# A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins

PEER BORK,*,† KAY HOFMANN,‡ PHILIPP BUCHER,‡ ANDREW F. NEUWALD,§ STEPHEN F. ALTSCHUL,§ AND EUGENE V. KOONIN§,1

*European Molecular Biology Laboratory, 69012 Heidelberg, Germany; †Max Delbrück Center for Molecular Medicine, Berlin-Buch, Germany; ‡Swiss Institute for Experimental Cancer Research, 1066 Epalinges, Switzerland; and §National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland 20894, USA

ABSTRACT **Computer analysis of a conserved domain, BRCT, first described at the carboxyl terminus of the breast cancer protein BRCA1, a p53 binding protein (53BP1), and the yeast cell cycle checkpoint protein RAD9 revealed a large superfamily of domains that occur predominantly in proteins involved in cell cycle checkpoint functions responsive to DNA damage. The BRCT domain consists of ~95 amino acid residues and occurs as a tandem repeat at the carboxyl terminus of numerous proteins, but has been observed also as a tandem repeat at the amino terminus or as a single copy. The BRCT superfamily presently includes ~40 nonorthologous proteins, namely, BRCA1, 53BP1, and RAD9; a protein family that consists of the fission yeast replication checkpoint protein Rad4, the oncoprotein ECT2, the DNA repair protein XRCC1, and yeast DNA polymerase subunit DPB11; DNA binding enzymes such as terminal deoxynucleotidyltransferases, deoxycytidyl transferase involved in DNA repair, and DNA-ligases III and IV; yeast multifunctional transcription factor RAP1; and several uncharacterized gene products. Another previously described domain that is shared by bacterial NAD-dependent DNA-ligases, the large subunits of eukaryotic replication factor C, and poly(ADP-ribose) polymerases appears to be a distinct version of the BRCT domain. The retinoblastoma protein (a universal tumor suppressor) and related proteins may contain a distant relative of the BRCT domain. Despite the functional diversity of all these proteins, participation in DNA damage-responsive checkpoints appears to be a unifying theme. Thus, the BRCT domain is likely to perform critical, yet uncharacterized, functions in the cell cycle control of organisms from bacteria to humans. The carboxyterminal BRCT domain of BRCA1 corresponds precisely to the recently identified minimal transcription activation domain of this protein, indicating one such function.—Bork, P., Hofmann, K., Bucher, P., Neuwald, A. F., Altschul, S. F., Koonin, E. V. A superfamily of conserved domains in DNA damage-responsive cell cycle checkpoint proteins. *FASEB J.* 11, 68–76 (1997)**

*Key Words: BRCT domain · database search · PARP · profile · motif*

CELL CYCLE CHECKPOINTS ARE molecular mechanisms for the negative control of DNA replication and mitosis (1–4). Checkpoints arrest DNA replication (the G1 checkpoint) or mitosis (G2 checkpoint) when the integrity of the genome is compromised either as a result of DNA damage or as part of programmed cellular events such as apoptosis, senescence, or immune cell development (5). The checkpoints involve the transmission of a signal from damaged DNA to effectors such as cyclin-dependent protein kinases via complex mechanisms dependent on a variety of proteins, including the universal tumor suppressors p53 and the retinoblastoma (RB)[2] protein (6–8). Impairment of the checkpoints is thought to play a critical role in cancer cell evolution (5).

Checkpoint proteins are highly diverse structurally, and no conserved domains have been found that are common to large groups of them. Very recently, we described a domain (dubbed BRCT) common to the breast cancer susceptibility protein BRCA1, a p53 binding protein (53BP1), the yeast checkpoint protein RAD9, and uncharacterized yeast and human proteins (9). Here we describe the results of further database searches for sequences similar to the BRCT domain using a variety of sensitive motif and profile detection methods (reviewed in refs 10, 11). The complementary use of these methods enabled us to identify a large BRCT domain superfamily that unites functionally diverse proteins from mammals, yeast, and bacteria, many of which play direct or indirect roles in DNA damage response and cell cycle checkpoints.

[1] Correspondence.

[2] Abbreviations: HMM, hidden Markov models; NR, nonredundant; TdT, terminal deoxynucleotidyl transferase; RF-C, replication factor C large subunits; PARP, poly(ADP-ribose) polymerases; RB, retinoblastoma.



0892-6638/97/0011-0068/$01.50 © FASEB

## MATERIAL AND METHODS

### General strategy for detection of conserved domains

Currently no single computer method ensures the optimal delineation of divergent protein superfamilies. Methods for detecting pairwise sequence similarity (e.g., BLAST and FASTA) and for motif or profile analysis often produce complementary results, and their iterative application improves the detection of distantly related domains (11). The BRCT superfamily was analyzed using such an iterative strategy.

An alignment of the originally described BRCT domains (9) was used to screen the sequence databases using profile and motif search methods, and reciprocal BLAST searches (12–14) were performed with all newly detected BRCT proteins. Before the BLAST searches, the sequences were partitioned into predicted globular and nonglobular domains using the SEG program with parameters adjusted for this task (15). When new statistically significant similarities were detected, the segments involved were added to the BRCT motifs or profiles, and a new round of database screening was performed. When distinct but divergent protein families were detected, separate profiles and motifs were constructed for each and used in reciprocal database searches. The whole process was repeated until no new superfamily members could be detected.

### Database searches with motifs and profiles

Nonredundant (NR) databases were searched for protein sequence segments related to a given motif using the MoST program (16). Briefly, an alignment block is converted to a position-dependent weight matrix using Dirichlet mixture prior distributions (17); any newly detected similar segments are added to the block, and the evolving matrix is used to scan the database iteratively until convergence. The ratio of the expected number of segments with a given score to the actually detected number is used as the cutoff at each iteration.

Generalized profiles were constructed from multiple sequence alignments by a procedure combining elements of the methods described in refs 18 and 19, with additional modifications (20). Input sequences were weighted using the algorithm of Sibbald and Argos (21). The BLOSUM45 matrix (22) was used to convert amino acid frequencies into match scores. To assess the significance of candidate matches to a given profile, a window-shuffled version of the SWISS-PROT database (release 30) was scanned with the same profile (for details, see ref 23).

Alternatively, profile construction and search were conducted using WiseTools (24). The standard parameters of PAIRWISE were used for weighting of the sequences in the alignment. The Gonnet250 Matrix was used for database searches with SEARCHWISE, using a cutoff score of 3700 (24). The profile alignment option in CLUSTALW (25) was used to add new members for the subsequent iterations.

A complementary method of database screening using multiple alignment included construction of hidden Markov models (HMM) with the HMMb program, with subsequent database search using the HMMls program. HMMls detects local matches between database sequences and the HMM with a modification of the Needleman-Wunsch algorithm (26). A cutoff score of $t = 20$ was used in all HMM searches.

### Motif detection in unaligned protein sequences

The presence and the number of shared motifs in sets of unaligned protein sequences were identified using Gibbs sampling (27, 28). This method detects and aligns multiple, diverged copies of a motif, then applies near-optimal sampling to estimate the predictive probability (28) that each of these copies matches the motif. By default, only those sequences with a probability greater than 0.50 are reported. Alternatively, motifs were derived directly form BLAST outputs using the CAP program (16).

### Other methods

Multiple alignments of protein sequences were constructed with MACAW (29) or CLUSTALW (25). Secondary structure predictions were made using the PHD program (30, 31) for all protein families that could be readily aligned.

