Tightening the Belt on Polymerases: Evaluating the Physical Constraints on Enzyme Substrate Size**

Miriam Frieden, Enrique Pedroso,* and Eric T. Kool*

Recent X-ray crystal structures have shown that the active site cleft of DNA polymerases holds approximately five base pairs of double-stranded DNA in a linear A-like conformation.[1] Polymerases are relatively large proteins consisting of about 600–1000 amino acids, and have dimensions of 50–100 Å on a side. Despite the large size of the proteins and the generally rigid, linear conformation of DNA substrates for these enzymes, recent studies have shown that DNA and RNA polymerases can be active on quite small, constrained templates.

For example, synthetic circular single-stranded (ss) oligodeoxynucleotides have been tested as substrates for polymerases.[2, 3] Interestingly, evidence suggests that DNA circles as small as 26 nucleotides can be active as templates for DNA polymerases.[2b] Similarly, studies with RNA polymerases (RNAPs) have shown that ssDNA circles as small as 34, or even 28, nucleotides (nt) may be active substrates for transcription.[3a, 4] Those surprising results raise an interesting question: How small can a circular DNA be and still act as a template for these enzymes? Here we show that polymerases can utilize circular templates that are remarkably small, which lends insight into the geometric constraints of these important enzymes, and also has practical implications in methods for the amplified synthesis of DNA and RNA.

To test this question, we prepared cyclic DNAs of 28, 23, 18, and 13 nucleotides (**1**–**4**, Figure 1) using recently developed methods.[5] We chose 28 nt as the upper limit because previous studies suggested that this largest circular DNA might be a substrate for both classes of enzymes.[2b, 4] The sequences were taken from the HIV-1 *gag* gene; if copied by polymerases, the resulting repeating DNA or RNA strands would be complementary (antisense) to the gene. The sequences are nested, in that the smaller circles are subsets of the larger ones. Also included is a recognition site (T▾CGA) for *Taq* Iα restriction endonuclease.

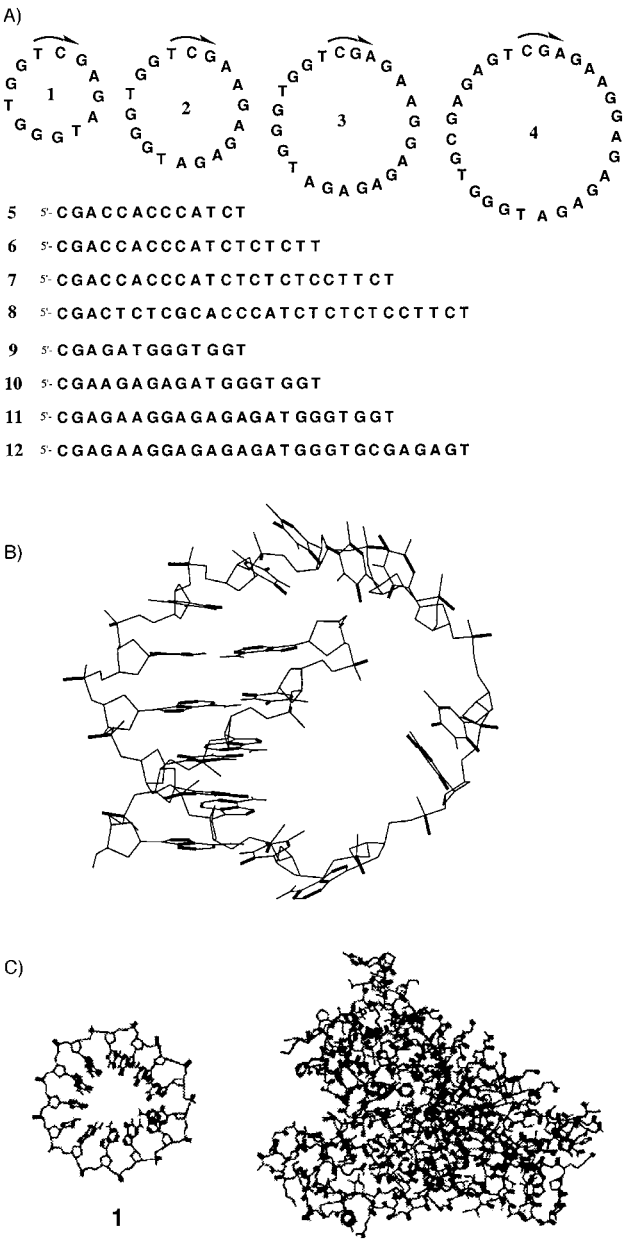

Figure 1. Sequences and structures used in this study. A) Sequences of circular DNA oligonucleotides **1**–**4**, complementary primers **5**–**8**, and linear controls **9**–**12**. Curved arrow denotes 5′→3′ strand orientation. B) Molecular model showing how a circular 13-mer might conceivably form four base pairs of duplex (as in a template–primer complex in a polymerase active site) without strongly distorting it. C) Models showing relative sizes of circle **1** and KF polymerase.

