SURVEY AND SUMMARY

Interaction of Ku protein and DNA-dependent protein kinase catalytic subunit with nucleic acids

William S. Dynan* and Sunghan Yoo

Program in Gene Regulation, Institute of Molecular Medicine and Genetics, Room CB-2803, Medical College of Georgia, 1120 15th Street, Augusta, GA 30912, USA

Received November 7, 1997; Revised and Accepted January 7, 1998

ABSTRACT

The Ku protein–DNA-dependent protein kinase system is one of the major pathways by which cells of higher eukaryotes respond to double-strand DNA breaks. The components of the system are evolutionarily conserved and homologs are known from a number of organisms. The Ku protein component binds directly to DNA ends and may help align them for ligation. Binding of Ku protein to DNA also nucleates formation of an active enzyme complex containing the DNA-dependent protein kinase catalytic subunit (DNA-PKcs). The interaction between Ku protein, DNA-PKcs and nucleic acids has been extensively investigated. This review summarizes the results of these biochemical investigations and relates them to recent molecular genetic studies that reveal highly characteristic repair and recombination defects in mutant cells lacking Ku protein or DNA-PKcs.

INTRODUCTION

Maintenance of chromosome integrity is a fundamental requirement in all living organisms. Chromosomal integrity is threatened when double-strand DNA breaks occur. Such breaks can be induced by ionizing radiation, chemical agents and certain recombination endonucleases. The DNA-dependent protein kinase (DNA-PK) system is part of an evolutionarily conserved pathway for repair of double-strand DNA breaks. The Ku protein component of the system is a heterodimer of 70 and 83 kDa subunits (referred to as Ku70 and Ku80 respectively). The catalytic component of the system is a single polypeptide of 470 kDa, termed the DNA-PK catalytic subunit (DNA-PKcs).

Ku protein was originally discovered because it was a target of autoantibodies in patients with autoimmune disease (1–4). (The name ‘Ku’ derives from the surname of the prototype Japanese patient.) Patients with systemic lupus erythematosus (SLE), scleroderma, polymyositis and Sjogren’s syndrome have been reported to have some level of anti-Ku antibodies (reviewed in 5). Estimates of the frequency of anti-Ku antibodies vary. For example, anti-Ku antibodies were detected in 39% of American SLE patients by ELISA (4), in 19% by immunoblotting (6) and in 4% by immunoprecipitation (7).

Early studies using Ku protein preparations immunopurified with patient antibodies showed that the Ku protein binds tightly to linear duplex DNA, but has a much lower affinity for circular DNA, denatured DNA, yeast transfer RNA and poly(rA)·poly(dT) (8). With restriction enzyme cleaved plasmids binding is proportional to the number of DNA ends in the reaction and DNase I footprinting shows protection of sequences near the DNA ends (8). These observations provided early evidence that Ku protein recognizes the ends of linear duplex DNA.

In contrast to the discovery of Ku protein via its reactivity with autoantibodies, the enzymatic activity called DNA-PK was first identified in a biochemical screen for kinases that were stimulated by double-stranded DNA (9). Conventional protein kinase assays were used to partially purify this activity and to characterize its properties (10,11). The unification of what had been two separate areas of inquiry came with the discovery that Ku protein is the regulatory component of DNA-PK (12,13). Binding of Ku protein to DNA nucleates assembly of an active complex containing the 470 kDa DNA-PKcs.

Although the authors of one of the original studies of Ku protein suggested, presciently, that Ku protein ‘is perhaps involved with DNA repair or transposition,’ (2) the role of this system in repair and related processes was not established experimentally until 8 years later, in 1994, when it was discovered that the Ku80 subunit was defective in X-ray sensitive mammalian cell mutants in the XRCC5 group (14–17). The in vivo role of Ku protein and DNA-PKcs in double-strand DNA break repair and in certain types of recombination has subsequently been investigated using mutants in mammalian cells, Drosophila and yeast. Recent work suggests that Ku protein may also have a role in telomere maintenance (18,19) and in the stress response (20–23). This article attempts to synthesize an understanding of the interaction of Ku protein and DNA-PKcs with nucleic acids, combining the insights obtained from both biochemical studies and molecular genetics.

STRUCTURE AND FUNCTION OF THE KU PROTEIN

Genes encoding Ku protein

Ku protein homologs have been identified in vertebrates, insects, Saccharomyces cerevisiae and Caenorhabditis elegans (24–39).
Table 1 summarizes the sequence information that is currently available. Although the Ku protein homologs in various organisms have diverged substantially in primary sequence, they are similar in overall size and subunit structure. Those homologs that have been characterized biochemically have similar DNA binding properties. Moreover, human and Drosophila Ku70 cDNAs will complement a Ku70 null mutation in yeast (40).

