

The promiscuous primase

Si-houy Lao-Sirieix, Luca Pellegrini and Stephen D. Bell

MRC Cancer Cell Unit, Hutchison MRC Research Centre, Hills Road, Cambridge, UK, CB2 2XZ

DNA primases are essential for the initiation of DNA replication and progression of the replication fork. Recent phylogenetic analyses coupled with biochemical and structural studies have revealed that the arrangement of catalytic residues within the archaeal and eukaryotic primase has significant similarity to those of the Pol X family of DNA-repair polymerases. Furthermore, two additional groups of enzymes, the ligase/primase of the bacterial nonhomologous end-joining machinery and a putative replicase from an archaeal plasmid have shown striking functional and structural similarities to the core primase. The promiscuous nature of the archaeal primases suggests that these proteins might have additional roles in DNA repair in the archaea.

Subunit architecture of archaeal and eukaryotic primases

DNA primase has a pivotal role in DNA synthesis by making a short oligoribonucleotide that acts as a primer for DNA polymerase. Thus, the primase is essential for both leading and lagging strand synthesis. In principle the primase needs to act once only on the leading strand, but is required for the initiation of every Okazaki fragment on the lagging strand. The DNA primase of eukaryotes and archaea is distinct in subunit composition, sequence and structure from the bacterial analogue, DnaG [1]. In contrast to monomeric DnaG, archaeal primases have two subunits [2], and in eukaryotes homologues of these two subunits further interact with two additional components to form the Pol α /primase assembly [1]. In this review, we shall refer to the two subunits shared by archaea and eukaryotes as the core primase and refer to the tetrameric eukaryotic assembly of core primase in complex with the B subunit and Pol α as the Pol α /primase (Table 1). The catalytic activity of the primase lies in the small subunit of the core primase [1]. Although the monomeric single catalytic subunit can be purified in recombinant form, interaction with the larger core primase subunits appears to both stabilize the catalytic subunit and also modulate its activity to some degree. Perhaps the most dramatic example of this has come from studies of the primase of *Pyrococcus* species. Initial studies of this hyperthermophilic enzyme focused on the catalytic subunit in isolation [3]. Remarkably, this enzyme was capable of synthesizing long (up to 6 kb) DNA strands in the absence of ribonucleotides, that is, it could both initiate and extend DNA chains. The isolated enzyme had little or no ability to synthesize RNA. However, when the heterodimeric enzyme was reconstituted it showed very

different properties: the DNA polymerase activity was substantially reduced and the RNA synthesis significantly stimulated [2]. An intriguing implication of this observation is that the large subunit might have the ability to modulate both processivity and substrate choice of the catalytic subunit. The mechanism by which this influence is exerted is currently unknown. The highly promiscuous nature of the archaeal primase is not restricted to the *Pyrococcus* enzyme; recent studies have revealed that the primase from the highly diverged archaeon *Sulfolobus solfataricus* also has the ability to initiate and extend both RNA and DNA chains for up to 1 kb or 7 kb, respectively [4].

In eukaryotes, the core primase synthesizes a short oligoribonucleotide primer (between 6 and 15 nt long, depending on the species studied), which is then extended by DNA synthesized by the Pol α component of the Pol α -primase complex [1]. In this light, it is tempting to speculate that the dual RNA and DNA synthesis capabilities of archaeal primases could fulfil primer synthesis and DNA extension functions, respectively. However, there is currently no evidence for this hypothesis. Indeed, experiments from the Ishino laboratory indicate that primers synthesized by archaeal primase are extended by the replicative polymerase in a reconstituted *in vitro* system [2]. Although archaeal primase has the capacity to make DNA primers *in vitro*, two pieces of evidence cast doubt on the *in vivo* significance of the property. First, it has been established that archaeal Okazaki fragments have RNA at their 5' end [5], and second, measurement of the apparent K_m of the *Sulfolobus* primase for nucleoside triphosphates (NTPs) and deoxynucleoside triphosphates (dNTPs) reveals that the affinity for dNTPs is at least three orders of magnitude lower than that for NTPs [4]. In addition to possessing primase and polymerase activities, the *Sulfolobus* enzyme has recently been demonstrated to have 3' nucleotidyl terminal transferase activity [4,6].

