Benchmarks

REFERENCES

1. Bentzen, P., A.S. Harris and J.M. Wright. 1991. Cloning of hypervariable minisatellite and simple sequence microsatellite repeats for DNA fingerprinting of important species of salmonids and tilapias, p. 243-262. In T. Burke, G. Dolf, A.J. Jeffreys and R. Wolff (Eds.), *DNA FingerPrinting: Approaches and Applications*. Birkhaeuser Verlag, Basel, Switzerland.
2. Griffin, H.G. and A.M. Griffin. 1993. DNA sequencing. Recent innovations and future trends. *Appl. Biochem. Biotechnol.* 38:147-159.

This work was supported by a grant from the U.S. Fish and Wildlife Service's Fisheries Stewardship Program. The mention of trade names or commercial products in this report does not constitute endorsement or recommendation for use by the federal government. Address correspondence to, Dr. Wallace G. Buchholz, U.S. Fish and Wildlife Service, 1011 E. Tudor Rd., Anchorage, AK 99503, USA. Internet: wally_buchholz@fws.gov

Received 19 April 1999; accepted 12 July 1999.

Wallace G. Buchholz, Steve J. Miller and William J. Spearman
*U.S. Fish and Wildlife Service
Anchorage, AK, USA*

Sequencing Errors in Reactions Using Labeled Terminators

BioTechniques 27:648-650 (October 1999)

Premature chain termination is a common problem in sequencing reactions and can result in multiple co-migrating bands that make accurate sequence determination difficult. The use of labeled dideoxynucleotides is generally assumed to alleviate this problem because only dideoxynucleotide-terminated chains are detected (1). However,

we have found that under some circumstances troublesome artifacts can still be seen when this class of terminators is used. In particular, we have observed such artifacts in sequencing G+C-rich templates when dITP is used in place of dGTP.

Templates that contain G+C-rich regions are prone to compression artifacts in standard reactions containing dGTP. These compressions are due to the formation of secondary structures on the nascent strand. The most stable, and therefore the most troublesome, structures are those that involve a number of G·C or G·G pairs. Figure 1 shows an example of compressions seen in standard reactions with the Thermo Sequenase™ Radiolabeled Terminator Cycle Sequencing Kit (Amersham Pharmacia Biotech, Piscataway, NJ, USA). Deoxyinosine triphosphate, an analog of dGTP, is commonly used to relieve compressions. Since dITP lacks an N²

amino group, it forms a less stable bond with cytosines and other guanines. Incorporation of inosine destabilizes the secondary structures that can form in the nascent DNA strand, thereby reducing compression artifacts. The use of the dITP master mixture, supplied with the Thermo Sequenase kit, eliminated the compressions but also produced a number of instances of two co-migrating bands terminated by a different dideoxynucleotide, (Figure 1, 1–9).

Because these bands are not seen when dGTP is used, nor are they seen at every position, they do not result from the presence of incomplete synthesis products in the primer. Most of the spurious products are not seen when a more distal primer is used (Figure 1, primer 2). Therefore, the sequencing artifacts observed with dITP and primer 1 probably do not result from a mixture of sequences, a problem associated with the preparation of the

Figure 1. Sequencing artifacts in cycle sequencing reactions containing radiolabeled terminators and dITP. Sequencing of a PCR fragment was carried out using the Thermo Sequenase radiolabeled terminator cycle sequencing kit using the conditions and buffers provided by the supplier, either without further modification or after the addition of 50 mM KCl. The brackets on the left-hand side of the figure indicate the region of compressions seen when dGTP is used. The solid lines numbered 1–9 indicate the location of the spurious products seen when dITP and primer 1 are used. The 5' end of primer 1 is located 40 bases 5' of the spurious product labeled 1, while the 5' end of primer 2 is 68 bases 5' of this product.

Benchmarks

template or a local effect of the presence of an inosine residue in the nascent strand. Primer-extension reactions on these templates using dATP, dCTP, TTP and dITP and end-labeled primer 1 produced a number of premature chain termination products that were one base shorter than a spurious product seen in the sequencing reactions (Figure 1). The formation of these termination products support the idea that polymerization through this region is compromised.