## RESULTS AND DISCUSSION

### The BRCT domain superfamily

To detect as many proteins as possible containing the BRCT domain in current databases, we applied an iterative strategy that included multiple rounds of database screening with methods for pairwise sequence similarity detection as well as motif and profile analysis (see Materials and Methods). Starting with the initially described set of six proteins that includes human and mouse BRCA1, 53BP1, two uncharacterized proteins from human and fission yeast, and yeast RAD9 (9), we detected the BRCT domain in ~40 nonorthologous proteins that form several distinct families with highly significant similarities among their members (**Fig. 1**). It is now possible to define the full complement of proteins with a particular domain that are encoded in yeast, the first eukaryote whose genome has been completely sequenced. To this end, we screened, in addition to general purpose databases, the database of yeast protein-coding sequences from the Saccharomyces Genome Database (http://genome-www.stanford.edu/Saccharomyces/). We found that, altogether, 10 yeast proteins contain readily detectable BRCT domains (Fig. 1). In time, the number of proteins with recognized BRCT domain will certainly increase as numerous ESTs from different organisms are also similar to the domain (data not shown).

In addition to the original six proteins, the BRCT superfamily includes: *1*) a previously described protein family consisting of the fission yeast protein Rad4(Cut5) and its homolog from budding yeast, the human ECT2 and XRCC1 proteins, and the yeast protein REV1 (32–35); *2*) yeast transcription factor RAP1; *3*) human DNA ligases III and IV; *4*) vertebrate terminal deoxynucleotidyl transferases (TdTs) and their homolog from fission yeast; *5*) a previously described domain (36) shared by bacterial NAD-dependent DNA ligases, eukaryotic replication factor C large subunits (RF-C), and eukaryotic poly(ADP-ribose) polymerases (PARP); and *6*) several functionally uncharacterized, putative proteins from yeast and nematode.

The BRCT domain in each protein sequence typically was detected by more than one method. Specifically, iterative profile searches with Pfsearch (the cutoff of 9.0, a score expected to occur once in a database containing $10^9$ amino acid residues, was well above the first clear false positive, with a score of 7.7) and SearchWise (the cutoff of 3300 was above the first false positive, with the score of 3200) detected all proteins shown in Fig. 1, with the exception of the RB family. In a complementary approach, database screening with the most highly conserved motif II (Fig. 1), using MoST with a relatively conservative cutoff ($r$=0.01), detected the majority of the same proteins with-

