DNA Polymerase Fidelity: From Genetics Toward a Biochemical Understanding

Myron F. Goodman and D. Kuchnir Fygenson

Department of Biological Sciences, Hedco Molecular Biology Laboratories, University of Southern California, Los Angeles, California 90089-1340

ABSTRACT

This review summarizes mutagenesis studies, emphasizing the use of bacteriophage T4 mutator and antimutator strains. Early genetic studies on T4 identified mutator and antimutator variants of DNA polymerase that, in turn, stimulated the development of model systems for the study of DNA polymerase fidelity *in vitro*. Later enzymatic studies using purified T4 mutator and antimutator polymerases were essential in elucidating mechanisms of base selection and exonuclease proofreading. In both cases, the base analogue 2-aminopurine (2AP) proved tremendously useful—first as a mutagen *in vivo* and then as a probe of DNA polymerase fidelity *in vitro*. Investigations into mechanisms of DNA polymerase fidelity inspired theoretical models that, in turn, called for kinetic and thermodynamic analyses. Thus, the field of DNA synthesis fidelity has grown from many directions: genetics, enzymology, kinetics, physical biochemistry, and thermodynamics, and today the interplay continues. The relative contributions of hydrogen bonding and base stacking to the accuracy of DNA synthesis are beginning to be deciphered. For the future, the main challenges lie in understanding the origins of mutational hot and cold spots.

THE development of molecular biology has been profoundly influenced by genetic and biochemical studies using the bacteriophage T4. In particular, T4 has served as an invaluable tool for testing new ideas and refining concepts of mutagenesis and DNA polymerase fidelity. Through his studies on T4 mutagenesis, Jan Drake, to whom this issue of Genetics is dedicated, played a central role in initiating the remarkably fertile area of research into the biochemistry of fidelity.

In 1968, Drake reported the surprising discovery of antimutagenic T4 polymerase mutants (Drake and Allen 1968). Until then, mutations in the structural gene coding for the T4 polymerase, gene *43* (de Waard *et al.* 1965), had only been reported to generate mutator phenotypes (Speyer 1965; Speyer *et al.* 1966; Freese and Freese 1967). The notion that a "defective" (*i.e.*, mutant) polymerase might replicate DNA with higher fidelity than the wild-type was revolutionary.

Reversion frequencies in the nonessential *rII* region of T4, used by Seymour Benzer in his classic studies on genetic fine structure (Benzer 1961), were the phenotype of choice for determining the effects of mutations in the T4 pol gene. The various T4 mutant polymerases exhibited very different mutation rates. While the effect depended somewhat on which *rII* reversion was investigated, for several of the *rII* alleles reversion frequencies in the *tsL56* mutator and *tsCB120* antimutator backgrounds differed by as much as 10^3–10^4-fold (Drake and Allen 1968; Drake *et al.* 1969; Speyer 1965). The mutations in *tsL56* are A89T+D363N, and the mutation in *tsCB120*, also known as *tsL141*, is A737V (Reha-Krantz 1988, 1989).

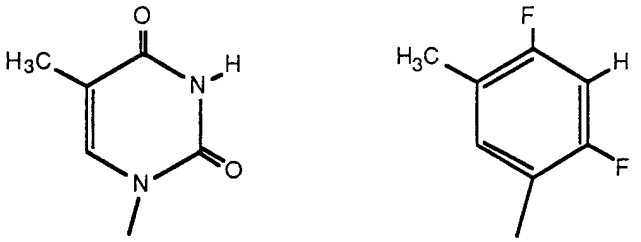
Such large variation in error frequencies suggested that the polymerase may play an active role in base selection during DNA synthesis. To quote Speyer's paper "Mutagenic DNA Polymerase" (Speyer 1965)

> . . . the replicating enzyme is involved more directly in the selection of the base . . . [such that] the information of the parental DNA strand is transmitted sequentially by the enzyme to an allosteric site where selection of the nucleotide . . . occurs. Such an enzymic mechanism may permit selection by criteria other than the relatively weak hydrogen bonds postulated in the template hypothesis and account for the high accuracy of DNA replication.

However, as the mechanisms of exonuclease editing and mismatch repair emerged, the contribution of the polymerase active site to fidelity was deemphasized. But, recently, Speyer's conclusion is regaining prominence. For example, Eric Kool has constructed a base analogue of T that is geometrically similar to T but cannot form H-bonds with A (Figure 1), and has shown that it is nevertheless incorporated opposite A almost as well as T by DNA polymerase I Klenow exo$^-$ (Moran *et al.* 1997).

Pioneers in genetic fidelity, such as Speyer, Drake, Freese and, of course, Watson and Crick, set the stage for three decades of ongoing research into the question of how DNA polymerases synthesize DNA with such exquisitely high accuracy. What follows is a review of key results from those decades and a personal assessment of how the fidelity field evolved from the early genetic experiments.

Corresponding author: Myron F. Goodman, University of Southern California, Department of Biological Sciences, SHS Rm. 172, University Park, Los Angeles, CA 90089-1340.
E-mail: mgoodman@mizar.usc.edu



T F

Figure 1.—Difluorotoluene, a non-hydrogen bonding base analogue of T. Chemical structures of thymine and difluorotoluene, an isosteric analog for thymine, used to demonstrate the relatively small influence of hydrogen bonding in DNA polymerase base selection (Goodman 1997; Moran *et al.* 1997).

Studies on the biochemical basis of mutation

The role of 3′-exonuclease proofreading in reducing polymerase errors: Two important papers published in 1972 suggested the existence of a polymerase-associated 3′→5′ exonuclease, which could increase fidelity by excising misincorporated nucleotides at their point of origin. Brutlag and Kornberg showed that *Escherichia coli* Pol I excised mispaired nucleotides in preference to correctly paired nucleotides from primer-3′-termini (Brutlag and Kornberg 1972). Bessman and co-workers, building on the work of Speyer and Drake, purified mutant and wild-type T4 polymerases and showed that the nuclease-to-polymerase (N/P) activity ratio was high for antimutator (*L141*), intermediate for wild type (*43*$^+$), and extremely low for mutator (*L56*) strains (Muzyczka *et al.* 1972).