Table 1. The subunit architecture of primase and primase/pol α from representative archaea and eukaryotes^a

	Eukaryotes		Archaea	
	Human	<i>S. cerevisiae</i>	<i>Sulfolobus</i>	<i>Pyrococcus</i>
Pol α	p165	Pol1 (180 kDa)	–	–
B subunit	p77	Pol12 (79 kDa)	–	–
Primase small	p50	Pri1 (48 kDa)	PriS (38 kDa)	Pfup41 (41 kDa)
Primase large	p59	Pri2 (58 kDa)	PriL (36 kDa)	Pfup46 (46 kDa)

^aThe names of the genes encoding the subunits are listed and the size of the protein products are indicated.

Corresponding author: Bell, S.D. (sb419@hutchison-mrc.cam.ac.uk). Available online 10 August 2005

Structural studies of archaeal primases

To date, crystal structures for the isolated catalytic subunit of archaeal primases from *P. horikoshii* and *P. furiosus* have been reported [7,8]. The crystallographic analysis showed that the catalytic subunit folds in a novel tertiary structure made up of two domains: a larger, mixed α/β domain, where the catalytic activity resides (prim domain), and a smaller α -helical domain of unknown function (Figure 1). Inserted onto the fold of the catalytic domain is a zinc-binding motif, which represents a conserved feature of archaeal and eukaryotic primases. The functional role of the zinc-binding motif is not yet known. Intriguingly, bacterial primases also contain a zinc-binding domain that has been implicated in binding single-stranded DNA [1,9]. It is, therefore, tempting to speculate that the zinc-binding motif in archaeal and eukaryotic primases might have an analogous role. However, it should be noted that, in contrast to the situation with the archaeal primase, the bacterial primase zinc-binding site lies in a domain distinct from the catalytic site.

Although the archaeal/eukaryotic primase catalytic subunit fold is novel, some features of the active site, as gleaned from the available structural data, suggest that the mode of catalysis is likely to be similar to the general mechanism of enzymatic synthesis of both RNA and DNA proposed earlier [10]. Thus, a triad of aspartate residues, necessary for catalytic activity and invariant across archaeal and eukaryotic primases, can be superimposed on the catalytic aspartates of the human DNA polymerase β [11], a member of the X family of polymerases (Figure 1 and below). It therefore appears that convergent evolution has driven primases to adopt the same mechanism of

catalysis, involving two divalent metal ions, proposed for other DNA and RNA polymerases [10].

The position of the active site was confirmed by diffusion of uridine 5'-triphosphate (UTP) in the *P. horikoshii* primase crystals [8]. The current model of primase catalysis postulates the existence of a rate-determining step when the enzyme is simultaneously bound to two nucleotides [1]. However, the structure of the primase-UTP complex provided no insight into the putative second nucleotide-binding event, which might therefore depend on the presence of the DNA template, or on the processing of the first nucleotide.

Relatives of primase

The family of archaeal/eukaryotic primase-related molecules has recently been extended with the characterization of the ORF904 product of the pRN1 plasmid from the hyperthermophilic archaeon *Sulfolobus islandicus* (Figure 2a). This large protein has multiple functional domains and possesses ATPase, primase and polymerase activity [12]. The primase and polymerase activities are found in the N-terminal domain of the protein and the ATPase activity is associated with a putative helicase domain in the C-terminal region of ORF904. Initial studies classified ORF904 as a member of a novel family of DNA polymerases, family E. Interestingly, the biochemical properties of ORF904 are reminiscent of the catalytic subunit of archaeal primase. More specifically, the primase can initiate both RNA- and DNA-strand synthesis, with a preference for DNA, and in the presence of dNTPs can extend DNA strands for several kilobases [12]. Although the ORF904 primase/polymerase domain shows little primary sequence homology with other nucleic

