Error-Prone Repair DNA Polymerases in Prokaryotes and Eukaryotes

Myron F. Goodman

Department of Biological Sciences and Chemistry, Hedco Molecular Biology Laboratory, University of Southern California, Los Angeles, California 90089-1340; e-mail: mgoodman@mizar.usc.edu

Key Words Y-family polymerases, SOS mutagenesis, low-fidelity DNA synthesis, somatic hypermutation, DNA repair factories

■ Abstract DNA repair is crucial to the well-being of all organisms from unicellular life forms to humans. A rich tapestry of mechanistic studies on DNA repair has emerged thanks to the recent discovery of Y-family DNA polymerases. Many Y-family members carry out aberrant DNA synthesis—poor replication accuracy, the favored formation of non-Watson-Crick base pairs, efficient mismatch extension, and most importantly, an ability to replicate through DNA damage. This review is devoted primarily to a discussion of Y-family polymerase members that exhibit error-prone behavior. Roles for these remarkable enzymes occur in widely disparate DNA repair pathways, such as UV-induced mutagenesis, adaptive mutation, avoidance of skin cancer, and induction of somatic cell hypermutation of immunoglobulin genes. Individual polymerases engaged in multiple repair pathways pose challenging questions about their roles in targeting and trafficking. Macromolecular assemblies of replication-repair "factories" could enable a cell to handle the complex logistics governing the rapid migration and exchange of polymerases.

CONTENTS

INTRODUCTORY PERSPECTIVE	18
SOS RESPONSE INDUCED BY DNA DAMAGE IN E. COLI	20
THREE E. COLI DNA POLYMERASES INDUCED BY SOS	23
Biochemical Basis of SOS Mutagenesis	23
Pol V Mut Catalysis of Error-Prone Translesion Synthesis	24
A Pivotal Role for Pol II in Error-Free Replication Restart.	28
Translesion Synthesis with Pol IV and Pol II	29
Pol IV Generates Untargeted and Adaptive Mutations.	31
EUKARYOTIC ERROR-PRONE DNA POLYMERASES	32
Rev1 and Pol ζ	32
Pol η	33
Pol ι	34
Human Pol κ	34

Trf4/pol σ	35
Pol μ	35
Pol λ	36
SOMATIC HYPERMUTATION	36
Error-Prone Polymerases as Somatic Hypermutation Generators.	38
FUTURE PERSPECTIVE	40
DNA Repair Factory	41

INTRODUCTORY PERSPECTIVE

"If it ain't broke don't fix it." That familiar saying has a corollary applicable to DNA damage repair—"If it is broke fix it." Base excision repair (BER) and nucleotide excision repair (NER) are responsible for fixing DNA damage by employing analogous biochemical pathways in prokaryotes and eukaryotes. But what happens if instead of fixing its DNA, an organism copies either damaged or undamaged DNA in a somewhat haphazard manner? This question emanates from the recent discoveries of enzymes called error-prone DNA polymerases (EP pols).

We define an EP pol as having one or possibly more of the following properties: (*a*) an ability to copy damaged DNA with high efficiency, either alone or in the presence of accessory proteins; (*b*) poor accuracy in nucleotide incorporation with base substitution error frequencies of $\sim 10^{-1}$ to 10^{-3} ; (*c*) a tendency to form base mispairs rather than correct Watson-Crick base pairs; and (*d*) a propensity to catalyze incorporation using aberrant DNA primer ends, including base mismatches, misaligned primer-template, and DNA damage sites. The term "error-prone" is meant to convey that EP pols behave differently from the more familiar replication and repair polymerases. It doesn't, however, imply a strict dichotomy between EP and normal polymerases, as some overlap in fidelity properties is inevitable.

Two EP pols from *Escherichia coli*, UmuD₂[']C (pol V) and DinB (pol IV), and two from *Saccharomyces cerevisiae*, Rev1 and Rad30, share conserved sequence motifs (Figure 1) and have been designated as charter members of the UmuC/DinB/Rev1/Rad30 family of polymerases. These EP-pol motifs bear little relationship to standard replication and repair polymerase motifs. An ever expanding number of UmuC/DinB/Rev1/Rad30 homologs (1), representing at least 57 separate phylogenetic groupings, have recently been renamed Y-family polymerases (2). The Y-family name originated ostensibly because it followed upon the heels of the previously described X-family pols, but the Y designation might just as easily have been used to ask "Why are they there?"

Although little is known about either the functions or properties of the vast majority of Y-family polymerases, what is known is surely remarkable. Each founding Y-family member exhibits a distinctive example of EP behavior during DNA synthesis. *E. coli* pol V (UmuD₂'C) copies a variety of DNA damage by



Figure 1 Domain structure of Y-family DNA polymerases. Conserved and unique domains are represented for *E. coli* pol V (UmuC), *E. coli* pol IV (DinB), human pol κ (DinB), human pol η (XPV or Rad30A), human pol ι (Rad30B), and *S. cerevisiae* Rev1. The highly conserved domains I–III (blue rectangles) contain catalytic residues, and IV–V (red ovals) contain helix-hairpin-helix motifs (HhH). Amino acid clusters involved in Mg²⁺ binding and catalysis, based on site-specific mutational analysis, are indicated above domains I–III in UmuC, the least conserved family member. The DinB subgroup contains a short conserved motif (orange square) present from *E. coli* to humans. The C₂HC zinc-binding motif (blue diamond) is involved in DNA binding, and the C₂H₂ zinc-binding motif (yellow diamond) is required for targeting to replication foci, perhaps via interactions with proliferating cell nuclear antigen (PCNA). (SHM, somatic hypermutation; aa, amino acids.)

leaving numerous mutations in its wake. *E. coli* pol IV (DinB) extends mismatched primer ends on undamaged DNA and also copies some types of DNA damage. Yeast Rev1 favors the exclusive incorporation of C opposite abasic (apurinic/apyrimidinic) template lesions. Yeast pol η (Rad30) copies UV-damaged DNA, but much more accurately than pol V. EP behavior may not be a common characteristic of all or perhaps even most Y-family members—it is still too early to tell. Despite this caveat, what is currently known concerning EP Y-family polymerases is worth recounting. This review is devoted primarily to a discussion of those Y-family polymerases that do exhibit error-prone behavior. These polymerases include the four founding Y-family members and their human homologs, along with errant fellow travelers such as pol ι , an enzyme that prefers to incorporate G rather than A opposite T. We also describe the properties of several repair polymerases that are not Y-family members; two examples are *E. coli* pol II and eukaryotic pol ζ .

Questions abound. Why do EP DNA polymerases even exist? Where are they found? When and how do they function? The potential benefit to the cell of using EP pols could come from their ability to replace normal replication complexes that stall when encountering DNA damage, or that disassemble occasionally while copying undamaged DNA (3). The price for using EP pols, an increased mutational load, may be more than offset by an increased relative fitness of cells growing in inhospitable environments, paving the road toward adaptation and evolution.

The value of EP pols is perhaps less obvious in more highly developed organisms because programmed cell death (apoptosis) provides a route for elimination of cells with damaged genomes. Even so, one critically important enzyme, human pol η , encoded by the structural gene XPV, (4, 5), plays an essential role in avoiding an especially ravaging type of sunlight-induced skin cancer, a variant form of xeroderma pigmentosum. EP polymerases in humans are candidates for roles in immunoglobulin hypermutation; pols η and ζ are almost surely involved, and pol ι is a suspected participant. Recent data suggest that EP Y-family members are engaged in a variety of biochemical pathways in dividing and quiescent cells, which may mean that tolerance of DNA damage is often preferable to cell death. Normal replicative polymerases have evolved to copy DNA accurately by imposing active-site geometric constraints strongly favoring incorporation of Watson-Crick base pairs (6-8), and by proofreading base mispairs that occasionally slip thorough the geometric sieve (6). In contrast, none of the EP pols appear to contain 3'-exonuclease proofreading activity, and most importantly, recent crystallographic data suggest that the active cleft architectures of EP pols are much less restrictive, accommodating non-Watson-Crick pairs along with distorted primer/template DNA (p/t DNA) caused by the presence of damaged DNA bases (9–12).

SOS RESPONSE INDUCED BY DNA DAMAGE IN E. COLI

E. coli responds to DNA damage by calling upon a sizable number of genes contained in the SOS regulon (13–15). Forty-three SOS genes inducible by DNA damage are transcriptionally up-regulated (16) following cleavage of the LexA repressor protein mediated by a RecA nucleoprotein filament



Figure 2 UV induction of SOS in *E. coli*. Binding of the LexA repressor (yellow) to regulatory operators upstream of SOS genes limits their expression under normal growth conditions. RecA (blue) is induced shortly after ($<1 \min$) irradiation with UV, and becomes activated by binding to regions of single-stranded DNA to form a nucleoprotein filament, RecA* (blue helix). RecA* acts as a coprotease in the autocleavage of LexA, allowing induction of the SOS genes. In a similar reaction taking place on RecA*, UmuD (green) is cleaved between residues 24 and 25 of its amino-terminal end to form the mutagenically active carboxy-terminal fragment UmuD' (26a). Two molecules of UmuD' combine with one UmuC (red) to form pol V (UmuD₂C). *PolB*, encoding pol II, is expressed early ($<1 \min$ post UV) and is involved in error-free replication restart (see Figure 5). Pol V appears much later (\sim 45 min post UV); it copies persisting UV lesions to generate targeted SOS mutagenesis.

(Figure 2). Many of the SOS genes are used in BER, NER, recombinational repair, control of cell division, and translesion synthesis (TLS) (17). There is an ~100-fold increase in mutations targeted primarily at DNA damage sites following exposure of *E. coli* to UV or to chemicals that damage DNA (17). Although UV is commonly thought of as an intrinsic mutagen, UV-induced mutations will not occur in the absence of either *umuC* or *umuD'* (18–20); the prefix umu refers to UV mutagenesis. A heterotrimer composed of one UmuC bound to two UmuD' molecules (UmuD'₂C) (21, 22) is an error-prone DNA polymerase (23), *E. coli* pol V (24, 25). Pol V exhibits the correct in vivo mutagenic specificity when copying TT *cis-syn* photodimers and TT (6–4) photoproducts in vitro (26).



Figure 3 *E. coli* replisome and DNA replication as proposed by the trombone model. The τ dimer (green) links two pol III core molecules (blue), one for leading- and the other for lagging-strand synthesis. The coupled leading- and lagging-strand reactions are interrupted in the presence of DNA damage (indicated by a distortion immediately ahead of the polymerase core on the leading-strand track), presumably causing disassembly of the replication complex. EP pols such as pol V or pol IV then take over from the pol III core to synthesize past the damage site. Reconstitution of the replisome with the pol III core occurs following translesion synthesis. Shown as part of the replisome are the pol III core (composed of three subunits: α polymerase, ϵ exonuclease proofreading, and θ subunits); the γ complex required for loading the β -dimer sliding clamp onto DNA (five subunits); and the DnaB helicase. Not shown is SSB (single-stranded-DNA-binding protein), which coats ssdNA regions ahead of the replication fork. The pol III holoenzyme (HE) is composed of a pol III core + β/γ + τ . Lagging-strand RNA Okazaki fragment primers appear in red.