The human Ku70 gene maps to chromosome 22q13 (41) and the Ku80 gene maps to 2q33–q34 (41,42). It has been suggested, based on cross-hybridization and protein sequencing data, that the human Ku70 protein might be part of a gene family (43,44). Recent data show that mice with an induced null mutation in Ku70 show a different and less severe V(D)J recombination phenotype than mice with a mutation in Ku80, which would be consistent with genetic redundancy in Ku70 (45). However, molecular clones corresponding to other human Ku70 family members have not been isolated to date. A splice variant of Ku80 mRNA encodes a protein with an additional 9 kDa sequence at the N-terminus (46). It has been proposed that this protein, termed KARP-1, regulates DNA-PK in a manner distinct from ordinary Ku80 (46).

The subunits of Ku protein are tightly associated and do not separate during protein purification. The dimerization interface has been mapped within the C-terminal 20 kDa of the Ku70 subunit and the C-terminal 32 kDa of Ku80 (47). Further studies have implicated a small region from amino acid 449 to 477 of Ku80 as particularly critical for subunit interaction (48). Although the Ku70 and Ku80 subunits are biochemically distinct, they appear to have arisen from a common ancestral gene. This relationship is most evident from comparison of the Ku70 and Ku80 subunits in yeast, which are 22% identical and 38% similar over a 258 amino acid region located in the C-terminal half of the protein (Fig. 1). This region of conservation includes at least part of the DNA binding domain (see below). The sequences of the two subunits of Ku protein are more divergent from each other in higher eukaryotes than in yeast and the relationship between the subunits was not apparent until yeast and other lower eukaryotic sequences became available.

The similarity between the Ku protein subunits implies that they are derived from a common ancestral gene, the product of which presumably functioned as a homodimer. If so, this suggests that the Ku protein–DNA complex is quasi-symmetrical, i.e. both subunits make similar DNA contacts via the homologous DNA binding domains located in the C-terminal half of each subunit.

Table 1.

<table>
<thead>
<tr>
<th>Organism</th>
<th>Size (amino acids)</th>
<th>References</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ku70</td>
<td>609</td>
<td>(24,25,30)</td>
<td>J04611, J04607</td>
</tr>
<tr>
<td>Ku80</td>
<td>732</td>
<td>(26,27)</td>
<td>M30938, J04977</td>
</tr>
<tr>
<td>DNA-PKcs</td>
<td>4127</td>
<td>(96,100,101)</td>
<td>U47077</td>
</tr>
<tr>
<td>Mouse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ku70 3′-end</td>
<td>608</td>
<td>(28)</td>
<td>M38700</td>
</tr>
<tr>
<td>Ku80 5′-end</td>
<td>732</td>
<td>Unpublished</td>
<td>Z14157</td>
</tr>
<tr>
<td>DNA-PKcs</td>
<td>4128</td>
<td>(97)</td>
<td>D87521</td>
</tr>
<tr>
<td>Syrian hamster</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ku80</td>
<td>732</td>
<td>(37)</td>
<td>U40570</td>
</tr>
<tr>
<td>Chinese hamster</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ku70</td>
<td>608</td>
<td>(36)</td>
<td>1836065 (sequence ID)</td>
</tr>
<tr>
<td>Ku80</td>
<td>732</td>
<td>(38)</td>
<td>L48606</td>
</tr>
<tr>
<td>Horse</td>
<td>DNA-PKcs</td>
<td>(98)</td>
<td>None</td>
</tr>
<tr>
<td>Toad (Xenopus laevis)</td>
<td>DNA-PKcs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tick (Rhipicephalus appendiculatus)</td>
<td>DNA-PKcs</td>
<td>Partial</td>
<td>AF001413</td>
</tr>
<tr>
<td>Worm (Caenorhabditis elegans)</td>
<td>DNA-PKcs</td>
<td></td>
<td>L41356</td>
</tr>
<tr>
<td>Fly (Drosophila melanogaster)</td>
<td>Ku70</td>
<td>600</td>
<td>(39)</td>
</tr>
<tr>
<td>Yeast (S.cerevisiae)</td>
<td>Ku70</td>
<td>602</td>
<td>(31)</td>
</tr>
<tr>
<td>Ku80</td>
<td>629</td>
<td>(19,34,35)</td>
<td>Z49702 (Q04437)</td>
</tr>
</tbody>
</table>
Alignment of *S. cerevisiae* Ku70 and Ku80 proteins

Identities = 62/280 (22%), Positives = 108/280 (38%), Gaps = 34/280 (12%)