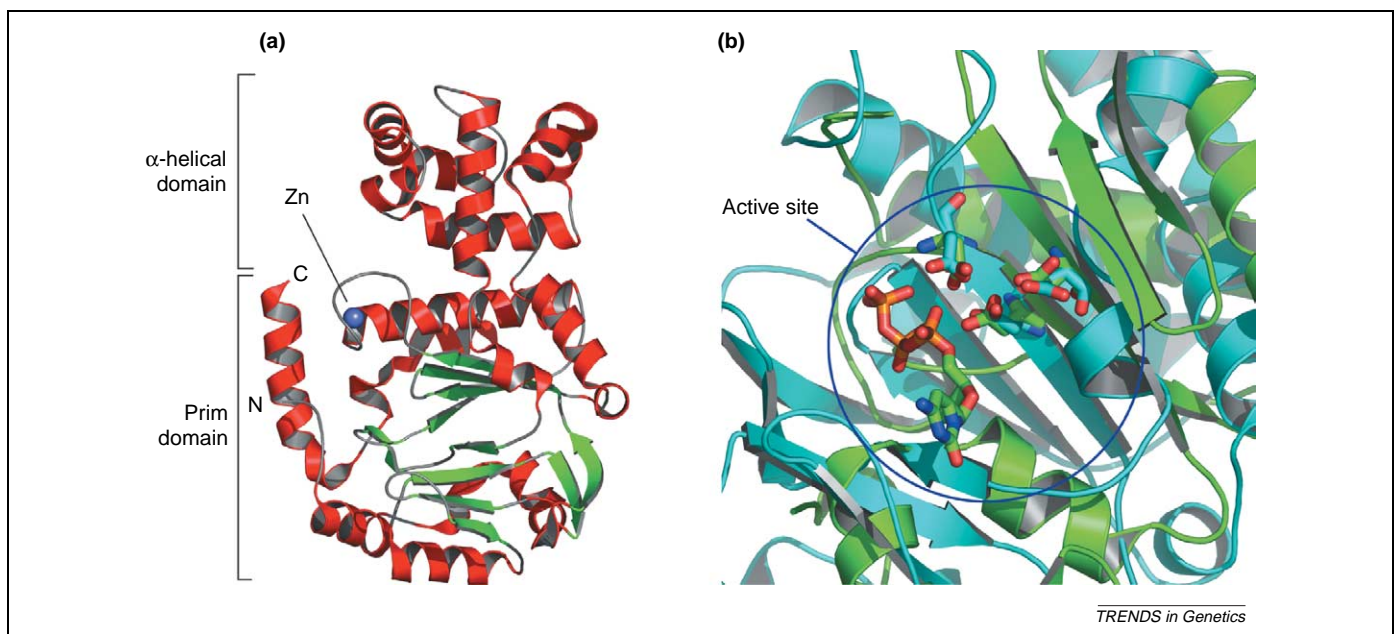


Figure 1. The catalytic subunit of the archaeal primase. **(a)** A cartoon representation of the crystallographic model for the catalytic subunit of the *Pyrococcus furiosus* primase [protein data bank (PDB) id: 1G71; <http://www.rcsb.org/pdb>]. Secondary structure elements are colour-coded (α helix is red and β strand is green). The position of the Zn ion is indicated as a blue sphere. The catalytic (prim) and α -helical domains are indicated by brackets. **(b)** Superposition of active site residues for the *P. horikoshii* primase (Asp95, Asp97 and Asp280) (PDB id: 1V33) and human DNA polymerase β (Asp190, Asp192 and Asp256) (PDB id: 2BPF). The side chains of the catalytic aspartic acid residues are shown as sticks, and the protein backbone as a ribbon (the primase is coloured in cyan and the pol β in green). The dideoxycytidine 5'-triphosphate present in the pol β active site is also shown.

