Chapter 4

Genome Instability in Bacteria and Archaea: Strategies for Maintaining Genome Stability

J.-E. Messling, A.B. Williams
University of Cologne, Cologne, Germany

Chapter Outline

1. INTRODUCTION

The genomes of all organisms are constantly challenged by DNA damaging forces, from their own internal metabolic byproducts to various outside forces. DNA, as the carrier of the genetic details of an organism, is undeniably the most important macromolecule, both for individual cells and for the species. Unlike RNA and protein, which can be replaced using the information stored in DNA, any loss of DNA sequence due to DNA damage is essentially irreplaceable without extraordinary measures. Besides this catastrophic outcome, DNA damage can also interfere with important ongoing cellular processes in individual cells, such as DNA replication and transcription, potentially leading to cell death. Because of the importance of maintaining both the fidelity of the heritable genetic information and the integrity of the DNA for ongoing cellular processes, a complex network of DNA-repair systems has evolved. In a testament to the primal need to maintain genetic integrity, these pathways are conserved in all known organisms. In some cases, the components of different DNA pathways are clearly orthologs derived from a common evolutionary ancestor; in other cases, the different players may have evolved independently to serve the same function. For comparison and reference, Table 4.1 compares orthologous and functionally homologous genes in bacteria, archaea, and the simple eukaryotic nematode *Caenorhabditis elegans*.

Like many biological processes, DNA repair proteins have been systematically organized into mostly linear pathways for the purpose of understanding how cells repair particular types of lesions; however, one must bear in mind that, in fact, DNA repair pathways represent a dazzlingly complex network in which proteins function in multiple DNA repair pathways, or even in entirely different biochemical processes. The following discussion addresses the fundamental processes that preserve genome integrity in bacteria and archaea, including how cells respond to DNA damage to support DNA repair,

Genome Stability. http://dx.doi.org/10.1016/B978-0-12-803309-8.00004-5

TABLE 4.1 DNA Repair Homologs in Selected Model Organisms

Function/Enzyme	*Escherichia coli*	*Archea*	*Caenorhabditis elegans*
Mismatch Repair (MMR)			
Mismatch recognition	MutS	MutS (TTHA1892) (*Thermus thermophilus*)	MutSα (MSH-1 + MSH-6)
			MutSβ (MSH-2 + MSH-3)
Match-making ATPase	MutL		
Endonuclease ATPase		MutL (TTHA1892) (*T. thermophilus*)	MutLα (MLH-1 + PMS-2)
			MutLβ (MLH-1 + PMS-1)
			MutLγ (MLH-1 + MLH-3)
Activation of MutL homolog		DNA Pol III β subunit (TTHA0001) (*T. thermophilus*)	PCN-1 (f.h.)[a]
		DNA Pol III δ,δ,γ,τ subunits (TTHA0001) (*T. thermophilus*)	RFC (RFC-1, -2, -3, -4, -5) (f.h.)
Crossover promotion during meiosis			Not identified
Mismatch repair during mitosis and meiosis			Not identified
Endonuclease	MutH	?	Not identified
DNA helicase	UvrD	UvrD (TTHA0244) (*T. thermophilus*)	?
Strand excision (5′–3′)	RecJ	RecJ (TTHA1892) (*T. thermophilus*)	
Strand excision (3′–5′)	ExoI	ExoI (TTHB187) (*T. thermophilus*)	
Strand excision (5′–3′)	ExoVII		
Strand excision (5′ –3′)	ExoX		
Strand excision (5′ –3′)			EXO-1
Single-strand binding	SSB	SSB (TTHA0244) (*T. thermophilus*)	?
DNA polymerase	Pol III holoenzyme (subunits: α,γ,δ,δ′,Ψ,λ,τ,ε,θ)	Pol III?	Pol δ (**F10C2.4, F12F6.7?**)
Ligation	LigA	?	LIG-1
Nucleotide Excision Repair (NER)			
Binding of damaged DNA			XPC-1
Endonuclease			XPG-1
Removal of transcription blocking lesions	Mfd	Mfd (TTHA0889) (*T. thermophilus*)	CSB-1 (f.h.)
			CSA-1
Replication factor C			RFC (consists of RFC-1 to -5)
			RPA-1 and RPA-2
Binding of DNA and proteins in preincision complex			XPA-1
3′-incision nuclease	Cho	Not identified	
3′-incision nuclease			ERCC-5
5′-incision nuclease			ERCC-1
5′-incision nuclease			ERCC-4

TABLE 4.1 DNA Repair Homologs in Selected Model Organisms—cont'd

Function/Enzyme	*Escherichia coli*	*Archea*	*Caenorhabditis elegans*
Nucleotide excision and basal transcription	No prokaryotic homolog	No prokaryotic homolog	
			XBP-1
			XPD-1
			R02D3.3
			T16H12.4
			Zk1128.4
			Y73F8A.24
			Y55B1AL.2
			CDK-7
			CYH-3
			MNAT-1
Involved in single strand break repair			Not identified
E3 ligase interaction			DDB-1
Recruitment of TLS polymerase in TC NER (proliferating cell nuclear antigen)			PCN-1
DNA synthesis	Pol III holoenzyme (subunits: α,γ,δ,δ′,Ψ,λ,τ,ε,θ)		**F10C2.4, F12F6.7?**
Binding of damaged DNA	UvrA	UvrA (*Halobacterium* sp.)	Not identified
DNA unwinding	UvrB	UvrB (*Halobacterium* sp.)	Not identified
3′- and 5′-incision nuclease	UvrC	UvrC (*Halobacterium sp.*)	Not identified
Base Excision Repair (BER)			
3MeA DNA glycosylase	AlkA, TagA	AlkA (TTHA0392) (*T. thermophilus*)	?
Uracil DNA glycosylase	Ung	UDGA (TTHA0718) (*T. thermophilus*)	UNG-1
Removal of Hydroxymethyl U			?
8 oxoguanine DNA glycosylase/AP lyase	MutY	MutY (TTHA1898) (*T. thermophilus*)	?
8 oxoguanine DNA glycosylase/AP lyase	Fpg (MutM)	MutM (TTHA1806) (*T. thermophilus*)	
Thymine glycol DNA glycosylase/AP lyase	Nth		?
TDG T:G mismatch DNA glycosylase			?
Exonuclease III	XthA		
Endonuclease IV	Nfo		?
Incision 3′ of hypoxanthine and uracil	Nfi (EndoV)	EndoV (TTHA1374) (*T. thermophilus*)	?
Removal of thymine glycol	Nei (EndoVIII)		?

Continued

TABLE 4.1 DNA Repair Homologs in Selected Model Organisms—cont'd

Function/Enzyme	*Escherichia coli*	*Archea*	*Caenorhabditis elegans*
Removal of oxidative products of C and U			?
Endonuclease IIIV like glycosylase 3			?
Flap endonuclease	PolI	Pol (TTHA1054) (*T. thermophilus*)	
Proliferating cell nuclear antigen			PCN-1
Replication factor C			RFC (consists of RFC-1 to -5)
DNA polymerases			
			Pol δ (**F10C2.4, F12F6.7 ?**)
			?
DNA ligase	LigA	LigA (TTHA1097) (*T. thermophilus*)	LIG-1
ATP-dependent DNA ligase			?
Poly (ADP ribose) Polymerase			PARP-1
ADPRT-like enzyme			PARP-2
Homologous Recombination (HR)			
Recombinase	RecA	RecA (TTHA1818) (*T. thermophilus*)	RAD-51
Branch migration complex	RuvA	RuvA (TTHA0291) (*T. thermophilus*)	
	RuvB	RuvB (TTHA0406) (*T. thermophilus*)	
DNA helicase	RecG	RecG (TTHA1266) (*T. thermophilus*)	
RecQ-like DNA helicase	RecQ		?
RecQ family DNA helicase			WRN-1
SMC-like ATPase		SbcC (TTHA1288) (*T. thermophilus*)	RAD-50
3′–5′-exonuclease (in complex with SbcC)	SbcD	SbcD (TTHA1289) (*T. thermophilus*)	MRE-11
Accessory protein for MR complex			
RecA-like ATPase	RadA/Sms	RadA/Sms (TTHA0541) (*T. thermophilus*)	
RadA paralog		RadC1 (*S. islandicus*)	
Rad51 like recombinase mediator			?
			?
			?
			?
RAD54 family DNA translocase, recombinase mediator			?
			RAD-54
Strand excision (5′–3′)	RecJ	RecJ (TTHA1892) (*T. thermophilus*)	EXO-1 (f.h.)
Strand excision (5′–3′)			

TABLE 4.1 DNA Repair Homologs in Selected Model Organisms—cont'd

Function/Enzyme	*Escherichia coli*	*Archea*	*Caenorhabditis elegans*
Helicase/nuclease complex	RecB		
	RecC		
	RecD		
5′–3′ exonuclease	RecE		
ssDNA annealing	RecT		
Single-strand binding	SSB	SSB (TTHA0244) (*T. thermophilus*)	?
DNA-binding complex	RecF	RecF (TTHA0264) (*T. thermophilus*)	
	RecO	RecO (TTHA0623) (*T. thermophilus*)	
	RecR	RecR (TTHA1600) (*T. thermophilus*)	
HJ resolvase	RuvC	RuvC (TTHA1090) (*T. thermophilus*)	
	RusA		
			SLX-1
			GEN-1
Recombinase inhibitor	RecX	RecX (TTHA0848) (*T. thermophilus*)	
DNA helicase	UvrD	UvrD (TTHA1427) (*T. thermophilus*)	
Accessory protein for structure-specific nucleases			HIM-18
Structure-specific endonuclease			MUS-81
		MutS2 (TTHA1645) (*T. thermophilus*)	
			Not identified
Complex with ERCC4 (Rad1)			ERCC-1

[a]*f.h. indicates functional homolog (ie, no obvious sequence homology).*

the actual molecular transactions at the DNA that lead to damage repair, and ways that cells preserve their genomes when challenged with parasitic foreign DNA.

