Seleno-Modified RNA

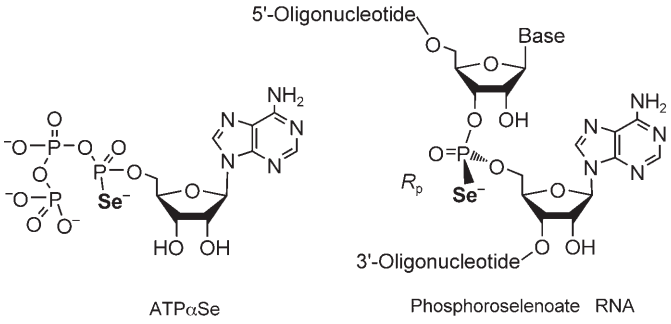
DOI: 10.1002/ange.200502215

Efficient Enzymatic Synthesis of Phosphoroselenoate RNA by Using Adenosine 5′-(α-*P*-Seleno)triphosphate**

*Nicolas Carrasco, Julianne Caton-Williams, Gary Brandt, Siming Wang, and Zhen Huang**

The discovery that RNA has a variety of biological functions in living organisms[1–3] has prompted the development of new methods to elucidate its structure and mechanism at the atomic level. X-ray crystallography is the method of choice for the elucidation of the three-dimensional structure of RNA macromolecules.[4,5] New developments in RNA X-ray crystallography have occurred due to advances in synchrotron radiation, diffraction data collection,[6,7] solid-phase synthesis of RNA oligonucleotides,[8] RNA crystallization,[9,10] and heavy-atom derivatization.[11] To further advance the field of RNA structure and function research, we have been working on the development of selenium derivatization for nucleic acid biochemistry and structure studies.[12–17]

The use of selenomethionyl proteins for multiwavelength anomalous-dispersion (MAD) phasing has revolutionized the field of protein X-ray crystallography.[18,19] This derivatization of proteins with Se has also been applied successfully in RNA structure determination through indirect derivatization of RNA that binds to the Se-derivatized protein.[11] Motivated by the phosphoroselenoate oligonucleotide structure and function studies with MAD phasing performed by Egli and co-workers,[20] we are developing the enzymatic synthesis of phosphoroselenoate nucleic acids for X-ray crystal structure studies. We report herein the first enzymatic synthesis of phosphoroselenoate RNA (Scheme 1), containing a selenium atom that replaces one of the nonbridging oxygen atoms on the phosphate group, by in vitro transcription with T7 RNA polymerase and adenosine 5′-(α-*P*-seleno)triphosphate (ATPαSe).



Scheme 1. Chemical structures of ATPαSe and phosphoroselenoate RNA. The suffix "P" indicates that the *R* or *S* nomenclature refers to the phosphorus center.

For this enzymatic synthesis, we first synthesized and characterized both diastereomeric monomers of adenosine triphosphate harboring the selenium functionality at the α-phosphate group (ATPαSe, Scheme 1) by using a modification of the procedures for the synthesis of nucleotide 5′-(α-*P*-thio)triphosphates (NTPαS)[21] and thymidine 5′-(α-*P*-seleno)triphosphate (TTPαSe).[17] The fast- and slow-moving ATPαSe isomers as represented by peaks on the reversed-phase (RP) HPLC profile were termed ATPαSe I and ATPαSe II, respectively. A DNA template (55 nucleotides) was designed to allow the incorporation of 12 A residues (Figure 1 a). The generated RNA transcript (35 nucleotides) was body-labeled by using α-[^{32}P]-cytidine 5′-triphosphate (α-[^{32}P]CTP) in the enzymatic reaction mixture for gel electrophoresis and autoradiography.

