5072–5076 *Nucleic Acids Research, 1997, Vol. 25, No. 24*

Test of the potential of a dATP surrogate for sequencing via MALDI-MS

Swanee Jacutin, Alex J. Zhang, David H. Russell, Richard A. Gibbs[1] and Kevin Burgess*

Department of Chemistry, Texas A&M University, College Station, TX 77843-3255, USA and [1]Department of Molecular and Human Genetics, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX 77030, USA

Received July 17, 1997; Revised and Accepted October 24, 1997

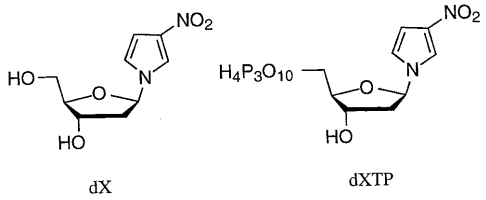
ABSTRACT

1-(2′-Deoxy-β-D-ribofuranosyl)-3-nitropyrrole phosphate was incorporated into a DNA decamer and analyzed via matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS). The extent and composition of the various fragment peaks were compared with those in the MALDI-MS spectrum of dT_4AT_5. The nitropyrrole-containing oligomer proved to be more robust. Two different DNA template assays were then used to attempt to identify DNA replicating enzymes that would incorporate the corresponding triphosphate, i.e. 1-(2′-deoxy-β-D-ribofuranosyl)-3-nitropyrrole triphosphate (dXTP). It was shown that dXTP was not incorporated by some enzymes and it inhibited others. However, DNA polymerase I Klenow fragment and avian myeloblastosis virus reverse transcriptase incorporated dXTP in place of dATP and then replicated the template overhang in the usual way. The potential of dXTP as a surrogate for dATP in DNA sequencing with MALDI-MS analysis is discussed.

INTRODUCTION

The Human Genome Project could be accelerated by methodologies for high throughput analyses of nested chain terminated DNA fragments that avoid gel electrophoresis. One of the most promising techniques in this regard is matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS). Strand fragmentation in the spectrometer contributes to the limitations currently encountered for MALDI-MS analyses of DNA mixtures. These fragmentation processes are thought to arise from protonation of basic sites on the aromatic nucleobase moieties. The extent of fragmentation is dependent on the sequence; T oligonucleotide residues are more stable than A, C and G, in that order. For this reason a surrogate for A, C and/or G that is less vulnerable to fragmentation, i.e. less basic, has potential for DNA sequencing.

With regard to incorporation of unnatural dNTP analogs by DNA replicating enzymes, relatively minor structural modifications, wherein the surrogates have similar structures to the parent purines or pyrimidines, have been investigated. For instance, DNA oligomers containing 7-deaza-dATP and 7-deaza-dGTP have been shown to give less fragmentation in MALDI-MS analyses than the corresponding natural strands. Moreover, 7-deaza-dGTP can be incorporated by DNA replicating enzymes to give nested chain terminated DNA fragments; this technique is frequently used to overcome 'GC compressions'. Biocatalytic replication using nucleobase surrogates with structures very different from those of the natural dNTPs, however, are relatively unexplored. It is not unreasonable to assume that very different structures can be incorporated, as indicated by the following recent observations. Bergstrom and co-workers indicated that 1-(2′-deoxy-β-D-ribofuranosyl)-3-nitropyrrole phosphate (dX) has characteristics that are desirable for a 'universal base'; small oligonucleotides containing dX bound with almost equal affinity to templates containing any of the four natural bases opposite the dX residue. Other work by Kool's group indicates that base stacking effects are more important than previously thought for matching two DNA strands and that unnatural dNTP analogs can do this well (11).



Scheme 1.

We reasoned that dX has no particularly basic centers and that the Bergstrom work showed that this nucleoside was functional in DNA complements. This paper describes experiments to explore the efficacy of 3-nitropyrrole as a surrogate for adenine in mass spectroscopic analyses of oligonucleotides. Screens for incorporation of 1-(2′-deoxy-β-D-ribofuranosyl)-3-nitropyrrole triphosphate (dXTP) by DNA replicating enzymes are also described. This work therefore represents a preliminary step in the search for nucleobase surrogates that can be incorporated via biocatalytic synthesis to give robust DNA complements for MALDI-MS analyses.

