# Polymerase Fidelity: What is it, and what does it mean for your PCR?

**The discovery and development of high-fidelity polymerases has for many years been a key focus at New England Biolabs (NEB). High-fidelity amplification is essential for experiments whose outcome depends upon the correct DNA sequence (e.g., cloning, SNP analysis, NGS applications). Whereas traditional fidelity assays are sufficient for *Taq* and other moderately faithful enzymes, Q5, an ultra high-fidelity enzyme, pushes the limits of current methods used to assess this critical feature of DNA polymerases.**

John A. Pezza, Ph.D., Rebecca Kucera, M.S., Luo Sun, Ph.D., New England Biolabs, Inc.

## Introduction: What is fidelity?

The fidelity of a DNA polymerase is the result of accurate replication of a desired template. Specifically, this involves multiple steps, including the ability to read a template strand, select the appropriate nucleoside triphosphate and insert the correct nucleotide at the 3′ primer terminus, such that Watson-Crick base pairing is maintained. In addition to effective discrimination of correct versus incorrect nucleotide incorporation, some DNA polymerases possess a 3′→5′ exonuclease activity. This activity, known as "proofreading", is used to excise incorrectly incorporated mononucleotides that are then replaced with the correct nucleotide. High-fidelity PCR utilizes DNA polymerases that couple low misincorporation rates with proofreading activity to give faithful replication of the target DNA of interest.

## When is fidelity important?

Fidelity is important for applications in which the DNA sequence must be correct after amplification. Common examples include cloning/subcloning DNA for protein expression, SNP analysis and next generation sequencing applications. Fidelity is less important for many diagnostic applications where the read-out is simply the presence or absence of a product.

## How does a high-fidelity polymerase ensure that the correct base is inserted?

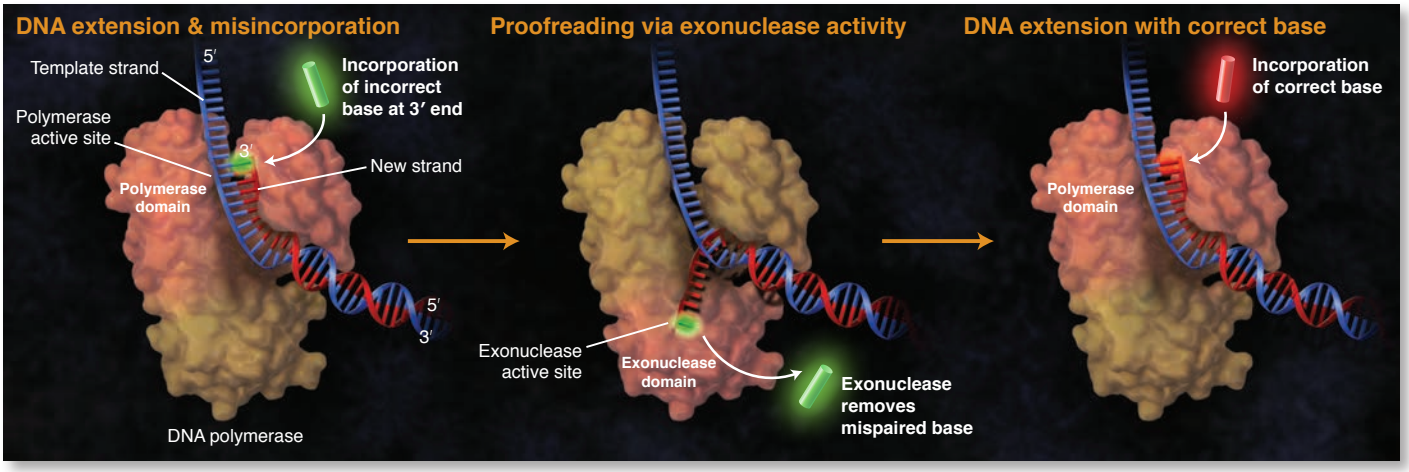
High-fidelity DNA polymerases have several safeguards to protect against both making and propagating mistakes while copying DNA. Such enzymes have a significant binding preference for the correct versus the incorrect nucleoside triphosphate during polymerization. If an incorrect nucleotide does bind in the polymerase active site, incorporation is slowed due to the sub-optimal architecture of the active site complex. This lag time increases the opportunity for the incorrect nucleotide to dissociate before polymerase progression, thereby allowing the process to start again, with a correct nucleoside triphosphate (1,2). If an incorrect nucleotide is inserted, proofreading DNA polymerases have an extra line of defense (Figure 1). The perturbation caused by the mispaired bases is detected, and the polymerase moves the 3′ end of the growing DNA chain into a proofreading 3′→5′ exonuclease domain. There, the incorrect nucleotide is removed by the 3′→5′ exonuclease activity, whereupon the chain is moved back into the polymerase domain, where polymerization can continue.