We first tested the incorporation of ATPαSe I and ATPαSe II. The results (Figure 1 b) indicated that ATPαSe I was incorporated into the RNA transcript as well as natural ATP; however, no full-length product was detected when ATPαSe II was used. Although the formation of short abortive fragments in in vitro transcription is normal,[22] surprisingly, ATPαSe I, which generated almost no short abortive sequences, led to a much cleaner reaction than natural ATP. A mixture of ATP and ATPαSe I also reduced the formation of nondesired short abortive sequences. These results suggest that ATPαSe I is an efficient substrate for T7 RNA polymerase, a finding that is consistent with an X-ray crystal structure study of the interaction between T7 RNA polymerase and NTP.[23] On the basis of the RNA polymerase recognition of the S_P diastereomer of nucleoside 5′-(α-*P*-thio)triphosphates as a substrate,[24,25] ATPαSe I is tentatively assigned as the S_P diastereomer and ATPαSe II as the R_P

[*] Dr. N. Carrasco, J. Caton-Williams, G. Brandt, S. Wang, Prof. Dr. Z. Huang
Department of Chemistry
Georgia State University
Atlanta, GA 30303 (USA)
and
Brooklyn College
Brooklyn, NY 11210 (USA)
Fax: (+1) 404-651-1416
E-mail: huang@gsu.edu

[**] This work was supported by the Georgia State University Research Program and the US National Institutes of Health (Grant no.: GM069703). We thank Dr. Martin Egli and Prof. Steven A. Benner for carefully reading this manuscript and Dr. Yanling Zhang and Sarah Shealy for assistance in MS data collection.

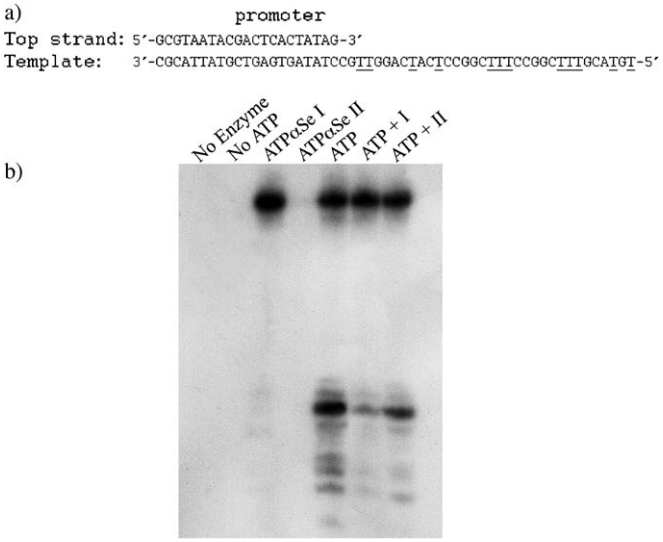


Figure 1. Enzymatic incorporation of ATPαSe into RNA by T7 RNA polymerase. a) The DNA sequences of the top strand and template. Underlined T residues in the template correspond to the locations of ATPαSe incorporated in the RNA transcript. b) Gel electrophoresis analysis of the transcription mixture after 1 h of incubation at 37 °C; each mix of ATP and ATPαSe I or II was prepared in a 1:1 ratio.

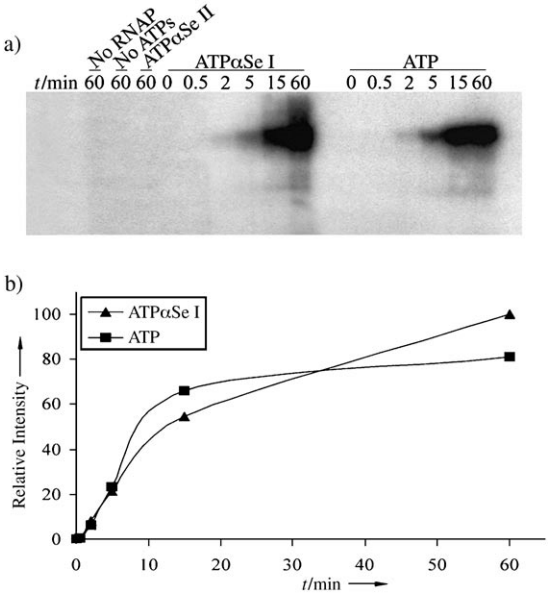


Figure 2. Time-course incorporation of ATPαSe I and ATPαSe II into RNA. Aliquots of the reaction mixture were taken at the indicated time points and quenched with 100 mM of ethylenediaminetetraacetate (EDTA); this was followed by immediate freezing in dry ice. The results are presented as a) gel electrophoresis autoradiography and b) the incorporation-versus-time plot.

diastereomer. Based on literature reports on the inversion of configuration at the phosphorous stereocenter during the enzymatic incorporation of NTPαS,[24–26] it is expected that the phosphoroselenoate RNA transcript is the R_P diastereomer.