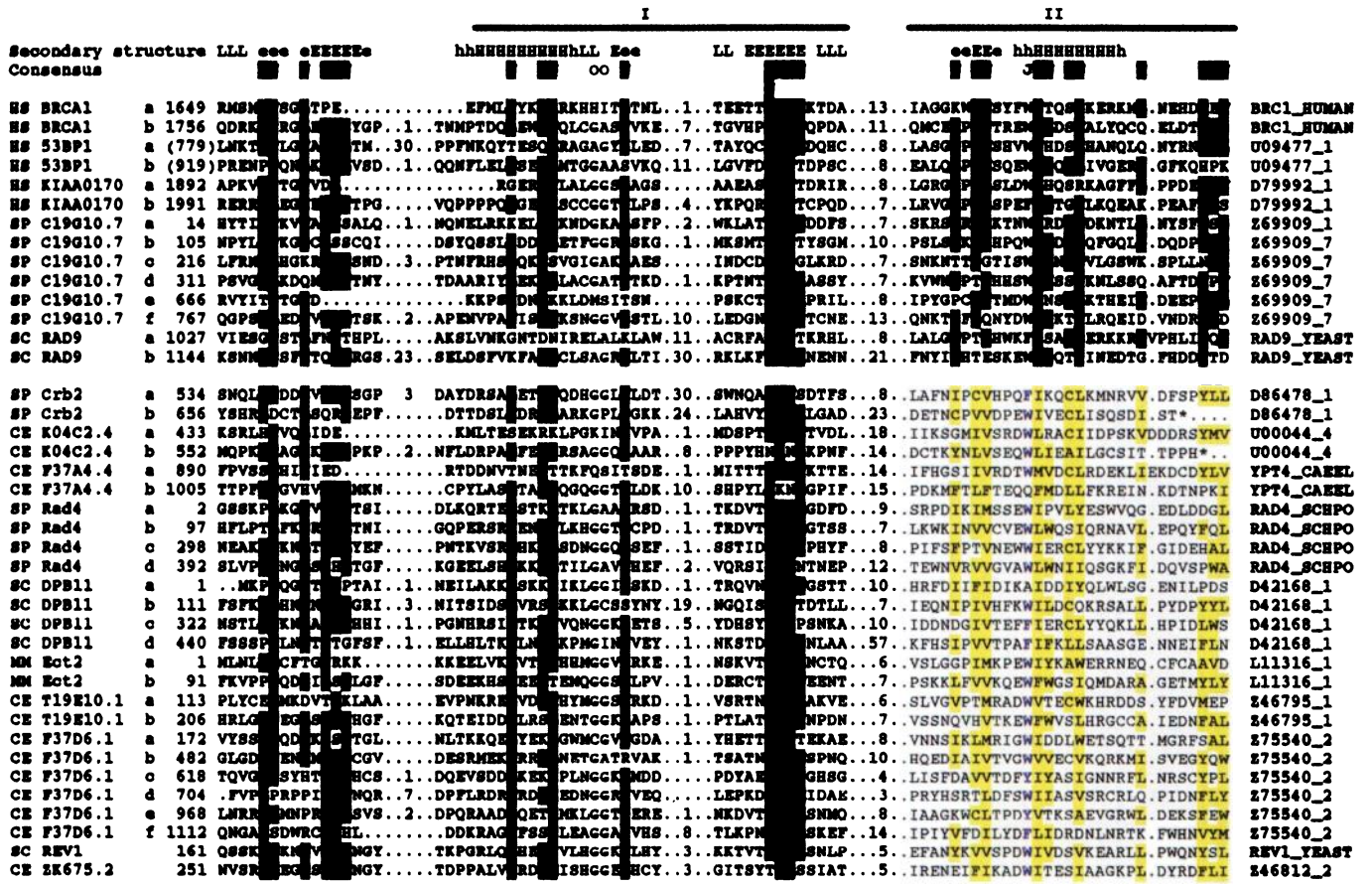

**Figure 1.** Multiple alignment of the most conserved motifs within the BRCT domains. The alignment is based on the results of database searching with profile alignment methods implemented in SEARCHWISE and Pftools (see Methods), with refinement on the basis of alignments for distinct protein families produced by the MACAW and CLUSTALW programs. The four sequence sets separated by blank lines are: the original group of five proteins (of the six proteins included in the alignment in ref 9, the mouse BRCA1 sequence, which is closely related to its human homolog, is omitted); the remaining core of the BRCT superfamily confidently detected by several methods; the RB family, for which the identification of the BRCT domain remains tentative; and the PARP/RF-C/bacterial DNA ligase group, which contains a distinct version of the BRCT domain. Typically, only one sequence from each set of highly conserved orthologs was included (e.g., only one TdT and one PARP). However, two RF-C sequences and two sequences from bacterial ligases are shown because these proteins contained a substantial number of differences. Consecutive copies of the BRCT domain from the same protein are designated a, b, etc. The distances from the protein termini to the aligned regions and the distances between the alignment blocks are indicated by numbers. The distance from the amino terminus of 53BP1 is shown in parentheses, as the available sequence of this protein is incomplete. Gaps introduced to optimize alignment are indicated by dots. Stop-codons are indicated by asterisks. The consensus includes amino acid residues conserved in the majority of the aligned sequences; the residues that conform with the consensus are shown in bold type and the respective color or shading; U indicates a hydrophobic residue, J indicates an aromatic residue, and O indicates a small residue. The predicted secondary structure elements are shown above the alignment, with E indicating extended conformation (β-strand), H indicating α-helix, and L indicating loop. Uppercase indicates the most reliable prediction (>82% accuracy); lowercase indicates prediction with ~72% accuracy (30). The leftmost column contains the abbreviated species name. HS, *Homo sapiens*; SP, *Schizosaccharomyces pombe*; SC, *Saccharomyces cerevisiae*; CE, *Caenorhabditis elegans*; MM, *Mus musculus*; CA, *Candida albicans*; EC, *Escherichia coli*; TT, *Thermus thermophilus*. The rightmost column contains sequence names from the SWISS-PROT database (ending with a five-letter organism name) or the GenBank database (ending with a number that indicates the number of the open reading frame in the given entry).

out obvious false positives. Not detected by this search were the RB protein, TdT, and the bacterial DNA ligase family, which appear to contain a distinct version of the BRCT domain. These sequences, however, were retrieved from the NR database without false positives when another conserved motif from the Rad4 protein family (motif I in Fig. 1) was used for screening the NR database (albeit with a liberal cutoff of $r=0.05$). The relationship between these proteins and the rest of the BRCT domain superfamily was corroborated by an HMM search that detected one of the bacterial DNA ligases with a score of 25.9, with the highest score of 22.9 for the first obvious false positive. Furthermore, control MoST searches with motif II extracted from distinct families, e.g., the Rad4 family (Fig. 1), specifically retrieved from the database the majority of the superfamily members. The significance of the relationships between some of the protein families containing the BRCT domain was also confirmed by statistically significant pairwise similarities detected by exhaustive BLAST searches. For example, the uncharacterized fission yeast protein (C19G10.7) from the original sequence set (9) was similar to Rad4, with a $P$ value of $2 \times 10^{-5}$ (13, 14).

The presence of the BRCT domain in the RB protein family could not be demonstrated as convincingly as for

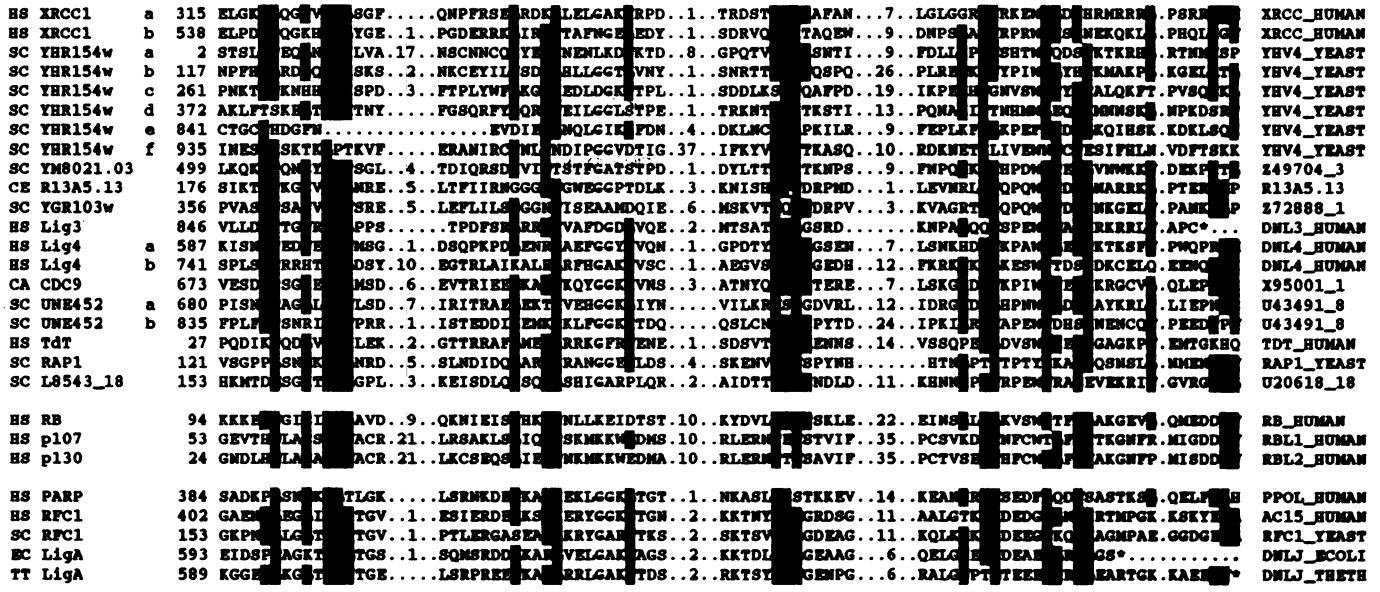

**Figure 1.** *Continued.*

most other proteins in the superfamily. The domain was detected in the RB sequence using motif II in a MoST search with a liberal cutoff ($r$=0.05); a few apparent false positives were also detected in this screening. The RB sequence scored only 5.5 with Pfsearch and 2600 with SEARCHWISE, which did not separate it from false positives. Nevertheless, the alignment of the RB sequence could be extended to include the whole BRCT domain (Fig. 1), and Gibbs near-optimal sampling indicated that the RB sequence matched the motif I statistical model as strongly as most of the other proteins in the BRCT domain superfamily (see below). Therefore, we believe that the RB family contains a highly diverged version of the BRCT domain.

The protein sequences containing candidate BRCT domains were further analyzed using Gibbs sampling in order to corroborate the presence of motifs I and II in each of them and to determine the number of copies of these motifs. Using a new optimization procedure (A. F. Neuwald, J. Liu, and C. Lawrence, unpublished results), the sampler converged on a 35-column aligned block for motif I and a 30-column block for motif II, with 23 and 20 columns, respectively, considered nonrandom. For each sequence in the superfamily (with the exception only of DNA ligase III, which contains a truncated motif II), the presence of at least one copy of either motif I or motif II or both was confirmed with an estimated confidence of 0.9 or greater.

## Features of BRCT domains

The BRCT domain does not contain a single invariant amino acid, but motif I centers on a conserved G[GA] doublet, and motif II centers on a tryptophan that is present in the great majority of the sequences. Furthermore, there is a clear consensus pattern of residues with conserved properties in other positions (Fig. 1). Note that motif II is the most conserved region in the majority of the BRCT domains, but is significantly modified in bacterial DNA ligases, RF-C, and PARP. Conversely, in the latter group of proteins, sequence conservation is more pronounced in motif I (Fig. 1).

The number of diverged copies of the BRCT domain per protein, identified by profile searches and by Gibbs sampling, varies between 1 and 6 (**Fig. 2**). The existence of proteins with a single BRCT domain (Fig. 2) and its predicted globular structure (Fig. 1) indicate that this domain may be an independent structural unit.