## How is fidelity measured?

A variety of polymerase fidelity assays have been described in the literature over the years, perhaps the most famous being that of Thomas Kunkel (3). The Kunkel method uses portions of the *lacZ*α gene in M13 bacteriophage to correlate host bacterial colony color changes with errors in DNA synthesis. Wayne Barnes built upon this assay and utilized PCR to copy the entire *lacZ* gene and portions of two drug resistance genes with subsequent ligation, cloning, transformation and blue/white colony color determination (4). In both assays, errors incorporated in the *lacZ* gene cause a disruption in β-galactosidase activity leading to a white colony phenotype. With these *lacZ*-based experimental approaches, the percentage of white colonies must be converted to the number of errors per base incorporated. As a more direct read-out of fidelity, Sanger sequencing of individual cloned PCR products can also score DNA polymerase fidelity and offers the advantage that all mutations will be detected. Using this method, the entire mutational spectrum of a polymerase can be determined and there is no need to correct for non-phenotypic changes.

A modification of the *lacZ* Barnes assay is commonly used at NEB for a relatively quick and inexpensive determination of DNA fidelity, as the 1,000 amino acid open reading frame affords a reasonable sequence window for the scoring of DNA polymerase errors (Figure 2). In this study, results from the *lacZ* assay were compared to Sanger sequencing to assess the fidelity of Q5, a new NEB DNA polymerase (for more info on Q5, please see page 6).

Figure 1. DNA Replication with a Proofreading Polymerase




*Extension proceeds along the template strand at the 3′ end of the newly synthesized strand. When the polymerase recognizes an error, the mismatched base is transferred to the exonuclease active site and the base is excised. The extended strand returns to the polymerase domain, re-anneals to the template strand, and replication continues.*

## Results

Here, Q5 was examined to determine its fidelity compared to *Taq* DNA polymerase using the two methods described below (Figure 2). A 3,874 bp target was PCR amplified with either *Taq* (Thermopol Buffer), Q5 (Q5 Reaction Buffer with or without GC enhancer) or Phusion® (Phusion HF Buffer) DNA Polymerase. Observed mutation rates were determined using both the blue/white selection method after 16 PCR cycles (4) and by Sanger sequencing after 25 PCR cycles (Table 1). The error rate per base incorporated was determined after calculating the effective number of amplification cycles for each experiment as described previously (4, 5). Comparing the data sets from *Taq* indicates that the two methods generate similar results with error rates of ~1 in 3,500 bases. Q5, on the other hand, yielded a significantly lower number of errors than *Taq* in both assay systems, consistent with an error rate of ~$10^{-6}$. The side-by-side evaluation of *Taq* and Q5 using the blue/white method suggests that Q5 is approximately 200x more faithful at replicating DNA than *Taq*. Similar results were observed for Q5 when the GC enhancer was added to the reactions (data not shown). For Phusion, the error rate was determined to be 80±39 times better than *Taq* using the blue/white method and 84 times *Taq* using the sequencing method.

For Sanger sequencing, only two mutations were detected in the Q5 data set (despite sequencing over 440.000 nucleotides). Although this speaks directly to the strikingly high fidelity of the Q5 enzyme, it is more difficult to make statistically significant conclusions about either the absolute Q5 error rate or the comparative fidelity rate versus *Taq* using this particular data set.