This type of stereospecificity of T7 RNA polymerase has also been reported on the RNA incorporation of nucleoside triphosphate boranophosphates.[27] In addition, our experimental results (Figure 1b) of ATPαSe II nonincorporation and with a mixture of natural ATP and ATPαSe II indicated that ATPαSe II is neither an inhibitor nor a substrate. These results are consistent with the literature report on ATPα S II.[25] Therefore, it is possible to use ATPαSe directly for the transcription without separation of the S_P and R_P diastereomers. Since both diastereomers underwent the same treatments during their synthesis, purification, and polymerization, it is unlikely that the RNA transcript observed in the ATPαSe I reaction was due to oxidation of ATPαSe I to ATP.

To obtain more insight into the incorporation of ATPα Se I into RNA, we carried out a time-course experiment (Figure 2) with ATP as a positive control. The results indicate that incorporation of ATPαSe I is as efficient as that of natural ATP, and ATPαSe I gives approximately 20% more RNA product than ATP, probably due to formation of many fewer nondesired short abortive sequences. The minimum formation of nondesired short abortive sequences and the formation of more desired RNA product suggest a favorable interaction between T7 RNA polymerase and ATPαSe I.

To confirm the incorporation of ATPαSe I into RNA, the RNA transcript was subjected to digestion with snake venom phosphodiesterase I, which is an exonuclease that degrades both DNA and RNA successively in the 3′→5′ direction.[28] The result (Figure 3) indicates that the formed phosphoroselenoate RNA can indeed resist the enzymatic digestion. Its digestion was four to five times slower than that of the

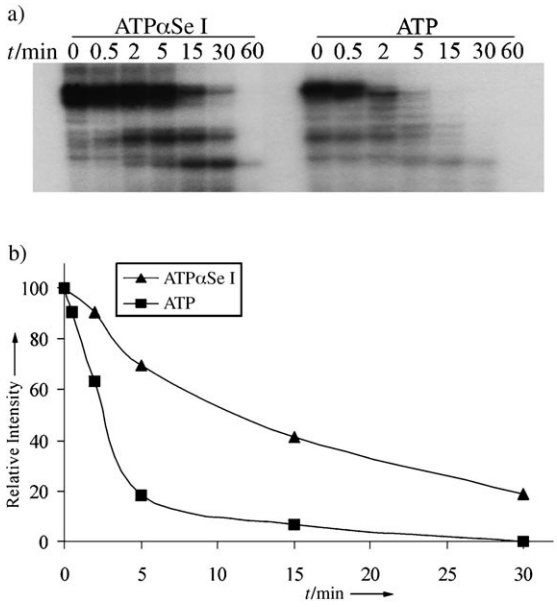


Figure 3. Time-course enzymatic digestion of phosphoroselenoate RNA with snake venom phosphodiesterase I. The results are presented as a) gel electrophoresis autoradiography and b) the digestion-versus-time plot.

corresponding nonmodified RNA. This finding is consistent with reports on the enzymatic digestion resistance of phosphoroselenoate DNA,[17] phosphorothioate DNA and RNA,[29] and boranophosphate DNA and RNA.[30]

In addition, the MALDI-TOF MS analysis[31] also confirms incorporation of 12 Se atoms per RNA molecule

(Figure 4). The measured average mass difference between the nonmodified RNA (11 605 Da) and the Se-modified RNA (12 359 Da, containing 12 Se-modified A units) is 754 Da. As the theoretical mass difference between the modified and