RecA is well-known for its two primary cellular roles, catalysis of DNA strand exchange during homologous recombination and initiation of the SOS mutagenic response (17). RecA also has a third, direct role in SOS mutagenesis revealed by the discovery of a RecA mutant that carries out both SOS induction and recombination but prevents UV mutagenesis (27–29). Biochemical data showing that TLS requires RecA and pol V (23, 26, 30–32) strongly support a direct role for RecA in mutagenic TLS through its interaction with pol V in the vicinity of DNA template damage.

THREE E. COLI DNA POLYMERASES INDUCED BY SOS

Three SOS genes encode DNA polymerases—pol II (33–35), pol IV (36), and pol V (23–25). Pol V and pol IV are charter members of the EP Y-family (Figure 1). Pol II is a high-fidelity B-family member (34, 37). The two major *E. coli* polymerases, pol I and pol III, are not under SOS control. The principal functions of pol I are to excise RNA primers while processing Okazaki fragments and to fill in short gaps during excision repair of DNA damage (38). Pol III carries out chromosomal replication as an integral component of the replication fork (Figure 3) (38).

A replication fork normally stalls when encountering a damaged DNA template base (Figure 3), causing an uncoupling of leading- and lagging-strand synthesis and the release of pol III core (38). Continued unwinding of the DNA ahead of the blocked replication fork could provide a region of single-stranded DNA (ssDNA), allowing assembly of an activated RecA nucleoprotein filament (RecA*) capable of inducing SOS (Figure 2). Gaining access to a wide variety of template lesions likely requires the combined action of EP and non-EP pols interacting with a slew of accessory proteins. Which polymerase is used and how it locates an intended target site are not understood. However, investigation into the mechanisms of SOS mutagenesis can be accomplished by using an in vitro reconstitution assay that measures pol V–catalyzed TLS.

Biochemical Basis of SOS Mutagenesis

SOS mutagenesis, also known as SOS induction of error-prone repair, was discovered in the mid-1970s by Witkin (14) and by Radman (13). SOS mutations are characterized as having base substitutions targeted directly opposite DNA template damage sites. For UV-induced mutations, the mutated sites correlate with adjacent pyrimidine bases. The two most common forms of UV damage are pyrimidine (6–4) pyrimidone photoproducts and cyclobutane dimers. An increase in SOS untargeted mutations is also found at sites not containing adjacent pyrimidines (17).

RECONSTITUTION OF SOS MUTAGENESIS IN VITRO Genetic data demonstrate that elevated mutational levels in cells exposed to UV light or to DNA-damaging chemicals require the presence of RecA, UmuC, and UmuD' proteins (14, 15, 39). UmuDC-dependent TLS was first studied by Echols and coworkers (40) by measuring the extension of a ³²P-labeled primer past a DNA template site containing a single abasic lesion. This earliest effort to reconstitute an in vitro biochemical TLS assay was undermined by the recalcitrant behavior of UmuC, which is insoluble in aqueous solution (21). To get around this difficulty, UmuC was purified as a denatured protein and then dialyzed into aqueous solution (21). This approach proved successful insofar as lesion bypass was detected (40), but the denaturation-renaturation procedure proved tenuous owing to the minuscule recovery of active UmuC (41).

We purified a soluble native UmuD'₂C protein complex using a plasmid to overexpress UmuC and UmuD' in the absence of chromosomal UmuC and UmuD (22), and observed UmuD'₂C-catalyzed TLS in the presence of RecA, ssDNA-binding protein (SSB), and β , γ complex (β/γ) (23). Unexpectedly, TLS occurred in the absence of pol III core, suggesting that UmuD'₂C was a new error-prone DNA polymerase (23). The observation that a mutant UmuD₂C104 (D101N) was unable to catalyze TLS activity proved conclusively that the UmuC subunit contains an intrinsic DNA polymerase activity (24). UmuD'₂C was subsequently designated as E. coli pol V (24). A different tack taken by Livneh and colleagues (30) used a recombinant UmuC protein linked at its N terminus to a maltose-binding protein (MBP). The MBP-UmuC protein was soluble in aqueous solution and catalyzed TLS on a gapped plasmid primer/template (p/t) DNA in the presence of pol III core, UmuD', RecA, and SSB. Although the pol III core was initially reported to be absolutely required for TLS (30), it was subsequently confirmed that MBP-tagged UmuC had polymerase activity and carried out TLS in conjunction with UmuD', RecA, and SSB in the absence of the pol III core (25).

Pol V Mut Catalysis of Error-Prone Translesion Synthesis

The term "replisome" refers to proteins assembled at the replication fork (Figure 3). By analogy, Echols coined the term "mutasome" to represent proteins assembled proximal to a template damage site (39, 42). In accordance with this suggestion, we use the term "pol V Mut" to mean pol V (UmuD'₂C) + RecA + SSB + β/γ (43). This designation is not meant to imply the existence of an actual physical complex involving any of these components either in the presence or absence of DNA.

A comparison of in vitro and in vivo data, using pol V Mut to copy a TT *cis-syn* photodimer, TT (6–4) photoproduct, and an abasic moiety supports a prominent role for pol V in UV-induced lesion-targeted mutagenesis (26). Each lesion is copied efficiently by pol V Mut, whereas synthesis by the pol III holoenzyme (HE) or pol IV $\pm \beta/\gamma$ is strongly inhibited (26). The nucleotide incorporation specificities for pol V Mut at each lesion site agree with in vivo data (44–47). The observation that pol V Mut favors misincorporation of G at the 3' T of the 6–4 photoproduct (26) agrees with in vivo data showing a T \rightarrow C transition hot spot at the 3'-T site (46, 47). In contrast, pol III HE and pol IV + β/γ weakly incorporate A at the 3'-T site (26).

By itself, pol V copies p/t DNA in a desultory manner. Synthesis is completely distributive (26)—less than one in a hundred p/t DNA encounters results in the incorporation of a single nucleotide. The presence of β/γ , RecA, and SSB stimulate pol V activity by 3-, 350-, and 1100-fold, respectively (31). RecA, SSB (48), and even pol V (22) bind avidly to ssDNA well removed from DNA damage sites. An experimental strategy aimed at confining mutasome-DNA

interactions to template sites proximal to a lesion uses short p/t oligomers (Figure 4*a*) in place of primed linear (23, 40) or gapped (30) plasmid DNA substrates as used previously.

A COWCATCHER MODEL FOR TRANSLESION SYNTHESIS RecA provides the key to understanding SOS mutagenesis. RecA filaments are assembled and disassembled in a 5' \rightarrow 3' direction on ssDNA in the presence of ATP (49, 50); the disassembly step requires ATP hydrolysis (51) (Figure 4*a*). RecA filaments form normally but disassemble much more slowly in the presence of ATP γ S, a slowly hydrolyzable ATP analog (48). The term "stabilized RecA filament" refers to filaments formed using ATP γ S. Although pol V–catalyzed TLS is stimulated in the presence of SSB, β/γ , or both, the pattern of synthesis is distributive when RecA filaments are assembled using ATP.

TLS patterns differ significantly with stabilized RecA filaments. SSB is now required both for synthesis and TLS, but the synthesis continues to be distributive in the absence of β/γ . However, a dramatic change takes place following the addition of β/γ , resulting in robust TLS accompanied by highly processive synthesis on the stabilized RecA filaments (31). The 35-Å opening in the β circular clamp is far too small to allow passage of a RecA nucleoprotein filament that has a diameter of 100 Å (48). Therefore, RecA must be removed from the template strand before replication can take place.

A mechanism to explain the removal of RecA is that pol V, acting in concert with SSB, strips RecA from the template track, in a manner loosely analogous to a locomotive cowcatcher (31) (Figure 4a). In this case, the advancing pol V + SSB facilitates filament disassembly in a $3' \rightarrow 5'$ direction on the DNA template, while maintaining contact with each next-to-be-removed RecA monomer located at the tip of the receding filament. Nuclease protection data support a $3' \rightarrow 5'$ disassembly process independent of ATP hydrolysis (31). The interaction of pol V with the 3' tip of the RecA filament is essential for TLS. Once contact with RecA is broken, pol V dissociates rapidly from the primer end and synthesis becomes distributive. At the other end (5' tip) of the filament, the 5' \rightarrow 3' disassembly reaction requires ATP hydrolysis and provides a means of eliminating the remaining downstream portion of the filament (Figure 4a), thus reducing the probability of pol V causing untargeted mutations downstream from the lesion. Therefore, bidirectional filament disassembly serves a dual role by facilitating pol V-catalyzed TLS while ensuring that mutations are primarily targeted at template damage sites.

A FLY IN THE FILAMENT OINTMENT That RecA protein is required for targeted mutagenesis by pol V Mut seems beyond doubt, but can the same be said for a RecA nucleoprotein filament? The $5' \rightarrow 3'$ filament assembly reaction offers a plausible way of targeting RecA to a lesion where it can then interact with pol V (Figure 4*a*). Filaments formed under conditions similar to those giving rise to TLS have been observed using electron microscopy (32, 52). Prior to the



Figure 4 Models of *E. coli* pol V catalysis of error-prone translesion synthesis. (*a*) A cowcatcher model involving a RecA nucleoprotein filament (RecA*). Replicative pol III stalls when encountering a template lesion (**X**), dissociates from the 3'-primer end, and is replaced by pol V. The activity and binding affinity of pol V are strongly stimulated by the presence of RecA, SSB, and β sliding clamp. The continued unwinding action of dnaB helicase (Figure 3) allows formation of a RecA nucleoprotein filament ahead of the lesion. The filament assembles in a 5' \rightarrow 3' direction on ssDNA, advancing to the DNA damage site. Pol V + SSB operate as a locomotive cowcatcher to strip RecA from the DNA template in a 3' \rightarrow 5' direction immediately ahead of an advancing pol V molecule. The cowcatcher stripping reaction does not involve ATP hydrolysis and takes place concurrently with the "standard" 5' \rightarrow 3' filament disassembly reaction requiring ATP hydrolysis. The p/t DNA is composed of a 30-nucleotide (nt) primer annealed to a 120-nt template. (*b*) Translesion synthesis requires the presence of RecA but not a RecA nucleoprotein filament. Pol V-cata-

27

discovery of pol V, Devoret and colleagues (53, 54) suggested that $UmuD'_2C$ binds to the 3' tip of a RecA filament adjacent to a DNA damage site to assist with lesion bypass. An observation consistent with this model is that pol V binds preferentially at RecA filament ends, as visualized by electron microscopy (55). TLS occurs at highest efficiency with a RecA/DNA nucleotide ratio of 1:5 in the short p/t DNA system (31) (Figure 4*a*), which is close to the stoichiometry of RecA binding to ssDNA (1 RecA monomer per 3 nucleotides) (48).

But what if TLS takes place under conditions where RecA filaments are unlikely to form? Cox has pointed out to us that with substoichiometric concentrations of RecA (1:50 nucleotides ssDNA), the likelihood of nucleoprotein filament formation on a short p/t DNA oligomer, even in the presence of ATP γ S, is rather low (Figure 4*a*). And yet pol V Mut–catalyzed TLS occurs under these conditions with an efficiency that is only about twofold lower than optimum (31). So perhaps a RecA–pol V complex (with help from SSB and β/γ) copies UV-damaged DNA, whether or not a RecA filament is present.