The BRCT domain is predicted to consist of four β-strands and two α-helices, with the β-strands probably forming a core sheet structure (Fig. 1). The observed pattern of amino acid conservation, the size of the domain, and the predicted secondary structure appear to be typical of domains involved in specific protein-protein interactions (37). It has been shown that a 270 amino acid, carboxy-terminal fragment of 53BP1, which consists largely of two BRCT domains, binds p53 as efficiently as the much larger carboxy-terminal portion of the protein translated from the longest isolated cDNA clone (38). Thus, p53 may be one of the BRCT domain ligands. Another potentially important observation is that two BRCT domain-containing proteins, namely, human DNA ligase III and XRCC1, form a complex (39, 40). This suggests the possibility of an interaction between BRCT domains in different proteins. All of the DNA ligase IV present in human cells has also been isolated in the form of a complex with another, as yet unidentified protein, and it has been proposed that the carboxy-terminal portion of the ligase that (as described here) contains the BRCT domain may be responsible for this interaction (41).

By contrast, it has been shown recently that the region of RF-C that we identified as the BRCT domain belongs to the DNA binding domain of this protein (42). The cen-

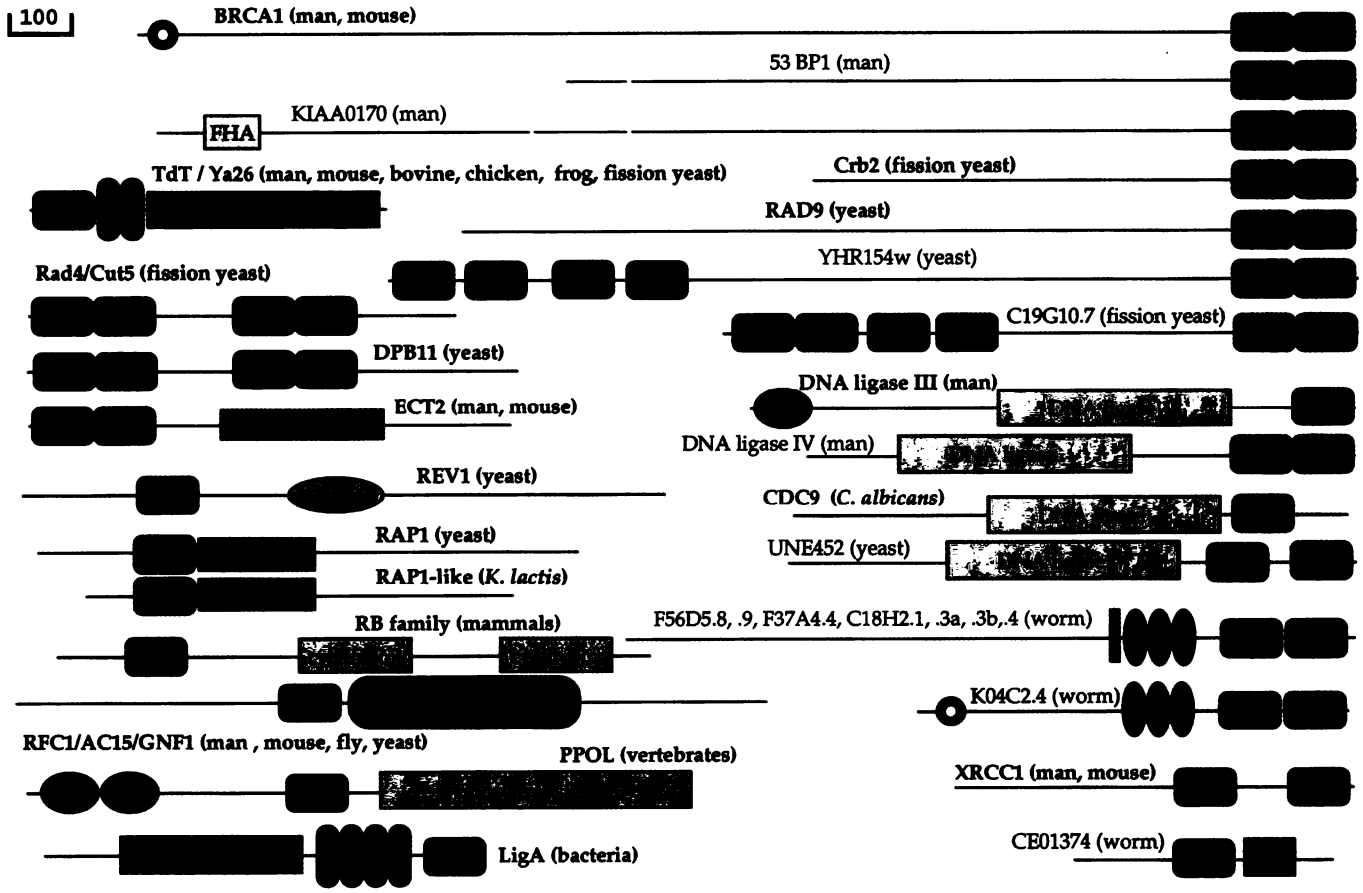


**Figure 2.** Domain organization of proteins containing BRCT domains. The proteins are shown roughly to scale as indicated by the bar in the upper left corner. The KIAA010 sequence is compressed as indicated by a broken line, and the 53BP1 sequence is incomplete at the amino terminus. The names of proteins that have been functionally characterized are in bold type. In addition to the BRCT domain, other domains detected experimentally or by computer analysis are indicated. FHA is a putative nuclear signaling domain (23); AZF is a specific Zn finger domain found in PARP (designated PPOL in the figure) and DNA ligase III; HhH is the recently identified helix-hairpin-helix DNA binding domain (93); S1 is a putative RNA binding domain shared by bacterial ribosomal protein S1, polynucleotide phosphorylase, and yeast splicing factors (P. Bork, unpublished observations); RB has been reported to contain two cyclin box domains (94, 95), but the observed sequence similarity is very low; ANK indicates ankyrin repeats, and a double circle in BRCA1 and K04C2.4 indicates a RING finger. Only one representative for each set of proteins with similar modular architecture is included, e.g., only one of six worm paralogs that contain a transmembrane region (gray box) and ankyrin repeat. The species range is indicated for each domain architecture. Only one representative for each set of orthologs is included. Note that some of the proteins do not correspond to the annotation in the databases or to translations obtained by automatic procedures. For example, the yeast genes UNE407 and UNE452 were fused because UNE407 contains the amino-terminal portion of the DNA ligase domain and UNE452 contains the carboxy-terminal portion. Translation of *C. elegans* genes obtained by genomic sequencing was modified in order to optimize the alignment within the family. Specifically, C18H2.3 (PID: g474199) was split into two ORFs; in C18H2.4 (PID: g474200), additional putative exons were introduced.

tral region of PARP, which contains its single BRCT domain with significant similarity to the BRCT domain of RF, has been implicated in the protein's dimerization, but is not involved in DNA binding (43). The difference between the results obtained with RF-C and PARP requires further clarification, even though it may be a reflection of the actual diversity of the BRCT domain binding affinities.