Many DNA polymerases, including *Taq* DNA polymerase, of which Thermo Sequenase is a variant, are known to add a non-templated base to the end of DNA strands. The identity of the added base depends on the polymerase and the last templated base in the chain (2). The sequencing artifacts that we have described here are consistent with

such a "terminal extendase" activity by Thermo Sequenase and with the observation that the dideoxynucleotide that is most frequently added in this way is ddA (6 ddAs compared to 2 ddT, 2 ddC and no added ddG) (2). "Extendase" activity is associated with the lack of a 3'→5' exonuclease proofreading activity (2). In theory, a polymerase that does not add a non-templated base should eliminate these artifacts. However, most exo-DNA polymerases thus far described have some "extendase" activity. Some thermolabile polymerases with a low 3'→5' exonuclease activity such as Sequenase (Version 2) do have a lower rate of non-templated base addition, perhaps because their polymerization rate significantly exceeds their "terminal extendase" activity. The use of thermostable polymerases with similar properties should reduce these artifacts in cycle-sequencing reactions, but such polymerases have yet to be identified.

Using a more distal primer is one alternative solution to the artifacts we have observed, although this is not always possible or desirable, and in some instances does not eliminate all of the artifactual bands (Kumari and Usdin, unpublished observations). The analog 7-deazadGTP is also known to reduce compression artifacts, but this analog poses a problem for most DNA polymerases and therefore should also be prone to the sorts of artifacts caused by dITP (4). We have found that addition of 50 mM KCl to the sequencing reactions reduces the appearance of the artifacts (Figure 1). Because salt increases duplex stability, this finding is consistent with the hypothesis that these artifacts result from the reduced stability of the hybrid formed between the template and the relatively short inosine-containing nascent strand. In reactions containing KCl, however, the amount of labeled products tends to drop off prematurely. This is due to the fact that K⁺ stabilizes an intrastrand tetraplex structure formed by this particular template. This tetraplex reduces the extent of readthrough by the polymerase (3). In most cases, the use of K⁺ should not be a problem since most templates do not form tetraplexes. This relatively simple modification to the reaction conditions makes it possible to

obtain accurate sequencing information close to the primer even in the presence of analogs such as dITP and can thus increase sequencing accuracy and efficiency. Awareness of, and if possible, the elimination of, the artifacts that we have described here could be particularly important when sequencing certain templates. For example, a small problematic region on a template generated by PCR directly from a diploid organism or from a pool of individuals might produce a small number of spurious products that could be misinterpreted as evidence for polymorphism or heterogeneity in the population.

REFERENCES

1. Fan, J., R.S. Ranu, C. Smith, C. Ruan and C.W. Fuller. 1996. DNA sequencing with [α -³²P]-labeled ddNTP terminators: a new approach to DNA sequencing with Thermo Sequenase DNA polymerase. *BioTechniques* 21:1132-1137.
2. Hu, G. 1993. DNA polymerase-catalyzed addition of nontemplated extra nucleotides to the 3' end of a DNA fragment. *DNA Cell Biol.* 12:763-770.
3. Usdin, K. and K.J. Woodford. 1995. CGG repeats associated with DNA instability and chromosome fragility form structures that block DNA synthesis in vitro. *Nucleic Acids Res.* 23:4202-4209.
4. Weitzmann, M.N., K.J. Woodford and K. Usdin. 1996. The development and use of a DNA polymerase arrest assay for the evaluation of parameters affecting intrastrand tetraplex formation. *J Biol Chem.* 271:20958-20964.

Address correspondence to Karen Usdin, Building 8, Room 202, National Institutes of Health, 8 Center Dr. MSC 0830, Bethesda, MD 20892-0830, USA. Internet: ku@helix.nih.gov

Received 18 March 1999; accepted 24 June 1999.

Daman Kumari and Karen Usdin

*National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health
Bethesda, MD, USA*