In the latter experiments, polymerase and 3′-exonuclease activities were measured on an oligo dT-polydA primer-template, using saturating dTTP substrate concentrations. Individual phosphocellulose column fractions of the three T4 pols showed N/P ratios that were constant across each chromatographic peak but varied between peaks. Wild-type T4 pol excised 1 molecule dTMP per 25 molecules inserted. In contrast, the *L141* antimutator T4 pol excised 10 out of 11 dTMPs inserted, while the *L56* mutator polymerase excised only one out of 200. The apparent correlation between N/P ratio and polymerase fidelity was very suggestive and demanded further substantiation.

In 1972, Linda J. Reha-Krantz joined Bessman's laboratory as a graduate student and embarked on a thesis project of heroic proportions. She grew T4 gene *43* amber mutants in *E. coli* suppressor strains and measured their mutation frequencies. She then purified the mutant polymerases and determined their N/P ratios. She observed a near-perfect correlation between antimutator and mutator behavior *in vivo* and correspondingly high and low N/P ratios (Reha-Krantz and Bessman 1977). These results were solid evidence that the balance between the polymerase and 3′-exonuclease reactions was fundamentally linked to the overall accuracy of DNA synthesis.

But was the N/P ratio actually determining the accuracy of DNA synthesis or was it merely correlated with increased accuracy in the individual polymerization and excision reactions? Bessman and co-workers addressed this question by measuring the specificity of the individual nuclease and polymerase reactions (Bessman *et al.* 1974). They showed that mutator, antimutator, and wild-type T4 pols (*L56*, *L141*, and *43*$^+$) inserted the mutagenic base analogue 2-aminopurine (2AP) opposite T with similar frequencies. What's more, the three polymerases were also similarly specific in removing 2AP: excising one correctly inserted A for every two to three "misinserted" 2AP molecules. The difference was in the overall activity of two reactions. The *L141* antimutator pol excised about 91% of the misinserted 2AP, resulting in a "low" net misincorporation frequency of about 3%, whereas the *L56* mutator excised only 20% of the 2APs, resulting a "high" error frequency of 10%.

The relevance of data using 2AP *in vitro* to the bacteriophage T4 system *in vivo* was documented in experiments showing that 2AP incorporation into T4 DNA *in vivo* was highest for *tsL56* mutator and very low for *tsL141* antimutator relative to *43*$^+$ (Goodman *et al.* 1977), and that the mutant and wild-type strains converted 2AP-free-base to 2AP-triphosphate with roughly similar efficiencies, giving rise to similar d(2AP)TP/dATP pool ratios for the three strains infecting *E. coli in vivo* (Hopkins and Goodman 1985).

Concurrently with experiments from Bessman's group, Nancy Nossal and her students at NIH were also using the T4 system to study polymerase fidelity (Hershfield 1973; Hershfield and Nossal 1972). Gillen and Nossal (1976) found that *L141* (CB120) polymerase had difficulty carrying out strand displacement, suggesting that an impediment to forward translocation may enable the enzyme to proofread more effectively. Indeed, it has been shown that the A737V mutation in *L141* causes an increase in exonuclease processivity at the expense of polymerase processivity (Spacciapoli and Nossal 1994). These results provide a mechanistic explanation for the increase in nuclease/polymerase ratio for the *L141* antimutator relative to wild-type polymerase.

To test and refine this mechanistic link between N/P ratio and polymerase fidelity, we carried out a kinetic analysis of the fidelity of *L141*, wild-type, and *L56* polymerases, comparing the incorporation of 2AP in direct competition with A opposite a template T (Clayton *et al.* 1979). We found that although 2AP misinsertion frequencies were the same for each enzyme at all dNTP concentrations, 2AP misincorporation frequencies were highly dependent on substrate concentration. At saturating dNTP concentrations, 2AP misincorporation frequencies were higher for mutator (*L56*) and lower for

antimutator (*L141*) compared to wild type but all three converged to the same value at low-dNTP concentrations. The effect of dNTP concentration was most pronounced for the relatively inactive *L56* exonuclease. The relatively active *L141* exonuclease was only marginally affected. Thus, we concluded that when low dNTP concentrations limit polymerase activity, even inactive exonucleases are able to edit out the majority of polymerase errors.

The logic can be seen by analogy to quality control along an assembly line. A polymerase is like a machine that makes widgets and sends them down the line at a certain rate. An exonuclease is like a worker responsible for removing defective widgets that come down the line. The worker sometimes removes perfect widgets by mistake. (The fewer such mistakes, the more "specific" the worker.) However, the number of defective widgets that get past the worker depends primarily on how many widgets the worker checks as the assembly line rolls by. If the assembly line slows down (*i.e.*, there arises an impediment to forward translocation), the worker will be able to check more widgets and therefore let fewer defective ones go by.

It should be noted, however, that N/P ratio is not a fail-safe indicator of a mutator phenotype. As Jan Drake has pointed out, it was fortunate that A·T→G·C mutations were investigated early on for the *tsL141* allele, otherwise it may not have been identified as an antimutator (Drake 1992).

> From the beginning (Drake and Allen 1968; Drake *et al.* 1969), it was clear that [antimutators] consistently reduce A·T→G·C transition rates (sometimes by more than 100-fold), reduce some but not all base-addition and base-deletion rates, but tend either not to affect or else to increase G·C→A·T transition rates.

Of course, antimutators will always exhibit some mutational specificity in the sense that they will only be found for alleles that are not well corrected in the wild type (Reha-Krantz 1995). And N/P ratio may not reflect on the ability to correct mutations templated by unusual (*e.g.*, slipped out) primer/template structures. *L141*, for example, exhibits an *increased* mutagenicity for simple frameshifts (Ripley and Shoemaker 1983), perhaps due to the altered processivity of its polymerase. Despite this lack of universality (Drake 1993), the N/P ratio continues to serve as an important enzymatic "marker" of polymerase fidelity.