## Modular architecture of BRCT domain-containing proteins

All members of the BRCT superfamily are large, multidomain proteins (Fig. 2). Many contain functionally characterized enzymatic domains, such as two unrelated types of DNA ligase, type X DNA polymerase (TdT), ADP-ribosyltransferase (PARP), and ATPase (RF-C). Other proteins in the superfamily contain additional common binding domains such as the RING finger in BRCA1 and an uncharacterized nematode protein, the DH domain in ECT2 and an uncharacterized yeast protein, the FHA domain in an uncharacterized human protein, the helix-hairpin-helix DNA binding domain in bacterial ligases and TdT, and ankyrin repeats in a family of uncharacterized nematode proteins. Yet other proteins contain highly conserved domains whose specific function is not known but are implicated in DNA repair, e.g., the UmuC domain in REV1. It is possible that some of these conserved domains possess yet uncharacterized enzymatic activities as demon-

strated by the recent discovery of a deoxycytidyltransferase activity for REV1 (44).

Notably, we did not detect any proteins consisting solely of BRCT domains even though multiple copies of it account for a large fraction of the amino acid sequences of several proteins, particularly C19G10.7, YHR154w, Rad4, and DPB11 (Fig. 2). Based on the observation that the BRCT domain is typically fused to other domains with a distinct activity in transcription, repair, or replication, one may speculate that it is involved in signal transduction, linking the activities of components of the cell cycle checkpoint machinery (see below).

**Functional implications**

BRCT domain containing proteins that have been functionally characterized appear to be directly or indirectly associated with DNA damage-responsive cell cycle checkpoints. Mutations in RAD9, Rad4, and DPB11 abolish the G1 and/or G2 checkpoints (1, 4, 32, 45–47). DPB11, which is a subunit of yeast DNA polymerase ε, triggers the S-phase checkpoint in response to replication blocks (35, 48). Like the DNA damage signal generated by RAD9 (49, 50), this response is transmitted by the MEC1/RAD53 cascade (48, 50). Perhaps the most straightforward observations implicating the BRCT domain in checkpoints involve Rad4, whose amino-terminal region, corresponding to a single BRCT domain, blocks fission yeast cell division in G2 when overexpressed (34). In accord with this result, amino-terminal truncation of ECT2 unmasks the transforming activity of this protein (51).

BRCA1 is a tumor suppressor specific for breast and ovarian cancers (52–55). Two recent, independent studies with BRCA1 have suggested for the first time a specific function for the BRCT domain. It has been shown that when fused to the GAL4 DNA binding domain, the carboxy-terminal domain of BRCA1 (amino acid residues 1560-1863), which encompasses both BRCT domains (Figs. 1, 2), activates transcription of a reporter gene in both yeast and mammalian cells (56, 57). Furthermore, both of these studies defined the minimal transactivation domain as the very carboxy-terminal portion of BRCA1 between residues 1760-1863; the left boundary of this minimal domain corresponds precisely to the beginning of the first predicted β-strand of the BRCT domain (Fig. 1). There is as yet no evidence for the specific mechanism of transcription activation by the carboxy-terminal fragments of BRCA1. It appears likely that this activation is mediated by the interaction between the BRCT domains and RNA polymerase or transcription factors.

The observations of transcription activation by the BRCT domains of BRCA1 may be relevant for the function of the BRCT domains in other proteins. In particular, p53 and RB, which appear to be the principal regulators of the G1 checkpoint in mammalian cells, operate primarily at the level of transcription (reviewed in refs 6, 58, 59). p53 typically activates the transcription of a number of specific genes (58), whereas RB seems primarily to repress transcription (59). However, opposite effects have been reported for each of these proteins (60, 61). In particular, RB expression may be down-regulated by p53 (62, 63). In addition to their roles as tumor suppressors, p53, RB, and very recently, BRCA1 have been shown to control cell proliferation in normal mammalian development (64–67). Thus, the presence of a BRCT domain in a p53 binding protein, located in a region sufficient for p53 interaction (36) and apparently in RB, is compatible with a critical role of this domain in DNA damage-responsive checkpoints, which may be mediated by protein-protein interactions leading to transcription regulation.

The transcription connection for the BRCT domain is further strengthened by the recent finding that RAD9, the classical yeast checkpoint gene, controls the expression of a number of coordinately regulated repair, recombination, and replication genes (68). Furthermore, RAP1 is a universal yeast transcription regulator that activates transcription of a variety of genes, but is also a repressor of genes at mating type loci and near telomeres (reviewed in ref 69). RAP1 binds yeast telomeres via its central DNA binding domain (70), and regulates their length both by protecting them from degradation (71–73) and by preventing their uncontrolled growth (74). Telomere degradation, in turn, activates the RAD9-mediated checkpoint (75), indicating that telomere length control may be one of the mechanisms of RAP1 participation in checkpoints. It has been demonstrated, however, that yeast strains carrying a deletion of the 5′-terminal portion of the RAP1 gene, coding for the BRCT domain, show no alteration in telomere length (76). Therefore, it appears likely that the BRCT domain in RAP1 modulates the effect of this protein on transcription. Whereas RAP1 is an essential protein, its amino-terminal portion, containing the BRCT domain, is not (69), which is compatible with such a regulatory role.

Although at the moment the indications are most direct for the involvement of the BRCT domain in transcription regulation, it appears likely that it also participates in checkpoints by directly affecting repair and replication of damaged DNA. Thus, human DNA ligase III is specifically involved in DNA repair: a mammalian protein complex (RC-1) containing DNA polymerase ε, exonuclease activities, and DNA ligase III has been described that repairs double-strand breaks and deletions by recombination (77). The interaction between XRCC1 and DNA ligase III mentioned above is required for the ligase activity, and the reduced ligase activity in XRCC1 mutants correlates with a deficiency in double-strand break repair (78).

Terminal nucleotidyl transferase (TdT) is involved in immunoglobulin gene somatic recombination (79, 80), a programmed cellular event that is thought to activate a checkpoint (5).

Considerable evidence of checkpoint functions is also available for RF-C and PARP, proteins with a distinct form of the BRCT domain. RF-C is a complex of five subunits that is essential for DNA replication and repair (81, 82). In yeast, the large RF-C subunit containing the BRCT domain is identical to CDC44, a pro-

tein that signals cell cycle arrest by the RAD9/MEC1/RAD53 checkpoint pathway (83). PARP is essential for efficient DNA base excision repair in mammalian cells (84, 85). Moreover, the participation of PARP in both the G1 and G2 checkpoints (86, 87) has been demonstrated.

The presence of a BRCT domain in bacterial DNA ligases is of particular interest, suggesting that similar checkpoint mechanisms may operate in eukaryotes and bacteria. Indeed, a simple DNA damage-responsive checkpoint appears to exist in *Escherichia coli*, ensuring that replication proceeds slowly after UV irradiation while the lesions are being repaired (88). DNA ligase is a major component of excision repair in bacteria (89), and therefore participation of its BRCT domain in the checkpoint is plausible. The BRCT domain in bacterial ligases is an especially attractive object for experimental studies because it seems to be the only BRCT domain encoded in bacterial genomes, and the effects of its disruption may be easily detectable. The transition from the single BRCT domain in bacteria to multiple domains in at least 10 functionally diverse proteins in yeast, and the apparently even greater number in multicellular eukaryotes, is striking. It appears that this diversification of BRCT domains correlates with the evolution in eukaryotes of much more elaborate checkpoint mechanisms than those existing in bacteria.