An experiment was carried out to determine if pol V in the presence of RecA and ATP γ S performs TLS in a short gap-filling reaction (Figure 4*b*). The answer is yes, when copying a gap as short as 3 nucleotides (nt) containing a central abasic lesion. However, TLS was not observed when the gap was shortened to 2 nt (P. Pham, S. Saveliev, M. Cox & M. F. Goodman, unpublished information). The 3-nt gap with an abasic lesion is not filled in unless pol V and RecA are both present. Because the binding site size for a RecA monomer is 3 nt, there is not much room for a filament to form on ssDNA, leaving aside the presence of pol V (a 72-kilodalton protein complex) bound to the 3'-end of the primer strand. A similar observation has been made when pol V replicates a short 5'-template overhang in the presence of RecA and ATP γ S — TLS occurs but only when the lesion is located three or more nucleotides from the 5'-end of the template strand (P. Pham, S. Saveliev, M. Cox & M. F. Goodman, unpublished information).

Does this experiment rule out a requirement for a Rec A nucleoprotein filament? Almost certainly. Does it demonstrate the presence of a RecA–pol V complex? Decidedly not. As the truism states, the devil is in the details, and experiments using the fluorescence reporter molecule 2-aminopurine are under way to investigate the details of the gap-filling reaction. If a RecA monomer were to straddle a lesion with pol V bound to the 3'-primer end, that would not be very different from Devoret's model with RecA at the tip of the filament (53, 54). A RecA monomer could, for all intents and purposes, serve the same function as a

lyzed TLS taking place within a 3-nt gap requires the presence of RecA and ATP γ S, but not SSB. The TLS gap-filling reaction does not occur within a 2-nt gap, although the 2-nt gap can be filled in the absence of a lesion by pol V alone. The p/t DNA is composed of a 120-nt template annealed to an oligonucleotide primer and "down-stream" oligonucleotide, forming a gap of either 3 or 2 nt.



Figure 5 Model of error-free replication restart involving *E. coli* pol II. Replicative DNA synthesis is blocked by DNA template damage (\mathbf{X}). Pol III dissociates from the damaged primer-template strand while synthesis continues on the undamaged strand. Regression of the replication fork occurs in the presence of RecF, O, and R proteins, thereby providing an undamaged template strand for pol II to copy that contains the correct coding information at the DNA damage site. RecG is required for progression of the fork past the lesion site and PriA is then involved in reconstituting the replication fork.

3' RecA tip in Devoret's filament model. If a RecA nucleoprotein filament were in fact to form on an extended ssDNA region downstream of a lesion, pol V + SSB could then act to disassemble the filament in accordance with the cowcatcher idea (Figure 4a).

A Pivotal Role for Pol II in Error-Free Replication Restart

Pol II, an orphan enzyme since its discovery in 1970 (56, 57), is just now beginning to see the light on the stage of *E. coli* replication. This B-family polymerase, harboring an active 3' exonuclease (58), synthesizes DNA accurately (59). Pol II is induced sevenfold in response to UV damage (33–35),

29

increasing from about 50 to 350 molecules (60), yet cells lacking pol II suffer no adverse consequences except when pol V is also missing (61). Double mutants of pols II and V exhibit increased UV sensitivity compared to cells lacking pol V alone (61). Pol II is induced within about 1 min after exposure to UV, but induction of pol V is delayed for about 45 min post UV (62) (Figure 2). The replication fork is blocked in the presence of UV damage, and DNA synthesis remains suppressed in the absence of pol II, until roughly 45 min later when pol V–catalyzed TLS occurs (61). A roughly similar 40-min delay in post-UV DNA synthesis occurs in pol II⁺ cells in the absence of either RecF, RecO, RecR, or PriA (63–65), proteins known to be involved in rescuing replication forks in UV-irradiated *E. coli* (63, 66).

A putative sequence of events giving rise to replication restart on a blocked leading strand (67, 68) are an uncoupling of leading- and lagging-strand synthesis, which generates regions of ssDNA; replacement of SSB-coated DNA by a RecA nucleoprotein filament mediated by RecOR proteins and stabilized by the RecFR complex (69–71); and induction of SOS, which turns on pols II and V (Figure 2). A collapsed fork undergoes a RecA nucleoprotein–mediated regression (72), forming a so-called chicken-foot structure (73) in which the uncoupled nascent lagging strand provides a template for synthesis by pol II, which then copies the correct information, avoiding TLS (Figure 5). After synthesis by pol II, RecG-dependent fork regression occurs in the opposite direction (67), followed by reestablishment of a bona fide replication fork using the PriA-dependent primosomal complex to load pol III HE (3). The upshot is that the lesion, which remains in the double-strand DNA, is bypassed accurately. Thus, pols II and V appear as opposite sides of a coin: Pol II plays a pivotal role during error-free replication restart (61), and pol V is responsible for error-prone TLS.

Translesion Synthesis with Pol IV and Pol II

Which of the three SOS polymerases replaces a displaced pol III core at the site of a replication-blocking lesion? That choice depends on the timing, polymerase availability, and the specific nature of the DNA damage encountered. The ability to gain access to a lesion is facilitated by the binding of pols II (74–76), IV (26, 77), and V (26) to the β processivity clamp.

Although pol IV fails to copy cyclobutane dimers or 6-4 photoproducts (26) and has no discernible effect on UV-induced mutagenesis (17), it is clearly involved in copying the bulky template adduct benzo(*a*)pyrene diol epoxide (BaP DE) (78, 78a). Both pol IV and pol V are required in order to carry out error-free and -1 frameshift TLS at a BaP DE adduct (78, 78a). Pol II is also used for TLS, albeit sparingly—acting in place of pol V, it copies abasic lesions when SOS is turned on in the absence of induction of the GroELS heat shock proteins (79). In a seemingly bizarre twist, the "high-fidelity" pol II is responsible for generating -2 frameshifts during TLS of *N*-2-acetylaminofluorene (AAF) guanine adducts, and EP pol V is responsible for error-free AAF bypass (78). This complex state of affairs all goes to demonstrate the likelihood that it is the



Figure 6 Biochemical properties of the EP pols. (*a*) Error-prone translesion synthesis (TLS) by *E. coli* pol V results in misincorporation of G opposite the 3' T of a TT (6–4) photoproduct, leading to A \rightarrow G transition mutations. (*b*) Misaligned primer-template ends are extended efficiently by *E. coli* pol IV, leading to frameshift mutations. (*c*) The DNA-dependent dCMP transferase activity of Rev1 protein incorporates C opposite an abasic template site. (*d*) Pol ζ , a B-family pol, efficiently incorporates two A nucleotides opposite a TT (6–4) photoproduct in vitro, resulting in error-free bypass of the lesion, dependent on the presence of Rev1 protein. (*e*) DNA polymerase η catalyzes error-free replication across a TT *cis-syn* photodimer by incorporating two A nucleotides, thereby avoiding mutation and offering protec-

31

rule, not the exception, that both EP and non-EP repair polymerases play seemingly disparate roles in the cell, sometimes causing mutations, oft-times not.

Pol IV Generates Untargeted and Adaptive Mutations

Replication forks are routinely inactivated during aerobic growth, even in the absence of DNA damage, perhaps as often as once per round of replication (3). For example, pol III may stall following insertion of a nucleotide on a transiently misaligned 3'-primer end (8). In the event that an ensuing slipped-base mispair is refractory to proofreading, pol IV may then be called upon to rescue a stalled replication fork by extending an aberrant primer end (Figure 6b). Extension at mismatched primer ends generates small untargeted frameshift mutations, and these have been attributed to the action of pol IV in vivo (80, 81). There is certainly plenty of pol IV present constitutively in E. coli to help rescue stalled replication forks—250 molecules per cell, which increases by 10-fold following SOS induction (82). Extension at mismatched primer ends is a reaction favored by pol IV in vitro (M. Valentine & M. F. Goodman, unpublished information). Rescue of stalled replication forks is also crucial in eukaryotic cells. This critical housekeeping function might be the primary raison d'etre of pol IV and is perhaps the reason why homologs of this enzyme appear in all organisms investigated to date (2). Pol V, on the other hand, has been identified only in prokaryotes (2).

Adaptive mutation is a process in which nonproliferating microbial populations generate mutations when placed under nonlethal selective pressure (83). Since microbes spend much of their time attempting to cope in hostile environments, adaptive mutations may play an important role in survival. And here again pol IV comes into play, while engaged in a mutational balancing act with pol II. Adaptive mutation rates, which increase by threefold in the absence of pol II (84), are attributed almost exclusively to pol IV (85, 86). Pol IV is responsible for about 85% of the *lacZ* adaptive frameshift mutations occurring on a plasmid in wild-type cells, and essentially all of the increased frameshifts in the absence of

tion against skin cancer; pol η is also responsible for error-prone incorporation within TAA motifs that generates mutations in the variable region of immunoglobulin geness (see Figure 7). (*f*) Pol ι misincorporates G in preference to A opposite T during a gap-filling reaction. This activity may function to avoid incorporating A opposite a template T generated by deamination of 5-methyl C at CpG sites (p refers to the 3'-5' phosphodiester bond linking C to G in the DNA strand), thus protecting against G \rightarrow T transitions. Pol ι may also be involved in generating mutations at G and C in RGYW hot spots during somatic hypermutation of immunoglobulin genes (see Figure 7). (*g*) DNA polymerase κ is involved in generating small lesion-targeted deletions, possibly by addition of a nucleotide using a transiently misaligned primer end. An extrahelical template lesion is denoted by the symbol X.

pol II (85, 86). The active 3'-exonuclease activity of pol II is responsible for keeping adaptive mutations somewhat in check because mutations are increased about fivefold in a pol II proofreading-deficient background (84).

EUKARYOTIC ERROR-PRONE DNA POLYMERASES

The expanding eukaryotic polymerase universe has yet to reach steady state (1). Biological roles can be assigned to several of the new polymerases: protection against skin cancer, pol η (4, 5); TLS, Rev1, pol ζ (87); sister chromatid cohesion, Trf4 (88); somatic hypermutation, pol η (89, 90), pol ζ (91), and possibly pol ι (92). On the basis of their presence in numerous different tissues, the likelihood is that many of the eukaryotic EP pols are involved in multiple DNA repair pathways in either a primary or a backup capacity, as observed for the *E. coli* SOS pols. A recently proposed revised nomenclature deals with the inevitable contradictory assignments accompanying the rapid rate of discovery of new eukaryotic polymerases (92a). In this section, we offer a current synopsis of the principal properties of the new eukaryotic EP pols (Figure 6).

Rev1 and Pol ζ

In Saccharomyces cerevisiae, the RAD6 epistasis group encodes genes that are involved in translesion synthesis (TLS) and spontaneous mutatagenesis. Three of these genes are known as REV1, REV3, and REV7 (87, 93, 94). Rev1 protein, the first-recognized Y-family member, acts as a deoxycytidyl transferase that incorporates dCMP opposite abasic sites (95) in yeast and humans (96) (Figure 6c). DNA pol ζ cooperates with Rev1 to accomplish TLS past abasic sites, with pol ζ extending from the mispaired C opposite the abasic site (95). Rev1 exhibits a second property in addition to deoxycytidyl transferase activity. It acts in combination with pol ζ to achieve predominantly error-free TLS past 6–4 TT photoproducts in vivo (97) (Figure 6d). A yeast REV1 Δ strain and a REV1–1 Δ strain (retaining 60% of deoxycytidyl transferase activity) exhibit a reduced ability to bypass both abasic sites and 6–4 TT photoproducts (97).