## Conclusion

The ultra low error rate of Q5 is extremely beneficial for many applications. However, the low number of identified errors makes absolute error rate quantitation difficult for this enzyme, even with extensive experimentation and analysis. From blue/white screening, we have observed that Q5 is approximately 200x more faithful at replicating DNA than, but results from Sanger sequencing hint that this value may actually be an underestimate.

Because the values generated from blue/white methods vary significantly between individual replicates and rely on a series of calculated extrapolations, we have chosen to conservatively represent the fidelity of Q5 High-Fidelity DNA Polymerase as >100x *Taq*, and Phusion High-Fidelity DNA Polymerase as >50x *Taq*. With ever-decreasing costs and extremely large datasets, next generation sequencing techniques may soon be able to provide direct, cost effective methods for more accurately quantitating error rates for an ultra high-fidelity polymerase like Q5.

Translating blue/white error rates into practical use, this data suggests that after using 25 PCR cycles to amplify a 400 bp fragment with *Taq*, several isolates should be screened since about half of the clones are predicted to have an error. For larger fragments of ~1000 bp, each clone amplified with *Taq* is likely to have an undesired mutation while the ultra low error rates of Q5 High-Fidelity DNA Polymerase predict that 199/200 clones amplified with this new enzyme will be correct.

Figure 2. Fidelity testing workflow

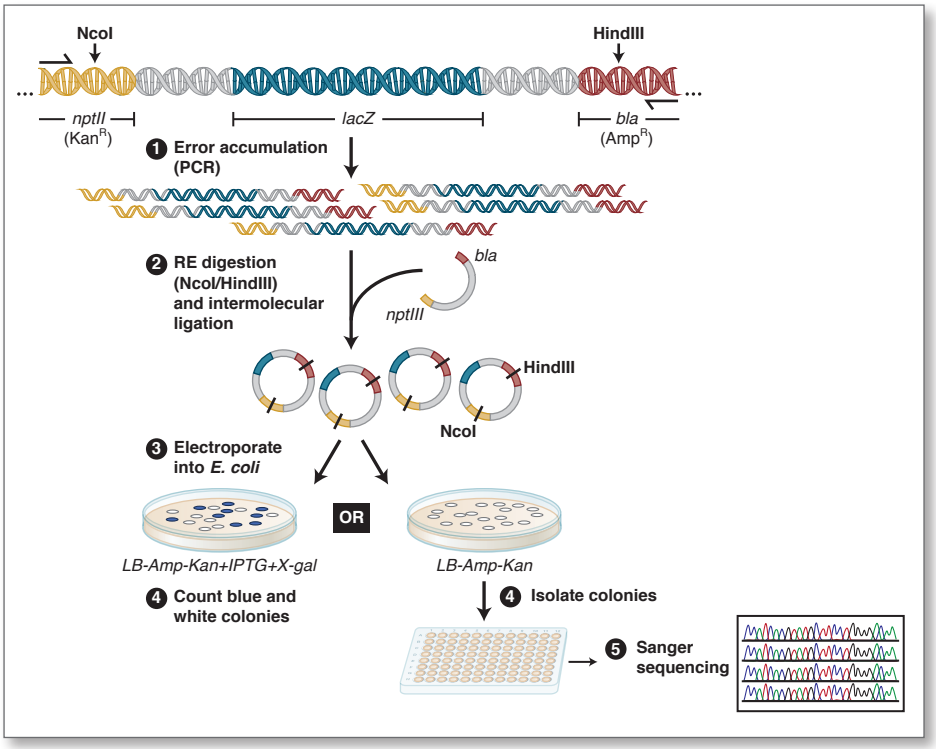



Table 1. Calculations for fidelity testing of *Taq* and Q5 DNA Polymerases

| METHOD | | NOTES | *Taq* | Q5 |
|---|---|---|---|---|
| Blue/white screening | Total colonies scored | | 30.192 | 22,296 |
| | White colonies scored | | 17,589 | 119 |
| | Error rate per base | Corrects for non-phenotypic changes and error propagation during PCR (4) | $2.7x10^{-4} \pm 0.8x10^{-4}$ (1 per 3,700 bases) | $1.4x10^{-6} \pm 0.6x10^{-6}$ (1 per 710.000 bases) |
| | Fold over *Taq* | | | 193 ± 101 |
| Sanger sequencing | Clones sequenced | | 340 (~215,000 nts) | 710 (~440.000 nts) |
| | Mutations detected | | 279 | 2 |
| | Error rate per base | Corrects for error propagation during PCR (5) | $\sim 3.0x10^{-4}$ (1 per 3,300 bases) | [$\sim 1.0x10^{-6}$] [~1 per 1,000.000 bases] |
| | Fold over *Taq* | | | [~300] |

