Frameshift Mutagenesis by Eucaryotic DNA Polymerases in Vitro*

(Received for publication, May 19, 1986)

Thomas A. Kunkel
From the Laboratory of Genetics, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

The frequency and specificity of frameshift errors produced during a single round of in vitro DNA synthesis by DNA polymerases-α, -β, and -γ (pol-α, -β, and -γ, respectively) have been determined. DNA polymerase-β is the least accurate enzyme, producing frameshift errors at an average frequency of one error for each 1,000-3,000 nucleotides polymerized, a frequency similar to its average base substitution accuracy. DNA polymerase-α is approximately 10-fold more accurate, producing frameshfts at an average frequency of one error for every 10,000-30,000 nucleotides polymerized, a frequency which is about 2- to 6-fold lower than the average pol-α base substitution accuracy. DNA polymerase-γ is highly accurate, producing on the average less than one frameshift error for every 200,000-400,000 nucleotides polymerized. This represents a more than 10-fold higher fidelity than for base substitutions.

Among the collection of sequenced frameshifts produced by DNA polymerases-α and β, both common features and distinct specificities are apparent. These specificities suggest a major role for eucaryotic DNA polymerases in modulating frameshift fidelity. Possible mechanisms for production of frameshfts are discussed in relation to the observed biases. One of these models has been experimentally supported using site-directed mutagenesis to change the primary DNA sequence of the template. Alteration of a pol-β frameshift hotspot sequence TTGT to CTCT reduced the frequency of pol-β-dependent minus-one-base errors at this site by more than 30-fold, suggesting that more than 97% of the errors at the TTGT run involve a slippage mechanism.

Frameshift mutations resulting from the removal or addition of small numbers of bases in DNA occur spontaneously in many organisms (1-9). Attempts to understand the mechanisms by which frameshifts are produced using in vivo genetic systems have demonstrated the importance of the primary DNA sequence in determining the frequency and specificity of frameshift errors, both at repeated base sequences (3, 5-9) and within quasipalindromic sequences (10). In addition, a clear role has been demonstrated for T4 DNA polymerase itself in controlling frameshift frequency and specificity in vivo (8).

Equivalent in vitro studies of frameshift fidelity have not been reported, primarily because a good system has not been available. Due to their repetitive nature, synthetic homopolymers or alternating copolymers are not useful frameshift templates for in vitro DNA synthesis reactions. More natural DNA templates have been used exclusively for base substitution mutagenesis in essential genes (for review, see Ref. 11) and have not permitted measurements of frameshift errors. In part to obviate this limitation and to provide in vitro information on possible mechanisms for the production of frameshift errors, a new system was developed (12). The assay measures a variety of errors, including frameshift errors, produced during in vitro DNA synthesis of the lacZα gene in M13mp2 DNA. This system was first used to analyze the fidelity of DNA polymerases-α, -β, and -γ (12-14). For each enzyme, frameshift errors were precisely defined by DNA sequence analysis of mutants, and the spectrum of single-base frameshfts was shown (12, 13). These analyses provide information on the actual target size for both total and specific subsets of frameshift errors within a 250-base target sequence. In conjunction with the heteroduplex expression experiments described here, this information can be used to calculate the frameshift error frequencies of DNA polymerases-α, -β, and -γ. These enzymes differ substantially in their relative accuracy, and the results further suggest that frameshift errors are produced by at least two mechanisms. One of these, the classical slippage mechanism first proposed by Streisinger et al. (1), is directly supported by experiments in which a frameshift hotspot is eliminated by site-directed mutagenesis.

EXPERIMENTAL PROCEDURES

All materials and procedures were as described in Refs. 12-15.

RESULTS

Description of the Target Sequence: Detectable Frameshift Sites—The mp series of M13 vectors was originally developed to allow easy detection of frameshifts resulting from insertion of DNA fragments into the coding sequence of the lacZα gene (16). The identification of mutants resulting from frameshift errors accumulated from several studies* (12, 13, 15) indicates that any nucleotide within at least the first 129 bases (43 codons) is a potential target. Most frameshifts result in a colorless phenotype, although certain errors in codons distal (3' direction, viral strand) to the start codon or in the regulatory sequence have a light blue phenotype. In addition to coding sequence, frameshifts* at (at least) 21 bases in the regulatory sequences for the lacZα gene produce a mutant phenotype. In most instances, minus-one-base frameshifts have been observed, while plus-one- or minus-two (or more)-base frameshifts at a limited subset of these same sites confirm that these too produce detectable phenotypic changes.