DNA pol ζ , a B-family polymerase (98), is composed of two subunits, Rev3 and Rev7 (99). The principal mutagenic role of pol ζ appears to be related to its remarkable promiscuity in extending mispaired primer ends. Rev3 serves as the catalytic subunit, but the function of Rev7 is unknown. Homologs of Rev1 and Rev3 have been found in human (100, 101), mouse (102), and *Drosophila* (103) cells, and putative homologs of Rev7 exist in human (104) and *Drosophila* (105) cells. In mice, disruption of REV3 confers embryonic lethality, which suggests a critical role for pol ζ during development (106–108).

In yeast, pol ζ is responsible for 50–70% of spontaneous mutations (105). Typical efficiencies of pol ζ for mismatch extension are $\sim 10^{-1}$ to 10^{-2} (109). When extending from a correctly paired primer terminus, pol ζ copies DNA

33

accurately, making base substitution errors in about 1 in $\sim 10^{-4}$ to 10^{-5} cases (109), a frequency comparable to high-fidelity polymerases lacking proofreading capability. Along with its ability to extend natural mismatched base mispairs, pol ζ alone can also carry out weak TLS when confronting TT *cis-syn* dimers (99) on its own. However, when in the presence of human pol ι , pol ζ was shown to bypass 6–4 TT photoproducts and abasic sites efficiently, probably by extending pol ι –catalyzed mismatches (109).

Pol η

Pol η homologs have been found in mouse, human (4, 5, 110), yeast (111, 112) and *Drosophila* (113) cells. The yeast *RAD30* gene encoding pol η also belongs to the *RAD6* epistasis group. In humans, pol η is encoded by the *XPV* gene, which if mutated induces a variant form of xeroderma pigmentosum (XP-V). XP-V individuals are UV sensitive and susceptible to a high incidence of skin cancer (110), emanating from the loss of pol η mediation of error-free TLS past UV damage (Figure 6*e*). In yeast, pol η is also responsible for suppressing UV mutations by copying TC (6–4) and CC (6–4) photoproducts accurately—the incidence of mutations is about fivefold higher in yeast strains lacking this EP pol (114).

Yeast and human pol η are extremely low fidelity polymerases, lack exonuclease activity, and have a misincorporation frequency of $\sim 10^{-1}$ to 10^{-3} on undamaged DNA (112, 115, 116). Yet both yeast and human pol η are able to bypass several bulky DNA lesions with relatively high fidelity, e.g., cisplatin G-G intrastrand cross-links (117), acetylaminofluorene-dG (117), 8-oxodeoxy-guanosine (118), and TT *cis-syn* dimers (119). In contrast, pol η is somewhat error-prone when bypassing O⁶-methylguanine lesions by incorporating either T or C residues (120). Once pol η incorporates an incorrect nucleotide, it tends to dissociate from the DNA. The pol η mismatch extension frequency is $\sim 10^{-2}$ to 10^{-3} (121).

Within the Y-family polymerases (2), the N-terminal region has five highly conserved motifs, I–V (Figure 1). The C-terminal region, however, is unique for each family member. Three highly conserved acidic amino acids essential for polymerase activity are located within motifs I and III (Figure 1). Yeast pol η activity is abolished when any one of these three, D30 or E39 in I or D155 in III, is replaced by an alanine. Presumably, these amino acids are involved with the interactions between divalent metal ions and the incoming deoxynucleoside triphosphate (dNTP) (122) (e.g., see Figure 1).

Deletion of the C terminus of pol η does not abolish polymerase activity in vitro, but the truncated yeast pol η is incapable of restoring UV resistance in $RAD30\Delta$ strains (122), arguing for the importance of the C terminus for interactions that may help target pol η activity. The last 100 C-terminal amino acids of human pol η are both necessary and sufficient for human pol η to localize in the nucleus and form foci upon exposure to UV radiation or carcinogens (123). Consistent with the presence of a bipartite nuclear localization

sequence present in the C terminus of both yeast and human pol η , C-terminal truncations of pol η fail to complement bypass deficiencies in XP-V cells (123). The C-terminal region of pol η (Figure 1) contains a C₂H₂ zinc finger motif, which may be required for targeting pol η to repair foci following DNA damage (123). It also contains proliferating cell nuclear antigen (PCNA) interaction sites (123), which allow PCNA to stimulate pol η activity in the presence of RFC and RPA (human single-stranded binding protein) (123a). Interaction with PCNA is essential for pol η function in yeast (123b).

Pol ı

Pol ι is one of two human homologs of the yeast *RAD30* gene (124). Other homologs of pol ι have been found in mouse and *Drosophila* cells, but not in yeast or other lower eukaryotes. Its function in vivo has not yet been determined. Pol ι is highly expressed in the testis, and ubiquitously expressed throughout all human tissues, with slightly higher levels in the heart and pancreas (125). Pol ι carries out low-processivity DNA synthesis, typically incorporating 1–3 nt (126). However, it is more active using gapped DNA substrates, synthesizing 7–10 nt with limited strand displacement (92, 125).

A common feature of DNA polymerases, even those that are highly error-prone, is that they still favor making Watson-Crick base pairs. But that is not true for pol ι . A truly unique feature of this EP pol is that it prefers making dGMP·T wobble mispairs rather than dAMP·T base pairs, by factors of 3- to 10-fold (109, 126, 127). It also makes T·T mispairs roughly 70% as well as A·T pairs (126). The base substitution fidelity is $\sim 10^{-2}$ for incorporation opposite template G and C, and $\sim 10^{-4}$ for incorporation opposite A (126). Remarkably, after extending a dGMP·T mispair, pol ι switches specificity to form a next correct dAMP·T pair (128, 129). Although pol ι is able to extend from all 12 possible mismatches, it is considerably less proficient at extending mismatched base pairs compared with pol ζ (109, 129). Pol ι copies abasic moieties and AAF adducts (130), but unlike pol η , it cannot copy past TT *cis-syn* dimers and TT (6-4) photoproducts—it is capable of incorporating an A opposite the 3' T of a 6-4 photoproduct, but then it dissociates (109).

Owing to an associated 5'-deoxyribose phosphate lyase activity, pol ι may be involved in base excision repair (BER) (128). It carries out BER reactions in vitro in the presence of uracil glycosylase, apurinic (AP) endonuclease, and DNA ligase I (128). A possible scenario during BER involves pol ι incorporating G opposite a template T that had been generated by deamination of 5-methyl C, thus protecting against G to T transitions (128, 130a).

Human Pol *k*

Pol κ , encoded by the *HDINB1* gene (131), is the human homolog of *E. coli* pol IV, and is capable of TLS. An abasic site is dealt with by pol κ through a

frameshift mechanism that uses the base downstream of the lesion as a template (132) (Figure 6g). This polymerase readily catalyzes extension of misaligned undamaged primer termini, resulting in -1 frameshift mutations (133). Two zinc fingers located in the C terminus of the enzyme are involved in pol κ processivity (Figure 1), which is reduced from ~25 nt to 1–2 nt when these domains are deleted from the protein (134). Pol κ bypasses benzopyrene G adducts in an error-free manner and copies AAF-modified G adducts by incorporation of either dCMP (error-free) or dAMP (error-prone) (132, 135). 8-oxo-dG and $1,N^6$ -ethenodeoxyadenosine are bypassed with low fidelity (132, 136). Pol κ is unable to bypass a cisplatin adduct, a TT dimer, or a TT (6–4) photoproduct (132, 137).

Misincorporation rates on the order of 10^{-3} to 10^{-4} have been reported for pol κ on undamaged DNA (133). Similar to pol η , pol κ appears to tolerate only certain types of DNA damage. There is evidence that pol κ may be up-regulated in lung tumors (138). The precise roles of pol κ in humans remain a mystery, but by analogy with its *E. coli* pol IV homolog, its primary functions may involve DNA damage tolerance and relief of stalled replication forks on undamaged DNA (3).

Trf4/pol σ

Trf4 protein (formerly called yeast pol κ , recently renamed yeast pol σ) (139), and its close homolog Trf5, contain highly conserved motifs loosely related to the nucleotidyl transferase domains of the β -like DNA X-family of polymerases (140). A His-tagged Trf4 protein has an intrinsic DNA polymerase activity that is relatively processive in the presence of high concentrations of deoxynucleoside triphosphate (dNTP) and is sensitive to dideoxynucleotides (88). Genetic studies indicate a requirement for pol σ during mitotic chromosome segregation (141) and a physical interaction with Smc1 (141), a protein involved in sister chromatid cohesion. Using fluorescence in situ hybridization (FISH) studies, TRF4 mutants were shown to have dramatic defects in sister chromatid cohesion both near centromeres and on chromosome arms (88). A double mutant of TRF4-ts/TRF5 is unable to completely replicate its genome and exhibits marked G_1/S transition delays. It has been suggested that pol σ /Trf4 and Trf5 work together to replicate the chromosome at cohesion sites that might otherwise fail to maintain cohesion if the pol δ - or pol ϵ -driven replication fork passes through these regions by switching between replicative and EP pols (139). PCNA and a modified replication factor C have also been implicated in the establishment of cohesion (142–144), but it remains to be seen if PCNA and pol σ /Trf4 orTrf5 interact. Trf5 protein has not been tested for polymerase activity.

Pol μ

Pol μ , closely related to terminal deoxynucleotidyl transferase (TdT), an X-family member, contains a BRCT (BRCA1 C-terminal) domain, and is expressed predominantly in lymphoid tissues including the thymus and lymph nodes (145,

146). Unlike TdT, pol μ acts as a partially processive template-directed DNA polymerase (146). Its transferase and polymerase activities are stimulated when Mn^{2+} replaces Mg^{2+} in vitro, but the relevance of this finding to the in vivo situation is unknown. Mn^{2+} is also reported to reduce pol μ fidelity in vitro (146). The tendency of pol μ in vitro is to make -1 frameshift errors in repeat sequences (146a).

Many mRNA splice variants are present at high levels for the *POLM* gene (encoding pol μ), 90% of which do not encode functional protein, which may reflect some form of regulation by alternative splicing (145). Splicing inhibition of pol μ mRNA occurs in response to DNA-damaging agents such as UV light, γ rays, and H₂O₂, possibly to prevent pol μ from acting on specific types of damage-induced lesions (145).

Pol λ

Another recent addition to the X-family of polymerases is pol λ . This polymerase is the closest homolog to pol β known (λ was designated originally as β 2), but it also contains an additional BRCT domain, absent in pol β , that is dispensable for polymerase function (147). Pol λ is expressed at very high levels in the testis and fetal liver and is present ubiquitously at low levels elsewhere. Pol λ is weakly processive and lacks detectable $3' \rightarrow 5'$ exonuclease activity (145). Pol λ contains an intrinisic deoxyribophosphate (dRP) lyase that can substitute for pol β in a BER reaction reconstituted in vitro (148). This activity depends largely on Lys³¹⁰, which when mutated eliminates 90% of the wild-type dRP lyase activity, suggesting this residue acts as the main nucleophile in the β -elimination reaction. Pol λ may play an active role in BER during spermatogenesis, and might be specifically recruited to this pathway via its BRCT domain. Pol λ exhibits limited strand displacement on a gapped substrate, perhaps allowing it to function in "long patch" BER as well (148), where typically two to fifteen nucleotides are excised and subsequently resynthesized in a gap (148a). Lyase activities have also been reported for pol ι (128) and for the mitochondrial pol γ (149, 150). Together, these polymerases, along with pol β , may each process different lesions during BER at unique locations within a cell or tissue type.