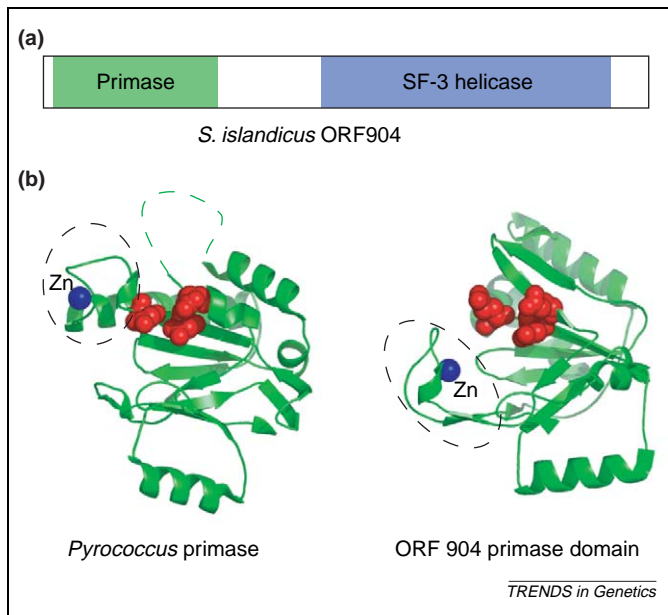


Figure 2. The ORF904 primase/helicase from *Sulfolobus islandicus* pRN1. (a) An illustration of the domain organization of ORF904, with regions homologous to primase and the superfamily 3 (SF-3) helicases indicated. (b) Comparison of the positions of the zinc-binding site in *Pyrococcus horikoshii* (left, PDB 1V33) and ORF904 primase (right, PDB 1RNI). The figures were produced using Pymol (<http://www.pymol.org>). The Zn ion is shown in blue and the catalytic residues shown in red. For clarity, residues 176–277 of *P. horikoshii* primase are not shown and are instead indicated by the green dashed line.

acid polymerizing enzymes, the recent elucidation of the crystal structure of this domain of ORF904 has revealed striking structural parallels with the architecture of the catalytic centre of the *Pyrococcus* archaeal primase [7,13]. In particular, the arrangement of metal coordinating acidic residues shows tight conservation within a region of β sheet. Both ORF904 and *Pyrococcus* enzymes have zinc-binding structural elements adjacent to the active centre of the enzyme. Surprisingly, however, these motifs are found in unrelated positions in the two enzymes (Figure 2b). Although it has been demonstrated that a positively charged cleft adjacent to the zinc-binding stem in ORF904 is important for template binding [13], it is not known what function the *Pyrococcus* zinc-binding motif performs. However, this observation suggests that ORF904 and archaeal/eukaryotic primase are derived from a common ancestor, and further implies that this common ancestor lacked a zinc-binding domain and that two separate insertion events have occurred and been selected for during the evolution of the ORF904 and primase families.

Bacterial mobile genetic elements encoding archaeal/eukaryotic primase proteins

The bacteriophage T7 encodes a protein that has an N-terminal domain related to the bacterial primase DnaG and a C-terminal helicase domain related to the bacterial replicative helicase DnaB [1]. A recent bioinformatics study has described the identification of a novel primase-helicase molecule found in an integrated phage in the bacterium *Bacillus cereus* (Figure 3). Remarkably, the primase and helicase domains show homology not to bacterial DnaG and DnaB, but rather to the archaeal and eukaryotic primase catalytic subunit and the mini

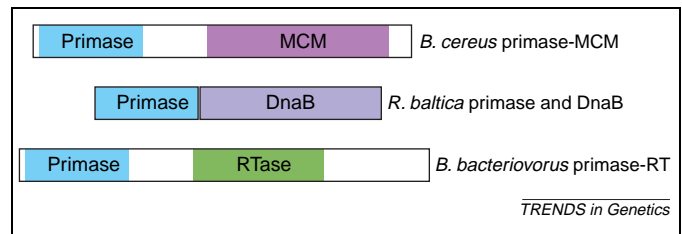


Figure 3. Novel primase-related proteins encoded in potentially mobile genetic elements in bacteria. In *Bacillus cereus* and *Bdellovibrio bacteriovorus*, the primase domain (blue) is fused to a MCM-related helicase domain and reverse transcriptase domain. In *Rhodospirellula baltica*, the primase is encoded immediately upstream of a gene encoding a phage DnaB-type helicase.

chromosome maintenance (MCM) presumptive replicative helicase [14]. Although the primase–MCM fusion is unique to this particular phage, it is interesting to note that the primase-like domain is found in several other contexts in other bacterial genomes. More specifically, as shown in Figure 3, *Rhodospirellula baltica* encodes an open reading frame (ORF) with high similarity to the *B. cereus* primase domain (BLAST score of $2 \times e^{-9}$). In *R. baltica*, this ORF encodes only the primase molecule. However, the gene encoding the primase lies immediately upstream of one of two *R. baltica* homologues of the helicase DnaB. Furthermore, this *dnaB* homologue appears to be related to the subfamily of phage DnaB-like helicases (Figure 3).