* T. A. Kunkel, unpublished observations.

* Despite the fact that the loss or gain of a base in a regulatory sequence is not in a strict sense a frameshift, here the term "frameshift" is meant to include mutants in both coding and noncoding sequence.

13581
DNA Polymerase Frameshift Mutagenesis

**TABLE I**

Summary of template bases as targets for minus-one-base frameshifts

This summary includes 128 bases of lacZa-coding sequence and 21 bases of the regulatory sequences at which addition and deletion events have been observed. As new regulatory sequences are found to be sensitive to length variation, the numbers shown below will increase somewhat.

<table>
<thead>
<tr>
<th>Number of consecutive bases</th>
<th>A</th>
<th>T</th>
<th>G</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of occurrences</td>
<td>No. of bases</td>
<td>No. of occurrences</td>
<td>No. of bases</td>
<td>No. of occurrences</td>
</tr>
<tr>
<td>1</td>
<td>16</td>
<td>16</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>12</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total bases</td>
<td>32</td>
<td>35</td>
<td>36</td>
<td>47</td>
</tr>
<tr>
<td>Total in runs</td>
<td>16</td>
<td>23</td>
<td>15</td>
<td>26</td>
</tr>
</tbody>
</table>

The assumption is therefore made that a site\(^3\) producing a minus-one-base frameshift is a potential target for other frameshift events as well (although for a limited number of sites in the regulatory sequences this may not be true). This 150-base frameshift target is described in Table I, with respect to the number of A, T, G, and C residues (template strand) and their location, either within runs of a common base, or at non-runs (where the one base has two different nearest neighbors). Thus, for example, there are 32 template A sites at which the loss or addition of bases results in a detectable phenotype, 16 of which are within runs of two or more consecutive A residues. Of these 16 bases, 12 are found in six 2-base runs and four are in one 4-base run.

The information in Table I is important for the analysis presented here, since all DNA polymerase error frequencies are expressed on a “per detectable nucleotide polymerized” basis. This correction factor is necessary in order to eliminate target-size biases when interpreting the results. The values in Table I will increase slightly if more regulatory sequences are found to be sensitive to frameshift errors. However, while spacing is obviously important in specific regulatory sequences, overall these sequences are not as sensitive to length variation as are coding sequences, and so the identification of the expected small number of additional sites should not alter the conclusions substantially.

**Expression of Frameshift Heteroduplex Molecules upon Transfection**—In order to determine if the relative occurrence of the various classes of frameshift events reflect the specificity of the DNA polymerase or perhaps reflect limitations in expression of various frameshift errors upon transfection, a variety of frameshift-containing heteroduplex molecules were constructed as previously described for single-base mispair heteroduplexes (14). The expression of the phenotype of the minus strand (equivalent to a new strand made by a DNA polymerase in vitro) was determined upon transfection of competent _Escherichia coli_ cells. The data for a typical experiment are shown in Table II for heteroduplexes containing -A or +A frameshift errors. To be consistent with the published spectra, frameshifts are described with reference to the template strand. The phenotype of the minus strand is expressed with about 30% efficiency for both plus-one and minus-one frameshifts. As for base mispair heteroduplexes, plagues are of only a single phenotype, both as initially observed on the plate and upon dilution and replating. Experiments similar to those in Table II were performed with both plus-one- and minus-one-base frameshifts at template T, G, and C sites as well as for plus- and minus-two-base frameshifts (results not shown). No biases were observed for minus strand expression of the various frameshifts. The average expression for all experiments was 26%. This value suggests that, on the average, only one of four frameshift errors made by a DNA polymerase is actually observed upon transfection of the products of an _in vitro_ reaction, and this value (0.26) has been used below to calculate error frequencies.