**Ku80**

268 QKERYVGVKSIIEYEHNEN6GK---------KNSVEDQQGS6SYIPVTSKID---SVTKAY 317
  +K QVGR Y G + ++ ++ S ++ +D K Y

**Ku70**

276 EKTNFIVGKTYGMYTKHEAKAVGKVLVYEHED1RQEAYSKRKLNPITGVEDTGTKVYK 335

**Ku80**

318 RYQADYVVLPSVLDQVTYESPPGLD----LYGFNLREALPFRFLT--SESSFIPATDRLG 372
  YG D + S DQ V E++ D + GF + YF +SSFI D

**Ku70**

336 PYS--DLDINLSDSQQMVAETQKDAFLKIIGFRSSKSIHYFNIDKSFFVIVP--EAK 393

**Ku80**

373 CQSDLMAFALVVMLENRRKIAAVARYVS6KDENVMCALCPVLI6HNNS6KFKVXSLT 432
  ++ + ++L+ ++ ++ KIA+ K+S ++ L P +S K + +

**Ku70**

394 YKGSRTLASSLLKILRKKVIAILWGRKLSHPSLVTLP-------SSVKNEN6FY 445

**Ku80**

433 LCRLPFAEDERVTPFDKLLDKTTSGVPKLKTGHQIDELMEQF--VD3MPTDELPERPI 490
  L +R+PF ++ + FP LL S L + ++M F D +

**Ku70**

446 L5YVPFLLDEE--IKFPPSLLSYDQ6SHEKLDYDNMK6KVUTQSIMG6NYL6GYP6FKNPL 503

**Ku80**

491 LGNYQPIGEV------T7TDTTLPLPSNKDOQENKKDPLR 525
  L +Y+ + + TT P+ KD+ + D LR

**Ku70**

504 LQKHYKVLHDYLLQITF6DETNPNTKDDMMRED6SSLR 543

**Figure 1.** Alignment of *S. cerevisiae* Ku70 and Ku80 proteins. Alignment was performed using the gapped BLAST algorithm with default parameters.

**Binding of Ku protein to DNA ends and other structural discontinuities**

In most cases the initial binding of Ku protein to DNA occurs via a free DNA end or other special structural feature. Ku protein does not ordinarily interact with closed circular DNAs. The binding of Ku protein to DNA ends is independent of the exact structure of the ends, i.e. binding occurs to blunt ends, to ends with 5′ and 3′ overhangs and to hairpin ends and is largely independent of DNA sequence (8,49,50). (One of these papers refers to the protein under study as ‘EBP-80, a transcription factor closely resembling the human autoantigen Ku’. For the purposes of this review we have considered EBP-80 and Ku protein to be identical. Ku protein has been independently rediscovered several times. Other proteins in the literature which appear to be identical to Ku protein include NFIV, TREF, PSE1 and CHBF.) Ku protein also binds to the chemically heterogeneous ends produced by ionizing radiation (51). Evaluation of quantitative binding data supports a *Kd* value for end binding in the range 1.5–4.0 × 10⁻¹⁰/M (50,52).

Double-stranded oligonucleotides of 14–18 bp length are sufficient for Ku protein binding (50,53).

Further studies have shown that Ku protein can interact with DNA by a two-step mechanism, where Ku protein first recognizes DNA ends, then translocates to internal sites (49,54,55). This allows binding of multiple Ku molecules to a single DNA fragment. The resulting complexes can be visualized by electrophoretic mobility shift assay, by protection assays and by electron microscopy. Each Ku molecule makes close contacts with 13–21 bp of DNA, as judged by methidiumpropyl EDTA–Fe(II) protection (54,56) and successive Ku molecules bind at 25–30 bp intervals when present in a multimeric array (54). Under some conditions binding of Ku protein to DNA induces molecules of Ku protein to bind to another, leading to formation of loop structures that can be visualized by atomic force microscopy (57).

DNA binding studies using recombinant protein have shown that full activity, defined as an ability to form stable complexes in an electrophoretic mobility shift assay, requires both subunits of Ku protein (47,58–60). Studies using *in vitro* translated Ku protein show that slightly more than half of each subunit is required for DNA binding in the electrophoretic mobility shift assay, i.e. the C-terminal 40 kDa of Ku70 and the C-terminal 45 kDa of Ku80 (47). The Ku70 subunit has some independent DNA binding capability, which can be demonstrated in Southwestern blotting (61) and immunoprecipitation assays (61–63), although not in an electrophoretic mobility shift assay. Amino acids 536–609 of Ku70 are necessary and sufficient for DNA binding in the Southwestern blotting assay (61), suggesting that these amino acids may comprise a core DNA binding domain in the Ku70 subunit.

The Ku70 protein can be cross-linked to DNA by UV light, indicating that it is in intimate contact with the DNA (13,55,56). Some of these contacts lie in close proximity to the 5′-end of DNA. Cross-linked Ku70 subunit protects a 5′-end label from nuclease digestion, whereas cross-linked Ku80 does not (55). Moreover, an iodouracil substitution at the second position of the DNA renders the protein susceptible to protease cleavage, indicating that binding is accompanied by a conformational change (64).