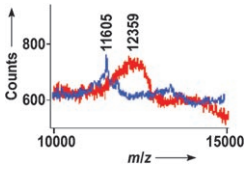


Figure 4. The MALDI-TOF MS analysis of Se-modified RNA transcription. Nonmodified RNA (blue) with a mass of 11 605 Da and Se-modified RNA (red) with an average mass of 12 359 Da. Their mass difference of 754 reflects incorporation of 12 Se-modified A units ($12\,Se-12\,O=12\times79-12\times16=756$).

nonmodified RNAs is 756 Da ($12\,Se-12\,O=12\times79-12\times16=756$), this experimental mass difference reflects the incorporation of 12 Se atoms per RNA molecule. The broad peak of the Se-modified RNA (average 12 359 Da) is probably caused by the satellite peaks from selenium isotopes (major ones: 76, 77, 78, 80, 82) and the formation of $N-1$ and $N+1$ minor products, which are commonly observed in RNA transcription.[22a] This minor 3′ heterogeneity of the transcribed RNA, also revealed by the gel electrophoresis analysis (Figure 2), can be removed by a ribozyme system.[32]

In summary, we have synthesized ATPαSe and enzymatically synthesized Se-modified RNA for the first time. We have found that only one diastereomer (ATPαSe I, presumably the S_P isomer) is efficiently recognized by T7 RNA polymerase for synthesizing diastereomerically pure phosphoroselenoate RNA (presumably R_P). As T7 RNA polymerase specifically recognizes just one diastereomer (ATPαSe I) and the other diastereomer (ATPαSe II) does not inhibit the transcription, prior separation of these two diastereomers is not necessary for preparation of diastereomerically pure RNA transcripts. In addition, our results suggest that ATPαSe I is an efficient substrate for T7 RNA polymerase and its use has shown advantages over ATP transcription: almost no short abortive sequences and a higher RNA transcription yield. Formation of almost no short abortive sequences also makes RNA purification by gel electrophoresis or HPLC easier. By following the established procedures,[9] this novel and efficient method can be used in particular for the preparation of long phosphoroselenoate RNA on a large scale (multimilligram quantity), which would be unattainable by using solid-phase synthesis. Introduction of multi-Se labels through transcription is very useful for studying RNA–protein complexes. Like the phosphorothioate[21] and boranophosphate[27] RNAs that have been previously synthesized, phosphoroselenoate RNAs are also expected to exhibit biochemical and biological properties that are closely related to those of natural RNA. Research in this area will undoubtedly open new frontiers in the fields of selenium chemistry and substrate recognition through RNA polymerase, in biochemistry studies of ribozyme catalysis, and in studies of the structures of functional RNAs and their complexes with proteins by using the MAD and/or single-wavelength anomalous-dispersion (SAD) methods.

Experimental Section

ATPαSe was synthesized by a modification of the procedures for the synthesis of NTPαS[21] and TTPαSe.[17] Briefly, 2′,3′-diacetyladenosine (35.13 mg, 0.10 mmol, Sigma) was placed in a 5-mL flask, dried under high vacuum overnight, and dissolved in a mixture of freshly distilled pyridine (0.10 mL) and dioxane (0.30 mL). The resultant solution was then injected dropwise over the course of about 5 min into a solution of 2-chloro-4*H*-1,3,2-benzodioxaphosphorin-4-one (21.0 mg, 0.10 mmol, 1 equiv, Aldrich) in dioxane (0.10 mL). The reaction was stirred at room temperature under dry argon for 10 min. A solution of tributylammonium pyrophosphate (64.1 mg, 0.14 mmol, 1.5 equiv, Sigma) in dry *N*,*N*-dimethylformamide (DMF, 0.28 mL) containing tributylamine (0.10 mL) was then injected, and the reaction was stirred for another 10 min. A solution of 3*H*-1,2-benzothiaselenol-3-one (43.0 mg, 0.20 mmol, 1.5 equiv) in dioxane (0.22 mL) was then added into the reaction mixture. Once selenization was completed (30 min, monitored by ^{31}P NMR spectroscopy), the reaction was quenched with water (1.0 mL) for 2 h.