**Average Frameshift Error Frequencies**—In combination with the information in Tables I and II, the published forward mutational spectra of errors produced by DNA polymerases-\(\beta\) (12), -\(\alpha\), and -\(\gamma\) (19) provide all the information needed to calculate the average frameshift error frequency of each of these enzymes (Table III). Averaged over 150 detectable sites, pol-\(\beta\) exhibits the lowest frameshift fidelity. Each of the two error prone pol-\(\beta\) preparations produces frameshifts at a frequency of 1/1,000 to 1/3,000, values similar to their base

---

3 The connotation of the word “site” is in fact different for a base loss, where the site is actually the base itself, versus a base addition, where the sites are the phosphodiester bonds on either side of the base itself.

---

4 In discussing average frameshift error frequencies, the values for minus-one-base events are used, since these represent by far the most frequent events. Inclusion of the plus-one events, which are shown separately in Table III for later discussion of specificity, does not alter the general conclusions about average error frequencies.

5 The abbreviations used are: pol-\(\beta\), polymerase-\(\beta\), pol-\(\alpha\), polymerase-\(\alpha\), pol-\(\gamma\), polymerase-\(\gamma\).
substitution accuracy (15). Each of the three pol-α preparations examined are more accurate than pol-β. Two of these (KB cell and calf) have a similar frameshift fidelity of approximately one error for every 9,000 bases polymerized. These two α-polymerases are on average about 2- to 3-fold more accurate for frameshifts than for base substitutions (error frequencies 1/3,100 and 1/4,300, respectively; see Ref. 15). The most accurate α-polymerase, from chick embryo, is 3- to 4-fold more accurate for frameshifts than the other two, and has 7-fold higher frameshift fidelity (1/33,000, Table III) than base substitution fidelity (1/4,700; see Ref. 15).

The most accurate of the DNA polymerases examined are the γ-polymerases. These enzymes produce frameshifts at a frequency of <1/200,000, a value less than one tenth their average base substitution frequency (15) and 100-fold lower than the pol-β frameshift frequency.

Specificity of Frameshift Errors by Pol-α and Pol-β—Both pol-α and pol-β produce minus-one-base errors at least 10-fold more frequently than plus-one-base errors (Table III) or minus-two-base errors (12, 13). The heteroduplex expression experiments (Table II) suggest that this bias is not due to differential biological expression of different types of frameshifts but rather is imposed in some way during the in vitro DNA synthesis.

Within the large collection of minus-one-base errors produced by pol-β and the smaller but substantial number produced by pol-α there are interesting biases in specific subsets of errors. The information in Table I has been used to correct the relative frequencies of these subclasses for target size in order to determine the true extent of these biases. The results are as follows (Table IV). Both pol-β preparations produce minus-one-base errors preferentially in runs of a common base, with a 10-fold preference for errors in template pyrimidine runs over purine runs. Single base losses are also produced in non-run sequences at frequencies well above background. These non-run events occur at frequencies only about 3-fold lower than in runs of purine bases.

DNA polymerase-α also produces minus-one frameshifts in runs of either pyrimidines or purines, but at lower frequency than pol-β, especially in pyrimidine runs. Purine run frameshifts occur at about the same frequency for both classes of polymerases, as do single base losses at non-run sites. Among all non-run single base losses, 31 of 41 (76%) were the loss of a template G residue. This bias is not due to target size, since there are similar numbers of each of the four bases in the target that are present as non-run positions (Table I).

Only three minus-one-base mutants were observed in the pol-γ spectrum. Strong conclusions cannot be made regarding specificity other than to note that all three frameshifts were
in runs and that, for all subclasses of errors in Table IV, pol-γ is more accurate than either pol-β or pol-α.

The influence of the DNA polymerase on specific subsets of errors is most apparent at two particular sites in the template sequence. Positions 70-73, a run of 4 consecutive T residues, is a hotspot for -T frameshifts for pol-β (12). The error rate (per template T residue) is 1/50 for rat pol-β. In contrast, chick pol-α and pol-γ produce -T events 64-fold and >80-fold less frequently at this run (error rates 1/3,200 and <1/34,000, respectively). At a second site, however, the relative accuracies of pol-β and pol-α are reversed. Both KB pol-α and calf pol-α produce -C errors at the run of 5 consecutive C residues at positions 132-136 at a frequency of about 1/1,000, while both rat and chick pol-β produce -C events at this same site with a 5-fold lower error frequency (1/5,000).