In some situations the initial binding of Ku protein can occur at structures other than ends. Ku protein has a low affinity for single-stranded DNA as such, but it can bind a variety of artificial
structures containing single-to-double strand transitions, including circular molecules containing a nick, a single-stranded region (a gap) or a double-stranded region of non-complementarity (a bubble) (50,52). Although the repair defect in DNA-PK-deficient cells appears to be limited to double-strand DNA breaks, the ability of Ku protein to recognize a diversity of structures in vitro leaves open the possibility that it could have some role in recognizing other types of DNA damage.

Further insight into the nature of the interaction between Ku protein and DNA is provided by experiments showing that Ku protein can be trapped on DNA in the form of a "sliding clamp". When Ku protein is allowed to bind to linear DNA and the DNA is converted to a closed circular form with DNA ligase, the Ku protein–DNA complex becomes highly resistant to dissociation, suggesting that Ku protein is trapped on DNA "as a bead threaded on a DNA string" (49). The formation of a topologically linked protein–DNA complex is reminiscent of yeast topoisomerase II (65) and of certain replication proteins (66) and may have important implications for the mechanistic role of Ku protein in repair and other processes.

Recent work has shown that Ku protein has an unexpected and impressive ability to transfer between different DNA molecules with cohesive ends (67). This may be a special case of the sliding clamp behavior, where Ku protein slides across a transient duplex structure formed by the cohesive ends. Atomic force microscopy studies show Ku protein bound at the junction of separate DNA fragments, apparently holding them together in a non-covalent complex (51). Other recent studies show that Ku protein stabilizes the association of DNA ends and stimulates joining of such ends by purified human DNA ligases I, III and IV (68).

The strong geometric constraints on Ku protein–DNA interaction imply a very specific geometry for the binding site. Apparently the protein has a deep channel, open at either end. The free end of a duplex DNA can thread through the channel, explaining why DNA ends are a common site of initial Ku protein–DNA interaction. After this initial interaction the duplex DNA can slide isoenergetically through the channel, permitting Ku protein to act as a sliding clamp.

To fully account for the observed DNA binding properties of Ku protein, however, one must postulate a slightly more complex model. The channel must also have an opening along one side, constrained in a way that allows lateral diffusion of single-stranded DNA, but not double-stranded DNA, in and out of the binding site. This additional feature of the binding site geometry is necessary to accommodate the experimental finding that Ku protein can load onto circular DNA at a site where there is a single-to-double strand transition (50). This is a potentially important property of Ku protein that may relate to its ability to interact properly with DNA repair intermediates, as discussed in a later section.

Ku protein has been reported to have associated DNA-dependent ATPase and ATP-dependent helicase activities (60,69–71). In principle helicase activity could actively facilitate formation of a repair complex dependent on regions of microhomology internal to the DNA ends, as proposed in a recent model of double-strand break repair (72). Although the model is attractive, the experimental evidence for these activities of Ku protein rests primarily on biochemical co-purification and co-immunoprecipitation. To provide further evidence that the activities are intrinsic to the Ku polypeptides themselves it would be helpful to identify Ku mutants with specific defects in ATPase or helicase activity. Proposed ATP binding motifs in the Ku70 and Ku80 subunits have been mutagenized, but no effect on repair has been detected (73,74).

### Binding of Ku protein to specific DNA sequences

A number of studies have attempted to address the question of whether Ku protein interacts with specific DNA sequences. Putative binding sites for Ku protein or Ku protein-containing complexes have been identified in a variety of genes, often in association with transcriptional regulatory elements. These genes include the transferrin receptor (75,76), a mouse retroviral-like element (50,77), U1 snRNA (56), gpr78 (78), T cell receptor β chain (79), collagen IV (80), parathyroid hormone (81) and c-myc (82). Binding has been reported to sequences in human T cell leukemia virus-I (83,84) and mouse mammary tumor virus (85). Binding of Ku or Ku-like proteins has also been reported to the immunoglobulin octamer motif (43), the AP-1 binding element (86) and the heat shock element (87). Finally, putative Ku protein binding sites have also been reported in a sequence containing a replication origin (88) and a sequence corresponding to a BCL2 major breakpoint (89). No general consensus recognition sequence has emerged from these studies.

In interpreting this literature it is important to be aware that the unique properties of Ku protein can influence the results of standard assay methods in unexpected ways. The ability of Ku protein to undergo facilitated transfer between DNA molecules with cohesive ends can confuse the results of binding competition assays (67) and the presence of DNA fragments in cell extracts can lead to artefactual co-immunoprecipitation of Ku protein with other transcription factors (90,91). Some early studies used preparations of Ku protein purified on DNA affinity columns and these preparations may have been contaminated with small amounts of other sequence-specific DNA binding proteins. Many of the unusual properties of Ku protein were not understood at the time that initial studies were performed and it is difficult to assess their impact in retrospect.