2. REPONSES TO DNA DAMAGE

2.1 The SOS Response: A Primitive Cell-Cycle Checkpoint

Since its discovery and early characterization by Evelyn Witkin and Miroslav Radman in the early to mid-1970s, the SOS response has become a paradigm for the bacterial DNA damage response (for a detailed review, see Ref. [1]). While the SOS pathway has proven to be extremely complex, at its core are only two proteins: the LexA repressor and the RecA activator. Under normal conditions, LexA binds to a special sequence, the SOS box, in the promoters of SOS-regulated genes and blocks their expression. One of the most common outcomes of DNA damage is the formation of single-strand DNA (ssDNA) through any number of possible processes. The ssDNA is rapidly coated with RecA protein to form the RecA filament; while this RecA filament may go on to participate in homologous recombination, it has another function: to induce the autocleavage of LexA, relieving its repressive activity, and allowing expression of the SOS regulon. As an aside, despite over 30 years of work, how the RecA filament forms is still an active area of research and discussion (eg, see Ref. [2]). The SOS response is rapid and within just a few minutes the amount of LexA decreases by nearly 10-fold. During this time, cell division is blocked by the cytokinesis inhibitor SulA (SfiA) (whose gene is expressed robustly after SOS induction) and cells form distinctive filaments as they grow without division. The outcome of SOS induction is a massive transcriptional

reprogramming. Quite remarkably, evolution has endowed the SOS response with a logical order. The more rapidly induced genes are from the nucleotide excision repair (NER) pathway (discussed in the following text), one of the most versatile DNA repair mechanisms. Subsequently, *recA* and other genes for homologous recombination are induced, supporting high-fidelity repair and amplification of the response. The *lexA* gene itself is also induced, preparing the cell to turn off the SOS response when the activating signal wanes, probably because the SOS response is demanding on cellular resources and that it blocks the formation of progeny cells. Finally, as late as 40 min into the SOS response, the error-prone DNA polymerase Pol V (encoded by the *umuDC* genes) is expressed and activated (the latter process requires a RecA filament-dependent autocleavage event of the UmuD subunit). A second error-prone polymerase, Pol IV (*dinB*), is also upregulated. These polymerases replicate damaged DNA to allow continued cell growth with the hope that other DNA repair pathways will catch up; however, these polymerases can also introduce mistakes that can be preserved as mutations that can have either deleterious, or sometimes advantageous, outcomes (see Chapter X for a discussion of the latter).

The importance of the SOS response for bacteria is highlighted by its vast conservation among divergent species; however, some notable deviations from the *Escherichia coli* model have been observed. There is a disconnection between RecA-ssDNA binding and activation of the SOS response in *Baccilus subtilus* [3]. In *Caulobacter crescentus*, following a similar cell-cycle arrest [4] (albeit by a different molecular mechanism), the SOS response also triggers a more sophisticated programmed cell death pathway, akin to eukaryotic apoptosis [5]. Mycobacteria have evolved a more complex regulatory system for DNA damage-responsive genes. In this species, most of the DNA damage genes are regulated by a second factor, ClpR [6–8]; however, conserved regulation does exist, as *recA* is controlled by LexA (in addition to ClpR). Large-scale transcriptional reprogramming occurs in the extremophilic species *Deinococcus radiodurans*, although it depends mostly on the PprI protein instead of the two LexA homologs [9]. PprI binds to damage-responsive promoters after exposure to ionizing radiation to induce expression, including the *recA* promoter. Some archaea cope quite differently with DNA damage. The hyperthermophile *Sulfolobus* does not induce a large number of genes following UV exposure; however, it has a sophisticated alternative coping mechanism (see later).

Growing evidence suggests that the SOS response has many other functions in addition to this basic checkpoint control, including functions in horizontal gene transfer, the development of antibiotic resistance, and pathogenesis (see Ref. [1]). Nevertheless, it is clear that the SOS response is a first line of defense in the preservation of genome stability for bacteria.

2.2 An Archaeal UV Response Based on DNA Sharing

Invocation of a large-scale transcriptional reprogramming after UV exposure does not appear to be universal in archaea, but a UV-induced stress response has been characterized in *Sulfolobus* (Fig. 4.1).

After UV exposure, the cells induce expression of the *ups* genes, which encode a specialized type IV pilus system that enables efficient DNA transfer between cells [10–12]. The relocated DNA can then be used as a template for homologous recombination (HR)-dependent DNA repair—a generally, but not exclusively, accurate DNA repair pathway. That this

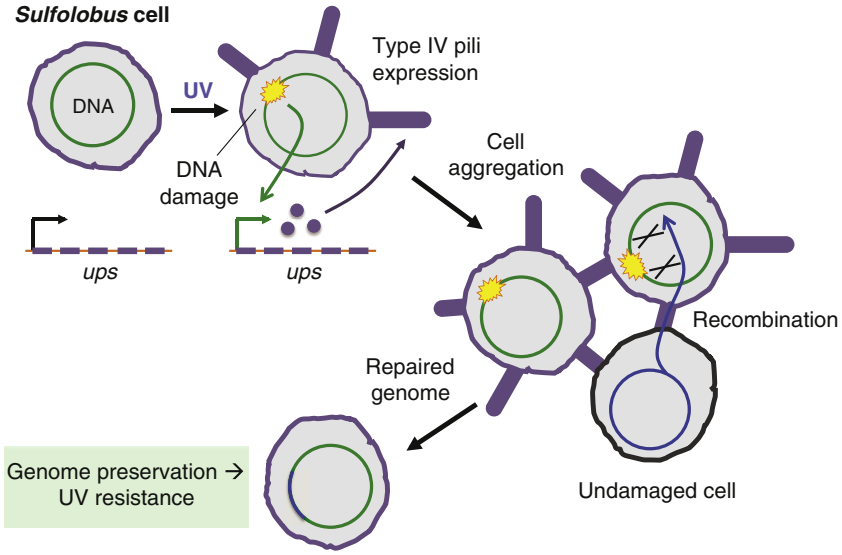

FIGURE 4.1 ***Sulfolobus*** **copes with UV-induced DNA damage via a type IV pilis-dependent DNA exchange pathway.** After DNA damage, *Sulfolobus* induces the *ups* operon to express type IV pili. After pili form, the cells aggregate and exchange DNA (both undamaged and damaged cells can act as DNA donors). Undamaged homologous DNA can be recombined to replace damaged segments, thus rescuing cells from the deleterious effects of DNA damage.

DNA sharing somehow protects the cells, presumably by dampening UV-induced genome instability, is supported by the observation that strains capable of expressing the type IV pili have higher survival rates after UV exposure. A 2015 work had further characterized this response by demonstrating additional involvement of four genes adjacent to the *ups* locus: an endonuclease III, a ParB-like protein, a glycosyltransferase, and a RecQ-like helicase [13]. With the exception of the ParB-like protein (which likely participates in the DNA transfer), these proteins are proposed to function in a homologous recombination-dependent DNA repair process downstream of the DNA transfer. While a bona fide SOS response is clearly absent in *Sulfolobus*, this system illustrates a novel genetic innovation for dealing with UV-induced DNA damage.

3. DNA REPAIR PATHWAYS

While the SOS response provides a genome-stabilizing function, it has no inherent DNA repair capacity. Instead, cells have evolved several intertwined molecular pathways comprised by the actual molecular transactions leading to damage repair: direct reversal (the only DNA synthesis–independent pathway); base excision repair (BER); NER; mismatch repair (MMR); and HR-dependent repair. The fundamentals of DNA repair have been most intensely studied in *E. coli*, thus its molecular biology forms the foundation of the discussion; however, important deviations in other species are also highlighted.