SOMATIC HYPERMUTATION

Behaving almost as if evolution were occurring on a time scale of days not millennia, a remarkably diverse group of antibodies is synthesized in higher eukaryotes to combat against antigenic invasion. Initially, low-affinity antibodies are generated in B cells by recombinational rearrangement of V, D, and J regions within the immunoglobulin genes. High-affinity antibodies are produced shortly thereafter in germinal center B cells (151). The synthesis of diverse numbers of antibodies emanates from mutations targeted to the variable (V) regions of



Figure 7 Somatic hypermutation break-repair model. Interactions between enhancerer-binding proteins and transcription-associated factors (TAF) bound at the promoter (P) are shown linking the enhancer (E) to the promoter to form a transcriptionally competent DNA structure. Single-strand nicks or double-strand breaks have been identified in the variable region of immunoglobulin genes proximal to RGYW mutational hot-spot motifs, caused presumably by an as yet unidentified endonuclease (Endo). The DNA nicks or breaks may be substrates for one or more EP pols to bind and generate mutations. Pol η is responsible for mutating A and T sites, primarily within TAA motifs (boldface indicates the favored mutational target). A different EP pol, possibly pol ι , is responsible for mutating G and C sites in RGYW motifs with G as the favored mutational target (R is A or G, Y is T or C, W is A or T). Pol ζ (not shown) is also involved in SHM, perhaps to extend mismatches made by pol η and by the second EP pol. MAR designates a matrix attachment region. The constant region and the 3' enhancer region of the immunoglobulin gene are not shown.

immunoglobulin genes. Because the rate of V-region mutations, $\sim 10^{-3}$ per base pair in one generation, is roughly a million times greater than normal somatic mutation, the process has come to be known as somatic hypermutation (SHM) (152).

A promoter immediately upstream of the V region and two distal downstream enhancer elements regulate the mutational process (153) (Figure 7). The mutations are concentrated within the V region of the immunoglobulin gene for a distance of about 1500 nt downstream of the promoter (154, 155). A few mutations also occur within the 5'-leader portion of the promoter. Although a promoter, target gene, and enhancers must all be present for mutations to occur, B-cell-specific elements are not required (156–158). The B-cell promoter and V region are fully replaceable using different promoter and target regions in cultured cells undergoing SHM (159).

Although the immunoglobulin genes are the dominant natural target for SHM in normal B cells, other genes can be targeted albeit at much lower mutation frequencies. *BCL-6* and *CD95* can mutate in germinal centers as a by-product of SHM (160–163), and recently, several proto-oncogenes have been identified as targets in diffuse large-cell lymphomas (158). This observation indicates that the SHM mechanism can lose control over its ability to target antibody V regions.

Perhaps the most distinctive hallmark of SHM is the nonrandom nature of the mutations (152). Approximately 20–50% of the mutations are targeted at RGYW motifs (R is A or G; Y is C or T; W is A or T), and TAA sites also show enhanced error rates (the "hottest" site is indicated in boldface in each sequence) (164, 165) (Figure 7). SHM is further characterized by an excess of transitions over transversions; A mutates considerably more often than T. Notably, the mutations that occur in nonimmunoglobulin genes are largely transitions favoring RGYW motifs and are limited to about 2 kilobases downstream from their respective promoters (158).

Most current models for SHM invoke a role for transcription based on the need for promoter and enhancer elements. An early model suggested a modified form of transcription-coupled repair in which an amplification in errors could arise if a normal DNA polymerase were to copy a short stretch of V-region DNA repeatedly, perhaps many thousands of times (166, 167). However, this multiple-pass mechanism was suggested prior to the discovery of EP polymerases. The discovery of the new errant polymerases suggests a way to mutate immunoglobulin genes when copying the V-region target just once, but the requirement for transcription-like DNA architecture remains a key element in mutational targeting (Figure 7).

Error-Prone Polymerases as Somatic Hypermutation Generators

Making errors at unprecedented rates, EP pols clearly satisfy the main SHM criterion. When copying undamaged DNA, base substitution rates of 10^{-1} to 10^{-3} are often observed for eukaryotic EP pols (109, 115, 116, 126, 134), leaving no shortage of potential mutator candidates. In fact, recent evidence suggests that there could be at least two, and perhaps even more, polymerases engaged in SHM, each providing a unique mutational signature (168–170).

POL η : A SOMATIC A-T MUTATOR Encoded by the human *XPV* gene, pol η keeps skin cancer at bay by copying UV-damaged DNA accurately. And yet its accuracy is remarkably poor when copying undamaged DNA, making errors at a rate of $\sim 10^{-1}$ in one study (116) and $\sim 10^{-2}$ to 10^{-3} in another (115). Although no obvious loss of immune function is observed in xeroderma pigmentosum patients lacking pol η , Gearhart and colleagues (89) measured a reduction in A and T mutations in variable genes obtained from the peripheral blood lymphocytes of XP patients. These data show that not only is pol η likely to be responsible for mutating A and T sites during SHM, but a second EP pol having G-C-mutator specificity is also probably involved. Another G-C mutator, Burkitt's lymphoma cell line CL-10, showed down-regulation of pol η following

SHM stimuli, supporting the notion that pol η contributes to A-T mutations in vivo (91).

A statistical analysis of multiple unselected somatic mutations in immunoglobulin loci from a variety of species confirmed the identity of RGYW and WA hot-spot motifs (90). Compared to SHM spectral data, a mutational spectrum obtained using pol η to copy the *lacZ* gene in vitro showed that almost all of the pol η hot spots are found in WA motifs throughout *lacZ* (90), thus reinforcing the idea that pol η is the principal A-T mutator during SHM and that another polymerase is responsible for causing mutations in the RGYW hot-spot motif.

The in vitro mutational data with *lacZ* show that pol η favors formation of dGMP·T mispairs immediately following an A-T base pair on the nontranscribed strand (90). But the targeting of mutations at WA motifs on the nontranscribed strand runs counter to the evidence of double-strand breaks (DSBs) in this region (170). However, mutational asymmetry could result from nicks, rather than DSBs, introduced into the nontranscribed strand acting as foci for pol η binding. An enrichment in ssDNA nicks in V regions was reported recently in cells undergoing SHM, suggesting the possibility for a mutational mechanism of breaks followed by error-prone repair (171) (Figure 7). This assay, designed to detect both nicks and DSBs, showed that many more nicks occurred than breaks (171). No strand bias has been reported for G-C mutations associated with RGYW motifs, arguing for an independent mechanism of nick and repair at these hot spots. Whether or not nicks, DSBs, or both are involved, the mechanism of SHM targeting is a challenging question.

SEEKING A SOMATIC G-C MUTATOR Although several EP pols could stand in as G-C mutator candidates, one early favorite, pol μ , may no longer be in the running. A mutant mouse homozygous for a *POLM* knockout exhibits both a normal immune response and mutational spectrum (172, 196). It seems likely, therefore, that pol μ , despite much early promise based in part on its preferential expression in germinal center lymphocytes (145, 146), is not involved in SHM. Pol β , a close homolog to pol μ , has also been tested for involvement in SHM. In this study, the immunological systems of irradiated mice were reconstituted with DNA pol β -deficient liver cells, and the mice mounted a normal immune response with no associated changes in mutation spectra (173).

Pol ι remains in contention as a G-C mutator. Cultured Burkitt's lymphoma BL2 cells show a 5- to 10-fold increase in heavy-chain V-region mutations targeted mainly to RGYW sequences (92, 174). The increased mutagenesis occurs only when BL2 cells are cocultured with human T cells and antigenically treated with anti-immunoglobulin M (anti-IgM) antibody to mimic antigenic challenge (92, 174). This increased mutation rate was accompanied by a 4-fold increase in levels of pol ι mRNA within 12 h of costimulation (92). mRNA levels for pols η , κ , ζ , and Rev1 also fluctuate in BL2; however, these changes do not correlate with the coculture requirements to observe SHM in the cell line (92). Slightly elevated levels of pol ι transcript were found in activated B cells from

XP-V patients (89), while roughly constant levels of the pol ι transcript were observed in a different Burkitt's lymphoma cell line, CL10, regardless of SHM stimulation (91). Despite the differences in levels of pol ι induction, the important point is that all of these polymerases appear to be present in the cell to some degree, prior to and during SHM triggering events.

Pol ζ is also a viable SHM candidate. Pol ζ , pol η , pol ι , pol κ , and pol μ , along with replicative pol α , pol δ , and pol ϵ , are expressed constitutively in cultured Burkitt's lymphoma CL10 cells after 12 h, based on RT (reverse transcriptase)-PCR mRNA analysis (91). In this cell line, pol ζ is up-regulated following costimulation with T cells and anti-B-cell receptor, while pol η is down-regulated concomitantly, in a dose-dependent manner with respect to the level of B-cell-receptor antibody (91). Antisense inhibition of the catalytic subunit Rev3 of pol ζ reduces the frequency of somatic mutation without affecting cell cycle or cell viability, but causes a slight delay in the generation of high-affinity antibodies (91, 175). Although pol ζ is not a Y-family polymerase, nor does it make base substitution errors nearly as facilely as some Y-family members, it nevertheless extends mismatched primer ends with relative ease (99, 109). By analogy to its role in yeast, where pol ζ acts in conjunction with Rev1 to catalyze TLS (Figure 6*d*), perhaps it is used during SHM to extend mismatches formed by pol η and perhaps by pol ι .

Mismatch repair proteins play a role in the SHM process, causing relatively small alterations in the overall spectra after clonal selection has taken place (176, 177). The spectra become altered in such a way that a strong G-C bias is observed in mice that are mutant for mismatch repair proteins (164, 177, 178) MSH2 (164) or MSH6 (178), but this effect is dependent upon whether the mice are receiving a primary or chronic antigenic stimulation.

Another important player in the pursuit of an SHM mechanism is the putative mRNA editing enzyme AID (179, 180). APOBEC-1, a structural homolog of AID, acts as a site-specific deaminase that converts dCMP to dUMP in mRNA coding for apolipoprotein B, resulting in a shorter protein with an altered physiological function (181, 182). Patients with defects in both alleles for AID exhibit type II hyper-IgM syndrome (180) and accumulate excess levels of IgM antibodies. Class switching and SHM can occur independently of one another, but both steps are required to cause positively selected IgM antibodies to undergo affinity maturation and convert to IgG. AID is required for both events to take place (179). AID appears to play a role upstream of both SHM and class switching, but the target transcript upon which it may act is currently unknown.

FUTURE PERSPECTIVE

Although the first two members of the error-prone Y-family polymerases were discovered in 1996 (95) and 1998 (23), with many other EP pols identified in 1999 (1,183), considerable progress has been made in determining roles for

these and other recently discovered polymerases. Aided by an impressive body of genetic data from the mid-1970s to the present, the two new EP pols in *E. coli* have found their niche—pol V in UV mutagenesis and pol IV in adaptive mutation and chemical mutagenesis. Significant progress has also been made in identifying roles for each of the four eukaryotic Y-family pols (2) plus a smattering of other new family B, X, and A members (98). Notably, a recently discovered B-family DNA polymerase, pol ϕ , is essential for viability in *S. cerevisiae* and appears to play a role in ribosomal RNA synthesis (183a). All of this raises questions of trafficking—how one polymerase can substitute for another—and of targeting—how a chosen polymerase winds up going where it is supposed to go.