**DNA Sequence Alteration to Increase Frameshift Fidelity**

The most clearly formulated model for frameshift errors is the Streisinger strand-slippage model (1), which accounts for frameshifts within runs of a common base. The high frequency with which pol-β produces -T frameshifts in a run of 4 T residues provides an opportunity to examine if indeed minus-one frameshifts at this site result from strand slippage within the run. The model predicts that if 1 or more T residues are changed to a different base, the frequency of minus-one frameshifts should decrease. This is because the stability of the predicted misaligned intermediates will be lower due to fewer correct base pairs to stabilize the intermediate (in Fig. 1, compare (a), 3 A-T pairs to (b), 1 A-T pair or (c), no A-T pairs, 5' to the extrahelical template base).

In order to examine this model, the DNA sequence of the TTTT run (nucleotides 70-73) was altered to TTCT and CTCT by site-directed mutagenesis. The change to TTCT does not result in an amino acid sequence change in the α-peptide, and thus plaques are still of wild-type blue color intensity. The change to CTCT results in production of an α-peptide containing a single amino acid difference, a valine to alanine change at codon 10. This results in a very slight reduction in blue color intensity which does not interfere with the detection of colorless frameshift mutants. Since no phenotypic selection for these changes was possible, the sequence alterations were produced using the recently developed method to change nucleotides 70-73, where position one is the first transcribed base of the lacZa gene in M13mp2. Misalignment a is only one of six possible intermediates that can be drawn (see Fig. 2).

**Effect of altering a run of T residues on forward mutation frequency by pol-β**

The analyses were performed as described previously (12) but using derivatives of wild-type M13mp2 having the indicated sequence alterations produced by site-directed mutagenesis as described (17).

<table>
<thead>
<tr>
<th>DNA sequence</th>
<th>Plaques Scored</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>Rat pol-β</td>
<td></td>
</tr>
<tr>
<td>TTTT</td>
<td>3372</td>
</tr>
<tr>
<td>TTCT</td>
<td>3613</td>
</tr>
<tr>
<td>CTCT</td>
<td>3278</td>
</tr>
<tr>
<td>Chick pol-β</td>
<td></td>
</tr>
<tr>
<td>TTTT</td>
<td>3397</td>
</tr>
<tr>
<td>TTCT</td>
<td>4162</td>
</tr>
<tr>
<td>HeLa pol-β</td>
<td></td>
</tr>
<tr>
<td>TTTT</td>
<td>4435</td>
</tr>
<tr>
<td>TTCT</td>
<td>1179</td>
</tr>
<tr>
<td>CTCT</td>
<td>1277</td>
</tr>
</tbody>
</table>

**Sequence-dependent fidelity of DNA synthesis by pol-β**

The analyses were performed as described previously (12) but using derivatives of wild-type M13mp2 having the indicated sequence alterations produced by site-directed mutagenesis as described (17).

<table>
<thead>
<tr>
<th>DNA sequence</th>
<th>Total forward mutation frequency</th>
<th>Total mutants sequenced</th>
<th>Minus-one-base mutants in target</th>
<th>Minus-one-base mutation frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x 10^-4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rat pol-β</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTTT</td>
<td>640 159</td>
<td>51</td>
<td>210</td>
<td></td>
</tr>
<tr>
<td>TTCT</td>
<td>430 137</td>
<td>18</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>CTCT</td>
<td>300 43</td>
<td>0</td>
<td>&lt;7</td>
<td></td>
</tr>
<tr>
<td>Chick pol-β</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTTT</td>
<td>440 37</td>
<td>9</td>
<td>110</td>
<td></td>
</tr>
<tr>
<td>TTCT</td>
<td>350 144</td>
<td>13</td>
<td>32</td>
<td></td>
</tr>
</tbody>
</table>

**DISCUSSION**

This report establishes the in vitro frameshift fidelity of purified DNA polymerases-α, -β, and -γ. The least accurate of the DNA polymerases is pol-β, which produces frameshifts about as frequently as base substitutions and is more than...
100-fold less accurate than pol-γ. The similar inaccuracy and frameshift specificity of two pol-β preparations implies similar discrimination for these two proteins. The frameshift accuracy of these enzymes may improve when associated with specific DNA-binding proteins (18, 19) or with accessory proteins known to bind to and stimulate β-polymerase (20, 21). To the extent that pol-β is involved only in short gap repair, the in vivo consequences of frameshift inaccuracy might not be too severe.