3.1 Direct Reversal of DNA Damage

One way to repair DNA damage is to simply undo the particular molecular changes, that is, to directly reverse the damage. Evolution has endowed life with (at least) three direct reversal pathways: photolyases, which repair UV-induced damage, and two mechanisms that repair alkylated bases, O^6-alkylguanine alkyl transferases (AGTs) and AlkB-family dioxygenases. While the molecular mechanisms vary drastically, the end result of all of these pathways is the restoration of the original molecular structure without the need for new DNA synthesis.

UV irradiation leads to two main types of DNA lesions that can disrupt many DNA-related processes, most importantly, replication and transcription: pyrimidine (6-4) photoproducts (6-4 PPs) and cyclobutane pyrimidine dimers (CPDs). Due to the different structures of these lesions, different photolyase enzymes are required for their repair; however, a common feature of photolyases is that they obtain energy from light to fuel the reaction (hence the classical name "light reactions") and use flavin adenine dinucleotide (FAD) for catalysis. In general, it is thought that direct reversal occurs via a multistep process in which light energy is harnessed to drive the full reduction of FAD to $FADH^-$, followed by an electron transfer to the lesion leading to the breakage of covalent bonds, and finally a retransfer of the free electron back to the FADH radical to produce $FADH^-$ (reviewed in [14] and [15]). In the end, the photolyase reaction is simply a stepwise transfer of energy that reconfigures the covalent bonds in the original bases to restore the original structure.

The *E. coli* K-12 photolyase is encoded by the *phr* gene (also known as *phrB*) [16] and the protein is maintained at low levels in stationary phase cells [17]. It is interesting to note that, despite its involvement in repairing UV-induced CPD dimers, *phr* is not regulated by the SOS response [18]. While CPD photolyases were one of the earliest characterized DNA repair mechanisms and have been found in all three domains of life, (6-4) PP photolyases remained elusive until only recently. The first bacterial (6-4) PP photolyase was reported in 2013 in *Agrobacterium tumefaciens* and is encoded by the *phrB* gene [19]. Photolyases have also been studied in various archaea, including halophiles, methanogenic species, and thermophiles [20–22]. Besides some structural differences and utilization of different chromophores for light collection, the functions of archaeal photolyases are conserved from their bacterial counterparts.

Alkylating agents interact with atoms in DNA bases leading to the formation of a variety of cytotoxic and potentially mutagenic adducts. These adducts can be as simple as methyl groups or larger bulky adducts. Without proper repair, these lesions represent a significant threat to genome stability. Lesions caused by alkylation are efficiently repaired by the BER pathway (discussed in the following section); however, they can also be directly repaired by alkyl transferases and AlkB-family dioxygenases (reviewed in Refs [23] and [24]).

Direct reversal of alkylation damage in *E. coli* is mediated by either the general housekeeping alkyltransferase Ogt, or an adaptive response controlled by the Ada protein that is mediated by its targets *alkA* and *alkB* [25]. O^6-methylguanine (6-meG) is mutagenic due to its ability to induce G:C to A:T transitions during DNA replication due to faulty base pairing. The Ada protein is a bifunctional alkyltransferase: the N-terminal domain (N-Ada) repairs methylphosphotriester lesions (damage to the DNA backbone that is generally innocuous to cells), while the C-terminal domain (C-Ada) repairs the much more potent alkyl lesions at the O^6 position of guanine. The direct reversal reaction occurs via the transfer of the alkyl group from the damaged base onto a reactive cysteine residue via an S_N2 reaction, thus permanently inactivating the protein.

The AlkB dioxygenase is similar to Ada in that it catalyzes the direct reversal of base alkylation damage. While the exact function of AlkB was difficult to determine (discussed in detail by Mishina et al. [24]), it was finally shown to catalyze the

direct reversal of 1-methyladenine (1-meA) and 3-methylcytosine (3-meC) to adenine and cytosine. The mechanisms of these two direct reversal pathways are different: while alkyltransferases use an S_N2 reaction, AlkB uses an oxidative dealkylation reaction that depends on an active site iron(II) atom.

At high temperatures, alkylated bases can be converted to abasic sites or DNA breaks that can cause irreversible chromosome fragmentation. From this perspective, it seems that alkylation repair should be highly developed and efficient in thermophilic archaea. In fact, however, the literature on alkylation repair in archaea is quite limited. Alkyltransferases from two thermophiles (*Methanococcus jannaschii* and *Sulfolobus taokadaii*) have been purified and crystalized and they show distinct structural similarities to the human homologs (PDB entry 1WRJ) [26]. Overexpression of a *Pyrococcus* methyltransferase in an *E. coli ogt* mutant can rescue its sensitivity to alkylating agents [27], confirming a functional conservation between bacterial and this archaeal methyltransferase. Other in vivo information on archaeal alkytransferase function has demonstrated that this functional conservation is probably a general feature (eg, see Refs [28] and [29]).

3.2 Base Excision Repair and Removal of Uracil from DNA

In addition to direct reversal, many organisms have another highly conserved (Table 4.1) pathway to repair damaged bases: BER (reviewed in detail in Ref. [30]). BER was first discovered in *E. coli* by Tomas Lindahl when he attempted to elucidate the pathway for the repair of genomic uracil, a byproduct of cytosine deamination [31]. Subsequent research over many years revealed that the cognate lesions for BER are base damage that does not cause major distortions in the DNA double helix, including oxidized bases (eg, 8-oxoguanine), alkylated bases (eg, 3-meA), deaminated bases (eg, hypoxanthine and xanthine), and uracil.

BER is initiated when a damage-specific DNA glycosylase recognizes a damaged base. *E. coli* has at least six glycosylases, while higher organisms tend to have more [32,33] (Table 4.1). The glycosylase activity hydrolyzes the *N*-glycosydic bond that connects the altered base to its sugar ring, leaving behind an abasic (AP) site. There are two types of glycosylases: monofunctional (such as *E. coli* AlkB) and bifunctional (such as *E. coli* Nei). Removal of bases by monofunctional glycosylases forms AP sites identical to those from spontaneous depurinations or depyrimidinations. These sites require further processing by an AP endonuclease. A bifunctional glycosylase also excises its cognate base (albeit by a different molecular mechanism), but can also incise the phosphodiester backbone at the abasic site to generate a single-strand break (SSB), excluding the need for a separate AP endonuclease. The incision in the phosphodiester backbone provides a 3′ hydroxyl group that is ultimately a substrate for DNA Pol I. The exonuclease function of Pol I removes the damaged strand and the polymerase activity synthesizes replacement DNA. Finally, the nick is sealed by DNA ligase.

An important consequence of BER is the suppression of mutagenesis due to biochemical properties of the damaged bases or uracil in the DNA. Occasional misincorporation of uracil-adjacent adenine during DNA replication is not inherently mutagenic; however, uracil formed via hydrolytic deamination of cytosine is unequivocally mutagenic. In *E. coli*, this misplaced uracil is removed mostly by the monofunctional uracil-DNA-glycosylase UNG (UDG) to avoid transition mutations.

BER enzymes have been found in archaea and the fundamental pathway is conserved [34], although archaeal BER components and their molecular biology are more similar to eukaryotes than to bacteria (for a detailed discussion of this topic, see Ref. [34]). The BER pathways of some archaea have novel features, while others use an additional mechanism to prevent mutation due to genomic uracil. *Ferroplasma acidarmanus* encodes a novel AGT protein (AGTendoV) that has an O^6-methyltransferase domain fused to an endonuclease V domain [35]. This bifunctional enzyme has been found in other archaeal genomes suggesting that it may be a general adaptation to their harsh environments. A more extreme deviation from the canonical BER pathway is the use of uracil-scanning DNA polymerases. In most cases, as noted earlier, genomic uracil is removed by uracil-DNA-glycosylases. Bacterial polymerases, in general, replicate past uracil by inserting an adenine (preserving the sequence); cases where uracil forms via cytosine deamination lead to CG-to-TA mutations [36]. In contrast, some archaeal replicative polymerases stall before misplaced uracils, representing a "read-ahead" proofreading function not found in bacteria or eukaryotes [37,38]. As uracil nucleotides are efficiently removed from these strains, the polymerase must somehow pass off the uracil to another protein for repair before continuing DNA replication. This idea was supported experimentally by Dionne and Bell when they demonstrated that in *Sulfolobus solfataricus,* a uracil glycosylase (UDG1) interacts with the DNA replication processivity factor PCNA (*E. coli* beta clamp), potentially recruiting it to replication forks [39]. One hypothesis to explain this unusual phenomenon is that it provides a "time stamp" to distinguish newly synthesized DNA strands from their templates, analogous to the GATC methylation used by *E. coli* for strand discernment during MMR (see later); however, this idea remains to be explored [40]. While in most respects archaeal BER seem unremarkable, these unique features hint that other interesting discoveries remain to be made.