MATERIALS AND METHODS

Synthesis of 1-(2′-deoxy-β-D-ribofuranosyl)-3-nitropyrrole and its phosphoramidite

1-(2′-Deoxy-β-D-ribofuranosyl)-3-nitropyrrole and its phosphoramidite were prepared via literature procedures. dX was obtained as a white solid (0.27 g, 99% yield). ^{1}H NMR (CD_3OD, 200 MHz): δ (p.p.m.) 7.89 (dd, J = 1.9 Hz, J = 2.4 Hz, 1H), 6.92 (dd, J = 2.4 Hz, J = 3.4 Hz, 1H), 6.61 (dd, J = 1.9 Hz, J = 3.4 Hz, 1H), 5.89 (t, J = 6.4 Hz, 1H), 4.41–4.29 (m, 1H), 3.93–3.82 (m, 1H), 3.69–3.52

*To whom correspondence should be addressed. Tel: +1 409 845 4345; Fax: +1 409 845 8839; Email: burgess@chemvx.tamu.edu

(m, 2H), 2.41–2.25 (m, 2H). ^{13}C NMR (CD_3OD, 50 MHz): δ (p.p.m.) 137.9, 121.6, 120.9, 106.1, 90.0, 88.8, 72.1, 62.9, 42.5. MS (FAB): *m/z* 229 $[M+H]^+$.

1-(2′-Deoxy-5′-dimethoxytrityl-β-D-ribofuranosyl)-3-nitropyrrole-3′-*O*-(2-cyanoethyl-*N*,*N*,-diisopropylphosphoramidite) was prepared via literature procedures. ^{31}P NMR (121 MHz, acetone-d_6): δ (p.p.m.) 150.0, 148.8. MS (FAB): *m/z* 731 $[M+H]^+$.

The phosphoramidite prepared as above was then incorporated into oligonucleotide strand dT_4XT_5 by phosphoramidite chemistry using an ABI 394 DNA synthesizer. The dT_4XT_5 and dT_4AT_5 (prepared via standard automated synthesis) samples were purified by RP-HPLC on a Vydac C18 column (5 µm, 4.6 × 250 mm). Buffer A was 0.1 M triethylammonium acetate (TEAA), pH 7.0, and buffer B was 30% 0.1 M TEAA and 70% acetonitrile. Separations were performed using a linear gradient of 90–50% A over 20 min, then 10% A for 1 min, then 90% B for 1 min with a flow rate of 1 ml/min.

Synthesis of 1-(2′-deoxy-β-D-ribofuranosyl)-3-nitropyrrole triphosphate (dXTP)

The nucleoside was converted to its triphosphate via the Kovács procedure. Thus, 1-(2′-deoxy-β-D-ribofuranosyl)-3-nitropyrrole (0.042 g, 0.18 mmol, 1 equiv.) and 1,8-bis(dimethylamino)naphthalene (0.058 g, 0.27 mmol, 1.5 equiv) were dissolved in trimethylphosphate (0.45 ml, 0.4 M) under a nitrogen atmosphere and stirred at 0°C for 10 min. Phosphorus oxychloride (18 µl, 0.2 mmol, 1.1 equiv) was added and the reaction was stirred at 0°C for 2 h. Tri-*n*-butylammonium pyrophosphate (0.49 g, 0.9 mmol, 5 equiv.) and tributylamine (21 µl) in DMF (70 µl) were then added to the solution. After 2 min the reaction was quenched with 1 M triethylammonium bicarbonate (TEAB), pH 7.5. The solution was concentrated under high vacuum and the triphosphate product was partially purified using a DEAE–cellulose column eluted with a linear gradient of 0.01–0.5 M aqueous TEAB. The nucleoside triphosphate was further purified by RP-HPLC on a Vydac C18 column (5 µm, 4.6×250 mm). Buffer A was 0.1 M TEAA, pH 7.0, and buffer B was 30% 0.1 M TEAA and 70% acetonitrile. Purification of a small sample was performed using the gradient conditions 100–80% A over 10 min, 80–0% A over 5 min, then 100% B for 10 min. Throughout the flow rate was 1 ml/min. The product was isolated as a white solid after lyophilization: ^{31}P NMR (121 MHz, D_2O): δ (p.p.m.) –7.10 (d, *J* = 20.1 Hz, 1P), –8.29 (d, *J* = 20.1 Hz, 1P), –20.06 (t, *J* = 20.1 Hz, 1P). MALDI/TOF: *m/z* 466.8 [M–H]. The reproducibility of the assays described below was dependent upon the purity of the triphosphate; one or more contaminants (possibly including the monophosphate) appeared to cause enzyme inhibition.