We then examined whether these could be enzyme substrates by incubating them with DNA polymerases, a complementary primer (**5**–**8**, Figure 1), and deoxynucleoside triphosphates, or with RNA polymerases and ribonucleoside triphosphates. We also tested the linear versions of the circular DNAs as controls (**9**–**12**, Figure 1). Three DNA polymerases were examined: The Klenow fragment of *E. coli* DNA polymerase I (KF), T4 DNA polymerase, and Sequenase 2.0 (a T7 DNA polymerase mutant). The results show that the T4 and Sequenase enzymes apparently do not accept any of the four circular DNAs as templates (data not shown). However, the KF enzyme gave relatively long products (ca.

[*] Prof. E. Pedroso, M. Frieden
Departament de Química Orgànica
Facultat de Química, Universitat de Barcelona
Barcelona, E-08028 (Spain)
Fax: (+34) 93-33-97-878
E-mail: pedroso@admin.qo.ub.es
Prof. E. T. Kool[+]
Department of Chemistry, University of Rochester
Rochester, NY 14627 (USA)
[+] Current address:
Department of Chemistry, Stanford University
Stanford, CA 94305-5080 (USA)
Fax: (+1) 650-725-0259
E-mail: kool@chem.stanford.edu

[**] This work was supported by the US National Institutes of Health (GM46625) and the US DOD Breast Cancer Research Program, administered by the US Army, and in Spain by the Ministerio de Educación (DGES, PB97-941) and the Generalitat de Catalunya (C. R. de Biotecnologia and SGR98-1).

  1433-7851/99/3824-3654 $ 17.50+.50/0

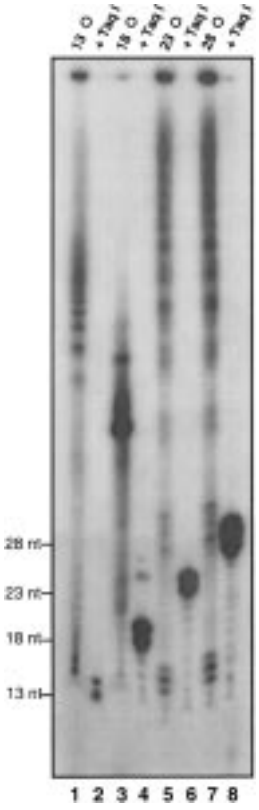

Figure 2. Products of DNA polymerase reactions on distorted circular templates **1**–**4**. The enzyme in these experiments is the Klenow fragment of *E. coli* DNA Polymerase I. Presented is an autoradiogram of a denaturing polyacrylamide gel, showing long DNA products synthesized on circular templates (lanes 1, 3, 5, 7) and the products of cleavage of these long DNAs with *Taq* Iα restriction enzyme (lanes 2, 4, 6, 8). Doubling of bands in the even lanes is expected because of the labeling method used.[6]

300 to greater than 1000 nt) with three of the four circles, including the smallest one (**1**), which is only 13 nt (Figure 2). By contrast, the linear controls gave no long DNA products, supporting the idea that circular structure is required for this activity (data not shown). The products of the 18 nt circle were shorter; it is as yet unclear why this is the case, but it is possible that this sequence adopts a secondary structure which disfavors primer binding. A regular banding pattern is visible for three of the cases; this we ascribe to pausing once per turn around the circle prior to incorporation of radiolabeled dGTP, which is kept at low concentrations. The expected products are single-stranded DNAs of a repeating sequence (complementary to the circle), and the regular banding pattern provides evidence that the products are indeed faithful copies of the circular templates.