3.3 Nucleotide Excision Repair: A Versatile DNA-Repair Pathway

NER is a tremendously versatile DNA-repair system that is highly conserved from bacteria to humans (Table 4.1). It consists of two subpathways: global-genome NER (GG-NER), which monitors the entire genome for damage; and transcription-coupled NER (TC-NER), which repairs damage that specifically interferes with transcription (reviewed in detail in [41]). As mentioned in the previous section, expression of NER components (with the exception of UvrC) is regulated by the SOS response and the genes are some of the earliest to be expressed after the detection of DNA damage. The versatility of NER is largely due to its mode of damage recognition. Unlike BER, which recognizes and repairs specific lesions that have little to no effect on the structure of the DNA double helix, the NER pathway monitors the DNA for even small structural distortions. From one perspective, it could be said that the NER pathway repairs distortions, and as a consequence, removes the causal damage (for a list of damage repaired by NER, see Table 4.2).

Both branches of NER consist of four distinct stages: damage detection, damage verification, excision, and ligation.

In *E. coli*, damage is detected via collaboration between UvrA and UvrB in the GG-NER pathway. Alternatively, if the damage is first encountered by RNA polymerase (RNAP), leading to transcription stalling, the Mfd protein (also known as the transcriptional-repair coupling factor, or TRCF) displaces the stalled RNAP and recruits UvrAB to the damage site (TC-NER) [42]. How exactly UvrA and UvrB bind to the damaged DNA remains a challenging experimental problem that is discussed in detail in [41].

After this initial detection step, UvrB takes over and separates the two strands to verify the position of the lesion, simultaneously, leading to the release of UvrA. UvrB remains tightly associated with the DNA and acts as scaffold for UvrC, an enzyme with two nuclease domains. UvrC makes two cuts in the damaged strand, one eight nucleotides 5′ to the lesion and the other four to five nucleotides 3′ to the lesion. *E. coli, Mycobacterium, Salmonella,* and *Clostridium* (at least) have a relatively recently discovered alternative endonuclease called Cho [43], which is also SOS regulated. Cho differs from UvrC, in that it has a single nuclease domain and makes the 3′ incision four nucleotides further from the lesion than UvrC. The exact biological basis for this redundancy is not clear; however, it has been speculated that Cho may be required for especially large lesions that can't be accommodated by UvrC [43]. After incision, UvrC is ejected by the helicase II UvrD and DNA polymerase I (Pol I). Together, UvrD and Pol I displace the damaged strand and Pol I synthesizes the replacement strand using the undamaged sequence as a template, leading to highly accurate repair. Finally, DNA ligase seals the nick at the end of the newly synthesized strand.

TABLE 4.2 Selected Lesions Recognized and Repaired by NER

Damaging Agent	Lesion
4-Nitroquinolone oxide	N^2-deoxyguanosine adducts, and others
Aflatoxin-B1	Purine adducts, N^7-guanine, formamidopyrimidine
Anthramycin	N^2-guanine
N-acetoxy-2-acetylaminofluorene (AAF), *N*-hydroxyaminofluorene (AF)	C^8-guanine
N′-methyl-*N*-nitronitrosguanidine (MNNG)	O^6-methylguanine
Polycyclic aromatic hydrocarbons	N^2-guanine, benzo(a)pyrene diol epoxide, and others
Psoralin	Monoadducts (8-methoxypsoralen, 8-MOP)
Cisplatin	N^7-guanine, interstrand crosslinks
Chemically induced	DNA-protein/DNA-peptide linkages
Mitomycin C	N^7-guanine, O^6-methylguanine, N^2-guanine
Nitrogen mustard	Alkylation
UV radiation	Pyrimidine dimers (6-4) photoproducts

Adapted from Truglio JJ, Croteau DL, Van Houten B, Kisker C. Prokaryotic nucleotide excision repair: the UvrABC system. Chem Rev 2006;106:233–52.

The NER pathways in bacteria and archaea are functionally similar; however, despite the ancient nature of archaea, some aspects of their NER pathways are more similar to eukaryotic versions than to bacterial versions and they may or may not have *uvr* homologs [40,44,45]. The presence of clear *uvr* homologs seems to coincide with lifestyle: mesophilic archaea tend to have *uvr* genes, while hyperthermophilic archaea (HA) do not. A universal feature, however, seems to be the presence of homologs of eukaryotic factors. In two mesophilic species, which have both eukaryotic-like proteins and *uvr* homologs, *Methanobacterium thermoautotrophicum*, and *Halobacterium salinarum*, experimental evidence suggests that they use mostly or entirely the prokaryotic NER proteins [46,47]. Beyond the DNA damage detection stage, most eukaryotic NER factors have additional nonrepair functions; thus, the conservation of these proteins in archaea may simply reflect other cellular functions [48]. The genomes of HA do not encode *uvr* homologs [40]; furthermore, deletion of any of their eukaryotic-like NER genes has little to no effect on UV resistance [40]. These observations force the question: How do archaea that thrive in such harsh environments cope with the absence of such a versatile DNA-repair pathway? One hypothesis proposed by Dennis Grogan is that HA do not attempt to remove lesions before DNA replication and, instead, rely on interactions between replication forks and lesions for repair [40]. Grogan suggests that upon the collision between the replication fork and a blocking lesion, a cut is made in the ssDNA liberated by the unimpeded helicase. This cut would result in the formation of a double-strand end that would be a substrate for end processing; degradation of the double-strand end would remove the lesion. Homologous recombination would then be used to restore the fork for continued replication. This model remains to be fully tested; however, if it is proven true, it would establish a novel paradigm for the repair of many types of lesions.

3.4 Correcting Mismatched Bases: Cleanup After DNA Replication

The primary function of MMR is to remove bases incorrectly inserted by DNA polymerase during DNA replication and its importance is emphasized by its cross-domain functional and homologous conservation (Table 4.1). In *E. coli,* MMR can improve the accuracy of DNA replication up to 400-fold [49]. The *E. coli* MMR pathway has been reconstituted in vitro with only three MMR-specific proteins: MutS, MutL, and MutH [50]. The initiating step of the MMR pathway is the recognition and binding of a mismatched base in the dsDNA helix by a MutS dimer. A MutL dimer subsequently binds to the MutS–DNA complex, thereby stabilizing it and activating the MutH restriction endonuclease. MutH then nicks the strand containing the incorrectly incorporated base. The errant strand is then removed via helicase (UvrD) and exonuclease activities (ExoI/ExoVII/RecJ) and a new strand is synthesized by DNA polymerase III using the undamaged strand as a template. Finally, the nick is sealed by DNA ligase.

An obvious challenge for MMR is to identify which DNA strand has the misincorporated base. Given that the bases themselves are not informative in this respect, *E. coli* meets this challenge by monitoring the methylation status of the two DNA strands (although there is ongoing discussion on the absolute necessity of this activity [51]). As the fork proceeds during DNA replication, the daughter strand is methylated at GATC sites by the DNA adenine methyltransferase Dam. During a transient period, the newly synthesized dsDNA is hemimethylated, that is, only one strand is methylated. Different values for how long this hemimethylated state persists have been obtained using different experimental systems (see Refs [52–54] and [55] and references therein); however, it seems clear that hemimethylation can exist for just minutes after the replication fork passes and the period of hemimethylation limits the window of opportunity for MMR to discern the daughter strand. While GATC sites are overrepresented in the *E. coli* genome, one may not be in the direct proximity of the mismatched base. How the MMR complex can discern the strands in this situation continues to be discussed and debated [51,56,57]; however, a consensus seems to be that the reading of distant GATC sites may occur by two mechanisms: a *cis*-model, in which MutS translocates along the DNA, or a *trans*-model, in which a loop forms between the sites. These details remain an open question and further work is required to fully understand this aspect of MMR.

In the preceding discussions, the *E. coli* repair pathways have generally been used as basic models; however, *E. coli* MMR may be the exception, rather than the rule. Homologs of MutS and MutL are widely distributed, but MutH seems to be rare in other bacteria and archaea. In bacteria that lack MutH, the MutH nuclease activity seems to be replaced by a nuclease activity in MutL [58]. In this way, the MMR of *mutH*-less bacteria are more reminiscent of eukaryotic MMR, where MutLα is required for the incision step. For detailed information on eukaryotic MMR and some comparisons with *E. coli* MMR, see Ref. [59].

Mesothermophilic archaea tend to have MMR pathways that mirror the canonical bacterial pathways [40,60], although they likely originated from horizontal gene transfer [61]. In contrast, the HA lack MutS and MutL homologs (the same group that lacks canonical *uvr* homologs); however, despite the lack of MutS and MutL, genome replication is accurate in these organisms [40]. This lack of increased mutagenesis suggests that the hyperthermophiles have some mechanism that accomplishes the same net outcome as MMR. Solutions to this problem have been hypothesized, but not experimentally validated [40]. One idea is that reconfiguration of progressing replication forks into “chicken foot” structures might expose

mismatched bases in the newly synthesized strands for removal by end-processing enzymes—in effect giving the replication fork a "do over" (analogous to the model proposed earlier to replace NER).