MALDI-MS

The HPLC-purified oligonucleotides were dissolved in 1:1 H_2O/MeCN to give an ~200 µM solution. The MALDI-TOF experiments were performed on a Voyager-Elite XL instrument (PerSeptive Biosystems) operated in the negative ion mode. In these experiments freshly prepared 0.15 M α-cyano-4-hydroxycinnamic acid in methanol was used as matrix. A 10 µl aliquot of matrix, 5 µl 0.5 M diammonium citrate and 1 µl analyte solution were briefly mixed. A 0.5 µl aliquot of this solution was applied to the mass spectrometer autosampler plate and dried in air; the plate was then inserted into the sampling chamber. The quantity of sample used in the MALDI experiments was in the 10 pmol region. Singly and doubly charged molecular ions of oligomer dT_{10} were used for mass calibration.

Polymerases and DNA templates

All the polymerases [DNA polymerase I Klenow fragment; exonuclease-free (exo-free) Klenow fragment; avian myeloblastosis virus reverse transcriptase (AMV-RT); Moloney murine leukaemia virus reverse transcriptase (M-MLV-RT); T7 Sequenase™ version 2.0 DNA polymerase; Δ*Taq* version 2.0 DNA polymerase] used in the assay described below were purchased from Amersham. T4 kinase was purchased from Pharmacia Biotech. M13mp18 single-stranded DNA template was purchased from BioRad Laboratories. The Universal Primer was provided by Mike Metzker (Baylor College of Medicine). The oligonucleotide template 5′-GCCAC-CTCCACTGGCCGTCGTTTTAC-3′ was synthesized using an ABI model 394T DNA/RNA synthesizer (underlining indicates that part which was not paired to a primer in the subsequent experiments). All non-radioactive nucleotides were purchased from Pharmacia Biotech and [γ-^{32}P]ATP was purchased from Amersham.

Polymerase incorporation assay

Two different templates were used to test incorporation of dXTP. In the first ^{32}P-labeled universal primer was annealed to single-stranded M13mp18 DNA (0.5 pmol to 0.2 µmol respectively per 5 µl) by heating to 80°C for 5 min and cooling slowly to 25°C (~30–45 min). Enzymatic extension of the primer–template complex was performed under standard conditions for these enzymes except that selected natural nucleotides were replaced by and mixed with various concentrations of the nucleotide analog. Control experiments were performed with natural nucleotides in the presence of dideoxynucleotides and water. For each reaction 5 µl mixtures of each enzyme, nucleotide and nucleotide analog, in the buffer solution used by the supplier to dissolve the enzyme, were added to 5 µl annealed primer–template samples. Final concentrations of nucleotides, enzymatic units and incubation temperature are shown in Table 1. The reaction mixtures were incubated for 25 min (M13mp18 template), then stopped by addition of 5 µl stop solution (consisting of 98% formamide, 10 mM EDTA, pH 8.0, 0.025% bromophenol blue and 0.025% xylene cyanol). The samples were heated to 85°C for 3 min, chilled on ice, then 3 µl were loaded on 10% polyacrylamide gels. Following electrophoresis the gel was fixed in an aqueous solution of 10% acetic acid, 10% methanol, dried for 1 h and autoradiographed on HyperfilmTM-MP (Amersham) or scanned on a Fuji Bas 2000 phosphorimager (Gene Technologies Laboratory, Department of Biology, Texas A&M University).