The available experimental evidence thus indicates that the BRCT domain may be involved in DNA damage-responsive checkpoints in all bacterial and eukaryotic cells. It is not strictly ubiquitous, however, as careful analysis of all proteins encoded in the first completely sequenced archaeal genome, that of *Methanococcus jannaschii* (90), failed to detect any sequences with significant similarity to the BRCT domain (E. V. Koonin, unpublished information). Clearly, the BRCT domain can operate at the level of transcription regulation, but probably also directly affects repair and replication. In terms of the functional classification of checkpoint machinery components (91), it seems plausible that the BRCT domain is a transducer that transmits the signal from DNA damage sensors such as, for example, the amino-terminal domain of PARP (87), to other components of the checkpoint machinery via specific protein-protein interactions. We do not know the specific ligand (or ligands) of the BRCT domain; p53, as shown for 53BP1 (38), may be only one of many, especially as no proteins with significant sequence similarity to p53 have been detected in yeast. A functionally analogous yeast protein is MEC1, and there are at least two yeast proteins (RAD9 and DPB11) containing BRCT domains that are thought to be active immediately upstream of MEC1 (50, 92). These results further support the hypothesis that the BRCT domain is a common element in the organization of checkpoint cascades in yeast and mammals. A variation on this theme is the interaction between BRCT domains in different proteins, e.g., DNA ligase III and XRCC1, which may be important for the formation of checkpoint protein complexes.

The sequence diversity of BRCT domains suggests that their targets may also be quite diverse, perhaps to the extent that protein-protein and protein-DNA interactions are both involved. The identification of these targets and the elucidation of the roles of BRCT domains in checkpoints may be crucial for understanding cell cycle control mechanisms in general and cancer cell evolution in particular.

*Note added in proof.* While this manuscript was being processed for publication, several papers that have important implications for the function of the BRCT domain have been published. A novel protein interacting with BRCA1 and called BARD1 (BRCA1-associated RING domain) has been isolated [Wu. L. C., Wang, Z. W., Tsan, J. T., Spillman, M. A., Phung, A., Xu, X. L., Yang, M.-C., Hwang, L.-Y., Bowcock, A. M., and Baer, R. (1996) Identification of a RING protein that can interact in vivo with the BRCA1 gene product. *Nature Genet.* **14,** 430–440]. BARD1 contains an N-terminal RING finger domain and a C-terminal duplication of the BRCT domain, thus resembling the domain organization of BRCA1 itself. Furthermore, the BRCT domains of BARD1 showed highly statistically significant similarity to those in BRCA1, which is compatible with a critical role of these domains in the functions of both proteins. Evidence has been presented that the XRCC1 protein interacts not only with DNA ligase III but also with DNA polymerase β and with PARP, and furthermore that DNA ligase III functions as a nick sensor [Caldecott, K. W., Aoufouchi, S., Johnson, P., and Shall, S. (1996) XRCC1 polypeptide interacts with DNA polymerase β and possibly poly (ADP-ribose) polymerase, and DNA ligase III is a novel molecular 'nick-sensor' *in vitro. Nucleic Acids Res.* **24,** 4387–4394]. Thus examples are accumulating of different proteins containing BRCT domains interacting with one another. Finally, it has been shown that BRCA1 is expressed in a cell-cycle dependent fashion, with the highest level of expression at the G1/S boundary [Rajan, J. V., Wang, M., Marquis, S. T., and Chodosh, L. A. (1996) Brca2 is coordinately regulated with Brca1 during proliferation and differentiation in mammary epithelial cells. *Proc. Natl. Acad. Sci. USA* **93,** 13078–13083], in accord with its proposed role in a cell-cycle checkpoint.

We thank Alvaro Monteiro and Richard Baer for communication of results before publication.

## REFERENCES

1. Hartwell, L. H., and Weinert, T. A. (1989) Checkpoints: controls that ensure the order of cell cycle events. *Science* **246,** 629–634
2. Murray, A. W. (1992) Creative blocks: cell-cycle checkpoints and feedback controls. *Nature (London)* **359,** 599–604
3. Hartwell, L. H. (1992) Defects in a cell cycle checkpoint may be responsible for the genomic instability of cancer cells. *Cell* **71,** 543–546
4. Kaufmann, W. K., and Paules, R. S. (1996) DNA damage and cell cycle checkpoints. *FASEB J.* **10,** 238–247
5. Hartwell, L. H., and Kastan, M. B. (1994) Cell cycle control and cancer. *Science* **266,** 1821–1828