Clearly, more work remains in order to understand how HA ensure the stability of their genome sequences under the harsh environmental conditions in which they live. It is difficult to understand why these organisms, which thrive in conditions that may pose the greatest threats to genome stability, lack some of the most universally conserved DNA-repair components (NER and MMR in particular), some of which have even evolved entirely independently in the three domains of life. Understanding the biological implications of this paradox may represent one of the greatest challenges in the DNA repair field and, while it is being addressed by only a relatively small number of groups, persistence may yield some of the most novel future breakthroughs in understanding the sources of genome stability.

A brief statement is necessary regarding the interaction between MMR and homologous recombination [62]. As both of these pathways function in tight association with the replication fork, they share both space and time. It is well established that MMR suppresses illegitimate recombination, especially highlighted by the observation that loss of MMR increases the frequency of interspecies DNA exchange between *E. coli* and *Salmonella* during conjugation [63,64]. Similar observations were also noted for transduction and transformation [65–67]. In this way, MMR can limit the impact of foreign DNA on genome stability, similar to restriction-modification systems (discussed in the following section). Despite the time since these observations, the molecular mechanisms underlying them were fully worked out in 2013 [68].

3.5 Recombination Repair: Dealing With Double-Strand Breaks

It is generally agreed that double-strand breaks (DSBs) in DNA represent the greatest threat to genome stability. Many exogenous and endogenous agents, including cosmic radiation and ionizing radiation; reactive oxygen species; replication fork malfunctions; and chemicals, can cause DSBs. In humans, defects in DSB repair can be potent precursors to cancer development. Because of the extreme importance of this pathway, it has been dissected in remarkable detail and reviewed extensively (for exhaustive reviews, see Refs [69,70]). This chapter presents a discussion focused on variations in the classical recombination pathways and highlights some interesting and important recent discoveries.

HR-dependent repair of DSBs can be distilled into discrete steps that are conserved from bacteria and archaea to eukaryotes (although the players in each step vary):

1. End resection. The broken ends of the DNA must be prepared for the subsequent molecular transactions.
2. Strand invasion. A single-stranded stretch of DNA terminating in a 3′-OH is guided into the duplex of a homologous molecule. This process is mediated by recombinases, including RecA (*E. coli*), Rad51 (many eukaryotes), and RadA (archaea).
3. Branch migration. Strand invasion leads to a four-strand branched intermediate. This intermediate is remodeled to facilitate new DNA synthesis and other molecular processes.
4. Holliday junction resolution. This step leads to the restoration of two DNA duplexes via strand cutting.

One of the most puzzling and unresolved aspects of homologous recombination is how homologous loci are located and brought together before the strand invasion can occur, and in particular, how far apart can the homologous molecules be before the homology search fails. This problem has been addressed in a study in *E. coli* that also highlights the power of superresolution microscopy to understand events in single bacterial cells. Lesterlin et al. [71] demonstrated that DSB-induced pairing of homologous sequences can occur even between distantly separated sister loci (already positioned for segregation into daughter cells). One interesting implication of this discovery is that HR-dependent repair of DSBs is not limited to the short period of time when two newly produced chromosomes are linked after DNA replication. A similar study in *C. crescentus* reinforced that this distant pairing ability is likely a general phenomenon on bacteria [72].

Resistance to radiation, which likely corresponds to an organism's ability to manage the deleterious outcomes of exposure, is not universal. The extremophilic bacteria *D. radiodurans* grows in environments with high levels of ionizing radiation (IR) supported by an exceptional DNA-repair faculty [73]. It was unclear for some time whether *D. radiodurans* possessed additional DNA-repair pathways, or whether it had more potent versions of known pathways. Michael Cox's lab set out to further understand the molecular basis of its radiation resistance by examining several evolved lines of *E. coli* obtained by repeatedly subjecting cultures to selection by IR exposure [74]. Quite remarkably, they were able to recover *E. coli* lines with three to four orders of magnitude higher resistance to 3000 Gy (a high dose) than the parental strain, on par with *D. radiodurans*. While these strains carried multiple mutations, it was ultimately determined that the increased resistance was conferred by mutations in only three genes: *recA*, *dnaB*, and *yfjK* [75]. The first two genes have well-understood functions in DNA repair demonstrating that extreme radiation resistance can arise via genetic innovations in existing pathways, and that additional protein machinery is not necessarily needed. Logically, increased DNA-repair activity would be

advantageous; however, this example shows that an organism may not realize its full potential in the absence of the proper selective pressures. In this case, *E. coli* has evolved a DNA-repair system that is just good enough to ensure the stability of its genome within its natural environments.

In most bacteria, mutations that completely abolish HR (eg, *recA* nulls in *E. coli*) are tolerated to varying degrees. Similarly, HR seems to be dispensable in some archaea; however, HR is an essential function in hyperthermophiles, as *radA* deletions are lethal [40]. Furthermore, archaeal homologs responsible for other central steps in HR are also essential (ie, Mre11, Rad50, HerA, and NurA). As discussed earlier, hyperthermophiles lack a conserved NER pathway; therefore, they should accumulate more DNA lesions that could inhibit DNA replication, transcription, or other DNA-related processes. It has been hypothesized that these functions are replaced by novel HR-dependent pathways [40] and their necessity in this pathway may underlie their essentiality.

Nonhomologous end joining (NHEJ) is another pathway for DSB repair. This pathway is error prone as it mediates the direct attachment of two DNA double-strand ends independent of extensive homology; thus, it is a last-resort effort as it almost certainly leads to heritable loss of significant amounts of genetic information. This pathway is perhaps best studied in eukaryotic models, but some bacteria have simplified versions (reviewed in Refs [76,77]). Not due to a lack of effort, a pathway for NHEJ remained elusive in *E. coli* and it was generally accepted for many years that no pathway exists. A 2010 work, however, has demonstrated that *E. coli* strains do possess an end-joining mechanism, now called alternative end joining (A-EJ) [78]. This novel pathway, which does not share conserved factors with canonical NHEJ pathways, depends on bidirectional strand resection, frequent use of microhomology, and nontemplated DNA synthesis. Although conserved components of NHEJ were readily identified in archaea [79–81], a functional repair pathway was identified in 2013 in a mesophilic archaeon [82]. Certainly, further study of end-joining in bacteria and archaea will yield further insight, and perhaps some additional surprises, into this complex DNA-repair pathway.

4. RESTRICTION-MODIFICATION SYSTEMS: PROTECTING THE GENOME FROM INVADERS

DNA damage at the atomic level can be catastrophic, potentially leading to mutations and loss of genetic information; however, genome stability is also threatened on a larger scale by various genetic parasites, including bacteriophage, plasmids, and other specialized genetic elements. In 1978, Werner Arber, Daniel Nathans, and Hamilton Smith won the Nobel Prize for Physiology or Medicine "for the discovery of restriction enzymes and their application to problems of molecular genetics." The first observations of the phenotypic readouts of restriction-modification systems were quite early in the 1950s when it was noticed that some *E. coli* strains were more resistant to bacteriophages (bacterial viruses) than others, leading to the use of the term "restriction" [83,84]. It was also noted that some bacteriophage escaped restriction and were able to infect their host. These modified strains could then be propagated on the original bacterial host, but the resistance was absent when the bacteriophages were transferred to new hosts. Thus, it was clear that the ability of bacteriophage to productively infect their host was controlled by a two-part process in which a pathway restricting infection competed with some type of modification that alleviated the restriction (Fig. 4.2).

Along with later work by Hamilton Smith, in which he purified the first restriction enzyme [85], these discoveries formed the foundation of our current understanding of restriction-modification systems: that cells can distinguish self and foreign DNA, and that the latter is destroyed. In this way, cells could protect their genomes from alteration due to the introduction of nonself-DNA. We now know that restriction-modification systems are widespread in prokaryotes and they have been found in both Bacteria and Archaea. Wild isolates of *E. coli* K-12 express the EcoKI enzyme, encoded by the *hsdRMS* genes (missing in most laboratory strains), as well as three other systems encoded by the *mcrA, mcrBC,* and *mrr* genes (one or more of which exist in many laboratory strains). The first restriction enzyme purified by Smith came from *Haemophilus influenzae* and the extreme thermophilic archaeon *Pyrococcus* encodes a thermostable restriction-modification system [86].