For the second assay the same set of conditions were applied except that the ^{32}P-labeled universal primer was annealed to the synthetic oligonucleotide template 5′-GCCACCTCCACTGGCC-GTCGTTTTAC-3′ (0.05–0.1 pmol respectively per 5 µl), the incubation time was 20 min and a 20% polyacrylamide gel was used. Enzymatic extensions were performed via the method described above.

RESULTS AND DISCUSSION

MALDI-MS

To test the relative proclivities of dA and dX residues in oligonucleotides to fragment, MALDI-MS spectra of dT_4AT_5 and dT_4XT_5 were recorded under the same conditions, then compared.

Table 1. Enzymatic conditions

Enzymatic conditions	Sequenase™ version 2.0	Δ*Taq*	Klenow, exo-free[a]	AMV-RT	M-MuLV-RT	Klenow, large fragment[a]
Incubation temperature (°C)	37	68	37	37	37	37
Units	3	8	4	4	200	5
dNTP (μM)	0.25	0.25	2.0	0.25	2.5	2.0
ddNTP (μM)	0.25	2.5	20	2.5	20	20
dXTP (μM)[b]	75	75	75	75	75	75

[a]M13 mp18 template was used.
[b]dXTP at 1× concentration, i.e. 0.075 mM; 5× = 0.375 mM; 10× = 0.75 mM.

Figure 1 shows these spectra. The spectrum of the dA-containing oligomer showed fragmentation associated with loss of the adenine base and cleavage of the strand, giving the following fragments: dT_4, pT_5 and pT_5 + matrix. Conversely, no strand fragmentation was observed for the dT_4XT_5 oligonucleotide. The only fragment peak detected corresponds to loss of an oxygen atom, presumably from the nitro group of the nitropyrrole moiety. A peak corresponding to the doubly charged molecular ion was present in both spectra.

Incorporation assays

The nucleotide analog dXTP was tested as a surrogate for dATP using several commercially available polymerases. An assay previously developed to detect incorporation of 3′-blocked nucleotides by DNA replicating enzymes was adapted for this purpose. The assay began with a DNA template annealed to a ^{32}P-labeled primer. The DNA replicating enzyme, natural dNTP (and/or dNTPs) and one unnatural dNTP were then added to the reaction tubes. In the critical experiments read-through beyond a dT residue in the template would only be possible by incorporation of dX or a mismatched natural dN. Control experiments were run in parallel using appropriate ddNTPs and blanks (water control) giving termination at this crucial site and at the preceding base respectively. This protocol differentiated between no incorporation of dX, incorporation with chain termination and incorporation of dX followed by read-through beyond this point.

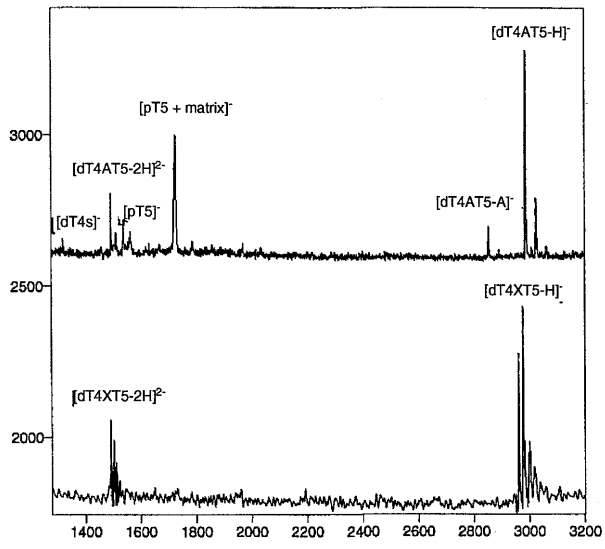


Figure 1. MALDI-MS of 5′-dT_4AT_5 (top) and 5′-dT_4XT_5 (bottom) using α-cyano-4-hydroxycinnamic acid matrix. The peaks are those associated with the molecular ion, the 3′-pT_5 fragment (where p is a 5′-phosphate) and the 5′-T_4s (where s is a 3′-anhydro-2-deoxyribofuranose sugar fragment).