The acetyl protecting groups were removed by hydrolysis with concentrated aqueous ammonia (3.0 mL) at 60 °C for 1.5 h. After most of the ammonia was removed by rotary evaporation, the pH value was adjusted to 7.0 by using an 80% acetic acid solution. A 100 mM solution of 1,4-dithiothreitol (DTT, 200 μL) was then added, and the crude product was transferred into a 15-mL centrifuge tube and spun for 3 min to remove the selenium metal. The supernatant was transferred into a 50-mL centrifuge tube, NaCl in water (3.0 M, 1.3 mL) was added, and the content was divided into two equal portions. Absolute ethanol (3 volumes, thoroughly purged with argon) was added to each portion, and the samples were placed in a −20 °C freezer for 10 min before centrifugation (10 min at 6000 rpm). After removal of the supernatant, the crude product was redissolved in water (500 μL) and purified by RP-HPLC with a Zorbax C18 column (9.4 × 250 mm). Samples were eluted (5 mL min^{-1}) with a linear gradient from buffer A (10 mM triethylammonium acetate (TEAAc), pH 7.0) to 20% buffer B (30% acetonitrile in water, 10 mM TEAAc, pH 7.0) over 20 min. The purified ATPαSe diastereomers were then analyzed by RP-HPLC (Figure 5a) and HR-MS (Figure 5b) and stored at −20 °C (13.78 mM in a solution of 10 mM tris(hydroxymethyl)aminomethane/HCl (Tris-HCl, pH 7.5) and 20 mM of DTT). The concentrations and quantities of the diastereomers were determined by UV analysis, which indicated 15% overall yield for each diastereomer (30% total yield), a satisfactory yield considering the many steps involved in the overall synthesis and HPLC purification. It is assumed that the selenium modification on the phosphate group does not alter the extinction coefficient of the modified nucleoside triphosphate.

Enzymatic synthesis of phosphoroselenoate 5′-*O*-adenosine-labeled RNA: The DNA top strand of the T7 RNA polymerase promoter (5′-GCGTAATACGACTCACTATAG-3′) and the DNA template (3′-CGCATTATGCTGAGTGATATCCGTTGGAC-TACTCCGGCTTTCCGGCTTTGCATGT-5′) were prepared by oligonucleotide solid-phase synthesis. AmpliScribe T7 Transcription Kits (Epicentre) were used for the in vitro transcription, where the DNA template and the top strand were added in equal molar amounts into a cocktail containing all the NTPs, except ATP, ATPαSe I, and ATPαSe II. α-[^{32}P]CTP (Perkin-Elmer) was included in the cocktail to body-label the RNA transcript. Equal amounts of the cocktail were then added to vials containing the appropriate amounts of ATPαSe I, ATPαSe II, ATP, or the negative control (water). The reaction was initiated by the addition of the T7 RNA polymerase solution (provided within the kit), and the mixtures were incubated at 37 °C.

For the time-course experiments, aliquots (4 μL) were withdrawn from the reaction mixtures at the appropriate time points and added to a loading dye containing 100 mM EDTA (4 μL); the samples were then placed on dry ice. The samples were analyzed by electrophoresis on a 12.5% gel, and quantitation was carried out on a BioRad phosphorimager (see, for example, Figure 2). A typical transcription