The precise molecular components of restriction-modification systems are diverse and they have been divided into four major groups (I–IV) based on several properties: structure, energy requirement, and cleavage mechanism. In general, all restriction-modification systems function on the same basic molecular principle to distinguish self and foreign DNA. One enzyme encodes a methyltransferase that modifies self-DNA via the addition of methyl groups to specific sequences. Another complementary enzyme recognizes the same sequences and, when they are unmodified, cuts the DNA by hydrolyzing the phosphodiester backbones of both strands. Depending on the group, some restriction-modification systems include additional factors for more complex biochemical activities. The genetic loci that encode restriction-modification systems can be quite complex and often encode variable but coregulated genes [87]—some of which may include additional protective functions (eg, the *E. coli* anticodon nuclease, a defense against T4 phage infection [88]).

The stability of prokaryotic genomes is challenged by three processes that allow the intercellular transfer of genetic material: transformation, transduction, and conjugation. It is clear from sequence analysis of bacterial genomes that genetic

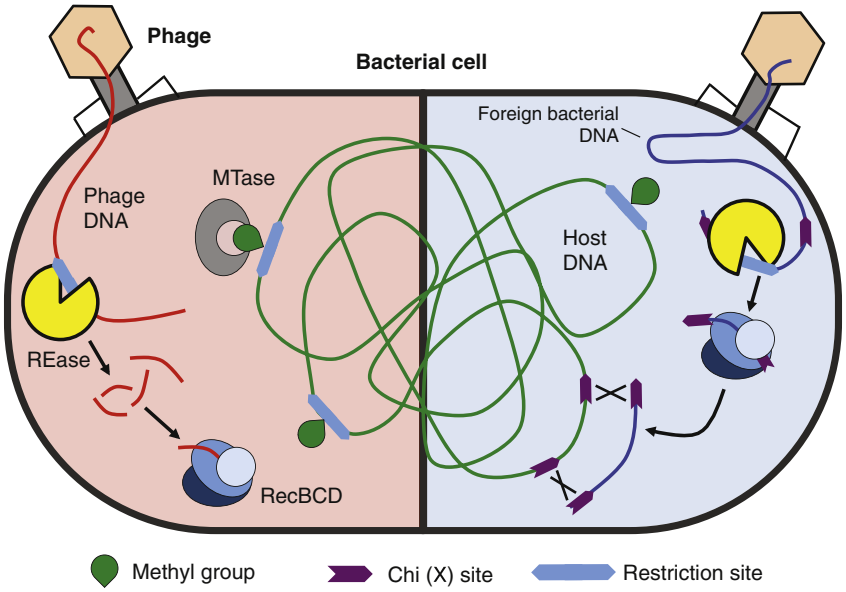

FIGURE 4.2 **Restriction-modification systems control the flux of foreign DNA.** *Red:* Unmethylated bacteriophage DNA is recognized and cleaved by restriction endonuclease (REase). The host DNA is not cut because the restriction enzyme recognition sites are methylated by a sequence-specific DNA methyltransferase (MTase). After cleavage by the REase, the DNA is degraded by the RecBCD complex. *Blue:* A bacteriophage injects nonphage DNA derived from other bacteria (transduction). Unless the DNA donor has a compatible restriction-endonuclease system, the DNA is cleaved by the donor REase. The cleaved DNA becomes a substrate for RecBCD, which degrades the DNA until it reaches a Chi site, a sequence-specific attenuator. DNA transferred from closely related species may recombine due to sequence homology. Since Chi sites are somewhat conserved between different bacterial species, homologous recombination between Chi sites in the foreign DNA and the host DNA can lead to integration of the foreign DNA into the genome.

exchange by these mechanisms has been extensive [89]. Invasions by foreign DNA can induce genome instability via interactions (ie, recombination) with the host chromosome. For example, upon infection, bacteriophages inject their genome into the host cell and, in many cases, the bacteriophage genome inserts into the host chromosome either at specific loci or nonspecifically. These insertions can represent powerful threats to genome integrity since they can disrupt coding or regulatory sequences, potentially disrupting genes or inducing potentially harmful gene expression changes (see Chapter X for further discussion).

Restriction-modification systems can protect cells from plasmids [90–92] and from DNA taken up by transformation (for an example, see Ref. [93]), but the best characterized example is the containment of incoming bacteriophage DNA. Upon entry into a host cell, bacteriophage DNA is, under normal conditions, unmethylated. At this junction, two outcomes are possible: the first option is that the DNA is rapidly cleaved by the restriction enzyme to prevent infection; the second option is that the DNA is methylated by the host methylase, preventing cleavage and supporting infection. As restriction enzymes tend to be more active than methylases, the balance is generally shifted in favor of protection. Support for the idea that restriction-modification systems protect host DNA against invading bacteriophage DNA is further provided by the presence of extensive countermeasures in bacteriophage to circumvent the protective functions [94].

DNA fragmentation by restriction enzymes can also stimulate recombination (eg, see Refs [95,96]), suggesting an alternative way that restriction-modification systems can influence genome stability. In this case, instead of limiting the effects of foreign genetic material on the genome, a restriction-modification system could support the incorporation of novel DNA via recombination [97]. McKane and Milkman [98] demonstrated that when chromosomal DNA from divergent *E. coli* strains was transduced in the laboratory strain K-12, the recombinational replacements were smaller (8–14 kb) than the fragment of DNA injected by the phage (about 100 kb) and that the foreign DNA was inserted in discrete units. These results suggested a model in which the incoming DNA was cleaved by the host restriction-modification system into smaller fragments that were subsequently integrated into the host genome via recombination. As restriction-modification-stimulated recombination in *E. coli* seems to be primarily mediated by the RecBCD end-processing complex [96], it is likely that the incorporation of foreign DNA via this mechanism may be limited to exchanges between closely-related species. One idea is that RecBCD degrades the fragmented DNA until it reaches its control element Chi. Since Chi sequences are well conserved among bacteria, they may serve as substrates for recombination. Ongoing research has revealed even greater complexity in the bacterial response to invading DNA, most notable the CRISPR/Cas system. This pathway represents a

primitive type of adaptive immunity and has led to rapid advancements in genome-editing capabilities in several model organisms.

These examples illustrate how restriction-modification systems could simultaneously protect the genome against foreign DNA and introduce genome instability by promoting the integration of foreign DNA. Since restriction-modification systems are present in most bacteria and archaea and the threats to genome stability from foreign parasitic DNA are ubiquitous, it is likely that these may be universal functions.

5. CONCLUSION

DNA has been successfully extracted from a number of ancient organic samples (in one study, as old as 13,000 years [99]) and, despite the lack of postmortem DNA-repair processes and exposure to harsh environmental conditions, some of this DNA has still been suitable for molecular analysis. That ancient DNA can be recovered and manipulated is a testament to its remarkable molecular stability. DNA in living cells is incomparably more stable than DNA in nonliving tissue, even though it is also constantly bombarded by both endogenous and exogenous insults. Since even small defects in DNA sequence can have profound deleterious effects, both on cellular function and preservation of species, natural selection has bestowed highly efficient DNA-repair pathways upon all known life forms. Within domains and across the three domains of life (bacteria, archaea, and Eukarya), a basic core of DNA-repair pathways exists. Remarkably, however, some functionally equivalent pathways appear to have evolved entirely independently. That organisms have converged on this common set of pathways affirms the concept that challenges to genome stability are universal. The examples discussed in this chapter specifically illustrate some variations on these common themes, suggesting that the evolution of DNA repair pathways in different species was influenced by specific challenges experienced in their environments. It is clear that there are exciting opportunities for continued research in the fields of DNA repair and genome stability, even the most humble organisms—bacteria and archaea. Their cells offer many riddles to solve and many possibilities for new, exciting, and beneficial discoveries.

GLOSSARY

Extremophilic bacteria Bacteria that thrive in unusually extreme environments.
Hyperthermophilic archaea Archaea that thrive in high-heat environments.
Mesothermophilic archaea Archaea that live in moderate temperatures.
Processivity The ability of an enzyme to catalyze sequential reactions without disassociating from its substrate.

LIST OF ABBREVIATIONS

1-meA 1-Methyladenine
3-meC 3-Methylcytosine
6-4PP Pyrimidine (6-4) photoproduct
6-meG O^6-methylguanine
A-EJ Alternative end joining
AGT Alkylguanine alkyl transferase
BER Base excision repair
CPD Cyclobutane pyrimidine dimer
DSB Double-strand break
FAD Flavin adenine dinucleotide
GG-NER Global-genome NER
HR Homologous recombination
IR Ionizing radiation
MMR (Methyl-directed) mismatch repair
NER Nucleotide excision repair
Pol I DNA polymerase I
Pol IV DNA polymerase IV
Pol V DNA polymerase V
RNAP RNA polymerase
TC-NER Transcription-coupled NER
TRCF Transcriptional-repair coupling factor
UNG/UDG Uracil-DNA-glycosylase

REFERENCES

[1] Kreuzer KN. DNA damage responses in prokaryotes: regulating gene expression, modulating growth patterns, and manipulating replication forks. Cold Spring Harb Perspect Biol 2013;5:a012674.