DNA Repair Factory

Leading- or lagging-strand DNA damage may block replication fork progression until the lesion is either repaired (BER, NER), avoided (replication restart; Figure 5), or copied (TLS; Figure 4 and Figure 6). The bottom line is that the catalytic subunit of a polymerase holoenzyme must be replaced by another polymerase, from which there are many to choose. The choice of which polymerase to swap with another is determined by the specific type of template damage encountered (Figure 6).

A newly emerging and rapidly growing catalog of protein-protein interactions might offer hints as to how polymerase swapping could in principle occur. For example, direct interactions have been observed between the *E. coli* "sloppier copier" pol V and the α pol, ϵ exonuclease, and β -dimer clamp subunits of the "fastidious" replicative pol III HE (24, 184, 185). These multiple interactions could facilitate replacement of pol III by pol V at a replication fork blocked by a lesion, and then help replace pol V with pol III once TLS has occurred. In *Schizosaccharomyces pombe*, the N-terminal region of pol ϵ encoding the polymerase and 3'-exonuclease activities is not required for cell survival. Yet mutant cells are exceedingly sensitive to DNA damage and undergo cell cycle delay, and their viability depends on genes that provide checkpoints for DNA damage control (186). It has been proposed that the C-terminal half of pol ϵ is needed for replication complex assembly at the beginning of S phase and for recruiting other DNA polymerases to the initiation site (186).

Based on its multiplicity of interactions with replication, repair, and cell cycle control proteins, a sliding processivity clamp could act as a replication traffic cop by helping to cue competing replication and repair reactions. The β dimer interacts with all five *E. coli* polymerases, and also with ligase and MutS mismatch recognition protein (187). PCNA interacts with the replication pols δ and ϵ , with the MSH2-MSH6 mismatch repair complex (188, 189), and, as shown recently, with human pol η (123a), pol ι (194), and pol κ (195).

The DNA replication trombone model (190) (Figure 3) provides an elegant description of how coupled leading- and lagging-strand DNA synthesis is coordinated at the replication fork when synthesis is unimpeded. However, a loss



Figure 8 DNA repair factory model. The cartoon depicts a stationary replicationrepair complex encountering damaged DNA rolled along as on a conveyer belt, in contrast to the more common textbook illustrations of DNA polymerase and accessory proteins moving along a stationary DNA molecule (Figure 3). A lesion is shown on the leading strand of the DNA template (green rod) as a distortion in the template track. The replication fork collapses when confronting the lesion, and leading- and lagging-strand synthesis become uncoupled. The leading-strand replicative polymerase (Repl pol, purple dumbbell) that was initially attached to the sliding clamp (yellow doughnut) cannot copy past the lesion and is replaced by an EP pol (red dumbbell). The bucket housing different EP pols symbolizes the high local concentrations of polymerases chosen at random to copy the damaged template strand. Lagging-strand synthesis continues on the undamaged template strand using the Repl pol. Okazaki fragments primed by RNA oligomers (red rods) are shown on the newly synthesized portion of the lagging strand.

of coupling occurs when the fork is blocked by damage to either strand, and an exchange of one polymerase for another is called for. This exchange process might proceed more efficiently if the DNA synthesis complex remains fixed, with the DNA passing through (Figure 8), rather than the familiar textbook depiction of a polymerase traversing a stationary DNA track (Figure 3). The idea of an immobile factory for DNA replication and repair is not new. Its current renaissance is based on seeking an efficient way to swap a variety of replication and repair polymerases on demand.

Intracellular fluorescence imaging, used in conjunction with multiple-hybrid screening and classical biochemical methods to identify protein-protein and protein-DNA interactions, should make it possible to determine whether or not there are structures that might serve as combined DNA replication-repair factories. Progress along these lines has been made using imaging in *Bacillus subtilis*, where PolC tagged with green fluorescent protein is visualized at discrete locations in the cell (191). The prokaryotic data suggest that DNA may indeed be moving through an anchored polymerase (191).

In eukaryotes the situation is likely to be far more complex because duplication of DNA is initiated at multiple origins activated at different times within S phase. Could multiple factories be present, perhaps even one per replicon? A recent fluorescence study using fibroblasts transformed by simian virus 40 (SV40) found that pol η , localized uniformly in the nucleus, becomes associated with replication foci in S phase and then accumulates at foci impeded by UV- and carcinogen-induced DNA damage (123). A key point is that 70 C-terminal amino acids are necessary for localization in the nucleus and an additional 50 are required for relocalization into foci, but these are not required for pol η activity. Colocalization of pol η with PCNA also occurs, presumably at DNA damage sites (123). The suggestion has been made that pol η is associated with the replication machinery (123, 192), perhaps not in the immediate vicinity of an unimpeded replication fork, but close enough to gain access to a blocked replication fork. In all likelihood, pol η along with all of the other EP pols are sequestered in DNA repair factories to be called upon when needed for TLS (123) (Figure 8).

A cartoon of the repair factory illustrates that the removal and replacement of a blocked replicative polymerase (Repl pol) by an EP pol is driven by increasing the local concentrations of the repair polymerases proximal to the replication fork (Figure 8). Random sampling can be used to select which polymerase is chosen, akin to drawing red, green, yellow, etc, dumbbells from the EP pol bucket; the selection probability is proportional to the relative concentration of each polymerase. A trafficking mechanism governed by the random selection of available EP pols implies that the "best" repair enzyme is not always chosen to copy a specific lesion. An excellent experimental illustration of random polymerase selection is that the relative copy numbers of the genes *polB* (*E. coli* pol II) and *umuDC* (*E. coli* pol V) determine whether AAF guanine adducts are copied in an error-prone or error-free manner in vivo (193). A preponderance of -2 frame-shifts occurs when pol II is expressed at higher levels than pol V and vice versa (193), demonstrating that pols II and V are in direct competition to copy the same lesion.

Biochemical triumphs taking place on the 3R (replication-recombinationrepair) front include model systems that depict DNA replication, generalized and site-specific recombination, mismatch repair, nucleotide excision repair, base excision repair, translesion replication, and replication restart. Assuming that the next goal is to elucidate the mysteries and complexities of polymerase trafficking and targeting, the next generation of biochemical model systems may require replicating basic elements from each of the 3R assays, judiciously recombined (pun intended). Not only that, but preserving multibody protein-protein interactions may require the presence of each protein component in its native state, free of the popular and ubiquitous but potentially interfering N- or C-terminal tags. It will be a formidable challenge to attain this gold standard.

43

ACKNOWLEDGMENTS

I want to thank Matthew D. Scharff, Brigette Tippin, and Ronda Bransteitter for their important contributions and generous help in shaping the eukaryotic DNA polymerase and somatic hypermutation sections; Erica Seitz, Phuong Pham, Jeffrey Bertram, and John Petruska for critically reading and editing the manuscript; and Nicholas Chelyapov, Jr., for his creative artistry in preparing the illustrations. I also want to acknowledge helpful comments made by Rick Fishel, Robert Fuchs, and Tom Ellenberger. Those friends and colleagues whose scientific contributions were primarily responsible for initiating the field of prokaryotic SOS error-prone repair are Evelyn Witkin and Miroslav Radman, and those who provided the inspiration for our studies include the aforementioned along with Keven McEntee, Roger Woodgate, Matty Scharff, Graham Walker, Chris Lawrence, and Bryn Bridges, with special gratitude accorded to my close friend Hatch Echols. This work was supported by grants from the National Institutes of Health, GM21422, GM42554, and AG17179.

The Annual Review of Biochemistry is online at http://biochem.annualreviews.org

LITERATURE CITED

- Goodman MF, Tippin B. 2000. Nat. Rev. Mol. Cell Biol. 1:101–9
- Ohmori H, Friedberg EC, Fuchs RPP, Goodman MF, Hanaoka F, et al. 2001. *Mol. Cell* 8:7–8
- Cox MM, Goodman MF, Kreuzer KN, Sherattt DJ, Sandler SJ, Marians KJ. 2000. *Nature* 404:37–41
- Masutani C, Kusumoto R, Yamada A, Dohmae N, Yokoi M, et al. 1999. *Nature* 399:700–4
- Johnson RE, Kondratick CM, Prakash S, Prakash L. 1999. Science 285:263–65
- Echols H, Goodman MF. 1991. Annu. Rev. Biochem. 60:477–511
- Goodman MF. 1997. Proc. Natl. Acad. Sci. USA 94:10493–95
- Kunkel TA, Bebenek K. 2000. Annu. Rev. Biochem. 69:497–529
- Trincao J, Johnson RE, Escalante CR, Prakash S, Prakash L, Aggarwal AK. 2001. *Mol. Cell.* 8:417–26
- Zhou B-L, Pata JD, Steitz TA. 2001. Mol. Cell. 8:427–37

- Ling H, Boudsocq F, Woodgate R, Yang W. 2001. Cell. 107:91–102
- Silvian LF, Toth EA, Pham P, Goodman MF, Ellenberger T. 2001. Nat. Struct. Biol. 8:984–89
- Radman M. 1975. In Molecular Mechanisms for the Repair of DNA, Part A, ed. P Hanawalt, RB Setlow, pp. 355– 67. New York: Plenum
- 14. Witkin EM. 1976. Bacteriol. Rev. 40:869–907
- 15. Walker GC. 1985. Annu. Rev. Biochem. 54:425–57
- Courcelle JA, Khodursky A, Peter B, Brown PO, Hanawalt PC. 2001. *Genetics* 158:41–64
- Friedberg EC, Walker GC, Siede W. 1995. In *DNA Repair and Mutagenesis*, pp. 407–522. Washington, DC: Am. Soc. Microbiol.
- Kato T, Shinoura Y. 1977. Mol. Gen. Genet. 156:121–31
- Steinborn G. 1978. Mol. Gen. Genet. 165:87–93
- 20. Sommer S, Knezevic J, Bailone A,

Devoret R. 1993. Mol. Gen. Genet. 239: 137–44

- Woodgate R, Rajagopalan M, Lu C, Echols H. 1989. Proc. Natl. Acad. Sci. USA 86:7301–5
- Bruck I, Woodgate R, McEntee K, Goodman MF. 1996. J. Biol. Chem. 271:10767–74
- 23. Tang MJ, Bruck I, Eritja R, Turner J, Frank EG, et al. 1998. *Proc. Natl. Acad. Sci. USA* 95:9755–60
- Tang MJ, Shen X, Frank EG, O'Donnell M, Woodgate R, Goodman MF. 1999. Proc. Natl. Acad. Sci. USA 96:8919–24
- Reuven NB, Arad G, Maor-Shoshani A, Livneh Z. 1999. J. Biol. Chem. 274: 31763–66
- Tang MJ, Pham P, Shen X, Taylor J-S, O'Donnell M, et al. 2000. *Nature* 404: 1014–18
- 26a. Nohmi T, Battista JR, Dodson LA, Walker GC. 1988. Proc. Natl. Acad. Sci. USA 85:1816–20
- Dutreix M, Moreau PL, Bailone A, Galibert F, Battista JR, et al. 1989. J. Bacteriol. 171:2415–23
- Sweasy JB, Witkin EM, Sinha N, Roegner-Maniscalco V. 1990. J. Bacteriol. 172:3030–36
- 29. Bailone A, Sommer S, Knezevic J, Dutreix M, Devoret R. 1991. *Biochimie* 73:479–84
- Reuven NB, Tomer G, Livneh Z. 1998. Mol. Cell 2:191–99
- Pham P, Bertram JG, O'Donnell M, Woodgate R, Goodman MF. 2001. *Nature* 409:366–70
- Reuven NB, Arad G, Stasiak AZ, Stasiak A, Livneh Z. 2001. J. Biol. Chem. 276:5511–17
- Bonner CA, Randall SK, Rayssiguier C, Radman M, Eritja R, et al. 1988. *J. Biol. Chem.* 263:18946–52
- Bonner CA, Hays S, McEntee K, Goodman MF. 1990. Proc. Natl. Acad. Sci. USA 87:7663–67
- 35. Iwasaki H, Nakata A, Walker G,