6. Wiman, K. G. (1993) The retinoblastoma gene: role in cell cycle and cell differentiation. *FASEB J.* **7**, 841–845
7. Slebos, R. J., Lee, M. H., Plunkett, B. S., Kessis, T. D., Williams, B. O., Jacks, T., Hedrick, L., Kastan, M. B., and Cho, K. R. (1994) p53-dependent G1 arrest involves pRB-related proteins and is disrupted by human papillomavirus 16 E7 oncoprotein. *Proc. Natl. Acad. Sci. USA* **91**, 5320–5324
8. Khanna, K. K., Beamish, H., Yan, J., Hobson, K., Williams, R., Dunn, I., and Lavin, M. F. (1995) Nature of G1/S cell cycle checkpoint defect in ataxia-telangiectasia. *Oncogene* **11**, 609–618
9. Koonin, E. V., Altschul, S. F., and Bork, P. (1996) BRCA1 protein products: functional motifs. *Nature Genet.* **13**, 266–268
10. Bork, P., and Gibson, T. (1996) Applying motif and profile searches. *Methods Enzymol.* **266**, 162–184
11. Koonin, E. V., Tatusov, R. L., and Rudd, K. E. (1996) Genome-scale comparison of protein sequences. *Methods Enzymol.* **266**, 295–322
12. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410
13. Karlin, S., and Altschul, S. F. (1990) Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc. Natl. Acad. Sci. USA* **87**, 2264–2268
14. Altschul, S. F., Boguski, M. S., Gish, W., and Wootton, J. C. (1994) Issues in searching molecular sequence databases. *Nature Genet.* **6**, 119–129
15. Wootton, J. C., and Federhen, S. (1996) Analysis of compositionally biased regions in sequence databases. *Methods Enzymol.* **266**, 554–571
16. Tatusov, R. L., Altschul, S. F., and Koonin, E. V. (1994) Detection of conserved segments in proteins: iterative scanning of sequence databases with alignment blocks. *Proc. Natl. Acad. Sci. USA* **91**, 12091–12095
17. Brown, M., Hughey, R., Krogh, A., Mian, I. S., Sjölander, K., and Haussler, D. (1993) Using Dirichlet mixture priors to derive hidden Markov models for protein families. *Intelligent Systems Mol. Biol.* **1**, 47–55
18. Lüthy, R., Xenarios, I., and Bucher, P. (1994) Improving the sensitivity of the sequence profile method. *Protein Sci.* **3**, 139–146
19. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Improved sensitivity of profile searches through the use of sequence weights and gap excision. *Comput. Appl. Biosci.* **10**, 19–29
20. Bucher, P., Karplus, K., Moeri, N., and Hofmann, K. (1996) A flexible motif search technique based on generalized profiles. *Comp. Chem.* **20**, 3–23
21. Sibbald, P. R., and Argos, P. (1990) Weighting aligned protein or nucleic acid sequences to correct for unequal representation. *J. Mol. Biol.* **216**, 813–818
22. Henikoff, S. , and Henikoff, J. (1992) Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. USA* **89**, 10915–10919
23. Hofmann, K., and Bucher, P. (1995) The FHA domain: a putative nuclear signalling domain found in protein kinases and transcription factors. *Trends Biochem. Sci.* **20**, 347–349
24. Birney, E., Thompson, J., and Gibson, T. (1996) PairWise and SearchWise: finding the optimal alignment in a simultaneous comparison of a protein profile against all DNA translation frames. *Nucleic Acids Res.* **24**, 2730–2739
25. Thompson, J. D., Higgins, D. G., and Gibson, T. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680
26. Eddy, S. R., Mitchison, G., and Durbin, R. (1995) Maximum discrimination models of sequence consensus. *J. Comput. Biol.* **2**, 9–23
27. Lawrence, C. E., Altschul, S. F., Boguski, M. S., Liu, J. S., Neuwald, A. F., and Wootton, J. C. (1993) Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. *Science* **262**, 208–214
28. Neuwald, A. F., Liu, J. S., and Lawrence, C. E. (1995) Gibbs motif sampling: detection of bacterial outer membrane repeats. *Protein Sci.* **4**, 1618–1632
29. Schuler, G. D., Altschul, S. F., and Lipman, D. J. (1991) A workbench for multiple alignment construction and analysis. *Proteins Struct. Funct. Genet.* **9**, 180–190
30. Rost, B., and Sander, C. (1994) Combining evolutionary information and neural networks to predict protein secondary structure. Proteins Struct. *Funct. Genet.* **19**, 55–72
31. Rost, B., Sander, C., and Schneider, R. (1994) PHD—an automatic mail server for protein secondary structure prediction. *Comput. Appl. Biosci.* **10**, 53–60
32. Saka, Y., and Yanagida, M. (1993) Fission yeast *cut5+*, required for S phase onset and M phase restraint, is identical to the radiation damage repair gene *rad4+*. *Cell* **74**, 383–393
33. Lehmann, A. R. (1993) Duplicated region of sequence similarity to the human XRCC1 DNA repair gene in the *Schizosaccharomyces pombe rad4/cut5* gene. *Nucleic Acids Res.* **21**, 5274
34. Saka, Y., Fantes, P., Sutani, T., McInerny, C., Creanor, J., and Yanagida, M. (1994) Fission yeast cut5 links nuclear chromatin and M phase regulator in the replication checkpoint control. *EMBO J.* **13**, 5319–5329
35. Araki, H., Leem, S. H., Phongdara, A., and Sugino, A. (1995) Dpb11, which interacts with DNA polymerase II(ε) in *Saccharomyces cerevisiae*, has a dual role in S-phase progression and at a cell cycle checkpoint. *Proc. Natl. Acad. Sci. USA* **92**, 11791–11795
36. Burbelo, P. D., Utani, A., Pan, Z.-Q., and Yamada,Y. (1993) Cloning of the large subunit of activator 1 (replication factor C) reveals homology with bacterial DNA ligases. *Proc. Natl. Acad. Sci. USA* **90**, 11543–11547
37. Bork, P., and Koonin, E. V. (1996) Protein sequence motifs. *Curr. Opin. Struct. Biol.* **6**, 366–376
38. Iwabuchi, K., Bartel, P. L., Li, B., Marraccino, R., and Fields, S. (1994) Two cellular proteins that bind to wild-type but not mutant p53. *Proc. Natl. Acad. Sci. USA* **91**, 6098–6102
39. Caldecott, K. W., McKeown, C. K., Tucker, J. D., Ljungquist, S., and Thompson, L. H. (1994) An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III. *Mol. Cell. Biol.* **14**, 68–76
40. Caldecott, K. W., Tucker, J. D., Stanker, L. H., and Thompson, L. H. (1995) Characterization of the XRCC1-DNA ligase III complex in vitro and its absence from mutant hamster cells. *Nucleic Acids Res.* **23**, 4836–4843
41. Robins, P., and Lindahl, T. (1996) DNA ligase IV from HeLa cell nuclei. *J. Biol. Chem.* **39**, 24257–24261
42. Fotedar, R., Mossi, R., Fitzgerald, P., Rousselle, T., Maga, G., Brickent, H., Messier, H., Kasibhatla, S., Hübscher, U., and Fotedar, A. (1996) A conserved domain of the large subunit of replication factor C binds PCNA and acts like a dominant negative inhibitor of DNA replication in mammalian cells. *EMBO J.* **15**, 4423–4433
43. De Murcia, J. M., Schreiber, V., Molinete, M., Saulier, B., Poch, O., Masson, M., Niedergang, C., and Menissier de Murcia, J. (1994) Structure and function of poly(ADP-ribose) polymerase. *Mol. Cell. Biochem.* **138**, 15–24
44. Nelson, J. R., Lawrence, C. W., and Hinkle, D. C. (1996) Deoxycytidyl transferase activity of yeast REV1 proteins. *Nature (London)* **382**, 729–731
45. Weinert, T. A., and Hartwell, L. H. (1988) The *RAD9* gene controls the cell cycle response to DNA damage in *Saccharomyces cerevisiae*. *Science* **241**, 317–322
46. Al-Khodairy, F., Fotou, E., Sheldrick, K. S., Griffiths, D. J., Lehmann, A. R., and Carr, A. M. (1994) Identification and characterization of new elements involved in checkpoint and feedback controls in fission yeast. *Mol. Biol. Cell* **5**, 147–160
47. Neff, M. W., and Burke, D. J. (1992) A delay in the *Saccharomyces cerevisiae* cell cycle that is induced by a dicentric chromosome and dependent upon mitotic checkpoints. *Mol. Cell. Biol.* **12**, 3857–3864
48. Navas, T.A., Zhou, Z., and Elledge, S.J. (1995) DNA polymerase epsilon links the DNA replication machinery to the S phase checkpoint. *Cell* **80**, 29–39
49. Weinert, T. A., Kiser, G. L., and Hartwell, L. H. (1994) Mitotic checkpoint genes in budding yeast and the dependence of mitosis on DNA replication and repair. *Genes & Dev.* **8**, 652–665
50. Lydall, D., and Weinert, T. (1995) Yeast checkpoint genes in DNA damage processing: implications for repair and arrest. *Science* **270**, 1488–1491
51. Miki, T., Smith, C. L., Long, J. E., Eva, A., and Fleming, T. P. (1993) Oncogene ect2 is related to regulators of small GTP-binding proteins. *Nature (London)* **362**, 462–465
52. Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L. M., Ding, W., et al. (1994) A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* **266**, 66–71
53. Futreal, P., Liu, Q., Shattuck-Eidens, D., Cochran, C., Harshman, K., Tavtigian, S., Bennett, L. M., Swensen, J., Miki, Y., et al. (1994) *BRCA1* mutations in primary breast and ovarian carcinomas. *Science* **266**, 120–122
54. Merajver, S. D., Pham, T. M., Caduff, R. F., Chen, M., Poy, E. L., Cooney, K. A., Weber, B. L., Collins, F. S., Johnston, C., and Frank, T. S. (1995) Somatic mutations in the *BRCA1* gene in sporadic ovarian tumors. *Nature Genet.* **9**, 439–443
55. Lee, E. Y. (1995) Tumor suppressor genes and their alterations in breast cancer. *Semin. Cancer Biol.* **6**, 119–125
56. Chapman, M. S., and Verma, I. M. (1996) Transcriptional activation by BRCA1. *Nature (London)* **382**, 678–679
57. Monteiro, A. N. A., August, A., and Hanafusa, H. (1996) Evidence for a transcriptional activation function of BRCA1 C-terminal region. *Proc. Natl. Acad. Sci. USA* In press
58. Ko, L. J., and Prives, C. (1996) p53: puzzle and paradigm. *Genes & Dev.* **10**, 1054–1072
59. Whyte, P. (1995) The retinoblastoma protein and its relatives. *Semin. Cancer Biol.* **6**, 83–90
60. Subler, M. A., Martin, D. W., and Deb, S. (1994) Overlapping domains of the p53 protein regulate its transcriptional activation and repression functions. *Oncogene* **9**, 1351–1359
61. Singh, P., Coe, J., and Hong,W. (1995) A role for retinoblastoma protein in potentiating transcriptional activation by the glucocorticoid receptor. *Nature (London)* **374**, 562–565
62. Shiio, Y., Yamamoto, T., and Yamaguchi, N. (1992) Negative regulation of Rb expression by the p53 gene product. *Proc. Natl. Acad. Sci. USA* **89**, 5206–5210
63. Osifchin, N. E., Jiang, D., Ohtani-Fujita, N., Carroza, M., Kim, S. J., Sakai, T., and Robbins, P. D. (1994) Identification of a p53 binding site in the human retinoblastoma susceptibility gene promoter. *J. Biol. Chem.* **269**, 6383–6389