The next most accurate enzymes are the α-polymerases. Two of these, purified from KB cells (22) or calf thymus (23, 24), have similar average frameshift error frequencies. This similarity suggests common mechanisms for discrimination against frameshift errors for two different proteins which share the common properties of α-polymerases. The KBα used in this study is presumably more complex than calf pol-α since it contains primase activity in addition to DNA polymerase activity (22). Despite this, KBα is no more accurate than calf pol-α, which has no associated primase activity. Thus, at least in the form used here, associated primase activity does not alter frameshift fidelity.

The most accurate α-polymerase, that from chick embryo (25), is 3- to 4-fold more accurate than the other two, implying additional discrimination against frameshifts. Since the specific activity of this enzyme is more than 10-fold lower than that of KBα or calf pol-α (see Ref. 13), this enzyme preparation may contain additional subunits or other proteins that provide such discrimination.

The most accurate of the three animal cell DNA polymerases is pol-γ. The high frameshift fidelity is remarkable since it is achieved by an enzyme which consists of a tetramer of four identical 47,000-dalton subunits (26). This implies that high frameshift accuracy does not necessarily require a complicated enzyme complex. This conclusion is further supported by observations with two single subunit procaryotic DNA polymerases (E. coli DNA polymerase I and T4 DNA polymerase); both exhibit high frameshift fidelity in vitro.

The M13mp2 assay permits detection of a variety of frameshift mutations in different target sequences. The observed mutants are the end points of the molecular interactions leading to their production rather than the actual misaligned intermediates themselves. Nevertheless, the specificity of the observed frameshifts provides information upon which to formulate models to describe the DNA:DNA and DNA:protein interactions that might be important. It is for this reason that the data in Tables III and IV have been arranged to show biases for minus-one over plus-one events, the ratio of run versus non-run frameshifts, and the template base preferences (pyrimidine runs and non-run G residues).

The production of a frameshift error requires a minimum of two steps: a misalignment must occur followed by subsequent incorporation events to “fix” the extrahelical base(s). The initial formation of the misalignment will at least depend upon the DNA sequence involved and may also depend upon the DNA polymerase. Thus pol-β may actively promote the formation of misalignments by, for example, tightly binding to the 3'OH terminus of the primer strand and pulling the helix apart as it dissociates during distributive DNA synthesis. Once formed, the misalignment must persist long enough for the extrahelical base, to provide increased stability to the misalignment as well as an acceptable (i.e. base-paired) primer terminus for continued incorporation.

**Fig. 2.** The possible intermediates leading to plus-one- or minus-one-base errors in the run of T residues are shown. The numbers in parentheses are the number of base pairs that must be disrupted to produce the intermediate shown. There are equal numbers of possible intermediates for either outcome, and this number varies with the run length as follows: two-base run, one intermediate (for each); three-base run, three intermediates; four-base run, six intermediates (shown); five-base run, 10 intermediates; six-base run, 15 intermediates, etc.
The evidence that this mechanism is operative during synthesis by DNA polymerases-α and β is 3-fold. First, each of these enzymes produces frameshifts in runs at a higher frequency than in non-runs (Table III). Second, the frequency of frameshifts within runs is usually (but not always) proportional to the length of the run (calculations not shown). Third, as predicted from a consideration of the possible intermediates (Fig. 1), the frequency of pol-β-dependent frameshift errors at a run of 4 T residues is decreased as the run is shortened (to TTCT) or eliminated (to CTCT) (Table VI).

DNA polymerases-α and β exhibit a very different ratio of run versus non-run frameshifts (2 or 3:1 for pol-α versus ~30:1 for pol-β). While pol-β produces many frameshifts in runs of two or three bases, pol-α produces most of its run frameshifts in four- or five-base runs. This may reflect different DNA polymerase requirements for utilization of the misaligned intermediate. In a short run, the number of base pairs between the 3′OH primer terminus and the extrahelical base is small (e.g. Fig. 2, e and h). Such an intermediate might be acceptable to pol-β, which has relaxed synthesis requirements (12, 18). In contrast, pol-α may require more extensive correct contacts, thus forcing misalignments into correct alignment in short runs but accepting such intermediates if the extrahelical base is several bases removed from the 3′OH terminus (e.g. Fig. 2, j and p) (13, 18).