Additionally, in *Bdellovibrio bacteriovorus* HD100, a homologue of the *B. cereus* phage primase domain is found in the N-terminal domain of a large ORF that has homology to bacterial retron reverse transcriptases. With the ever-increasing number of bacterial genome sequences it will be interesting to determine how ubiquitous the primase domain is. The fact that thus far all three homologues appear associated with mobile genetic elements or phage strongly implies that this could be a broadly distributed molecule.

The primase–Pol X connection

The first hint that there could be an architectural relationship between the Pol X family of DNA polymerases and primases emerged from sequence comparisons of a range of eukaryotic DNA Pol Xs and the small subunit of archaeal and eukaryotic primases [15]. With the elucidation of the structure of human DNA Pol β and the first crystal structure of an archaeal primase, it became apparent that the primary sequence similarity was mirrored by a degree of structural similarity [7,11]. More specifically, as mentioned above, the positioning of the highly conserved catalytic aspartate residues of archaeal primase was superimposable on the catalytic core of DNA Pol β . Interestingly, however, the secondary structural contexts in which the aspartates are found are dissimilar between primase and DNA Pol β , suggesting divergent evolution. Because these aspartates are also structurally conserved among polymerases and are implicated in the two-metal-ion mechanism of DNA polymerization [10], Kirk and Kuchta suggested that the archaeal primase might possess similar catalytic mechanisms to the family X polymerases [15].

To date, five Pol X family members have been described in mammalian cells: Pol β , Pol λ , Pol μ , Pol σ and terminal

deoxynucleotidyl transferase (TdT) [16]. These enzymes have various roles in DNA replication, repair and recombination processes [17]. Pol β , the reference member of the Pol X family, is a template-dependent DNA polymerase that also possesses a 5'-deoxyribose 5'-phosphate (dRP) lyase activity. The principal role for Pol β is in base excision repair (BER). TdT has a 22% amino acid identity with Pol β . It is a template-independent DNA and RNA polymerase that has a key role in the generation of antibody diversity by the V(D)J recombination process. In pre-T-cell extracts, TdT associates with Ku, a heterodimeric protein required for V(D)J recombination and for nonhomologous end joining (NHEJ), suggesting a role for TdT in double-strand break (DSB) repair [18]. More recently, two unusual novel polymerases were added to the Pol X family: Pol λ , which has a high amino acid identity (32%) to Pol β [19,20] and Pol μ , which is more closely related to TdT (41% amino acid identity) [21]. Like the archaeal primases, both enzymes possess template-dependent and template-independent polymerase activities [22,23] and can synthesize DNA *de novo* [17]. Additionally, these Pol X family members are prone to frameshift synthesis [24] and can fill in short gaps [19], supporting a role in NHEJ. Finally, Pol λ has a dRP lyase activity used for BER *in vitro* although its role in this process *in vivo* is still unclear [25].

Additional primase-like molecules in DNA repair

A further extension to the family of primase-like molecules has been made with the exciting discovery of a pathway for NHEJ in prokaryotes. Initial bioinformatics studies by the Koonin, Jackson and Doherty laboratories revealing the presence of prokaryotic homologues of the essential NHEJ factor, Ku [26,27], have been extended with a series of biochemical and genetic analyses, primarily focusing on bacterial organisms encoding this pathway [28–30]. In contrast to the highly complicated eukaryotic NHEJ machinery, the prokaryotic apparatus appears to require just two gene products [28]. One is the homologue of Ku, and the other has been named as LigD. Intriguingly, like the ORF904 protein described above, LigD is a multifunctional enzyme. In the case of LigD, the enzyme has ligase, polymerase, nuclease and primase activities [28,30]. Significantly, examination of the sequence of this protein has revealed homology with the archaeal/eukaryotic primase catalytic subunit. Further, biochemical experiments have shown that, like archaeal primase, LigD can synthesize RNA primers, has DNA-dependent DNA and RNA polymerase activity and, like *Sulfolobus* primase, has 3' nucleotidyl terminal transferase activity [28,31].