Shinagawa H. 1990. J. Bacteriol. 172: 6268-73

45

- Wagner J, Gruz P, Kim SR, Yamada M, Matsui K, et al. 1999. *Mol. Cell* 40:281–86
- Braithwaite DK, Ito J. 1993. Nucleic Acids Res. 21:787–802
- Kornberg A, Baker TA. 1992. DNA Replication. New York: Freeman. 2nd ed.
- Echols H, Goodman MF. 1990. *Mutat. Res.* 236:301–11
- Rajagopalan M, Lu C, Woodgate R, O'Donnell M, Goodman MF, Echols H. 1992. Proc. Natl. Acad. Sci. USA 89:10777–81
- Petit M-A, Bedale W, Osipiuk J, Lu C, Rajagopalan M, et al. 1994. J. Biol. Chem. 269:23824–29
- 42. Echols H. 1982. Biochimie 64:571-75
- 43. Goodman MF. 2000. *Trends Biochem. Sci.* 25:189–95
- Lawrence CW, Borden A, Banerjee SK, LeClerc JE. 1990. Nucleic Acids Res. 18:2153–57
- Lawrence CW, Banerjee SK, Borden A, LeClerc JE. 1990. Mol. Gen. Genet. 166:166–68
- LeClerc JE, Borden A, Lawrence CW. 1991. Proc. Natl. Acad. Sci. USA 88:9685–89
- Smith CA, Wang M, Jiang N, Che L, Zhao XD, Taylor J-S. 1996. *Biochemistry* 35:4146–54
- Kuzminov A. 1999. Microbiol. Mol. Biol. Rev. 63:751–813
- Kowalczykowski SC, Dixon DA, Eggleston AK, Lauder SD, Rehrauer WM. 1994. *Microbiol. Rev.* 58:401–65
- Roca AI, Cox MM. 1997. Prog. Nucleic Acid Res. Mol. Biol. 56:129– 223
- 51. Arenson TA, Tsodikov OV, Cox MM. 1999. J. Mol. Biol. 288:391–401
- 52. Livneh Z. 2001. J. Biol. Chem. 276: 25639-42
- 53. Sommer S, Bailone A, Devoret R. 1993. *Mol. Microbiol.* 10:963–71

- 54. Boudsocq F, Campbell M, Devoret R, Bailone A. 1997. *J. Mol. Biol.* 270: 201–11
- Frank EG, Cheng NQ, Do CC, Cerritelli ME, Bruck I, et al. 2000. J. Mol. Biol. 297:585–97
- 56. De Lucia P, Cairns J. 1969. *Nature* 224:1164–66
- 57. Knippers R. 1970. Nature 228:1050-53
- Cai H, Yu H, McEntee K, Goodman MF. 1995. *Methods Enzymol.* 262:13–21
- Cai H, Yu H, McEntee K, Kunkel TA, Goodman MF. 1995. J. Biol. Chem. 270:15327–35
- Qiu ZH, Goodman MF. 1997. J. Biol. Chem. 272:8611–17
- Rangarajan S, Woodgate R, Goodman MF. 1999. Proc. Natl. Acad. Sci. USA 96:9224–29
- Sommer S, Boudsocq F, Devoret R, Bailone A. 1998. *Mol. Microbiol.* 28:281–91
- Courcelle J, Carswell-Crumpton C, Hanawalt PC. 1997. Proc. Natl. Acad. Sci. USA 94:3714–19
- Pham P, Rangarajan S, Woodgate R, Goodman MF. 2001. Proc. Natl. Acad. Sci. USA 98:8350–54
- Rangarajan S, Woodgate R, Goodman MF. 2002. Mol. Microbiol. 43:617–28
- 66. Courcelle J, Crowley DJ, Hanawalt PC. 1999. J. Bacteriol. 181:916–22
- McGlynn P, Lloyd RG. 2000. *Cell* 101: 35–45
- 68. Cox MM. 2001. Annu. Rev. Genet. 35:53-82
- Umezu K, Kolodner RD. 1994. J. Biol. Chem. 269:30005–13
- Webb BL, Cox MM, Inman RB. 1997. *Cell* 91:347–56
- 71. Shan Q, Bork JM, Webb BL, Inman RB, Cox MM. 1997. J. Mol. Biol. 265: 519–40
- Robu ME, Inman RB, Cox MM. 2001. Proc. Natl. Acad. Sci. USA 98:8211–18
- 73. Postow L, Ullsperger C, Keller RW, Bustamante C, Vologodskii AV, Coz-

zarelli NR. 2001. J. Biol. Chem. 276: 2790–96

- 74. Wickner S, Hurwitz J. 1974. *Proc. Natl. Acad. Sci. USA* 71:4120–24
- Hughes AJ, Bryan SK, Chen H, Moses RE, McHenry CS. 1991. J. Biol. Chem. 266:4568–73
- Bonner CA, Stukenberg PT, Rajagopalan M, Eritja R, O'Donnell M, et al. 1992. J. Biol. Chem. 267:11431–38
- 77. Wagner J, Fujji S, Gruz P, Nohmi T, Fuchs RPP. 2000. EMBO Rep. 1:484–88
- Napolitano R, Janel-Bintz R, Wagner J, Fuchs RPP. 2000. *EMBO J.* 19: 6259–65
- 78a. Shen X, Sayer J, Kroth I, Pontén I, O'Donnell M, et al. 2002. J. Biol. Chem. 277:5265–74
- Tessman I, Kennedy MA. 1993. Genetics 136:439–48
- Brotcorne-Lannoye A, Maenhaut-Michel G. 1986. Proc. Natl. Acad. Sci. USA 83:3904–8
- Kim S-R, Maenhaut-Michel G, Yamada M, Yamamoto Y, Matsui K, et al. 1997. Proc. Natl. Acad. Sci. USA 94:13792–97
- Kim S, Matsui K, Yamada M, Gruz P, Nohmi T. 2001. Mol. Genet. Genomics 266:207–15
- Foster PL. 1993. Annu. Rev. Microbiol. 47:467–504
- Foster PL, Gudmundsson G, Trimarchi JM, Cai H, Goodman MF. 1995. Proc. Natl. Acad. Sci. USA 92:7951–55
- Foster PL. 2000. Cold Spring Harbor Symp. Quant. Biol. 65:21–29
- McKenzie GJ, Lee PL, Lombardo M-J, Hastings PJ, Rosenberg SM. 2001. *Mol. Cell* 7:571–79
- Haracska L, Unk I, Johnson RE, Johansson E, Burgers PMJ, et al. 2001. *Genes Dev.* 15:945–54
- Wang ZH, Castano IB, De Las Penas A, Adams C, Christman MF. 2000. Science 289:774–79
- 89. Zeng XM, Winter DB, Kasmer C,

Kraemer KH, Lehman AR, Gearhart PJ. 2001. *Nat. Immunol.* 2:537–41

- Rogozin IB, Pavlov YI, Bebenek K, Matsuda T, Kunkel TA. 2001. Nat. Immunol. 2:530–36
- Zan H, Komori A, Li ZD, Cerutti A, Schaffer A, et al. 2001. *Immunity* 14:643–53
- Poltoratsky V, Woo CJ, Tippin B, Martin A, Goodman MF, Scharff MD. 2001. Proc. Natl. Acad. Sci. USA 98:7976–81
- 92a. Burgers PMJ, Koonin EV, Bruford E, Blanco L, Burtis KC, et al. 2001. J. Biol. Chem. 276: 43487–90
- Lawrence CW, Christensen RB. 1978.
 J. Mol. Biol. 122:1–21
- 94. Lawrence CW, Das G, Christensen RB. 1985. Mol. Gen. Genet. 200:80-85
- Nelson JR, Lawrence CW, Hinkle DC. 1996. *Nature* 382:729–31
- 96. Lin WS, Xin H, Zhang YB, Wu XH, Yuan FH, Wang ZG. 1999. Nucleic Acids Res. 27:4468–75
- Nelson JR, Gibbs PEM, Nowicka AM, Hinkle DC, Lawrence CW. 2000. *Mol. Microbiol.* 37:549–54
- Ito J, Braithwaite DK. 1991. Nucleic Acids Res. 19:4045–57
- Nelson JR, Lawrence CW, Hinkle DC. 1996. Science 272:1646–49
- 100. Gibbs PEM, McGregor WG, Maher VM, Nisson P, Lawrence CW. 1998. Proc. Natl. Acad. Sci. USA 95:6876–80
- Gibbs PEM, Wang X-D, Li Z, McManus TP, McGregor WG, Lawrence CW. 2000. Proc. Natl. Acad. Sci. USA 97:4186–91
- Van Sloun PP, Romejin RJ, Eeken JC.
 1999. *Mutat. Res.* 433:109–16
- 103. Eeken JC, Romejin RJ, de Jong AW, Pastink A, Lohman PH. 2001. *Mutat. Res.* 485:237–53
- 104. Murakumo Y, Roth T, Ishii H, Rasio D, Numata S, et al. 2000. J. Biol. Chem. 275:4391–97
- 105. Lawrence CW, Maher VM. 2001. Phi-

los. Trans. R. Soc. London Ser. B 356: 41–46

47

- 106. Bemark M, Khamlichi AA, Davies SL, Neuberger MS. 2000. Curr. Biol. 10:1213–16
- Wittschieben J, Shivji MKK, Lalani E, Jacobs MA, Marini F, et al. 2000. *Curr. Biol.* 10:1217–20
- Esposito G, Godin I, Klein U, Yaspo ML, Cumano A, Rajewsky K. 2000. *Curr. Biol.* 10:1221–24
- 109. Johnson RE, Washington MT, Haracska L, Prakash S, Prakash L. 2000. *Nature* 406:1015–19
- 110. Yamada A, Masutani C, Iwai S, Hanaoka F. 2000. Nucleic Acids Res. 28:2473–80
- McDonald JP, Levine AS, Woodgate RW. 1997. *Genetics* 147:1557–68
- Washington MT, Johnson RE, Prakash S, Prakash L. 1999. *J. Biol. Chem.* 274: 36835–38
- Ishikawa T, Uematsu N, Mizukoshi T, Iwai S, Iwasaki H, et al. 2001. J. Biol. Chem. 276:15155–63
- 114. Yu S, Johnson RE, Prakash S, Prakash L. 2000. Mol. Cell. Biol. 20:8001–7
- Johnson RE, Washington MT, Prakash S, Prakash L. 2000. J. Biol. Chem. 275: 7447–50
- 116. Matsuda T, Bebenek K, Masutani C, Hanaoka F, Kunkel TA. 2000. Nature 404:1011–13
- 117. Masutani C, Kusumoto R, Iwai S, Hanaoka F. 2000. EMBO J. 19:3100–9
- 118. Haracska L, Yu S, Johnson RE, Prakash L, Prakash S. 2000. Nat. Genet. 25:458–61
- 119. Johnson RE, Prakash S, Praskah L. 1999. Science 283:1001–4
- 120. Haracska L, Prakash S, Prakash L. 2000. Mol. Cell. Biol. 20:8001–7
- Washington MT, Johnson RE, Prakash S, Prakash L. 2001. J. Biol. Chem. 276: 2263–66
- 122. Kondratick CM, Washington MT, Prakash S, Prakash L. 2001. Mol. Cell. Biol. 21:2018–25