Some of the reported Ku binding sites may be specific in the sense that they are preferred resting sites for Ku protein when it is interacting with duplex DNA in its "sliding clamp" mode. It is unlikely that any protein will bind completely randomly with respect to DNA sequence and it is therefore not surprising when experiments show that some sequences are preferred over others. More puzzling are certain elements that appear to permit binding of Ku protein to circular plasmids. As described in the preceding section, the initial loading of Ku protein onto DNA usually requires a free end or other special structure. Certain sequences in mouse mammary tumor virus, in the c-myc gene and in HTLV-I seem to bypass this constraint, allowing entry of Ku protein onto circular plasmids (82,85). Interestingly, these sequences contain 4–6 nt mirror repeats, with purines on one strand and pyrimidines on the other. Although such sequences have some potential for triple helix formation, there is no evidence that a stable non-B-form DNA structure is present prior to Ku binding (85). The structural basis for interaction between Ku protein and these sites in circular DNA therefore remains unknown.

Despite the uncertainties associated with some of the studies describing the interaction of Ku protein with specific sites in DNA, this issue is nevertheless important. There have been many suggestions in the literature that Ku protein has roles in the regulation of gene expression and in other regulatory processes.
If the Ku protein–DNA-PK system indeed has biological functions other than at DNA ends, then it is likely that there is some mechanism for interaction with sequences internal to intact chromosomes.

**Binding of Ku protein to RNA**

There is evidence that Ku protein interacts with RNA, although this interaction has been less well-characterized than the interaction with DNA. Antibodies to Ku protein stain both the nucleoplasm and the nucleolus of mammalian cells. Nucleolar staining is sensitive to RNase treatment, suggesting that association of Ku protein with the nucleolus is RNA-dependent (4). Ku protein does not bind to tRNA (8) and binds very weakly to total HeLa cell RNA (92). Ku protein has been reported, however, to bind selectively to an RNA containing the HIV transactivation response (TAR) element (92).

Recently the RNA binding properties of Ku protein have been systematically investigated using SELEX (systematic evolution of ligands by exponential enrichment) technology (93). Small RNAs were identified that bind to Ku protein as tightly as double-stranded DNA fragments (94). The Ku binding RNAs share primary sequence motifs. Many of the RNAs inhibit DNA-PK activity by competing for the DNA binding site in Ku protein, making them potentially useful as tools for sensitizing cells to ionizing radiation (94).

**THE CATALYTIC SUBUNIT OF THE DNA-DEPENDENT PROTEIN KINASE**

The gene encoding DNA-PKcs is located near the C-terminus, between amino acids 3719 and 4127 (96, 101). The kinase domain of DNA-PKcs is located near the C-terminus, between amino acids 3719 and 4127 (96, 101). The kinase domain is a member of a specific subfamily within the phosphatidylinositol 3-kinase family (96, 101). Unlike PI3-K, DNA-PK has been observed to have only protein kinase and not lipid kinase activity (96). A pharmacological inhibitor, OK-1035 [3-cyano-5-(4-pyridyl)-6-hydroazomethylyl-2-pyridone] has been described that selectively inhibits the phosphorylation activity of DNA-PK in an ATP-competitive manner (106). This compound appears promising for studies of the role of DNA-PK in intact cells (107).

The kinase domain of DNA-PKcs is located near the C-terminus, between amino acids 3719 and 4127 (96, 101). The kinase domain is a member of a specific subfamily within the phosphatidylinositol 3-kinase family (96, 101). Unlike PI3-K, DNA-PK has been observed to have only protein kinase and not lipid kinase activity (96). A pharmacological inhibitor, OK-1035 [3-cyano-5-(4-pyridyl)-6-hydroazomethylyl-2-pyridone] has been described that selectively inhibits the phosphorylation activity of DNA-PK in an ATP-competitive manner (106). This compound appears promising for studies of the role of DNA-PK in intact cells (107).

Sequences near the kinase homology domain, between amino acids 3002 and 3850, have been implicated both in interaction with the Ku protein and with the c-AbI tyrosine kinase (108). DNA-PKcs is capable of autophosphorylation, which inhibits activity by causing dissociation of DNA-PKcs from the Ku–DNA complex (109). Cleavage in the central part of DNA-PKcs, between amino acid 2712 and 2713 in the human sequence, occurs during apoptosis and is correlated with loss of catalytic activity (110–115).