Cleavage of the products with *Taq* Iα resolved them into the predicted short products (Figure 2),[6] serving as a second confirmation of accurate polymerase copying. Finally, dye-terminator sequencing of the products shows that all four cases give evidence of the expected repeating sequences (see Figure 3 for results with the 13-mer circle). Perhaps significantly, there are a number of errors observed in these sequences; this is likely due to standard sequencing errors and to artifacts such as fold-back and slipped synthesis on such repetitive polydisperse products, or possibly to a decreased level of fidelity on these unusual templates.

We then tested two different RNA polymerases, T7 RNA Polymerase and *E. coli* RNA polymerase, with these four very small circular templates. The data show that with the linear templates only very short RNAs are produced (Figure 4). However, cyclization of the DNAs results in the synthesis of quite long RNAs. For *E. coli* RNAP, the results show that this polymerase is active on the 28 nt circle, producing RNAs longer than 1000 nt. This same polymerase also utilizes the 23 nt circular template, although to a lesser extent. This appears to be the size limit for this enzyme, as the shorter 18 nt and 13 nt DNAs are apparently not transcribed.

Interestingly, the results for the smaller T7 RNA polymerase are quite different. This enzyme appears to actively synthesize long RNAs on all four templates (Figure 4). The

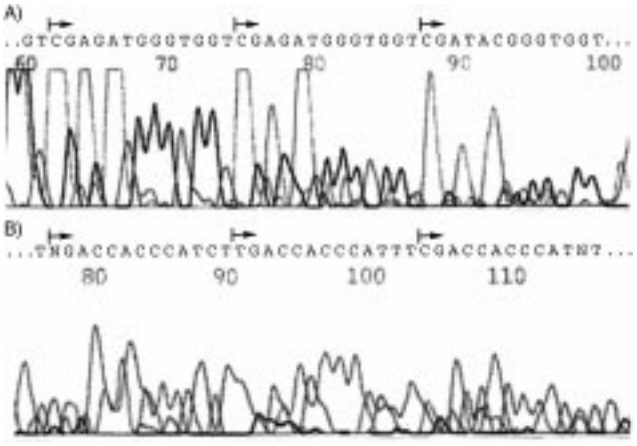

Figure 3. Dye-terminator sequencing results for repeating nucleic acids produced on 13-mer circular template **1**. Arrows denote the observed repeats encoded by the circle. A) Sequencing of DNA produced with the KF DNA polymerase. B) Sequencing of RNA produced with T7 RNA polymerase. Note that only short reads can be obtained because the primer can bind anywhere in the body of the repeating sequence.

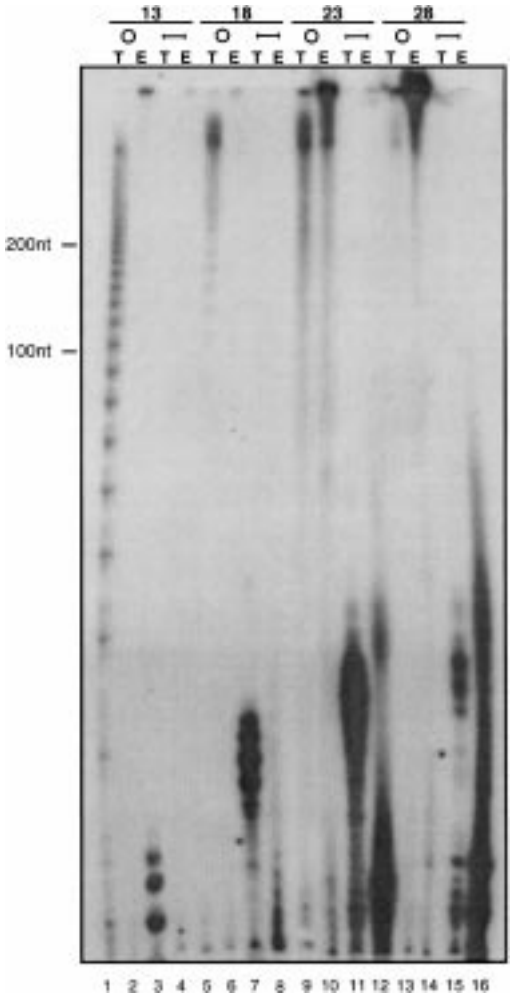

Figure 4. Products of RNA polymerase reactions on circular templates **1**–**4**. The enzyme used is denoted by E (*E. coli* RNAP) or T (T7 RNAP); the template topology (circular or linear) is denoted by circles and lines.

