tomita, *J. Antibiot*. **1987**, *40*, 668.

[4] a) C. Bailly, X. Qu, D. E. Graves, M. Prudhomme, J. B. Chaires, *Chem. Biol*. **1999**, *6*, 277.

[5] a) M. J. Slater, R. Baxter, R. W. Bonser, S. Cockerill, K. Gohil, N. Parry, J. E. Robinson, R. Randall, C. Yeates, W. Snowden, A. Walters, *Bioorg. Med. Chem. Lett*. **2001**, *11*, 1993; b) M. J. Slater, S. Cockerill, R. Baxter, R. W. Bonser, K. Gohil, C. Gowrie, J. E. Robinson, N. Parry, R. Randall, W. Snowden, *Bioorg. Med. Chem*. **1999**, *7*, 1067.

[6] For some recent examples, see: a) M. Ohkubo, K. Kojiri, H. Kondo, S. Tanaka, H. Kawamot, T. Nishimura, I. Nishimura, T. Yoshinari, H. Arakawa, H. Suda, H. Morishima, S. Nishimura, *Bioorg. Med. Chem. Lett*. **1999**, *9*, 1219; b) P. Moreau, M. Sancelme, C. Bailly, S. Leonce, A. Pierre, J. Hickman, B. Pfeiffer, M. Prudhomme, *Eur. J. Med. Chem*. **2001**, *36*, 887; c) K. S. Lam, D. R. Schroeder, J. M. Veitch, K. L. Colson, J. A. Matson, J. W. C. Rose, T. W. Doyle, S. Forenza, *J. Antibiot*. **2001**, *54*, 1; d) A. Voldoire, M. Sancelme, M. Prudhomme, P. Colson, C. Houssier, C. Bailly, S. Leonce, S. Lambel, *Biorg. Med. Chem*. **2001**, *9*, 357; e) A. Marotto, Y. S. Kim, E. Schulze, U. Pindur, *Pharmazie* **2002**, *57*, 194.

[7] B. A. Pfeifer, S. J. Admiraal, H. Gramajo, D. A. Cane, C. Khosla, *Science* **2001**, *291*, 1790.

[8] T. Ohuchi, T. A. Ikeda-Araki, A. Watanabe-Sakamoto, K. Kojiri, M. Nagashima, M. Okanishi, H. Suda, *J. Antibiot*. **2000**, *53*, 393.

[9] C.-G. Hyun, T. Bililign, J. S. Thorson, *ACS Natl. Meet., Abstr. Pap*. **2001**, abstract Carb. 67.

[10] C. Sanchez, I. A. Butovich, A. F. Brana, J. Rohr, C. Mendez, J. A. Salas, *Chem. Biol*. **2002**, *9*, 519.

[11] a) P. R. August, T. H. Grossman, C. Minor, M. P. Draper, I. A. MacNeil, J. M. Pemberton, K. M. Call, D. Holt, M. S. Osburne, *J. Mol. Microbiol. Biotechnol*. **2000**, *2*, 513; b) P. Ruhul, A. Z. M. Momen, T. Hoshino, *Biosci. Biotechnol. Biochem*. **2000**, *64*, 539.

[12] C. Albermann, J. Liao, J. S. Thorson, unpublished results.

[13] a) C. J. Pearce, T. W. Doyle, S. Forenza, K. S. Lam, D. R. Schroeder, *J. Nat. Prod*. **1988**, *51*, 937; b) D. Meksuriyen, G. A. Cordell, *J. Nat. Prod*. **1988**, *51*, 893; c) K. S. Lam, S. Forenza, T. W. Doyle, C. J. Pearce, *J. Ind. Microbiol*. **1990**, *6*, 291; d) S.-W. Yang, G. Cordell, *J. Nat. Prod*. **1997**, *60*, 788.

[14] a) J. Jiang, J. B. Biggins, J. S. Thorson, *J. Am. Chem. Soc*. **2000**, *122*, 6803; b) J. Jiang, J. B. Biggins, J. S. Thorson, *Angew. Chem*. **2001**, *113*, 1550 – 1553; *Angew. Chem. Int. Ed*. **2001**, *40*, 1502; c) W. A. Barton, J. B. Biggins, J. Lesniak, P. D. Jeffrey, J. Jiang, K. R. Rajashankar, J. S. Thorson, D. B. Nikolov, *Nat. Struct. Biol*. **2001**, *8*, 545; d) J. S. Thorson, T. J. Hosted, Jr., J. Jiang, J. B. Biggins, J. Ahlert, M. Ruppen, *Curr. Org. Chem*. **2001**, *5*, 139; e) W. A. Barton, J. B. Biggins, J. Jiang, J. S. Thorson, D. B. Nikolov, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 13397 – 13402; f) H. Losey, J. Jiang, J. B. Biggins, M. Oberthur, X.-Y. Ye, S. D. Dong, D. Kahne, J. S. Thorson, C. T. Walsh, *Chem. Biol*. **2002**, in press.

Received: October 23, 2002 [Z515]