Studies on the biophysical basis of mutation

Models of DNA polymerase fidelity: The discovery of proofreading spurred the development of theoretical models to account for polymerase fidelity. John Hopfield proposed that polymerases might rely on "kinetic proofreading" to edit out miscreant base pairs (Hopfield 1974). The key idea was that, after binding a dNTP in the polymerase active site, the enzyme might irreversibly enter an activated state, perhaps driven by hydrolysis of ATP. Discrimination between correct and incorrect nucleotides could then occur twice: first, upon entering the active site, where difference in the free energy of binding of right vs. wrong dNTPs would favor the correct nucleotide, and again upon leaving the activated state, where the reaction rates of hydrolysis or unbinding might also distinguish between correct and incorrectly bound nucleotides. Jacques Ninio proposed a similar model, invoking a "time delay" that facilitated nonproductive hydrolysis of a wrongly bound nucleotide (Ninio 1975). These models offered a means for reducing the number of nucleotides misinserted by a DNA polymerase without resorting to "brute force" excision by a dedicated proofreading exonuclease (Hopfield 1974).

We now know that Nature has found "brute force" acceptable, however, and a model which explicitly invokes a 3′→5′ exonuclease to excise polymerase insertion errors has proven most useful. The model was proposed by Galas and Branscomb (1978) in the context of analyzing the data of Bessman and co-workers (Bessman *et al.* 1974) for the incorporation and proofreading of 2AP using T4 *L56* mutator, *43*$^+$, and *L141* antimutator polymerases. A simplified sketch of this polymerase-proofreading model is presented in Figure 2.

The model treats polymerization and proofreading as two possible outcomes of a series of random events, which take place after a dNTP (right or wrong) binds to the enzyme. In the sketch, polymerization occurs in the lower reaction pathway and proofreading takes place in the upper pathway. Connecting the two pathways are the states (A) and (M), referring to annealed and melted primer-3′-termini, respectively. No distinction is made between right and wrong base pairs, except to recognize that Watson-Crick (WC) pairs favor the annealed state, $k_A > k_M$, while non-WC pairs tend to be melted out, $k_M > k_A$. However, a non-WC pair may, with low probability, be in the annealed state and get incorporated into a growing DNA chain, while a proper WC pair may be melted out and get excised. The model therefore suggests that it is the equilibrium between melted and annealed primer-3′-termini, rather than any intrinsic/geometric difference between WC and non-WC base pairs, that determines whether proofreading is likely to occur.

It was originally assumed that following either an incorporation or excision the system was constrained to begin a new polymerization-proofreading cycle starting from the annealed state (A). This assumption led to the prediction that saturating concentrations of a next-correct dNTP (complementary to the template base immediately downstream from the initial dNMP incorporation site) would completely suppress proofreading. However, the experimental data clearly showed that, although the excision of dNMP by the proofreading exonuclease diminished at saturating next-nucleotide

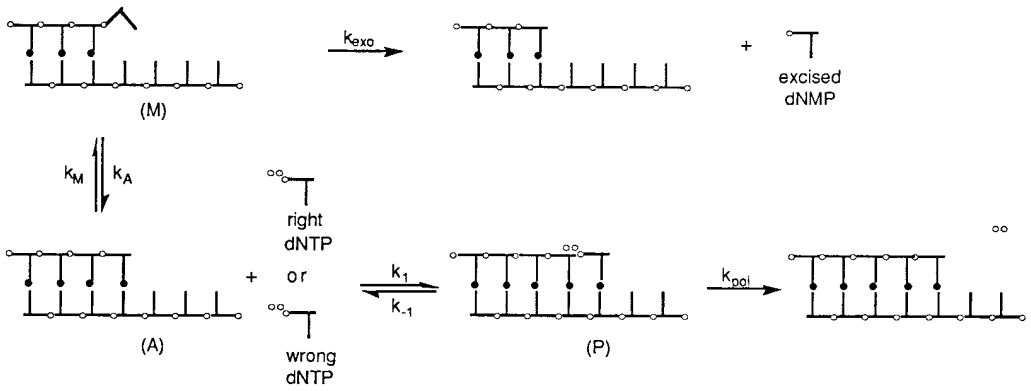


Figure 2.—Polymerase-proofreading model. Sketch of a simple model illustrating insertion and 3′-exonuclease proofreading of right and wrong nucleotides. State (M) refers to a melted primer terminus from which exonucleolytic excision takes place; state (A) refers to an annealed primer terminus along the polymerization pathway. Selective hydrolysis of misincorporated nucleotides results from the ratio, k_M/k_A, being much larger for mismatches than for correct matches. Polymerization from state (P) is favored over proofreading from state (M) as the concentration of rescue dNTP is increased. Following either excision or insertion, a shift occurs one base backward or forward to allow the cycle to repeat. When cycling occurs, the terminal base is assumed to reach an equilibrium distribution between states (A) and (M), explaining why proofreading is not entirely suppressed even at saturating concentrations of rescue dNTP (Clayton *et al.* 1979).

concentrations, it was nevertheless present to a significant extent (Clayton *et al.* 1979). A refinement of the model, allowing the system to reach an equilibrium distribution of melted and annealed primer termini following nucleotide incorporation and excision, resolved the problem. Partial suppression of proofreading in the presence of high dNTP concentrations is referred to as the "next-nucleotide effect" (Clayton *et al.* 1979; Fersht 1979), and has come to be recognized as a basic hallmark of proofreading (Echols and Goodman 1991; Goodman *et al.* 1993).

The Galas-Branscomb model highlights the importance of the interactions between polymerases, proofreading exonucleases, and primer-template DNA. It has served as a starting point for investigations into why mutational spectra and error rates differ substantially among polymerases in different sequence contexts.

Sequence context effects on DNA polymerase fidelity: One of the most general and important sequence context effects can be understood by examining the influence of local DNA stability on N/P ratios. Simply stated, stable regions are less frequently melted out, and so less available to exonuclease. Consequently, base substitution mutations tend to occur more frequently in more stable (*e.g.*, G·C-rich) sequences and less frequently in less stable (A·T-rich) regions. For example, it has been shown that T4 mutation frequencies *in vivo* and misincorporation of 2AP by T4 pol *in vitro* decrease with increasing temperature (Bessman and Reha-Krantz 1977). Were it not that higher temperatures made stable regions more accessible to exonuclease proofreading, one might expect mutations to increase because of higher rates of deamination and depurination reactions. The same study also showed that sites on DNA which are relatively insensitive to temperature also did not show an appreciable difference in mutation comparing *43*$^+$ and antimutator *L141* alleles.