[2] Indiani C, O'Donnell M. A proposal: source of single strand DNA that elicits the SOS response. Front Biosci (Landmark Ed) 2013;18:312–23.

[3] Simmons LA, et al. Comparison of responses to double-strand breaks between *Escherichia coli* and *Bacillus subtilis* reveals different requirements for SOS induction. J Bacteriol 2009;191:1152–61.

[4] Modell JW, Hopkins AC, Laub MT. A DNA damage checkpoint in *Caulobacter crescentus* inhibits cell division through a direct interaction with FtsW. Genes Dev 2011;25:1328–43.

[5] Bos J, Yakhnina AA, Gitai Z. BapE DNA endonuclease induces an apoptotic-like response to DNA damage in *Caulobacter*. Proc Natl Acad Sci USA 2012;109:18096–101.

[6] Rand L, et al. The majority of inducible DNA repair genes in *Mycobacterium tuberculosis* are induced independently of RecA. Mol Microbiol 2003;50:1031–42.

[7] Gamulin V, Cetkovic H, Ahel I. Identification of a promoter motif regulating the major DNA damage response mechanism of *Mycobacterium tuberculosis*. FEMS Microbiol Lett 2004;238:57–63.

[8] Wang Y, Huang Y, Xue C, He Y, He Z-G. ClpR protein-like regulator specifically recognizes RecA protein-independent promoter motif and broadly regulates expression of DNA damage-inducible genes in mycobacteria. J Biol Chem 2011;286:31159–67.

[9] Lu H, Chen H, Xu G, Shah AM-U-H, Hua Y. DNA binding is essential for PprI function in response to radiation damage in *Deinococcus radiodurans*. DNA Repair 2012;11:139–45.

[10] Fröls S, et al. Response of the hyperthermophilic archaeon *Sulfolobus solfataricus* to UV damage. J Bacteriol 2007;189:8708–18.

[11] Ajon M, et al. UV-inducible DNA exchange in hyperthermophilic archaea mediated by type IV pili. Mol Microbiol 2011;82:807–17.

[12] Grogan DW. Exchange of genetic markers at extremely high temperatures in the archaeon *Sulfolobus acidocaldarius*. J Bacteriol 1996;178:3207–11.

[13] van Wolferen M, Ma X, Albers S-V. DNA processing proteins involved in the UV-induced stress response of *Sulfolobales*. J Bacteriol 2015. http://dx.doi.org/10.1128/JB.00344-15.

[14] Müller M, Carell T. Structural biology of DNA photolyases and cryptochromes. Curr Opin Struct Biol 2009;19:277–85.

[15] Brettel K, Byrdin M. Reaction mechanisms of DNA photolyase. Curr Opin Struct Biol 2010;20:693–701.

[16] Husain I, Sancar A. Photoreactivation in *phr* mutants of *Escherichia coli* K-12. J Bacteriol 1987;169:2367–72.

[17] Harm W, Harm H, Rupert CS. Analysis of photoenzymatic repair of UV lesions in DNA by single light flashes. II. In vivo studies with *Escherichia coli* cells and bacteriophage. Mutat Res 1968;6:371–85.

[18] Payne NS, Sancar A. The LexA protein does not bind specifically to the two SOS box-like sequences immediately 5' to the *phr* gene. Mutat Res 1989;218:207–10.

[19] Zhang F, Scheerer P, Oberpichler I, Lamparter T, Krauß N. Crystal structure of a prokaryotic (6-4) photolyase with an Fe-S cluster and a 6,7-dimethyl-8-ribityllumazine antenna chromophore. Proc Natl Acad Sci USA 2013;110:7217–22.

[20] Fitt PS, Sharma N, Castellanos G. A comparison of liquid-holding recovery and photoreactivation in halophilic and non-halophilic bacteria. Biochim Biophys Acta 1983;739:73–8.

[21] Kiener A, Gall R, Rechsteiner T, Leisinger T. Photoreactivation in *Methanobacterium thermoautotrophicum*. Arch Microbiol 1985;143:147–50.

[22] Wood ER, Ghané F, Grogan DW. Genetic responses of the thermophilic archaeon *Sulfolobus acidocaldarius* to short-wavelength UV light. J Bacteriol 1997;179:5693–8.

[23] Yi C, He C. DNA repair by reversal of DNA damage. Cold Spring Harb Perspect Biol 2013;5:a012575.

[24] Mishina Y, Duguid EM, He C. Direct reversal of DNA alkylation damage. Chem Rev 2006;106:215–32.

[25] Kleibl K. Molecular mechanisms of adaptive response to alkylating agents in *Escherichia coli* and some remarks on O(6)-methylguanine DNA-methyltransferase in other organisms. Mutat Res 2002;512:67–84.

[26] Roberts A, Pelton JG, Wemmer DE. Structural studies of MJ1529, an O^6-methylguanine-DNA methyltransferase. Magn Reson Chem 2006;44. Spec No, S71–S82.

[27] Leclere MM, et al. The O6-methylguanine-DNA methyltransferase from the hyperthermophilic archaeon *Pyrococcus* sp. KOD1: a thermostable repair enzyme. Mol Gen Genet 1998;258:69–77.

[28] Kanugula S, Pegg AE. Alkylation damage repair protein O^6-alkylguanine-DNA alkyltransferase from the hyperthermophiles *Aquifex aeolicus* and *Archaeoglobus fulgidus*. Biochem J 2003;375:449–55.

[29] Perugino G, et al. Activity and regulation of archaeal DNA alkyltransferase: conserved protein involved in repair of DNA alkylation damage. J Biol Chem 2012;287:4222–31.

[30] Krokan HE, Bjørås M. Base excision repair. Cold Spring Harb Perspect Biol 2013;5:a012583.

[31] Lindahl T. An N-glycosidase from *Escherichia coli* that releases free uracil from DNA containing deaminated cytosine residues. Proc Natl Acad Sci USA 1974;71:3649–53.

[32] Fromme JC, Banerjee A, Verdine GL. DNA glycosylase recognition and catalysis. Curr Opin Struct Biol 2004;14:43–9.

[33] Jacobs AL, Schär P. DNA glycosylases: in DNA repair and beyond. Chromosoma 2012;121:1–20.

[34] Grasso S, Tell G. Base excision repair in Archaea: back to the future in DNA repair. DNA Repair 2014;21:148–57.

[35] Kanugula S, Pauly GT, Moschel RC, Pegg AE. A bifunctional DNA repair protein from *Ferroplasma acidarmanus* exhibits O^6-alkylguanine-DNA alkyltransferase and endonuclease V activities. Proc Natl Acad Sci USA 2005;102:3617–22.