Concluding remarks

Thus, the archaeal/eukaryotic primases and a range of repair and mobile genetic element DNA polymerases are clearly related in both the arrangement of the catalytic aspartates within the active centre of the enzymes and by a common promiscuity in the reactions they can carry out. It appears therefore that this arrangement of the catalytic site has a degree of functional malleability that is ideally suited to performing nucleic acid synthetic functions on a diverse range of substrates. This adaptability could be

significant in the context of the broad range of damaged DNA substrates that are dealt with by the repair polymerases. Interestingly, the archaeal primases appear to be capable of a broader spectrum of activities than their eukaryotic primase counterparts.

In light of the absence of members of the DNA-Pol-X family in archaeal species (with the exception of a single archaeal species, *Methanobacterium thermoautotrophicum*), it is tempting to speculate that the archaeal primases might have dual functionality, working both as *bona fide* primases but also playing roles in DNA-repair processes. Whether the relative efficiencies of the archaeal primases in DNA and RNA synthesis could be modulated by interaction with as yet unidentified factors remains an open question. If the archaeal primases also act as repair polymerases, then it is tempting to speculate that as eukaryotic cells evolved, the primase became dedicated to DNA replication and novel polymerases (perhaps co-opted from mobile genetic elements) adopted specific roles in DNA repair.

Acknowledgements

S-h.L.-S. and S.D.B. are funded by the Medical Research Council. L.P. is funded by the Wellcome Trust. We would like to thank Adam McGeoch for informative discussions.

References

- 1 Frick, D.N. and Richardson, C.C. (2001) DNA primases. *Annu. Rev. Biochem.* 70, 39–80
- 2 Liu, L.D. *et al.* (2001) The archaeal DNA primase – biochemical characterization of the p41-p46 complex from *Pyrococcus furiosus*. *J. Biol. Chem.* 276, 45484–45490
- 3 Bocquier, A.A. *et al.* (2001) Archaeal primase: bridging the gap between RNA and DNA polymerases. *Curr. Biol.* 11, 452–456
- 4 Lao-Sirieix, S.H. and Bell, S.D. (2004) The heterodimeric primase of the hyperthermophilic archaeon *Sulfolobus solfataricus* possesses DNA and RNA primase, polymerase and 3'-terminal nucleotidyl transferase activities. *J. Mol. Biol.* 344, 1251–1263
- 5 Matsunaga, F. *et al.* (2003) Identification of short 'eukaryotic' Okazaki fragments synthesized from a prokaryotic replication origin. *EMBO Rep.* 4, 154–158
- 6 De Falco, M. *et al.* (2004) The DNA primase of *Sulfolobus solfataricus* is activated by substrates containing a thymine-rich bubble and has a 3'-terminal nucleotidyl-transferase activity. *Nucleic Acids Res.* 32, 5223–5230
- 7 Augustin, M.A. *et al.* (2001) Crystal structure of a DNA-dependent RNA polymerase (DNA primase). *Nat. Struct. Biol.* 8, 57–61
- 8 Ito, N. *et al.* (2003) Crystal structure of the *Pyrococcus horikoshii* DNA primase-UTP complex: implications for the mechanism of primer synthesis. *Genes Cells* 8, 913–923
- 9 Pan, H. and Wigley, D.B. (2000) Structure of the zinc-binding domain of *Bacillus stearothermophilus* DNA primase. *Structure* 8, 231–239
- 10 Steitz, T.A. *et al.* (1994) A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science* 266, 2022–2025
- 11 Sawaya, M.R. *et al.* (1994) Crystal-structure of rat DNA-polymerase-beta – evidence for a common polymerase mechanism. *Science* 264, 1930–1935
- 12 Lipps, G. *et al.* (2003) A novel type of replicative enzyme harbouring ATPase, primase and DNA polymerase activity. *EMBO J.* 22, 2516–2525
- 13 Lipps, G. *et al.* (2004) Structure of a bifunctional DNA primase-polymerase. *Nat. Struct. Mol. Biol.* 11, 157–162
- 14 McGeoch, A.T. and Bell, S.D. (2005) Eukaryotic/archaeal primase and MCM proteins encoded in a bacteriophage genome. *Cell* 120, 167–168
- 15 Kirk, B.W. and Kuchta, R.D. (1999) Arg304 of human DNA primase is a key contributor to catalysis and NTP binding: primase and the family X polymerases share significant sequence homology. *Biochemistry* 38, 7727–7736