The ambiguous base pairing properties of 2AP make it a useful compound for studying fidelity *in vitro* and mutagenesis *in vivo* (Echols and Goodman 1991; Ronen 1979). However, 2AP has another extremely useful property; it is moderately fluorescent and can therefore be used to study polymerase mechanisms by observing its insertion by polymerase and excision by exonuclease on a pre-steady-state time scale (Bloom *et al.* 1993; Bloom *et al.* 1994; Frey *et al.* 1995).

Further evidence for the effect of local DNA stability on mutagenesis came from such pre-steady-state measurements. Excision of 2AP was measured on a millisecond time scale by its increase in fluorescence upon excision from a primer-3′-terminus and a concomitant increase in rotation, as measured by fluorescence depolarization (Bloom *et al.* 1994). 2AP was placed at a primer-3′-terminus opposite template T, C, A or G, while maintaining a constant surrounding sequence context. The observed excision rate correlated inversely with the stability of the base pair. Thus, removal of 2AP was slowest when paired opposite T, with the order of excision being 2AP·T < 2AP·A < 2AP·C < 2AP·G. Measurements were then made of the hydrolysis of 2AP·N base pairs placed proximal to either A·T- or G·C-rich neighboring sequences. It was found that a proper Watson-Crick 2AP·T base pair in an A·T-rich environment was actually excised faster than a wobble 2AP·C mispair in a G·C-rich environment (Bloom *et al.* 1994).

Another important sequence context effect on fidelity comes from the influence of base-stacking interactions. Ronen and Rahat (1976) first showed that neighboring base pairs influenced 2AP-induced base substitution mutation rates. Later, Pless and Bessman (1983) cre-

ated an extensive data set for misincorporation of 2AP at 57 different template T sites on ϕX DNA, using T4 wild-type and *L141* DNA polymerases. These data showed that 2AP misincorporation frequencies varied from 0 to 20% when T was located at the primer-3′-terminus, 0 to 14% for C nearest-neighbors, and from 0 to ~7% for both G and A primer-termini. At first, the 2AP misincorporation frequency did not appear to correlate strongly with the identity of the nearest-neighbor base-stacking partners on the primer-3′-terminus. However, John Petruska and M.F.G. showed that base-stacking interactions between an incoming 2dAPTP or dATP and the base at the primer-3′-terminus were a significant factor in explaining the data, once the influence of DNA duplex stability on exonuclease activity in the vicinity of the primer terminus was taken into account (Petruska and Goodman 1985). The data for the *L141* antimutator fit best to a model in which five base pairs, both upstream and downstream from the site of 2AP misincorporation, contributed to neighboring sequence stability. The requirement to include contributions from downstream DNA implied that synthesis can continue for as many as five base pairs beyond the misincorporation site before the enzyme enters a processive peelback mode to excise the errant base pair, as previously observed *in vitro* (Goodman *et al.* 1974).

While duplex stability and base-stacking are certainly fundamental, sequence context effects on DNA synthesis fidelity can be considerably more complex, leading to differences that persist irrespective of proofreading (Echols and Goodman 1991). For example, Miller and co-workers (Coulondre *et al.* 1978) showed that deamination of 5-methyl cytosine gives rise to strong mutational hot spots in *E. coli*. Koch (1971), Streisinger *et al.* (1966), and Kunkel and colleagues (Kunkel 1985, 1986; Kunkel and Soni 1988) showed that frameshift and base substitutions could result from local primer-template slippage. We have shown that polymerase-dNTP interactions can stabilize slipped primer-template sequences and lead to misincorporations (Bloom *et al.* 1997, 1998; Efrati *et al.* 1997). Polymerase-catalyzed primer-template slippage is also a plausible explanation for the occurrence of triplet repeat expansions that are implicated in causing neurodegenerative diseases (Mitas 1997).

Studies on the physical chemical basis of mutation

Kinetics of fidelity: Insight into nucleotide misinsertion and proofreading mechanisms has come from enzyme kinetic analysis on steady state (Bloom *et al.* 1997, 1998; Goodman *et al.* 1993) and pre-steady-state (Bloom *et al.* 1994; Johnson 1993; Kuchta *et al.* 1987) time scales. Pre-steady-state analysis provides detailed mechanistic information on individual steps in polymerization and proofreading pathways, while steady-state measurements are used to determine transition and transversion misincorporation rates for polymerases in different sequence contexts. In one complementary series of studies, mutational spectra obtained *in vitro* (Kunkel 1985, 1986; Kunkel and Soni 1988) and *in vivo* (Schaaper 1988) offer a large-scale view of base substitution mutations and small and large additions and deletions, which occur spontaneously in different target genes.

A "gold standard" measurement of fidelity is made by having ^{3}H-labeled right and ^{32}P-labeled wrong dNTPs compete directly for incorporation into DNA. While this method has proven to be excellent for measuring fidelity using base analogues such as 5BU (Trautner *et al.* 1962) and 2AP (Bessman *et al.* 1974), which are incorporated with reasonably high efficiencies, it is difficult, if not impossible, to measure misincorporations occurring at frequencies less than 10^{-4}, which are characteristic of natural base mispairs (Echols and Goodman 1991).

We have replaced the nucleotide competition method to measure fidelity by a kinetic approach originally suggested by Fersht (1985). In the kinetic method, incorporation of right and wrong nucleotides are measured in separate reactions as a function of dNTP concentration to obtain V_{max}/K_m values for incorporating each nucleotide. The ratio of V_{max}/K_m's for right and wrong incorporations measures polymerase fidelity in either the presence or absence of proofreading (Creighton and Goodman 1995; Fersht 1985). We have built on the ideas of Fersht to develop a gel-kinetic fidelity assay in which polymerases, including those with proofreading, processivity clamps, and clamp loading proteins can be included in the assay (Bloom *et al.* 1997, 1998; Creighton and Goodman 1995). Recently, we have used the assay to determine the base substitution error frequency of the *E. coli* pol III holoenzyme complex and found it to fall within in a range from 5×10^{-6} to 4×10^{-7} (Bloom *et al.* 1997, 1998).