Three scenarios arose in these assays, as shown in Figure 2. The sequence complementary to the template is shown to the right of each gel. The critical point in this sequence is the first A in the complement, because no dATP was added in all the experiments described below. Figure 2A shows a gel pattern indicating that the unnatural analog (dXTP) was not incorporated. This is shown in lanes 1–3, wherein increasing concentrations of dXTP were added and the band terminates at the same position as the blank/water control (lane 5). This was one residue lower than in lane 4, where ddATP was added. Figure 2B is illustrative of the second scenario. This displays a gel pattern wherein incorporation was observed. In all experiments where dXTP was added (lanes 1–5) there was a band corresponding to termination opposite the first T in the template, but no bands above that. This indicates that dXTP was incorporated but no read-through beyond this incorporation point occurred. Figure 2C depicts the third scenario, i.e. incorporation and read-through. In this case incorporation occurred at the crucial template T residue and the enzyme was able to carry on replication for several bases after this. In fact, there were two consecutive T residues in this template, so the enzyme incorporated beyond this TT site.

Throughout these studies relatively high concentrations of dXTP were required to obtain incorporation or incorporation/read-through. For instance, when DNA polymerase I Klenow large fragment was tested in conjunction with the M13mp18 DNA template, 75 μM dXTP gave mainly termination; a read-through band was detected when the concentration was doubled and predominant read-through was observed at a dXTP concentration of 0.75 mM. Consequently, it appears that high concentrations of dXTP can act as a surrogate for dATP.

These assays also provided some evidence that dXTP is not an effective surrogate for TTP. It appears that dXTP can be incorporated in place of T, but read-through after this was not observed. Thus when DNA polymerase I Klenow large fragment was used in conjunction with the M13mp18 DNA template high concentrations of dXTP caused read-through beyond the template TT region but not past the first template A encountered, i.e. incorporation/termination. These findings were confirmed in experiments in which dATP was added (Fig. 2C, lanes 8–12).

Each enzyme was tested using two different templates: a long one, M13mp18 DNA, and a short synthetic 26mer. The purpose of this was to determine if the template length had an effect on