[36] Lindahl T. Instability and decay of the primary structure of DNA. Nature 1993;362:709–15.
[37] Greagg MA, et al. A read-ahead function in archaeal DNA polymerases detects promutagenic template-strand uracil. Proc Natl Acad Sci USA 1999;96:9045–50.
[38] Wardle J, et al. Uracil recognition by replicative DNA polymerases is limited to the archaea, not occurring with bacteria and eukarya. Nucleic Acids Res 2008;36:705–11.
[39] Dionne I, Bell SD. Characterization of an archaeal family 4 uracil DNA glycosylase and its interaction with PCNA and chromatin proteins. Biochem J 2005;387:859–63.
[40] Grogan DW. Understanding DNA repair in hyperthermophilic archaea: persistent gaps and other reasons to focus on the fork. Archaea 2015:1–12. http://dx.doi.org/10.1155/2015/942605.
[41] Kisker C, Kuper J, Van Houten B. Prokaryotic nucleotide excision repair. Cold Spring Harb Perspect Biol 2013;5:a012591.
[42] Selby CP, Sancar A. Molecular mechanism of transcription-repair coupling. Science 1993;260:53–8.
[43] Moolenaar GF, van Rossum-Fikkert S, van Kesteren M, Goosen N. Cho, a second endonuclease involved in *Escherichia coli* nucleotide excision repair. Proc Natl Acad Sci USA 2002;99:1467–72.
[44] White MF. Archaeal DNA repair: paradigms and puzzles. Biochem Soc Trans 2003;31:690–3.
[45] Grogan DW. The question of DNA repair in hyperthermophilic archaea. Trends Microbiol 2000;8:180–5.
[46] Ogrünç M, Becker DF, Ragsdale SW, Sancar A. Nucleotide excision repair in the third kingdom. J Bacteriol 1998;180:5796–8.
[47] Crowley DJ, et al. The *uvrA, uvrB* and *uvrC* genes are required for repair of ultraviolet light induced DNA photoproducts in *Halobacterium* sp. NRC-1. Saline Syst 2006;2:11.
[48] Rouillon C, White MF. The evolution and mechanisms of nucleotide excision repair proteins. Res Microbiol 2011;162:19–26.
[49] Schaaper RM. Base selection, proofreading, and mismatch repair during DNA replication in *Escherichia coli*. J Biol Chem 1993;268:23762–5.
[50] Lahue RS, Au KG, Modrich P. DNA mismatch correction in a defined system. Science 1989;245:160–4.
[51] Sneppen K, Semsey S. Mismatch repair at stop codons is directed independent of GATC methylation on the *Escherichia coli* chromosome. Sci Rep 2014;4:7346.
[52] Razin A, Szyf M. DNA methylation patterns. Formation and function. Biochim Biophys Acta 1984;782:331–42.
[53] Szyf M, et al. DNA methylation pattern is determined by the intracellular level of the methylase. Proc Natl Acad Sci USA 1984;81:3278–82.
[54] Szyf M, Gruenbaum Y, Urieli-Shoval S, Razin A. Studies on the biological role of DNA methylation: V. The pattern of *E.coli* DNA methylation. Nucleic Acids Res 1982;10:7247–59.
[55] Barras F, Marinus MG. The great GATC: DNA methylation in *E. coli*. Trends Genet 1989;5:139–43.
[56] Wang H, Hays JB. Signaling from DNA mispairs to mismatch-repair excision sites despite intervening blockades. EMBO J 2004;23:2126–33.
[57] Kolodner RD, Mendillo ML, Putnam CD. Coupling distant sites in DNA during DNA mismatch repair. Proc Natl Acad Sci USA 2007;104:12953–4.
[58] Fukui K, Nishida M, Nakagawa N, Masui R, Kuramitsu S. Bound nucleotide controls the endonuclease activity of mismatch repair enzyme MutL. J Biol Chem 2008;283:12136–45.
[59] Jiricny J. Postreplicative mismatch repair. Cold Spring Harb Perspect Biol 2013;5:a012633.
[60] Grogan DW. Stability and repair of DNA in hyperthermophilic Archaea. Curr Issues Mol Biol 2004;6:137–44.
[61] Lin Z, Nei M, Ma H. The origins and early evolution of DNA mismatch repair genes–multiple horizontal gene transfers and co-evolution. Nucleic Acids Res 2007;35:7591–603.
[62] Spies M, Fishel R. Mismatch repair during homologous and homeologous recombination. Cold Spring Harb Perspect Biol 2015;7:a022657.
[63] Rayssiguier C, Dohet C, Radman M. Interspecific recombination between *Escherichia coli* and *Salmonella typhimurium* occurs by the RecABCD pathway. Biochimie 1991;73:371–4.
[64] Rayssiguier C, Thaler DS, Radman M. The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch-repair mutants. Nature 1989;342:396–401.
[65] Zahrt TC, Maloy S. Barriers to recombination between closely related bacteria: MutS and RecBCD inhibit recombination between *Salmonella typhimurium* and *Salmonella typhi*. Proc Natl Acad Sci USA 1997;94:9786–91.
[66] Zahrt TC, Mora GC, Maloy S. Inactivation of mismatch repair overcomes the barrier to transduction between *Salmonella typhimurium* and *Salmonella typhi*. J Bacteriol 1994;176:1527–9.
[67] Majewski J, Zawadzki P, Pickerill P, Cohan FM, Dowson CG. Barriers to genetic exchange between bacterial species: *Streptococcus pneumoniae* transformation. J Bacteriol 2000;182:1016–23.
[68] Tham K-C, et al. Mismatch repair inhibits homeologous recombination via coordinated directional unwinding of trapped DNA structures. Mol Cell 2013;51:326–37.
[69] Kowalczykowski SC, Dixon DA, Eggleston AK, Lauder SD, Rehrauer WM. Biochemistry of homologous recombination in *Escherichia coli*. Microbiol Rev 1994;58:401–65.
[70] Persky NS, Lovett ST. Mechanisms of recombination: lessons from *E. coli*. Crit Rev Biochem Mol Biol 2008;43:347–70.
[71] Lesterlin C, Ball G, Schermelleh L, Sherratt DJ. RecA bundles mediate homology pairing between distant sisters during DNA break repair. Nature 2014;506:249–53.
[72] Badrinarayanan A, Le TBK, Laub MT. Rapid pairing and resegregation of distant homologous loci enables double-strand break repair in bacteria. J Cell Biol 2015;210:385–400.
[73] Cox MM, Battista JR. *Deinococcus radiodurans* – the consummate survivor. Nat Rev Microbiol 2005;3:882–92.

[74] Harris DR, et al. Directed evolution of ionizing radiation resistance in *Escherichia coli*. J Bacteriol 2009;191:5240–52.
[75] Byrne RT, et al. Evolution of extreme resistance to ionizing radiation via genetic adaptation of DNA repair. eLife 2014;3:e01322.
[76] Shuman S, Glickman MS. Bacterial DNA repair by non-homologous end joining. Nat Rev Microbiol 2007;5:852–61.
[77] Pitcher RS, Brissett NC, Doherty AJ. Nonhomologous end-joining in bacteria: a microbial perspective. Annu Rev Microbiol 2007;61:259–82.
[78] Chayot R, Montagne B, Mazel D, Ricchetti M. An end-joining repair mechanism in *Escherichia coli*. Proc Natl Acad Sci USA 2010;107:2141–6.
[79] Weller GR, et al. Identification of a DNA nonhomologous end-joining complex in bacteria. Science 2002;297:1686–9.
[80] Nair PA, Smith P, Shuman S. Structure of bacterial LigD 3'-phosphoesterase unveils a DNA repair superfamily. Proc Natl Acad Sci USA 2010;107:12822–7.
[81] Aravind L, Koonin EV. Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. Genome Res 2001;11:1365–74.
[82] Bartlett EJ, Brissett NC, Doherty AJ. Ribonucleolytic resection is required for repair of strand displaced nonhomologous end-joining intermediates. Proc Natl Acad Sci USA 2013;110:E1984–91.
[83] Luria SE, Human ML. A nonhereditary, host-induced variation of bacterial viruses. J Bacteriol 1952;64:557–69.
[84] Bertani G, Weigle JJ. Host controlled variation in bacterial viruses. J Bacteriol 1953;65:113–21.
[85] Smith HO, Wilcox KW. A restriction enzyme from *Hemophilus influenzae*. I. Purification and general properties. J Mol Biol 1970;51:379–91.
[86] Morgan R, Xiao JP, Xu SY. Characterization of an extremely thermostable restriction enzyme, PspGI, from a *Pyrococcus* strain and cloning of the PspGI restriction-modification system in *Escherichia coli*. Appl Environ Microbiol 1998;64:3669–73.
[87] Makarova KS, Wolf YI, Snir S, Koonin EV. Defense islands in bacterial and archaeal genomes and prediction of novel defense systems. J Bacteriol 2011;193:6039–56.
[88] Kaufmann G. Anticodon nucleases. Trends Biochem Sci 2000;25:70–4.
[89] Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. Nature 2000;405:299–304.
[90] Pinedo CA, Smets BF. Conjugal TOL transfer from *Pseudomonas putida* to *Pseudomonas aeruginosa*: effects of restriction proficiency, toxicant exposure, cell density ratios, and conjugation detection method on observed transfer efficiencies. Appl Environ Microbiol 2005;71:51–7.
[91] Purdy D, et al. Conjugative transfer of clostridial shuttle vectors from *Escherichia coli* to *Clostridium difficile* through circumvention of the restriction barrier. Mol Microbiol 2002;46:439–52.
[92] Moser DP, Zarka D, Kallas T. Characterization of a restriction barrier and electrotransformation of the cyanobacterium *Nostoc* PCC 7121. Arch Microbiol 1993;160:229–37.
[93] Stein DC, Gregoire S, Piekarowicz A. Restriction of plasmid DNA during transformation but not conjugation in *Neisseria gonorrhoeae*. Infect Immun 1988;56:112–6.
[94] Labrie SJ, Samson JE, Moineau S. Bacteriophage resistance mechanisms. Nat Rev Microbiol 2010;8:317–27.
[95] Thaler DS, Stahl MM, Stahl FW. Tests of the double-strand-break repair model for red-mediated recombination of phage lambda and plasmid lambda dv. Genetics 1987;116:501–11.
[96] Stahl MM, Kobayashi I, Stahl FW, Huntington SK. Activation of Chi, a recombinator, by the action of an endonuclease at a distant site. Proc Natl Acad Sci USA 1983;80:2310–3.
[97] Price C, Bickle TA. A possible role for DNA restriction in bacterial evolution. Microbiol Sci 1986;3:296–9.
[98] McKane M, Milkman R. Transduction, restriction and recombination patterns in *Escherichia coli*. Genetics 1995;139:35–43.
[99] Pääbo S. Ancient DNA: extraction, characterization, molecular cloning, and enzymatic amplification. Proc Natl Acad Sci USA 1989;86:1939–43.
[100] Truglio JJ, Croteau DL, Van Houten B, Kisker C. Prokaryotic nucleotide excision repair: the UvrABC system. Chem Rev 2006;106:233–52.