Relating kinetics to thermodynamics: Perhaps the most important feature of the Galas-Branscomb model is that it yields estimates of the differences between free energy of matched and mismatched base pairs in the polymerase active site and in the exonuclease active site. These $\Delta\Delta G_{pol}$ and $\Delta\Delta G_{exo}$ parameters can be compared with free energy differences, $\Delta\Delta G^0$, obtained using van't Hoff (Aboul-ela *et al.* 1985; Petruska *et al.* 1988) or calorimetric analyses (Breslauer 1995). The comparisons reveal a sizeable discrepancy.

Measurements of 2AP insertion opposite template T in competition with A (Clayton *et al.* 1979) or insertion of C opposite 2AP in competition with T (Watanabe and Goodman 1981, 1982) yielded free energy differences too large by about a factor of two compared with $\Delta\Delta G^0$'s obtained from melting heteroduplex DNA containing 2AP·T and 2AP·C base pairs (Law *et al.* 1996). Similar conclusions were reached using the gel kinetic assay to measure the fidelity of DNA polymerases for a

wide variety of naturally occurring transition (Pu·Py) and transversion (Pu·Pu, Py·Py) mispairs (Mendelman *et al.* 1989; Petruska *et al.* 1988).

The notion that thermodynamically derived $\Delta\Delta G^0$ values might govern polymerase insertion specificity was based on the fact that on-off rates for binding of dNTP to the polymerase-DNA complex are extremely rapid compared to phosphodiester bond formation, thereby allowing right and wrong nucleotides to reach an effective equilibrium at the pol active site (Clayton *et al.* 1979; Galas and Branscomb 1978). Similarly, the melting, k_M, and annealing, k_A, of primer-3-termini (Figure 2) are rapid relative to excision of a previously inserted nucleotide or insertion of a next-correct nucleotide, so that both pol and exo sites sample an equilibrium population of melted and annealed primer termini (Clayton *et al.* 1979; Galas and Branscomb 1978).

The discrepancy in the thermodynamic and kinetically determined values of the free energy differences prompted Hatch Echols and M.F.G to propose that geometric selection imposed at the pol active site could account for the high nucleotide insertion fidelity (Echols and Goodman 1991; Sloane *et al.* 1988). This post-dNTP binding step strongly selects for insertion of bases most closely approximating W-C base pairing geometries (Echols 1982; Sloane *et al.* 1988). Geometric selection can be accounted for by an induced-fit mechanism derived from pre-steady-state kinetic analyses (Kuchta *et al.* 1987; Kuchta *et al.* 1988; Wong *et al.* 1991). A commentary on the use of geometric selection as a principal determinant of polymerase insertion specificity has recently been published (Goodman 1997).

In considering various ways that polymerases might achieve high fidelity by taking advantage of free energy differences between right and wrong base pairs, we investigated two aspects of DNA melting in relation to the pol active site. These are: (1) the partial exclusion of water at the pol active site (Petruska *et al.* 1986), and (2) the effect of enthalpy-entropy compensation in suppressing free entropy differences between right and wrong base pairs (Petruska and Goodman 1995). With regard to exclusion of water at the pol active site, we argued that when bound dNTPs and template bases confront each other in a lower dielectric medium that acts to partially exclude water, the enthalpic component, $\Delta\Delta H_{pol}$, is likely to be amplified over the value derived from the melting of DNA in aqueous solution. In aqueous solution, water-base H-bonding interactions act to suppress intrinsic base-base interactions, but when dNTPs enter the polymerase active site, the partial removal of water around the bases may restore the H-bonding and stacking interactions at the p/t terminus, and thereby amplify the free energy differences between matched and mismatched base pairs (Petruska *et al.* 1986).

When double-stranded DNA is melted in aqueous solution, there is a concomitant increase in both the enthalpy and entropy components of $\Delta G^0 = \Delta H^0 - T\Delta S^0$. The correlation between the energy required to melt DNA base pairs and their increase in entropy is referred to as enthalpy-entropy compensation (Lumry and Rajender 1970). It takes more energy to melt stable, rigidly constrained base pairs than it does to melt unstable, weakly constrained base pairs. At the same time, rigid base pairs having fewer degrees of freedom in the double helix will gain a greater number of degrees of freedom after melting than do less stable base pairs. Thus, if the polymerase active site can, via geometric selection, constrain the movement of the bound dNTPs, it may be possible to reduce the value of $\Delta\Delta S$ so that $\Delta\Delta G$ is much closer in magnitude to $\Delta\Delta H$. $\Delta\Delta H^o$ values as measured in aqueous solution are almost large enough to account for pol insertion fidelity (Law *et al.* 1996; Petruska *et al.* 1988). The effects of suppressing $\Delta\Delta S$ coupled with the partial exclusion of water at the pol active site to increase $\Delta\Delta H$ may be sufficient to account for the ability of polymerases to discriminate between WC and non-WC base pairs.

Perspectives and Conclusions

In the last five years, progress has proceeded apace. Pre-steady-state kinetic experiments have succeeded in breaking down polymerization and proofreading pathways into individual steps (Capson *et al.* 1992; Eger and Benkovic 1992; Johnson 1993). Genetic studies have succeeded in isolating T4 mutators and antimutators in which switching between pol and exo sites is impeded (Reha-Krantz and Nonay 1994; Stocki *et al.* 1995), and direct measurements of the rates at which primer-template switches from exonuclease to polymerase active clefts of T4 pol have been reported (Bloom *et al.* 1994; Capson *et al.* 1992). X-ray crystallographic studies have yielded crystal structures for two proofreading polymerases (Beese *et al.* 1993; Derbyshire *et al.* 1988; Wang *et al.* 1997), which promise new insights into the physical processes which transport p/t DNA between pol and exo active sites. And, theoretical calculations have elucidated the roles of divalent cations in transition state stabilization and proton transfer at the exonuclease active site (Fothergill *et al.* 1995).