Among the minus-one frameshifts within runs there is a distinct bias for losses in template pyrimidine runs rather than purine runs. This is obvious for both pol-β preparations (Table IV), where the bias is approximately 10:1. While the effect is less obvious for pol-α when all runs are considered (Table IV), this smaller bias may be deceiving. The analysis in Table IV includes many pol-α frameshifts in two-base runs, which could occur by a non-run mechanism. Considering only longer (three, four, or five bases) runs then, the pol-α spectra also exhibit a bias for loss of bases in pyrimidine rather than purine runs. This bias is not as pronounced as for pol-β, being 5:1 for KBα and 4:1 for calf pol-α. The simplest explanation is that bases within pyrimidine runs unstack to become extrahelical more easily than do bases within purine runs, where stacking is known to be stronger (27, 28). Alternatively, there may be an inherent bias for utilization of misalignments containing an extrahelical pyrimidine.

Among the one-base frameshifts at all sites there is a >10-fold bias for loss rather than addition of a base (Table III). In considering only those events within runs (and for pol-α, only those in three-, four-, or five-base runs), the minus-one versus plus-one mutant ratios are: rat β, 157:2; chick β, 44:3; KBα, 25:3; calf pol-α, 17:1; chick pol-α, 6:2; and chick pol-γ, 3:0. There are at least two possible explanations for these biases. As shown in Fig. 2, while an equal number of potential misalignments can be drawn for production of plus-one versus minus-one events, each plus-one misalignment requires one additional base pair to be disrupted in order to produce an extrahelical base in the primer. The extra energy required to form the intermediate may thus explain less frequent formation. Alternatively, or in addition, the bias may result from more stringent polymerase constraints imposed upon extrahelical bases in the primer rather than the template strand.

While several features of the frameshifts in runs support a strand-slipage mechanism, site-specific polymerase-mediated effects on frequencies are apparent that complicate simple interpretations. For example, within the rat pol-β spectrum, several exceptions to the frequency versus length concept have been observed (12). Also, while overall pol-β is less accurate than pol-α for frameshifts, pol-β is actually 5-fold more accurate than pol-α at a run of 5 C residues. While the exact nature of the interactions leading to such enzyme-specific events remains to be determined, the eucaryotic DNA polymerases are known to exhibit distinctive and specific interactions with defined DNA sequences of various compositions (for review, see Ref. 18). These data imply a major role for the DNA polymerases in influencing frameshift fidelity and are consistent with T4 polymerase-mediated effects on frameshift fidelity in vivo (19).

Several mechanisms may operate to produce frameshifts at noniterated bases. In addition, some of the frameshifts within runs, especially two-base runs, may also be produced by a non-run mechanism. Assuming that extrahelical bases cannot be stabilized by additional immediately adjacent base pairs in the same way as in runs, the stability necessary for subsequent incorporation may result from other interactions. One possibility is that a non-run base can assume a stable conformation in which it cannot template an incorporation event itself but at the same time does not interfere with its neighbor’s ability to do so. There is some indication that this might be possible from NMR studies of nucleic acids, demonstrating that an extra base in an otherwise fully self-complementary oligonucleotide does not interfere with the hydrogen bonding potential of the adjacent base pairs (29). Perhaps such a base conformation could be stabilized by stacking interactions with dNTP’s or with amino acids in the polymerase in a manner similar to that suggested by Streisinger for base stacking with intercalators (the sandwich model, see Ref. 9).

A second possible mechanism to produce frameshifts at noniterated sequences is transient dislocation of the primer to a new position in order to template the frameshift event with a limited amount of incorporation, followed by realignment to the original position and continued synthesis. Such an explanation has been previously proposed for base substitutions by pol-β (14) and could account for the non-run frameshifts observed in this and several in vivo studies (10, 30, 31).

The formation provided by the forward mutational spectra of frameshift errors by polymerases-α, β, and γ provide a starting point for the search for additional fidelity components and to continue to formulate and test concepts concerning mechanisms for the production of frameshifts.

Acknowledgments—I would like to thank John Drake and Roel Schaper for helpful comments and critical evaluation of the manuscript.

REFERENCES


DNA Polymerase Frameshift Mutagenesis
DNA Polymerase Frameshift Mutagenesis