**Binding of DNA-PKcs to DNA**

The central role of Ku protein in activation of DNA-PK was first demonstrated in two independent studies, which showed that Ku protein and DNA-PKcs can be separated biochemically, that activity can be restored by mixing the Ku protein and DNA-PKcs fractions and that the Ku protein component physically recruits DNA-PKcs to DNA (12, 13). This recruitment can be demonstrated by formation of distinctive complexes in an electrophoretic mobility shift assay (12) and by UV cross-linking (13). In general the ability of a given DNA structure to bind to Ku protein is closely correlated with its ability to stimulate DNA-PK catalytic activity (116). The only reported exception is cisplatin-damaged DNA, which binds Ku protein but does not stimulate DNA-PK (117). DNA-PKcs appears to make direct contact with DNA in the active complex, as evidenced by its UV cross-linking properties (13) and by its ability to bind to DNA independently of Ku protein when present at high concentrations (53).

Interestingly, binding of purified DNA-PKcs to a linear DNA probe can be competed by excess linear DNA but not by supercoiled circular DNA, suggesting that DNA ends are required for binding (53). Thus, as with Ku protein, binding of DNA-PKcs to DNA appears to be subject to geometric constraints. Atomic force microscopy showed that DNA-PKcs bound to DNA predominantly at ends in one study (53) and at ends and internal positions in another (57). As with Ku protein, some structures are observed where DNA-PKcs appears to tether two DNA ends in a non-covalent complex (53).

Free Ku protein and DNA-PKcs do not appear to form a stable complex. They elute independently in size exclusion chromatography (118) and they do not co-immunoprecipitate, provided that the endogenous DNA in crude lysates is removed (91). These biochemical studies suggest that much of the DNA-PK in intact and undamaged cells is in a latent form, where regulatory and catalytic components are not associated with each other.

**What makes DNA-PK ‘DNA-dependent’?**

Experiments using highly purified preparations of DNA-PKcs show that DNA dependence is attributable, at least in part, to the direct physical interaction between DNA-PKcs and DNA. Purified DNA-PKcs retains a low level of enzymatic activity in the absence of Ku protein. Under some assay conditions this residual activity is DNA-dependent, indicating that DNA-PKcs can be activated directly by contact with DNA. In one study this DNA-dependent activity was observed with a DNA-bound protein substrate, but not with a free substrate under the same conditions (118). Presumably the DNA-bound substrate served to increase the local concentration of DNA in the vicinity of the DNA-PKcs active site. In a subsequent study DNA-dependent activity of isolated DNA-PKcs was observed with free substrate under conditions where relatively high concentrations of DNA-PKcs and DNA were present (53). These data suggest that although the interaction of DNA-PKcs with DNA is weak, it is nonetheless effective in promoting an allosteric change to the active state.
It is significant that in both studies where the enzymatic activity of isolated DNA-PKcs was characterized addition of Ku protein stimulated activity 5- to-10 fold beyond what could be attained by addition of DNA alone. These data suggest that Ku protein may stimulate activity by two mechanisms. One mechanism involves recruitment of DNA-PKcs to the DNA, as described in the preceding section. This mechanism is important when the concentration of DNA ends is low and other mechanisms of recruiting DNA-PKcs to DNA are not operative. The other mechanism by which Ku stimulates DNA-PKcs is through direct protein–protein contact.

The importance of Ku protein in regulating DNA-PK activity is underscored by the identification of proteins that modulate DNA-PK activity by altering the ability of Ku protein to interact with DNA-PKcs. One of these is the c-Abl tyrosine kinase, which binds to DNA-PKcs at a site near the Ku binding site (108,119). c-Abl DNA-PKcs. One of these is the c-Abl tyrosine kinase, which binds to DNA-PKcs at a site near the Ku binding site (108,119). c-Abl kinase phosphorylates DNA-PKcs in this region and it is proposed that this phosphorylation dissociates Ku protein from DNA-PKcs (108). A contrasting example of regulation is provided by the transcription factor HSF1, which stimulates DNA-PK activity (22). HSF1 binds specifically to both Ku protein and DNA-PKcs (23) and may stimulate activity by enhancing formation of an active complex between the two components of DNA-PK.

DNA-PK activity is also stimulated by high mobility group proteins 1 and 2. These proteins appear to facilitate binding of DNA-PK to DNA, but the detailed mechanism of stimulation remains to be investigated (120).

FUNCTION OF KU PROTEIN AND DNA-PK IN DNA REPAIR

Mammalian cells that are deficient in Ku protein or the catalytic subunit of DNA-PK show highly characteristic defects. They are sensitive to ionizing radiation, which induces double-strand DNA breaks (14–17,121–125). The degree of sensitivity varies, because many cells have an alternative recombinational pathway of repair of double-strand DNA breaks that is Ku-independent. In mammalian cells mutations in Ku protein or DNA-PKcs cause an ~10-fold increase in radiation sensitivity, although this varies with position in the cell cycle (126). Yeast cells that are defective in the Ku70 or Ku80 subunits are temperature-sensitive for growth and sensitive to various DNA damaging treatments (19,34,35,40,127–130). Because yeast have an efficient recombinational repair pathway, the radiation-sensitive phenotype of Ku mutants is much more evident when this alternative pathway is inactivated, i.e. in a rad52 mutant background (19,34,127,129,130).