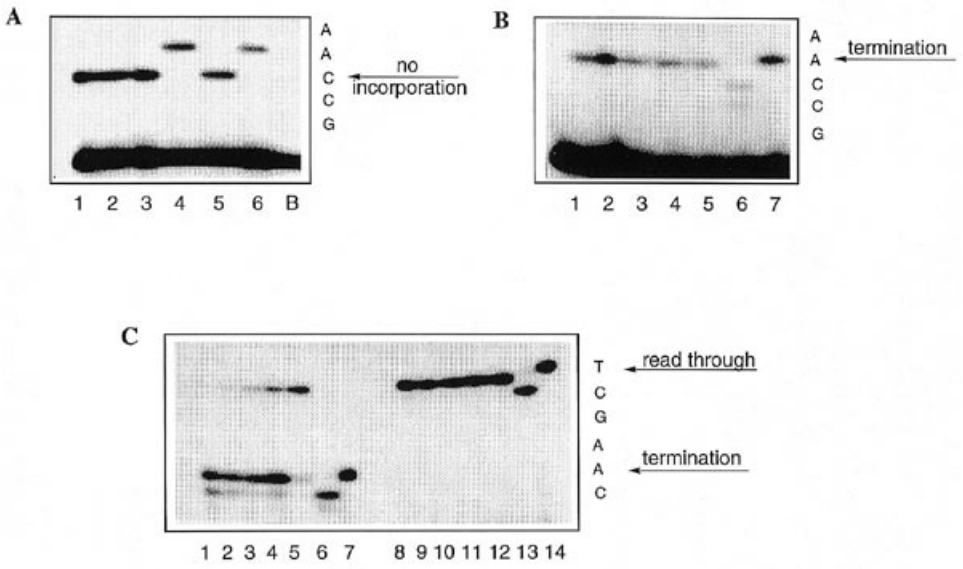


Figure 2. Attempted incorporation of dXTP by: (**A**) Sequenase version 2.0; (**B**) exo-free Klenow fragment; (**C**) Klenow large fragment of DNA polymerase I. The sequence complementary to the template is shown on the right of each gel. ^{32}P-Labeled universal primer (–20) was annealed to single-stranded M13mp18 DNA template. (A) Lanes 1–3 contained 2.0 µM dGTP and dCTP with 0.075, 0.15 and 0.225 mM dXTP respectively. Lane 4 contained 2.0 µM dGTP and dCTP with 20 µM ddATP. Lane 5 was identical to lane 4 except that it did not contain ddATP. Lane 6 is identical to lane 4 with 0.75 mM dXTP. For (B) and (C) the conditions were as follows: lanes 1–5 contained 2.0 µM dGTP and dCTP with 0.075, 0.15, 0.225, 0.30 and 0.75 mM dXTP respectively; lane 6 contained 2.0 µM dGTP and dCTP; lane 7 was identical to lane 6 except that it also contained 20 µM ddATP. For (C) only: lanes 8–12 contained 2.0 µM dGTP, dCTP and dATP with 0.075, 0.15, 0.225, 0.30 and 0.75 mM dXTP respectively; lane 13 contained 2.0 µM dGTP, dCTP and dATP; lane 14 was identical to lane 13 except that it also contained 20 µM ddTTP.

incorporation and, in fact it did, *vide infra*. For the short template the overhang sequence was 3′-CC*T*CCACCG-5′, consequently the only natural nucleotide used was dGTP (to complement the leading CC template sequence). Combinations of dGTP and dCTP were used in the experiments with the longer template.

Table 2. Activity of nitropyrrole triphosphates against commercially available DNA polymerases

	M13 mp18 DNA	Synthetic 26mer
T7 Sequenase^{TM}k	No incorporation	No incorporation
Δ*Taq*	No incorporation	Termination
Exo-free Klenow fragment	Termination	Not assayed
AMV-RT	No incorporation	Read-through
M-MuLV-RT	No incorporation	No incorporation
Klenow large fragment	Read-through	Inhibition

Table 2 summarizes the data gathered from the enzymatic screens. T7 Sequenase^{TM} and M-MuLV-RT showed no evidence of incorporation of dX with either template (cf. Fig. 2A). Δ*Taq* and exo-free Klenow fragment gave termination with no read-through (cf. Fig. 2B). DNA polymerase I Klenow large fragment, gave incorporation with read-through (cf. Fig. 2C) but only with the longer template (M13mp18). Conversely, AMV-RT showed incorporation with read-through only for the shorter template (a synthetic 26mer).

Possible misincorporation of natural bases was a concern throughout these experiments, consequently the relative amounts of the dNTPs and dXTP were critical. For each enzyme the minimum concentrations of dNTPs and ddNTPs appropriate for replication were determined via a published procedure. Use of these minimum concentrations ensured efficient incorporation of natural bases at sites in which they would be paired in a Watson–Crick fashion but reduced the likelihood of incorporation of a natural dNTP at the critical site to give a mismatch. Evidence that misincorporation of a natural dNTP did not occur in the experiments with dXTP was inferred from the controls wherein no dXTP was added. Replication strands observed in these experiments terminated at the base immediately preceding the proposed site of incorporation of dXTP. Further evidence that the read-through observed is not due to a forced mismatch of a natural base is inferred from the experiments in which dATP was added to test if dXTP could act as a TTP surrogate (*vide supra*). Only incorporation/termination occurred under these conditions, even though dATP, dCTP and dGTP were present. If the observations we attribute to dXTP incorporation as a dATP surrogate (wherein only dGTP and/or dCTP were present) were due to forced mismatches, then incorporation would also be expected in the TTP surrogate probes and it was not observed.

CONCLUSION

We have shown that a nucleobase that has no close structural similarity to adenosine is more robust with respect to MALDI-MS analyses and can act as a surrogate for dATP in enzymatic DNA syntheses. It was incorporated by two naturally occurring enzymes.