However, even though properties of polymerases and proofreading exonucleases have been examined in considerable detail and DNA structural features are well-understood, mutational hot and cold spots remain, for the most part, unpredictable. What is clear is that our basic understanding is incomplete and perhaps even flawed. Recent results showing that difluorotoluene, an isosteric analogue of T that cannot form H-bonds, is incorporated opposite A "almost" as well as is T (Moran *et al.* 1997) call for a re-examination of the role of H-bonding aiding in the selection of WC over non-WC base pairs (Goodman 1997). But counterintuitive results, such as this one or the one Jan Drake reported

three decades ago, are just the sort that energize a field and leave remarkable revelations in their wake.

M.F.G. thanks Maurice J. Bessman for providing an exceptionally rich intellectual environment in which to learn and, more importantly, to appreciate the field of molecular biochemistry. M.F.G. is also deeply grateful to Jan Drake for providing numerous invaluable contributions to the mutagenesis community and to John Petruska and Hatch Echols for sharing their deep scientific insights and unbridled enthusiasm. This work was supported by National Institutes of Health grant GM-21422. D.F.K. is a Fellow of The Jane Coffin Childs Memorial Fund for Medical Research.

LITERATURE CITED

Aboul-ela, F., D. Koh, I. J. Tinoco and F. H. Martin, 1985 Base-base mismatches: thermodynamics of double helix formation for $dCA_3XA_3G + dCT_3YT_3G$ (X, Y = A,C,G,T). Nucleic Acids Res. **13:** 4811–4825.

Beese, L. S., V. Derbyshire and T. A. Steitz, 1993 Structure of DNA polymerase I Klenow fragment bound to duplex DNA. Science **260:** 352–355.

Benzer, S., 1961 On the Topography of the Genetic Fine Structure. Proc. Natl. Acad. Sci. USA **47:** 403–415.

Bessman, M. J., and L. J. Reha-Krantz, 1977 Studies on the biochemical basis of spontaneous mutation. V. Effects of temperature on mutation frequency. J. Mol. Biol. **116:** 115–123.

Bessman, M. J., N. Muzyczka, M. F. Goodman and R. L. Schnaar, 1974 Studies on the biochemical basis of spontaneous mutation. II. The incorporation of a base and its analogue into DNA by wild-type, mutator, and antimutator DNA polymerases. J. Mol. Biol. **88:** 409–421.

Bloom, L. B., M. R. Otto, J. M. Beechem and M. F. Goodman, 1993 Influence of 5′-nearest neighbors on the insertion kinetics of the fluorescent nucleotide analog 2-aminopurine by Klenow fragment. Biochemistry **32:** 11247–11258.

Bloom, L. B., M. R. Otto, R. Eritja, L. J. Reha-Krantz, M. F. Goodman *et al.*, 1994 Pre-steady-state kinetic analysis of sequence dependent excision of the fluorescent nucleotide analog, 2-Aminopurine, by T4 DNA polymerase. Biochemistry **33:** 7576–7586.

Bloom, L. B., X. Chen, D. Kuchnir Fygenson, J. Turner, M. O'Donnell *et al.*, 1997 Fidelity of *Escherichia coli* DNA polymerase III holoenzyme: the effects of β, γ complex processivity proteins and ε proofreading exonuclease on nucleotide misincorporation efficiencies. J. Biol. Chem. **272:** 27919–27930.

Breslauer, K. J., 1995 Extracting thermodynamic data from equilibrium melting curves for oligonucleotide order-disorder transitions. Methods Enzymol. **259:** 221–242.

Brutlag, D., and A. Kornberg, 1972 Enzymatic synthesis of deoxyribonucleic acid. XXXVI. A proofreading function for the 3′ → 5′ exonuclease activity in deoxyribonucleic acid polymerases. J. Biol. Chem. **247:** 241–248.

Capson, T. L., J. A. Peliska, B. F. Kaboord, M. W. Frey, C. Lively *et al.*, 1992 Kinetic characterization of the polymerase and exonuclease activities of the gene 43 protein of bacteriophage T4. Biochemistry **31:** 10984–10994.

Clayton, L. K., M. F. Goodman, E. W. Branscomb and D. J. Galas, 1979 Error induction and correction by mutant and wild type T4 DNA polymerases: kinetic error discrimination mechanisms. J. Biol. Chem. **254:** 1902–1912.

Coulondre, C., J. H. Miller, P. J. Farabaugh and W. Gilbert, 1978 Molecular basis of base substitution hotspots in *Escherichia coli.* Nature **274:** 775–780.

Creighton, S., and M. F. Goodman, 1995 Gel kinetic analysis of DNA polymerase fidelity in the presence of proofreading using bacteriophage T4 DNA polymerase. J. Biol. Chem. **270:** 4759–4774.

Derbyshire, V., P. S. Freemont, M. R. Sanderson, L. Beese, J. M. Friedman *et al.*, 1988 Genetic and crystallographic studies of the 3′,5′-exonucleolytic site of DNA polymerase I. Science **240:** 199–201.

de Waard, A., A. V. Paul and I. R. Lehman, 1965 The structural gene for deoxyribonucleic acid polymerase in bacteriophage T4 and T5. Proc. Natl. Acad. Sci. USA **54:** 1241–1248.

Drake, J. W., 1992 Mutation rates. BioEssays **14:** 137–140.

Drake, J. W., 1993 General antimutators are improbable. J. Mol. Biol. **229:** 8–13.

Drake, J. W., and E. F. Allen, 1968 Antimutagenic DNA polymerases of bacteriophage T4. Cold Spring Harbor Symp. Quant. Biol. **33:** 339–344.

Drake, J. W., E. F. Allen, S. A. Forsberg, R. Preparata and E. O. Greening, 1969 Spontaneous mutation. Genetic control of mutation rates in bacterophage T4. Nature **221:** 1128–1131.

Echols, H., 1982 Mutation rate: some biological and biochemical considerations. Biochimie **64:** 571–575.