RNA products are shorter on the 13-mer template than on the longest two; this may reflect some inefficiency with utilization of this most distorted template. Once again, a regular banding

  1433-7851/99/3824-3655 $ 17.50+.50/0

pattern is observed; this we ascribe to preferential initiation with GTP at the single C in the sequence combined with partial termination at each round of synthesis.[4] Again, these regular banding patterns serve as evidence that the polymerase is indeed utilizing these very small circles as viable templates. Further, sequencing of the RNAs shows in general that the chief identifiable sequences are indeed the predicted repeating sequences encoded by these circles. Even the 13 nt circle shows evidence for at least eleven repeats (implying eleven trips around the circular template). This is consistent with the gel evidence (Figure 4), which shows that at least 15 repeats can be synthesized with this smallest circle.

Such small constrained DNAs put significant topological constraints on polymerase activity. For DNA polymerases, a primer oligonucleotide must bind for several adjacent nucleotides, which in itself may be problematic with a circle only 13 nt in size. The polymerase must then bind this complex before nucleoside triphosphates are added to the primer end. X-ray crystal structures of DNA polymerases bound to double-stranded DNAs show that in the enzyme the DNA is straight and in an A-like conformation for roughly five base pairs.[1] If one assumes that five base pairs of duplex are formed in the active site, this causes additional topological difficulties for a circular 13-mer. Half a helical turn (5 bp) of DNA forces the two ends of one strand to be oriented nearly on opposite faces of the duplex. Models suggest (Figure 1B) that the remaining eight nucleotides could conceivably bridge this distance; however, it is possible that fewer base pairs (perhaps two to four) are formed in the active site at one time with these constrained templates. It is worth noting that since the 13-mer circle is active as a template, the actual size limit may be even smaller than this for the KF and T7 RNAP enzymes, although our models suggest that this may be unlikely.

Although there are fewer crystal structures of RNA polymerases for comparison,[7a] RNA polymerases likely face many of the same issues.[7b] One significant difference is that RNA polymerases do not require a primer for initiation; rather, they initiate immediately downstream of a conserved double-stranded promoter sequence. Thus the present results are unusual both because these templates are strongly structurally distorted and because they contain no promoters.

Replication and transcription of naturally occurring DNAs often requires processing of topologically closed (i.e., circular) templates, such as those from bacteria and viruses. On a local scale, however, bacterial and viral genomes are essentially linear, as they are quite large. For example, the genome of *E. coli* has 4.6 million base pairs,[8] and even the much smaller circular genome of the T7 bacteriophage has 40 000 base pairs.[9] We find it remarkable that polymerases from these organisms can process a circular 13-mer.

Finally, small synthetic circular DNAs have been investigated in several laboratories recently for their unusual ability to act as templates for amplified DNA and RNA synthesis.[2–4] These rolling circle replication (RCR) and rolling circle transcription (RCT) reactions have utility in detection of nucleic acids and for in vitro synthesis of amplified quantities of DNAs and RNAs. The current results indicate that even very short DNAs could in principle be synthesized enzymatically on a monomer-length circular template. For example, restriction cleavage of the products from the present 23-mer template (Figure 2) produces a linear DNA very close in sequence to an oligonucleotide which has undergone clinical trials for treatment of HIV infection.[10]

Experimental Section

Cyclic oligodeoxynucleotides: The 28-, 23-, 18-, and 13-mer cyclic oligodeoxynucleotides **1**–**4** were synthesized as reported elsewhere[5] on a 2-μmol scale. The crude products were obtained in 7, 9, 8, and 7% yields, respectively. The 13-mer was purified by semipreparative HPLC, and the rest by preparative denaturing 20% polyacrylamide gel electrophoresis (PAGE). The product band was isolated by excision, crushing, and elution into 2 M ammonium acetate. Salts were removed with a Sep Pak C18 cartridge (Waters), and the DNA was quantitated by absorbance at 260 nm. The circular structures were confirmed by their altered size dependence of gel mobility, by their partial chemical and/or enzymatic (S1 nuclease) degradation (which produced bands migrating as expected for linear products of nicking),[11] and by MALDI-TOF mass spectrometry, (showing in all cases the expected mass). Linear DNA controls and primers were synthesized on an ABI 392 DNA synthesizer using the standard DNA cycle. Deprotection in concentrated ammonia was carried out at 55 °C for 12 h. After lyophilization, the DNAs were purified by preparative denaturing PAGE and quantitated by UV absorbance at 260 nm.