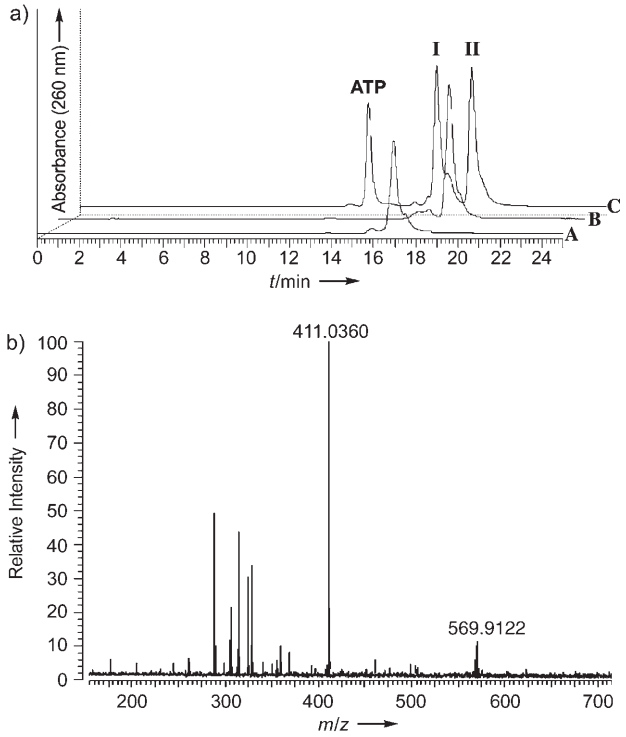


Figure 5. a) HPLC analysis of ATPαSe diastereomers: A) ATPαSe I; B) ATPαSe II; C) a mixture of ATP, ATPαSe I, and ATPαSe II. b) HR-MS analysis of ATPαSe I. (The spectrum for ATPαSe II is almost the same.) Molecular formula: $C_{10}H_{16}N_5O_{12}P_3Se$; calculated mass $[M-H^+]^-$: 569.9101; measured mass: 569.9122.

reaction contained the top strand (0.1 μM), the template (0.1 μM), all NTPs (1.0 mM, including ATPαSe I and II), α-[^{32}P]CTP, the polymerase (0.06 μL of T7 RNA polymerase per 1 μL of transcription reaction, Epicentre), and the buffer provided. By using the Epicentre kit, large-scale RNAs (6 mg) were prepared under the typical conditions of the template (1 μM) and with all NTPs (7.5 mM, including ATPαSe) in a transcription reaction (1 mL reaction volume) that took 2 h.

MALDI-TOF analysis of the RNAs: All mass spectra (with crude transcribed samples) were recorded on a biomass spectrometer in linear-negative mode with delayed extraction. 3-hydroxypicolinic acid (3-HPA)/diammonium citrate (9:1) in water was used as the matrix. A mixture (1 μL) of sample and matrix (1:30) was spotted and dried naturally before the analysis. 25 kV were applied as the acceleration voltage. A mass range of 1000–20000 was scanned. Each spectrum was summed from multiple spectra at different spots. Proteins, such as insulin, thioredoxin, and apomyoglobin, were used as external standards.

Exonuclease digestion analysis: The phosphoroselenoate RNA transcripts were digested with snake venom phosphodiesterase I (USB), following Se-modified RNA transcription by using the same protocol as that described above. The transcribed RNA was desalted by centrifugation (by using a membrane with a cut-off of 3000 Da, 3 times), and the RNA was digested with snake venom phosphodiesterase I (0.001 U μL^{-1}, USB) in its buffer over 60 min. The digested RNA samples were analyzed by PAGE (Figure 3).

Received: June 24, 2005
Revised: September 9, 2005
Published online: November 22, 2005

Keywords: nucleic acids · nucleotides · RNA synthesis · RNA transcription · selenium