Dissection of the in vivo role of the Ku–DNA-PK system has been possible using substrates in which defined double-strand DNA breaks have been introduced by site-specific recombination enzymes or by restriction endonucleases. In mammalian cells the RAG1/RAG2 endonuclease creates paired double-strand breaks at recombination signal sequences in antigen receptor genes. This initiates so-called V(D)J recombination, allowing combinatorial assembly of immunoglobulin and T cell receptor genes with an immense number of different specificities (reviewed in 131). When the genes encoding Ku protein or DNA-PKcs are mutated RAG1/RAG2-dependent breaks are made, but they are not properly rejoined (14,16,38,121,122,124,125).

The V(D)J recombination-deficient phenotype has been analyzed in detail in different mutants. The mutation in scid mice, which produces an 83 amino acid C-terminal truncation of DNA-PKcs (97,132,133), leads to a defect in V(D)J coding joint formation, but not signal joint formation (134). In contrast, a SCID mutation in Arabian horses, which causes a 967 amino acid deletion in DNA-PKcs, leads to a defect in both coding and signal joint formation (98). Ku70 knockout mice are impaired in immunoglobulin rearrangements, but not T cell receptor rearrangements (45), whereas Ku80 knockout mice are impaired in both types of rearrangement (135,136). The basis of these phenotypic differences remains under investigation.

In addition to V(D)J recombination, two other eukaryotic recombination systems make use of the Ku protein. In Drosophila P element transposition involves a double-strand break intermediate. Mutation of a gene encoding the Drosophila Ku70 subunit affects the processing of double-strand breaks after P element excision and increases the incidence of deletions (137). In yeast mating type switching involves a double-strand break induced by HO endonuclease. This break is normally rejoined via the RAD52 recombinational pathway, but can be rejoined via the Ku protein system. Mutation of the Ku70 gene reduces the rate of induced mating type switching (127). Moreover, mutation of either the Ku70 or the Ku80 gene reduces the frequency of survivors when HO endonuclease is induced in a rad52 background (34).

As an alternative to recombination endonucleases, another way to introduce defined double-strand DNA breaks into a repair substrate is to use restriction endonucleases. Repair of restriction enzyme-cleaved plasmids in yeast is Ku-dependent. The small amount of repair that occurs in the absence of Ku protein results in large deletions (19,34,129). An analogous reduction in the ability to repair I-SceI endonuclease-induced breaks is seen in mammalian cells that lack functional Ku protein (138). Interestingly, the ability of Ku protein to repair restriction enzyme-cleaved DNA is dependent on the structure and sequence of the ends; cohesive ends are repaired, but blunt GC-rich ends generated by SmaI are not (129). In separate studies, however, Ku protein did promote repair of blunt PvuII ends (34). Because Ku protein is believed to bind to all types of free DNA ends, the failure to repair certain types of blunt ends in vivo presumably reflects a difficulty in performing a later step in the repair reaction (137).

Recent studies have begun to map the sequences within the Ku protein subunits that are required for repair function (73,74). In Ku70 all mutations that impair dimerization, DNA end binding and DNA-PK activation also affect repair. In addition, several mutations that do not affect known biochemical activities of Ku protein nevertheless fail to complement repair deficiency in vivo. This suggests that Ku70 participates in interactions with other, yet to be identified, proteins in the repair pathway.

It remains to be determined whether DNA-PKcs has a direct role in the double-strand break repair mechanism or whether its primary function is in the signaling processes surrounding DNA damage and repair. An interesting model has been proposed where DNA-PKcs assembled on one DNA end transphosphorylates DNA-PK components bound on the opposite side of a double-strand DNA break (72). Such phosphorylation can occur only when two ends have been brought into physical proximity and could provide a signal that enables the next step in the repair reaction to occur.

DNA-PKcs could also interact with other components of the intracellular signal transduction machinery, perhaps to generate a ‘survival signal’ indicating that repair is in progress or has been completed. As noted previously, DNA-PKcs has a physical and functional interaction with the c-Abl proto-oncogene product, which supports the idea that DNA-PKcs is involved in intracellular signaling (119). In addition, Ku protein binds to the p95vav
motifs characteristic of signal transduction proteins (139).