DNA polymerase reactions: Conditions for DNA synthesis using the Klenow fragment (KF) of DNA polymerase I (Pharmacia) were as follows: 1 μM circle or linear control, 1 μM primer, 3 U enzyme, 0.5 mM each of dATP, dTTP, and dCTP, 60 μM dGTP (Boehringer Mannheim), 0.30 μCi [α-^{32}P]dGTP in a buffer (pH 7.5) containing 50 mM Tris·HCl, 10 mM $MgCl_2$, 1 mM dithiothreitol (DTT), and 50 μg mL^{-1} acetylated bovine serum albumin (BSA), in 15 μL total volume. – Conditions for T4 DNA polymerase (Gibco BRL) were: 1 μM circle or linear control, 1 μM primer, 3.75 U enzyme, 0.5 mM each of dATP, dTTP, and dCTP, 60 μM dGTP (Boehringer Mannheim), 0.30 μCi [α-^{32}P]dGTP in a buffer (pH 7.9) containing 33 mM Tris·acetate, 66 mM NaOAc, 10 mM $Mg(OAc)_2$, 100 μg mL^{-1}acetylated BSA, and 0.5 mM DTT, in a volume of 15 μL. – Conditions with T7 Sequenase version 2.0 (Pharmacia) were: 1 μM circle or linear control, 1 μM primer, 4.5 U enzyme, 0.5 mM each of dATP, dTTP, and dCTP, 60 μM dGTP (Boehringer Mannheim), 0.30 μCi [α-^{32}P]dGTP in a buffer (pH 7.5) containing 40 mM Tris·HCl, 20 mM $MgCl_2$, and 50 mM NaCl, in a total volume of 15 μL. – Reaction mixtures were incubated at 37 °C for 1.5 h and stopped by addition of one volume of stop solution (30 mM EDTA, 8 M urea). Reaction mixtures were heated to 90 °C for 2 min, and then chilled on ice before being loaded on a 10% polyacrylamide denaturing gel (0.4 mm thick).

RNA polymerase reactions: Conditions for an internally labeled rolling circle transcription reaction were: 1 μM circle or linear control, 3 U *E. coli* RNA polymerase holoenzyme (Boehringer Mannheim) or 25 U T7 RNA polymerase (New England Biolabs), 0.5 mM ATP, GTP, and CTP (Pharmacia Biotech), 60 μM UTP (Boehringer Mannheim), 0.30 μCi [α-^{32}P]UTP in a buffer (pH 8.1) containing 25 mM Tris·HCl, 20 mM NaCl, 12 mM $MgCl_2$, 0.4 mM spermine·HCl (Sigma), 100 μg mL^{-1} acetylated BSA (New England Biolabs), 10 mM DTT, and 12.5 U mL^{-1} RNase inhibitor (Promega), in a total reaction volume of 15 μL. Reaction mixtures were incubated at 37 °C for 1.5 h and stopped by addition of one volume of stop solution. They were then analyzed by PAGE as above.

Cleavage of DNA Multimers: Rolling circle reactions were carried out with internal labeling so that cleaved products would retain their labels. The reaction mixture was the same as described above for DNA synthesis. DNA products after 1.5 h were precipitated with 2.5 volumes of cold ethanol and 0.2 volumes of NaOAc (3 M, pH 5.2) and resuspended in a buffer (pH 8.4) containing 100 U of restriction enzyme *Taq* Iα (20 U μL^{-1}) (New England Biolabs), 100 μg mL^{-1} BSA, 100 mM NaCl, 10 mM Tris·HCl, and 10 mM $MgCl_2$ in a total volume of 20 μL. After 12 h incubation at 65 °C, the products were extracted with phenol/chloroform (1/1) and separated from dNTPs and other small molecules with Micro Bio-Spin 6 column (Bio-Rad). Products were analyzed by denaturing PAGE.