- 123. Kannouche P, Broughton BC, Volker M, Hanaoka F, Mullenders LHF, Lehmann AR. 2001. Genes Dev. 15:158–72
- 123a. Haracska L, Johnson RE, Unk I, Phillips B, Hurwitz J, et al. 2001. Mol. Cell. Biol. 21:7199–206
- 123b. Haracska L, Kondratick CM, Unk I, Prakash L, Prakash S. 2001. Mol. Cell 8:407–15
- McDonald JP, Rapic-Otrin V, Epstein JA, Broughton BC, Wang XY, et al. 1999. *Genomics* 60:20–30
- 125. Frank EG, Tissier A, McDonald JP, Rapic-Otrin V, Zeng XM, et al. 2001. *EMBO J.* 20:2914–22
- 126. Tissier A, McDonald JP, Frank EG, Woodgate R. 2000. Genes Dev. 14:1642–50
- 127. Zhang YB, Yuan FH, Wu XH, Wang ZG. 2000. Mol. Cell. Biol. 20:7099– 108
- 128. Bebenek K, Tissier A, Frank EG, McDonald JP, Prasad R, et al. 2001. Science 291:2156–59
- 129. Vaisman A, Tissier A, Frank EG, Goodman MF, Woodgate R. 2001. J. Biol. Chem. 276:30615–22
- Zhang YB, Yuan FH, Wu XH, Taylor J-S, Wang ZG. 2001. Nucleic Acids Res. 29:928–35
- 130a. Vaisman A, Woodgate R. 2001. *EMBO* J. 20:6520–9
- Gerlach VL, Aravind L, Gotway G, Schultz R, Koonin EV, Friedberg EC.
 1999. Proc. Natl. Acad. Sci. USA 96:11922–27
- Zhang YB, Yuan FH, Wu XH, Wang M, Rechkoblit O, et al. 2000. Nucleic Acids Res. 28:4138–46
- 133. Johnson RE, Prakash S, Prakash L. 2000. Proc. Natl. Acad. Sci. USA 97:3838–43
- 134. Ohashi E, Bebenek K, Matsuda T, Feaver WJ, Gerlach VL, et al. 2000. J. Biol. Chem. 275:39678-84
- 135. Ohashi E, Ogi T, Kusumoto R, Iwai S,

Masutani C, et al. 2000. *Genes Dev.* 14:1589–94

- Levine RL, Miller H, Grollman A, Ohashi E, Ohmori H, et al. 2001. J. Biol. Chem. 276:18717–21
- Gerlach VL, Feaver WJ, Fischhaber PL, Friedberg EC. 2001. J. Biol. Chem. 276:92–98
- 138. O-Wang J, Kawamura K, Tada Y, Ohmori H, Kimura H, et al. 2001. Cancer Res. 61:5366–69
- Carson DR, Christman MF. 2001. Proc. Natl. Acad. Sci. USA 98:8270–75
- 140. Aravind L, Koonin EV. 1999. Nucleic Acids Res. 27:1609–18
- Castano IB, Brzoska PM, Sadoff BU, Chen HY, Christman MF. 1996. Genes Dev. 10:2564–76
- Skibbens RV, Corson LB, Koshland D, Hieter P. 1999. Genes Dev. 13:307–19
- 143. Mayer ML, Gygi SP, Aebersold R, Hieter P. 2001. *Mol. Cell* 7:959–70
- 144. Hanna JS, Kroll ES, Lundblad V, Spencer FA. 2001. Mol. Cell. Biol. 21:3144–58
- 145. Aoufouchi S, Flatter E, Dahan A, Faili A, Bertocci B, et al. 2000. Nucleic Acids Res. 28:3684–93
- 146. Dominguez O, Ruiz JF, de Lera LT, Garcia-Diaz M, Gonzalez MA, et al. 2000. *EMBO J.* 19:1731–42
- 146a. Zhang YB, Wu XH, Yuan FH, Xie ZW, Wang ZG. 2001. Mol. Cell. Biol. 21:7995–8006
- 147. Nagasawa K-i, Kitamura K, Yasui A, Nimura Y, Ikeda K, et al. 2000. J. Biol. Chem. 275:31233–38
- 148. Garcia-Diaz M, Bebenek K, Kunkel TA, Blanco L. 2001. J. Biol. Chem. 276: 34659–63
- 148a. Prasad R, Lavrik OI, Kim S-J, Kedar P, Yang X-P, et al. 2001. J. Biol. Chem. 276:32411–14
- 149. Longley MJ, Prasad R, Srivastava DK, Wilson SH, Copeland WC. 1998. Proc. Natl. Acad. Sci. USA 95:12244–48
- 150. Pinz KG, Bogenhagen DF. 2000. J. Biol. Chem. 275:12509-14

- 151. Zhang J, MacLennan IC, Liu YJ, Lane PJL. 1988. *Immunol. Lett.* 18: 2393–400
- 152. Berek C, Milstein C. 1988. Immunol. Rev. 105:5–26
- Neuberger MS, Erhenstein MR, Klix N, Jolly CJ, Yelamos J, et al. 1998. *Immunol. Rev.* 162:107–16
- 154. Kim S, Davis M, Sinn E, Patten P, Hood L. 1981. *Cell* 27:573–81
- 155. Gearhart PJ, Bogenhagen DF. 1983. Proc. Natl. Acad. Sci. USA 80:3439-43
- 156. Betz AG, Milstein C, Gonzalez FA, Pannell R, Larson T, Neuberger MS. 1994. Cell 77:239–48
- 157. Yelamos J, Klix N, Govenechea B, Lozano F, Chui YL, et al. 1995. *Nature* 376:225–29
- Pasqualucci L, Neumeister P, Goossens T, Nanjangud G, Chaganti RSK, et al. 2001. *Nature* 412:341–46
- 159. Winter DB, Gearhart PJ. 1995. Curr. Biol. 5:1345–46
- Migliazza A, Martinotti S, Chen W, Fusco C, Ye BH, et al. 1995. Proc. Natl. Acad. Sci. USA 92:12520–24
- Shen HM, Peters A, Baron B, Zhu XD, Storb U. 1998. Science 280:1750–52
- 162. Pasqualucci L, Migliazza A, Fracchiolla N, William C, Neri A, et al. 1998. *Proc. Natl. Acad. Sci. USA* 95:11816–21
- Muschen M, Re D, Jungnickel B, Diehl V, Rajewsky K, Kuppers R. 2000. J. Exp. Med. 192:1833–40
- Rogozin IB, Kolchanov NA. 1992. Biochim. Biophys. Acta 1171:11–18
- 165. Reynaud CA, Garcia C, Hein WR, Weill JC. 1995. *Cell* 80:115–25
- 166. Storb U. 1996. Curr. Opin. Immunol. 8:206–14
- 167. Storb U. 1998. Immunol. Rev. 162:5-11
- Spencer J, Dunn M, Dunn-Walters DK. 1999. J. Immunol. 162:6596–601
- Poltoratsky V, Goodman MF, Scharff MD. 2000. J. Exp. Med. 192:F27–30

- 170. Storb U. 2001. Nat. Immunol. 2:484–85
- 171. Kong Q, Maizels N. 2001. *Genetics* 158:369-78
- 172. Weill J-C, Bertocci B, Faili A, Aoufouchi S, Frey S, et al. 2001. Adv. Immunol. 80:183–202
- 173. Esposito G, Texido G, Betz UAK, Gu H, Muller W, et al. 2000. Proc. Natl. Acad. Sci. USA 97:1166–71
- 174. Denepoux S, Razanajaona D, Blanchard D, Meffre G, Capra JD, et al. 1997. *Immunity* 6:35–46
- 175. Diaz M, Verkoczy LK, Flajnik MF, Klinman NR. 2001. J. Immunol. 167: 327–35
- 176. Wood RD. 1998. *Curr. Biol.* 8:R757–60
- 177. Reynaud C-A, Bertocci B, Frey S, Delbos F, Quint L, Weill J-C. 1999. *Immunol. Today* 20:522–27
- 178. Wiesendanger M, Kneitz B, Edelmann W, Scharff MD. 2000. J. Exp. Med. 191:579–84
- Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Hongo T. 2000. *Cell* 102:553–63
- Revy P, Muto T, Levy Y, Geissmann F, Plebani A, et al. 2000. *Cell* 102: 565–75
- Teng B, Burant CF, Davidson NO. 1993. Science 260:1816–19
- Mehta A, Kinter MT, Sherman NE, Driscoll DM. 2000. Mol. Cell. Biol. 20:1846–54
- 183. Woodgate R. 1999. *Genes Dev.* 13: 2191–95
- 183a. Shimizu K, Kawasaki Y, Hiraga S-I, Tawaramoto M, Nakashima N, Sugino A. 2002. *Cell.* In press
- 184. Sutton MD, Opperman T, Walker GC. 1999. Proc. Natl. Acad. Sci. USA 96:12373–78
- 185. Sutton MD, Walker GC. 2001. Proc. Natl. Acad. Sci. USA 98:8342–49
- 186. Feng W, D'Urso G. 2001. Mol. Cell. Biol. 21:4495–504

- 187. Lopez de Saro FJ, O'Donnell M. 2001. Proc. Natl. Acad. Sci. USA 98:8376–80
- 188. Umar A, Buermeyer AB, Simon JA, Thomas DC, Clark AB, et al. 1996. Cell 87:65–73
- 189. Flores-Rozas H, Clark D, Kolodner RD. 2000. Nat. Genet. 26:375–78
- Alberts BM, Barry J, Bedinger P, Formosa T, Jongeneel CV, Kreuzer KN. 1983. Cold Spring Harbor Symp. Quant. Biol. 47:655–68
- 191. Lemon KP, Grossman AD. 1998. Science 282:1516–19

- 192. Limoli CL, Giedzinski E, Morgan WF, Cleaver JE. 2000. Proc. Natl. Acad. Sci. USA 97:7939–46
- 193. Becherel OJ, Fuchs RPP. 2001. Proc. Natl. Acad. Sci. USA 98:8566–71
- 194. Haracska L, Johnson RE, Unk I, Phillips BB, Hurwitz J, et al. 2001. Proc. Natl. Acad. Sci. USA 98:14256-61
- 195. Haracska L, Unk I, Johnson RE, Phillips BB, Hurwitz J, et al. 2002. Mol. Cell. Biol. 22:784–91
- 196. Bertocci B, De Smet A, Flatter E, Dahan A, Bories JC, et al. 2002. J. Immunol. In press