The results described here lay foundations for several related studies. These include screening mutated enzymes with higher substrate tolerances (e.g. AmpliTaq® FS and Taquenase®), systematic variations in the reaction conditions (e.g. substitution of magnesium in the buffer medium with manganese salts) and screens of other nucleobase analogs (e.g. those derived from 5-nitroindole).

We intend to continue this screening process, then focus on the issues related to kinetics of incorporation and enzyme processivity for the most efficacious combinations. Refined systems for incorporation of unnatural bases could offer considerable advantages in analyses of artificial DNA nested chain terminated strands.

ACKNOWLEDGEMENTS

We would like to thank Drs Mike Metzker and M.Ali Ansari for very helpful discussions and Dr Aroonsiri Shitangkoon for assistance with some HPLC purifications. Financial support was provided by The Texas Advanced Technology Program and The Robert A.Welch Foundation and K.B. thanks the NIH for a Research Career Development Award and the Alfred P.Sloan Foundation for a fellowship. NMR facilities were provided by the NSF via the chemical instrumentation program. The MALDI-MS research (financial support for A.J.Z.) is funded by the US Department of Energy, Division of Chemical Sciences, Office of Basic Energy Sciences.

REFERENCES

1 Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
2 Smith,L.M. (1993) *Science*, **262**, 530–532.
3 Zhu,L., Parr,G.R., Fitzgerald,M.C., Nelson,C.M. and Smith,L.M. (1995) *J. Am. Chem. Soc*., **117**, 6048–6056.
4 Nordhoff,E., Karas,M., Cramer,R., Hahner,S., Hillenkamp,F., Kirpekar,F., Lezius,A., Muth,J., Meier,C. and Engels,J.W. (1993) *Rapid Commun. Mass Spectrom*., **7**, 99–112.
5 Schneider,K. and Chait,B.T. (1995) *Nucleic Acids Res*., **23**, 1570–1575.
6 Kirpekar,F., Nordhoff,E., Kristiansen,K., Roepstorff,P., Hahner,S. and Hillenkamp,F. (1995) *Rapid Commun. Mass Spectrom*., **9**, 525–531.
7 Nordhoff,E., Cramer,R., Karas,M., Hillenkamp,F., Kirpekar,F., Kristiansen,K. and Roepstorff,P. (1993) *Nucleic Acids Res*., **21**, 3347–3357.
8 Schneider,K. and Chait,B.T. (1993) *Org. Mass Spectrom*., **28**, 1353–1361.
9 Bergstrom,D.E., Zhang,P., Toma,P.H., Andrews,P.C. and Nichols,R. (1995) *J. Am. Chem. Soc*., **117**, 1201–1209.
10 Nichols,R., Andrews,P.C., Zhang,P. and Bergstrom,D.E. (1994) *Nature*, **360**, 492–493.
11 Kuckian,K.M., Schweitzer,B.A., Ren,R.X.-F., Sheils,C.J., Paris,P.L., Tahmassebi,D.C. and Kool,E.T. (1996) *J. Am. Chem. Soc*., **118**, 8182–8183.
12 Caruthers,M.H. (1991) *Acc. Chem. Res*., **24**, 278.
13 Kovács,T. and Ötvös,L. (1988) *Tetrahedron Lett*., **29**, 4525–4528.
14 Metzker,M.L., Raghavachari,R., Richards,S., Jacutin,S.E., Civitello,A., Burgess,K. and Gibbs,R.A. (1994) *Nucleic Acids Res*., **22**, 4259–4267.
15 Tabor,S. and Richardson,C.C. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 6339–6343.
16 Tabor,S. and Richardson,C.C. (1989) *Proc. Natl. Acad. Sci*., *USA*, **86**, 4076–4080.
17 Loakes,D., Hill,F., Linde,S. and Brown,D.M. (1995) *Nucleosides Nucleotides*, **14**, 1001–1003.
18 Loakes,D., Brown,D.M., Linde,S. and Hill,F. (1995) *Nucleic Acids Res*., **23**, 2361–2366.