[1] G. Storz, *Science* **2002**, *296*, 1260–1263.
[2] M. Mandal, R. R. Breaker, *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 451–463.
[3] M. T. McManus, P. A. Sharp, *Nat. Rev. Genet.* **2002**, *3*, 737–747.
[4] M. Egli, *Curr. Opin. Chem. Biol.* **2004**, *8*, 580–591.
[5] S. R. Holbrook, S. H. Kim, *Biopolymers* **1997**, *44*, 3–21.
[6] W. A. Hendrickson, *Trends Biochem. Sci.* **2000**, *25*, 637–643.
[7] W. A. Hendrickson, *J. Synchrotron Radiat.* **1999**, *6*, 845–851.
[8] W. S. Marshall, R. J. Kaiser, *Curr. Opin. Chem. Biol.* **2004**, *8*, 222–229.
[9] A. Ke, J. A. Doudna, *Methods* **2004**, *34*, 408–414.
[10] A. R. Ferre-D'Amare, K. Zhou, J. A. Doudna, *J. Mol. Biol.* **1998**, *279*, 621–631.
[11] A. R. Ferre-D'Amare, K. Zhou, J. A. Doudna, *Nature* **1998**, *395*, 567–574.
[12] N. Carrasco, D. Ginsburg, Q. Du, Z. Huang, *Nucleosides Nucleotides Nucleic Acids* **2001**, *20*, 1723–1734.
[13] Q. Du, N. Carrasco, M. Teplova, C. J. Wilds, M. Egli, Z. Huang, *J. Am. Chem. Soc.* **2002**, *124*, 24–25.
[14] M. Teplova, C. J. Wilds, Z. Wawrzak, V. Tereshko, Q. Du, N. Carrasco, Z. Huang, M. Egli, *Biochimie* **2002**, *84*, 849–858.
[15] Y. Buzin, N. Carrasco, Z. Huang, *Org. Lett.* **2004**, *6*, 1099–1102.
[16] N. Carrasco, Y. Buzin, E. Tyson, E. Halpert, Z. Huang, *Nucleic Acids Res.* **2004**, *32*, 1638–1646.
[17] N. Carrasco, Z. Huang, *J. Am. Chem. Soc.* **2004**, *126*, 448–449.
[18] W. A. Hendrickson, J. R. Horton, D. M. LeMaster, *EMBO J.* **1990**, *9*, 1665–1672.
[19] S. E. Ealick, *Curr. Opin. Chem. Biol.* **2000**, *4*, 495–499.
[20] C. J. Wilds, R. Pattanayek, C. Pan, Z. Wawrzak, M. Egli, *J. Am. Chem. Soc.* **2002**, *124*, 14910–14916.
[21] a) F. Eckstein, *Annu. Rev. Biochem.* **1985**, *54*, 367–402; b) J. Ludwig, F. Eckstein, *J. Org. Chem.* **1989**, *54*, 631–635; c) F. Eckstein, *Biochimie* **2002**, *84*, 841–848.
[22] a) J. F. Milligan, D. R. Groebe, G. W. Witherell, O. C. Uhlenbeck, *Nucleic Acids Res.* **1987**, *15*, 8783–8798; b) G. M. Cheetham, T. A. Steitz, *Science* **1999**, *286*, 2305–2309; c) G. M. Cheetham, D. Jeruzalmi, T. A. Steitz, *Nature* **1999**, *399*, 80–83.
[23] a) Y. W. Yin, T. A. Steitz, *Cell* **2004**, *116*, 393–404; b) D. Temiakov, V. Patlan, M. Anikin, W. T. McAllister, S. Yokoyama, D. G. Vassylyev, *Cell* **2004**, *116*, 381–391.
[24] P. M. Burgers, F. Eckstein, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 4798–4800.
[25] A. D. Griffiths, B. V. Potter, I. C. Eperon, *Nucleic Acids Res.* **1987**, *15*, 4145–4162.
[26] F. Eckstein, H. Gindl, *Eur. J. Biochem.* **1970**, *13*, 558–564.
[27] a) K. He, A. Hasan, B. Krzyzanowska, B. R. Shaw, *J. Org. Chem.* **1998**, *63*, 5769–5773; b) A. H. S. Hall, J. Wan, E. E. Shaughnessy, B. R. Shaw, K. A. Alexander, *Nucleic Acids Res.* **2004**, *32*, 5991–6000.
[28] L. Dolapchiev, E. Sulkowski, M. Laskowski, *Biochem. Biophys. Res. Commun.* **1974**, *61*, 274–279.
[29] F. Eckstein, H. Gindl, *FEBS Lett.* **1969**, *2*, 262–264.
[30] K. He, B. R. Shaw, *Nucleic Acids Res. Symp. Ser.* **1999**, *41*, 99–100.
[31] Y.-S. Kwon, K. Tang, C. R. Cantor, H. Koester, C. Kang, *Nucleic Acids Res.* **2001**, *29*, e11.
[32] A. R. Ferre-D'Amare, J. A. Doudna, *Nucleic Acids Res.* **1996**, *24*, 977–978.