- 16 Hubscher, U. *et al.* (2002) Eukaryotic DNA polymerases. *Annu. Rev. Biochem.* 71, 133–163
- 17 Ramadan, K. *et al.* (2004) *De novo* DNA synthesis by human DNA polymerase lambda, DNA polymerase mu and terminal deoxyribonucleotidyl transferase. *J. Mol. Biol.* 339, 395–404
- 18 Mahajan, K.N. *et al.* (1999) Association of terminal deoxynucleotidyl transferase with Ku. *Proc. Natl. Acad. Sci. U. S. A.* 96, 13926–13931
- 19 Garcia-Diaz, M. *et al.* (2002) DNA polymerase lambda, a novel DNA repair enzyme in human cells. *J. Biol. Chem.* 277, 13184–13191
- 20 Garcia-Diaz, M. *et al.* (2000) DNA polymerase lambda (Pol lambda), a novel eukaryotic DNA polymerase with a potential role in meiosis. *J. Mol. Biol.* 301, 851–867
- 21 Dominguez, O. *et al.* (2000) DNA polymerase mu (Pol mu), homologous to TdT, could act as a DNA mutator in eukaryotic cells. *EMBO J.* 19, 1731–1742
- 22 Ruiz, J.F. *et al.* (2003) Lack of sugar discrimination by human Pol mu requires a single glycine residue. *Nucleic Acids Res.* 31, 4441–4449
- 23 Ramadan, K. *et al.* (2003) Human DNA polymerase lambda possesses terminal deoxyribonucleotidyl transferase activity and can elongate RNA primers: implications for novel functions. *J. Mol. Biol.* 328, 63–72
- 24 Bebenek, K. *et al.* (2003) The frameshift infidelity of human DNA polymerase lambda. Implications for function. *J. Biol. Chem.* 278, 34685–34690
- 25 Garcia-Diaz, M. *et al.* (2001) Identification of an intrinsic 5'-deoxyribose-5-phosphate lyase activity in human DNA polymerase lambda: a possible role in base excision repair. *J. Biol. Chem.* 276, 34659–34663
- 26 Aravind, L. and Koonin, E.V. (2001) 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.* 11, 1365–1374
- 27 Doherty, A.J. *et al.* (2001) Identification of bacterial homologues of the Ku DNA repair proteins. *FEBS Lett.* 500, 186–188
- 28 Della, M. *et al.* (2004) Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. *Science* 306, 683–685
- 29 Weller, G.R. *et al.* (2002) Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* 297, 1686–1689
- 30 Zhu, H. and Shuman, S. (2005) A primer-dependent polymerase function of *Pseudomonas aeruginosa* ATP-dependent DNA ligase (LigD). *J. Biol. Chem.* 280, 418–427
- 31 Gong, C. *et al.* (2005) Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C. *Nat. Struct. Mol. Biol.* 12, 304–312

AGORA initiative provides free agriculture journals to developing countries

The *Health Internetwork Access to Research Initiative* (HINARI) of the WHO has launched a new community scheme with the UN Food and Agriculture Organization.

As part of this enterprise, Elsevier has given 185 journals to *Access to Global Online Research in Agriculture* (AGORA). More than 100 institutions are now registered for the scheme, which aims to provide developing countries with free access to vital research that will ultimately help increase crop yields and encourage agricultural self-sufficiency.

According to the Africa University in Zimbabwe, AGORA has been welcomed by both students and staff. 'It has brought a wealth of information to our fingertips' says Vimbai Hungwe. 'The information made available goes a long way in helping the learning, teaching and research activities within the University. Given the economic hardships we are going through, it couldn't have come at a better time.'

For more information visit:
<http://www.healthinternetwork.net>