  1433-7851/99/3824-3656 $ 17.50+.50/0 *Angew. Chem. Int. Ed.* **1999**, *38*, No. 24

Sequencing of Products: Unlabeled DNA was prepared from rolling circle replication reaction as above but with all four dNTPs at 0.5 mM. After 1.5 h incubation at 37 °C, the mixture was heated to 90 °C for 5 min and partially purified with Micro Bio-Spin 30 columns (Bio-Rad). Sequencing was carried out using a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) following the recommended procedure. The linear controls **9**–**12** were used as sequencing primers. Unlabeled RNA was prepared from rolling circle transcription reactions as above but with all four rNTPs at 0.5 mM. After 1.5 h incubation at 37 °C, the mixture was heated to 90 °C for 5 min and the repetitive RNA products were partially purified by size exclusion as above. The complementary DNA was obtained by using these conditions: partially purified RNA, 50 pmol linear primer DNA (either **9**, **10**, **11**, or **12**), 50 U AMV Reverse Transcriptase (Pharmacia), 10 mM dATP, dGTP, dCTP and dTTP, 20 U RNAse inhibitor (Promega) in a buffer (pH 8.3) containing 100 mM Tris·HCl, 10 mM $MgCl_2$, 10 mM DTT, and 50 mM KCl, in a total volume of 20 μL. After 1 h incubation at 42 °C, the reaction mixture was heated to 90 °C for 5 min and cleaned up with a Micro Bio-Spin 30 column. The complementary DNA was then sequenced as above, using primers **5**–**8**.

Received: March 22, 1999
Revised version: September 8, 1999 [Z13195IE]
German version: *Angew. Chem.* **1999**, *111*, 3870–3874

Keywords: DNA recognition · polymerases · RNA · transcription

[1] a) S. Doublié, S. Tabor, A. M. Long, C. C. Richardson, T. Ellenberger, *Nature* **1998**, *391*, 251–258; b) J. R. Kiefer, C. Mao, J. C. Braman, L. S. Beese, *Nature* **1998**, *391*, 304–307; c) H. Huang, R. Chopra, G. L. Verdine, S. C. Harrison, *Science* **1998**, *282*, 1669–1675; d) C. A. Brautigam, S. Sun, J. A. Piccirilli, T. A. Steitz, *Biochemistry* **1999**, *38*, 696–704.

[2] For DNA polymerase activity on small circular DNAs, see a) A. Fire, S. Q. Xu, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 4641–4645; b) D. Liu, S. L. Daubendiek, M. A. Zillmann, K. Ryan, E. T. Kool, *J. Am. Chem. Soc.* **1996**, *118*, 1587–1594; c) P. M. Lizardi, X. Huang, Z. Zhu, P. Bray-Ward, D. C. Thomas, D. C. Ward, *Nat. Genet.* **1998**, *19*, 225–232; d) J. Baner, M. Nilsson, M. Mendel-Hartvig, U. Landegren, *Nucleic Acids Res.* **1998**, *26*, 5073–5078.

[3] For RNA polymerase activity on small circular DNAs, see a) S. L. Daubendiek, K. Ryan, E. T. Kool, *J. Am. Chem. Soc.* **1995**, *117*, 7818–7819; b) S. L. Daubendiek, E. T. Kool, *Nature Biotechnology* **1997**, *15*, 273–277; c) A. M. Diegelman, E. T. Kool, *Nucleic Acids Res.* **1998**, *26*, 3235–3241; d) A. M. Diegelman, S. L. Daubendiek, E. T. Kool, *BioTechniques* **1998**, *25*, 754–758; e) A. M. Diegelman, E. T. Kool, *Chem. Biol.* **1999**, *6*, 569–576.

[4] S. L. Daubendiek, PhD thesis, University of Rochester, **1997**.

[5] E. Alazzouzi , N. Escaja, A. Grandas, E. Pedroso, *Angew. Chem.* **1997**, *109*, 1564–1567; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 1506–1508.

[6] There are two expected oligonucleotide products of *Taq* I cleavage, with length n and $n+1$ (where n is the circle length). The n-length product arises from normal cleavage away from the ends of the repeating sequence; the $n+1$ product arises from cleavage near the 3′ terminus of a given strand. The 3′ terminus for most strands is the C immediately after the *Taq* I site (TCGA); this is because dGTP concentration is kept low during radiolabeling, leading to termination prior to insertion of the G.

[7] a) R. Sousa, Y. J. Chung, J. P. Rose, B. C. Wang, *Nature* **1993**, *364*, 593–599; b) T. A. Steitz, S. J. Smerdon, J. Jager, C. M. Joyce, *Science* **1994**, *266*, 2022–2025.

[8] F. R. Blattner, G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, Y. Shao, *Science* **1997**, *277*, 1453–1474.