Genetic studies in yeast have helped to identify other proteins that are involved in the Ku-dependent double-strand break repair pathway. These include: a DNA ligase IV homolog, LIG4/DNL4; the silent information regulatory proteins, SIR2, SIR3 and SIR4; a radiation-sensitivity gene, RAD50 (140–142). All of these have an epistatic relationship with yeast Ku70, i.e. single mutants in either gene have the same phenotype as the double mutant, indicating that they function in a common repair pathway. The potential role of a DNA ligase in double-strand break repair is obvious. The potential function of the other proteins is less obvious. As at telomeres, the SIR proteins might be involved in assembly of a large macromolecular complex (143).

**SUMMARY AND CONCLUSIONS**

How do the nucleic acid binding properties of DNA-PK relate to its biological function in double-strand DNA break repair? Figure 2 illustrates the possible sequence of events that occur during double-strand break repair.

Rapid and avid binding of Ku protein to the ends of broken DNA is likely to be the first step in the repair process. The striking affinity of Ku protein for DNA ends is the feature which first suggested a role for Ku protein in DNA repair. The next step may be synopsis of the DNA ends in a repair complex, as shown in Figure 2B. Biochemical studies suggest that Ku protein, DNA-PKcs or both may have a direct role in this process. In particular, the extended region of methidiumpropyl EDTA–Fe(II) protection and the sliding clamp behavior of Ku protein imply the existence of a channel in Ku protein that may be well-suited to hold the ends of two DNA molecules aligned in preparation for ligation.

In the simplest case a double-strand break may be the result of two closely staggered single-strand cuts, creating ends that can anneal directly. Once these ends have been aligned in a repair complex they can be ligated without further processing. In many cases, however, the ends may be blunt or bases may be missing as a result of DNA damage. A current model draws on observations that repair of non-cohesive ends often involves deletions back to regions of microhomology (144,145). In this model (72) DNA ends are proposed to unwind to the point where a region of microhomology is encountered. A helicase activity intrinsic to Ku protein could facilitate this unwinding. Complementary strands then anneal to form a short duplex, with mismatched single-stranded segments protruding from each end of the duplex (Fig. 2C). After processing of the mismatched ends by nucleases, ligation occurs (Fig. 2D).

If a primary function of Ku protein is to bind to DNA ends, why is the protein also designed in such a way as to bind to DNA containing nicks and single-to-double strand transitions and, even more strikingly, to allow DNA containing such features to associate and dissociate from the nucleic acid binding site independently of DNA ends? This characteristic may relate to the need for Ku protein to dissociate from the repair complex as the reaction proceeds. After the broken ends have annealed free DNA ends are no longer available. Ku can nevertheless exit the complex because nicks or other short unpaired regions are present. After Ku exits the complex ligation can be completed without leaving a molecule of Ku protein trapped on the DNA!

One of the important needs in the DNA-PK field is to better define the structure and composition of the active complex that mediates DNA double-strand break repair. Although current models emphasize the importance of a synaptic complex where the ends of DNA are brought into apposition, our knowledge remains incomplete. When DNA ends come together, what is the stoichiometry of the different components and what is the precise geometry of the protein–DNA contacts? In electrophoretic mobility shift assays DNA-PKcs has some preference for binding to DNA fragments containing two or more bound Ku proteins (12), but the stoichiometry of binding in the active repair complex remains unknown. Does an intrinsic helicase activity in Ku protein facilitate DNA unwinding to allow annealing through regions of microhomology? What is the role of DNA-PK-mediated protein phosphorylation in promoting the repair reaction? Are there other relevant targets of DNA-PK phosphorylation that provide signaling functions with respect to DNA repair? Are the ends of broken DNA sequestered in a discrete structure within the nucleus? These and other questions remain to be addressed.

Although this review has focused on the interaction of DNA-PK components with DNA ends and the biological role of DNA-PK in repair and recombination, there are strong indications that DNA-PK is involved in other cellular processes. In HeLa
cells DNA-PKcs is very abundant, comprising 0.1–1.0% of nuclear DNA-PKcs is very abundant, comprising 0.1–1.0% of DNA repair proteins (146). Genetic studies in yeast implicate DNA-PK in a complex with RNA polymerase II and other DNA repair proteins (147). Genetic studies in yeast implicate DNA-PK in telomere maintenance (18,19). DNA-PK is also strongly implicated in the heat shock or stress response (20–23). It remains to be learned which of these roles are biologically essential and how the unique nucleic acid binding properties of Ku protein and DNA-PKcs may be involved.

ACKNOWLEDGEMENTS

We thank John A. Hardin for helpful discussions, Paul Labhart for comments on DNA sequence alignments and Rhea-Beth Markowitz for critical reading of the manuscript. W.S.D. is an Eminent Scholar of the Georgia Research Alliance and is the recipient of an American Cancer Society Faculty Research Award. The author’s work on Ku protein and DNA-PK is supported by a NIH grant GM35866 and by a Biomedical Research Grant from the Arthritis Foundation.

REFERENCES