Echols, H., and M. F. Goodman, 1991 Fidelity mechanisms in DNA replication. Annu. Rev. Biochem. **60:** 477–511.

Efrati, E., G. Tocco, R. Eritja, S. H. Wilson and M. F. Goodman, 1997 Abasic translesion synthesis by DNA polymerase b violates the "A-Rule": novel types of nucleotide incorporation by human DNA polymerase b at an abasic lesion in different sequence contexts. J. Biol. Chem. **272:** 2559–2569.

Eger, B. T., and S. J. Benkovic, 1992 Minimal kinetic mechanism for misincorporation by DNA polymerase I (Klenow fragment). Biochemistry **31:** 9227–9236.

Fersht, A. R., 1979 Fidelity of replication of phage ϕX174 DNA by DNA polymerase III holoenzyme; spontaneous mutation by misincorporation. Proc. Natl. Acad. Sci. USA **76:** 4946–4950.

Fersht, A. R., 1985 *Enzyme Structure and Mechanism.* W. H. Freeman & Co., New York.

Fothergill, M., M. F. Goodman, J. Petruska and A. Warshel, 1995 Structure-energy analysis of the role of metal ions in phosphodiester bond hydrolysis by DNA polymerase I. J. Am. Chem. Soc. **117:** 11619–11627.

Freese, E. B., and E. F. Freese, 1967 On the specificity of DNA polymerase. Proc. Natl. Acad. Sci. USA **57:** 650–657.

Frey, M. W., L. C. Sowers, D. P. Millar and S. J. Benkovic, 1995 The nucleotide analog 2-A minopurine as a spectroscopic probe of nucleotide incorporation by the Klenow fragment of Escherichia coli polymerase I and bacteriophage T4 DNA polymerase. Biochemistry **34:** 9185–9192.

Galas, D. J., and E. W. Branscomb, 1978 Enzymatic determinants of DNA polymerase accuracy: theory of coliphage T4 polymerase mechanisms. J. Mol. Biol. **88:** 653–687.

Gillen, F. D., and N. G. Nossal, 1976 Control of mutation frequency by bacteriophage T4 DNA polymerase I. The CB120 antimutator DNA polymerase is defective in strand displacement. J. Biol. Chem. **251:** 5219–5224.

Goodman, M. F., 1997 Hydrogen bonding revisited: geometric selection as a principal determinant of DNA replication fidelity. Proc. Natl. Acad. Sci. USA **94:** 10493–10495.

Goodman, M. F., W. C. Gore, N. Muzyczka and M. J. Bessman, 1974 Studies on the biochemical basis of spontaneous mutation. III. Rate model for DNA polymerase-effected nucleotide misincorporation. J. Mol. Biol. **88:** 423–435.

Goodman, M. F., R. Hopkins and W. C. Gore, 1977 2-aminopurine-induced mutagenesis in T4 bacteriophage: a model relating mutation frequency to 2-aminopurine incorporation in DNA. Proc. Natl. Acad. Sci. USA **74:** 4806–4810.

Goodman, M. F., S. Creighton, L. B. Bloom and J. Petruska, 1993 Biochemical basis of DNA replication fidelity. Crit. Rev. Biochem. Mol. Biol. **28:** 83–126.

Hershfield, M. S., 1973 On the role of deoxyribonucleic acid polymerase in determining mutation rates: characterization of the defect in the T4 deoxyribonucleic acid polymerase caused by the *ts*L88 mutation. J. Biol. Chem. **248:** 1417–1423.

Hershfield, M. S., and N. G. Nossal, 1972 Hydrolysis of template and newly synthesized deoxyribonucleic acid by the 3′ to 5′ exonuclease activity of the T4 deoxyribonucleic acid polymerase. J. Biol. Chem. **247:** 3393–3404.

Hopfield, J. J., 1974 Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. Proc. Natl. Acad. Sci. USA **77:** 5248–5252.

Hopkins, R. L., and M. F. Goodman, 1985 Ribonucleoside and deoxyribonucleoside triphosphate pools during 2-Aminopurine mutagenesis in T4 mutator-, wild type-, and antimutator-infected *Escherichia coli.* J. Biol. Chem. **260:** 6618–6622.

Johnson, K. A., 1993 Conformational coupling in DNA polymerase fidelity. Annu. Rev. Biochem. **62:** 685–713.

Koch, R. E., 1971 The influence of neighboring base pairs upon base-pair substitution mutation rates. Proc. Natl. Acad. Sci. USA **68:** 773–776.

Kuchta, R. D., V. Mizrahi, P. A. Benkovic, K. A. Johnson and S. J. Benkovic, 1987 Kinetic mechanism of DNA polymerase I (Klenow). Biochemistry **26:** 8410–8417.

Kuchta, R. D., P. Benkovic and S. J. Benkovic, 1988 Kinetic mechanism whereby DNA polymerase I (Klenow) replicates DNA with high fidelity. Biochemistry **27:** 6716–6725.

Kunkel, T. A., 1985 The mutational specificity of DNA polymerase-β during *in vitro* DNA synthesis. J. Biol. Chem. **260:** 5787–5796.

Kunkel, T. A., 1986 Frameshift mutagenesis by eucaryotic DNA polymerases *in vitro.* J. Biol. Chem. **261:** 13581–13587.

Kunkel, T. A., and A. Soni, 1988 Mutagenesis by transient misalignment. J. Biol. Chem. **263:** 14784–14789.

Law, S. M., R. Eritja, M. F. Goodman and K. J. Breslauer, 1996 Spectroscopic and calorimetric characterizations of DNA duplexes containing 2-Aminopurine. Biochemistry **35:** 12329–12337.

Lumry, R., and S. Rajender, 1970 Enthalpy-entropy compensation phenomena in water solutions of proteins and small molecules: a ubiquitous property of water. Biopolymers **9:** 1125–1227.

Mendelman, L. V., M. S. Boosalis, J. Petruska and M. F. Goodman, 1989 Nearest neighbor influences on DNA polymerase insertion fidelity. J. Biol. Chem. **264:** 14415–14423.