[9] F. W. Studier, *J. Mol. Biol.* **1983**, *166*, 477–535.

[10] a) J. Lisziewicz, D. Sun, F. F. Weichold, A. R. Thierry, P. Lusso, J. Tang, R. C. Gallo, S. Agrawal, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 7942–7946; b) S. Agrawal, Q. Zhao, *Curr. Opin. Chem. Biol.* **1998**, *2*, 519–528.

[11] E. Rubin, S. Rumney, E. T. Kool, *Nucleic Acids Res.* **1995**, *23*, 3547–3553.

Tetrametallic Reduction of Dinitrogen: Formation of a Tetranuclear Samarium Dinitrogen Complex**

Tiffany Dubé, Sabrina Conoci, Sandro Gambarotta,* Glenn P. A. Yap, and Giuseppe Vasapollo

Examples of dinitrogen complexes are rare in lanthanide chemistry and have been documented for only four cases.[1, 2] In the case of samarium in particular, the two existing complexes display a completely different extent of dinitrogen reduction.[1] Thus, in an attempt to evaluate the role played by the nature of the ligand in promoting dinitrogen reduction, we have now prepared a novel diphenylmethyldipyrrolide dianion and investigated its ability to stabilize Sm^{II} complexes.[3] Herein we describe the first example in which a four-electron dinitrogen reduction is achieved through the cooperative one-electron oxidation of four metal centers. To the best of our knowledge this type of process has been observed only in the nitrogenase enzyme.[4]

The reaction of $[SmI_2(thf)_2]$ with one equivalent of diphenylmethyldipyrrolide dianion (dipyrr; either as disodium or dipotassium salt) at room temperature under N_2 yielded a dark purple-brown solution from which dark brown crystals of **1** (Figure 1) were isolated in an analytically pure form in good

$[\{[\mu\text{-}Ph_2C(\eta^1{:}\eta^5\text{-}C_4H_3N)_2]Sm\}_4(\mu\text{-}\eta^1{:}\eta^1{:}\eta^2{:}\eta^2\text{-}N_2)]$ **1**

yield on a reproducible basis. Unlike $[\{N(CH_2CH_2NSitBuMe_2)_3U\}_2(\mu\text{-}\eta^2{:}\eta^2\text{-}N_2)]$[5] or $[\{(\eta^5\text{-}C_5Me_5)_2Sm\}_2(\mu\text{-}\eta^2{:}\eta^2\text{-}N_2)]$,[1a] and similar to the case of the calix-tetrapyrrole derivative,[1b] complex **1** showed no sign of reversible coordination of dinitrogen and remained unchanged upon prolonged exposure to heat or vacuum.

The presence of a molecule of coordinated dinitrogen in the molecular structure was demonstrated by an X-ray crystal structure analysis. The complex is composed of four $[\{Ph_2C(C_4H_3N)_2\}Sm]$ units; the four metal atoms arrange to form a slightly distorted lozenge, in the center of which is located a dinitrogen molecule that is coordinated *side-on* to two metal atoms and *end-on* to the other two metal atoms. The N–N distance and the fact that degradation with anhydrous HCl did not release N_2 indicates that the Sm atoms are present in the formal trivalent state.[6]

The same reaction carried out under an argon atmosphere afforded dark red crystals of a new ionic compound formulated as **2** (Figure 1). The ionic structure is formed by a

$[\{[\mu\text{-}Ph_2C(\eta^1{:}\eta^5\text{-}C_4H_3N)_2]Sm\}_5(\mu_5\text{-}I)]^-$-
$[\{[\mu\text{-}Ph_2C(\eta^1{:}\eta^5\text{-}C_4H_3N)_2]_2[Sm(thf)]_3\}(\mu_3\text{-}I)]^+$ **2**

[*] Prof. S. Gambarotta, T. Dubé, Dr. G. P. A. Yap
Department of Chemistry, University of Ottawa
Ottawa, ON, K1N 6N5 (Canada)
Fax: (+1) 613-562-5170
E-mail: sgambaro@oreo.chem.uottawa.ca
Dr. S. Conoci, Prof. G. Vasapollo
Department of Material Science, Faculty of Engineering
University of Lecce (Italy)

[**] This work was supported by the Natural Sciences and Engineering Council of Canada (NSERC) and by NATO (travel grant).

  1433-7851/99/3824-3657 $ 17.50+.50/0