*±= 95% confidence*
*Numbers in brackets have limited statistical significance as only 2 mutations were detected after sequencing 441,670 nucleotides.*

References

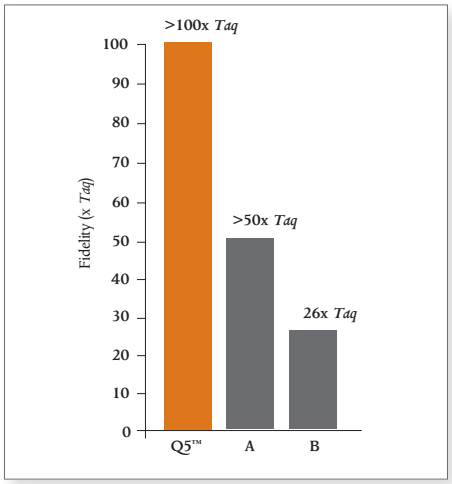

1. Johnson, K.A. (2010) *Biochemica et Biophysica Acta,* 1804, 1041–1048.
2. Joyce, C.M. and Benkovic, S.J. (2004) *Biochemistry,* 43, 14317–14324.
3. Kunkel, T.A. and Tindall, K.R. (1988) *Biochemistry,* 27, 6008–6013.
4. Barnes, W.M. (1992) *Gene,* 112, 29–35.
5. Eckert, K.A. and Kunkel, T.A. (1991) *PCR Methods and Applications,* 1, 17–24.

# Q5™ High-Fidelity DNA Polymerase
# NEB's best PCR Polymerase ever!

Q5 provides the highest fidelity of all PCR polymerases available!

## Five Quality Features of Q5 DNA Polymerase

- **Fidelity** – the highest-fidelity amplification available (>100x higher than *Taq*)
- **Robustness** – high specificity and yield with minimal optimization
- **Coverage** – superior performance for a broad range of amplicons (from high AT to high GC)
- **Speed** – short extension times
- **Amplicon length** – robust amplifications:
  - up to 20 kb for simple templates
  - up to 10 kb for complex templates




[A] Phusion®
[B] AccuPrime™ *Pfx*

Ordering information:

| PRODUCT | NEB # | SIZE | PRICE |
|---|---|---|---|
| Q5 High-Fidelity DNA Polymerase | M0491S/L | 100/500 u | 90 € / 399 € |
| Q5 Hot Start High-Fidelity DNA Polymerase | M0493S/L | 100/500 u | 105 € / 480 € |
| Q5 High-Fidelity 2X Master Mix | M0492S/L | 100/500 rxn | 150 € / 600 € |
| Q5 Hot Start High-Fidelity 2X Master Mix | M0494S/L | 100/500 rxn | 175 € / 700 € |

**Request a Free sample at ww.neb-online.de**

NEW ENGLAND BioLabs® GmbH

New England Biolabs GmbH
Brüningstr. 50; Geb B852
D-65926 Frankfurt am Main

Telefon: +49 (0) 69/305-23 14-0
Telefax: +49 (0) 69/305-23 14-9
E-Mail: info@de.neb.com

**Hotline:**

**Deutschland:** 0800/246 5227

**Österreich:** 00800/246 522 77

**www.neb-online.de**

DNA CLONING
DNA AMPLIFICATION & PCR
EPIGENETICS
RNA ANALYSIS
SAMPLE PREP FOR NEXT GEN SEQUENCING
PROTEIN EXPRESSION & ANALYSIS
CELLULAR ANALYSIS