Mitas, M., 1997 Trinucleotide repeats associated with human disease. Nucleic Acids Res. **25:** 2245–2253.

Moran, S., R. X.-F. Ren and E. T. Kool, 1997 A thymidine triphosphate shape analog lacking Watson-Crick pairing ability is replicated with high sequence selectivity. Proc. Natl. Acad. Sci. USA **94:** 10506–10511.

Muzycyka, N., R. L. Poland and M. J. Bessman, 1972 Studies on the biochemical basis of spontaneous mutation. I. A comparison of the deoxyribonucleic acid polymerase of mutator, antimutator, and wild type strains of bacteriophage T4. J. Biol. Chem. **247:** 7116–7122.

Ninio, J., 1975 Kinetic amplification of enzyme discrimination. Biochimie **57:** 587–595.

Petruska, J., and M. F. Goodman, 1985 Influence of neighboring bases on DNA polymerase insertion and proofreading fidelity. J. Biol. Chem. **260:** 7533–7539.

Petruska, J., and M. F. Goodman, 1995 Enthalpy-entropy compensation in DNA melting thermodynamics. J. Biol. Chem. **270:** 746–750.

Petruska, J., L. C. Sowers and M. F. Goodman, 1986 Comparison of nucleotide interactions in water, proteins, and vacuum: model for DNA polymerase fidelity. Proc. Natl. Acad. Sci. USA **83:** 1559–1562.

Petruska, J., M. F. Goodman, M. S. Boosalis, L. C. Sowers, C. Cheong *et al.*, 1988 Comparison between DNA melting thermodynamics and DNA polymerase fidelity. Proc. Natl. Acad. Sci. USA **85:** 6252–6256.

Pless, R. C., and M. J. Bessman, 1983 Influence of local nucleotide sequence on substitution of 2-aminopurine for adenine during deoxyribonucleic acid synthesis *in vitro.* Biochemistry **22:** 4905–4915.

Reha-Krantz, L. J., 1988 Amino acid changes coded by bacteriophage T4 DNA polymerase mutator mutants: relating structure to function. J. Mol. Biol. **202:** 711–724.

Reha-Krantz, L. J., 1989 Locations of amino acid substitutions in bacteriophage T4 *ts*L56 DNA polymerase predict an N-terminal exonuclease domain. J. Virol. **63:** 4762–4766.

Reha-Krantz, L. J., 1995 Learning about DNA polymerase function by studying antimutator DNA polymerases. TIBS **20:** 136–140.

Reha-Krantz, L. J., and M. J. Bessman, 1977 Studies on the biochemical basis of spontaneous mutation. IV. Effect on amino acid substitution on the enzymatic and biological properties of the T4 DNA polymerase. J. Mol. Biol. **116:** 99–113.

Reha-Krantz, L. J., and R. L. Nonay, 1994 Motif A of bacteriophage T4 DNA polymerase: role in primer extension and DNA replication fidelity. J. Biol. Chem. **269:** 5635–5643.

Ripley, L. S., and N. B. Shoemaker, 1983 A major role for bacteriophage T4 DNA polymerase in frameshift mutagenesis. Genetics **103:** 353–366.

Ronen, A., 1979 2-aminopurine. Mutat. Res. **75:** 1–47.

Ronen, A., and A. Rahat, 1976 Mutagen specificity and position effects on mutation in T4 *rII* nonsense sites. Mutat. Res. **34:** 21–34.

Schaaper, R. M., 1988 Mechanisms of mutagenesis in the *Escherichia coli* mutator mutD5: role of DNA mismatch repair. Proc. Natl. Acad. Sci. USA **85:** 8126–8130.

Sloane, D. L., M. F. Goodman and H. Echols, 1988 The fidelity of base selection by the polymerase subunit of DNA polymerase III holoenzyme. Nucleic Acids Res. **16:** 6465–6475.

Spacciapoli, P., and N. G. Nossal, 1994 A single mutation in bacteriophage T4 DNA polymerase (A737V, *ts*L141) decreases its processivity as a polymerase and increases its processivity as a 3′→5′ exonuclease. J. Biol. Chem. **269:** 438–446.

Speyer, J. F., 1965 Mutagenic DNA polymerase. Biochem. Biophys. Res. Comm. **21:** 6–8.

Speyer, J. F., J. D. Karam and A. B. Lenny, 1966 On the role of DNA polymerase in base selection. Cold Spring Harbor Symp. Quant. Biol. **31:** 693–697.

Stocki, S. A., R. L. Nonay and L. J. Reha-Krantz, 1995 Dynamics of bacteriophage T4 DNA polymerase function: identification of amino acid residues that affect switching between polymerase and 3′ to 5′ exonuclease activities. J. Mol. Biol. **254:** 15–28.

Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita *et al.*, 1966 Frameshift mutations and the genetic code. Cold Spring Harbor Symp. Quant. Biol. **31:** 77–84.

Trautner, T. A., M. N. Swartz and A. Kornberg, 1962 Enzymatic synthesis of deoxyribonucleic acid. X. Influence of bromouracil substitutions on replication. Proc. Natl. Acad. Sci. USA **48:** 449–455.

Wang, J., A. K. Sattar, C. C. Wang, J. D. Karam, W. H. Konigsberg *et al.*, 1997 Crystal structure of a pol alpha family replication DNA polymerase from bacteriophage RB69. Cell **89:** 1087–1099.

Watanabe, S. M., and M. F. Goodman, 1981 On the molecular basis of transition mutations: frequencies of forming 2-aminopurine·cytosine and adenine·cytosine base mispairs *in vitro.* Proc. Natl. Acad. Sci. USA **78:** 2864–2868.

Watanabe, S. M., and M. F. Goodman, 1982 Kinetic measurement of 2-aminopurine·cytosine and 2-aminopurine·thymidine base pairs as a test of DNA polymerase fidlity mechanisms. Proc. Natl. Acad. Sci. USA **79:** 6429–6433.

Wong, I., S. S. Patel and K. A. Johnson, 1991 An induced-fit kinetic mechanism for DNA replication fidelity: direct measurement by single-turnover kinetics. Biochemistry **30:** 526–537.