64. Rotter, V., Aloni-Grinstein, R., Schwartz, D., Elkind, N. B., Simons, A., Wolkowitz, R., Lavigne, M., Beserman, P., Kapon, A., and Goldfinger, N. (1994) Does wild-type p53 play a role in normal cell differentiation? *Semin. Cancer Biol.* **5**, 229–236
65. Lee, E. Y., Hu, N., Yuan, S. S., Cox, L. A., Bradley, A., Lee, W. H., and Herrup, K. (1994) Dual roles of the retinoblastoma protein in cell cycle regulation and neuron differentiation. *Genes & Dev.* **8**, 2008–2021
66. Hakem, R., De la Pompa, J. L., Sirard, C., Mo, R., Woo, M., Hakem, A., Wakeham, A., Potter, J., et al. (1996) The tumor suppressor gene *Brca1* is required for embryonic cellular proliferation in the mouse. *Cell* **85**, 1009–1023
67. Gowen, L. C., Johnson, B. L., Latour, A. M., Sulik, K. K., and Koller, B. H. (1996) rca1 deficiency results in early embryonic lethality characterized by euroepithelial abnormalities. *Nature Genet.* **12**, 191–194
68. Aboussekhra, A., Vialard, J. E., Morrison, D. E., de la Torre-Ruiz, M. A., Cernáková, L., Fabre, F., and Lowndes, N. F. (1996) A novel role for the budding east *RAD9* checkpoint gene in DNA damage-dependent transcription. *EMBO J.* **5**, 3912–3922
69. Shore, D. (1994) RAP1: a protein regulator in yeast. *Trends Genet.* **10**, 408–412
70. König, P., Giraldo, R., Chapman, L., and Rhodes, D. (1996) The crystal structure of the DNA-binding domain of yeast RAP1 in complex with telomeric DNA. *Cell* **85**, 125–136
71. Conrad, M. N., Wright, J. H., Wolf, A. J., and Zakian, V. A. (1990) RAP1 protein nteracts with yeast telomeres in vivo: overproduction alters telomere structure and decreases chromosome stability. *Cell* **63**, 739–750
72. Lustig, A. J., Kurtz, S., and Shore, D. (1990) Involvement of the silencer and UAS binding protein RAP1 in regulation of telomere length. *Science* **250**, 549–553
73. Kyrion, G., Boakye, K. A., and Lustig, A. J. (1992) C-terminal truncation of RAP1 results in the deregulation of telomere size, stability, and function in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **12**, 5159–5173
74. Sandell, L. L., and Zakian, V. A. (1993) Loss of a yeast telomere: arrest, recovery, and chromosome loss. *Cell* **75**, 729–739
75. Krauskopf, A., and Blackburn, E. H. (1996) Control of telomere growth by interactions of RAP1 with the most distant telomeric repeats. *Nature (London)* **382**, 354–357
76. Müller, T. Gilson, E., Schmidt, R., Giraldo, R., Sogo, J., Gross, H., and Gasser, S. M. (1994) Imaging the asymmetrical DNA bend induced by repressor activator rotein 1 with scanning tunneling microscopy. *J. Struct. Biol.* **113**, 1–12
77. Jessberger, R., Podust, V., Hübscher, U., and Berg, P. (1993) A mammalian protein complex that repairs double-strand breaks by recombination. *J. Biol. Chem.* **268**, 15070–15079
78. Ljungquist, S., Kenne, K., Olsson, L., and Sandstrom, M. (1994) Altered DNA ligase III activity in the CHO EM9 mutant. *Mutat. Res.* **314**, 177–186
79. Landau, N. R., Schatz, D. G., Rosa, M., and Baltimore, D. (1987) Increased frequency of N-region insertion in a murine pre-B-cell line infected with a terminal deoxynucleotidyl transferase retroviral expression vector. *Mol. Cell. Biol.* **7**, 3237–3243
80. Gilfillan, S., Dierich, A., Lemeur, M., Benoist, C., and Mathis, D. (1993) Mice lacking TdT: mature animals with an immature lymphocyte repertoire. *Science* **61**, 1175–1178
81. Tsurimoto, T., and Stillman, B. (1989) Purification of a cellular replication factor, RF-C, that is required for coordinated synthesis of leading and lagging strands during simian virus 40 DNA replication in vitro. *Mol. Cell. Biol.* **9**, 609–619
82. Yoder, B. L., and Burgers, P. M. J. (1991) *Saccharomyces cerevisiae* replication factor C. I. Purification and characterization of its ATPase activity. *J. Biol. Chem.* **33**, 2689–22697
83. Howell, E. A., McAlear, M. A., Rose, D., and Holm, C. (1994) CDC44: a putative nucleotide-binding protein required for cell cycle progression that has homology to subunits of replication factor C. *Mol. Cell. Biol.* **14**, 255–267
84. Durkacz, B. W., Omidiji, O., Gray, D. A., and Shall, S. (1980) (ADP-ribose)n participates in DNA excision repair. *Nature (London)* **283**, 593–596
85. Shall, S. (1994) The function of poly(ADP-ribosylation) in DNA breakage and rejoining. *Mol. Cell. Biochem.* **138**, 71–75
86. Masutani, M., Nozaki, T., Wakabayashi, K., and Sugimura, T. (1995) Role of poly(ADP-ribose) polymerase in cell-cycle checkpoint mechanisms following gamma-irradiation. *Biochimie* **77**, 462–465
87. Schreiber, V., Hunting, D., Trucco, C., Gowans, B., Grunwald, D., De Murcia, G., and De Murcia, J. M. (1995) A dominant-negative mutant of human poly(ADP-ribose) polymerase affects cell recovery, apoptosis, and sister chromatid exchange following DNA damage. *Proc. Natl. Acad. Sci. USA* **92**, 4753–4757
88. Bridges, B. A. (1995) Are there DNA damage checkpoints in *E. coli? BioEssays* **17**, 63–70
89. Dianov, G., and Lindahl, T. (1994) Reconstitution of the DNA base excision-repair pathway. *Curr. Biol.* **4**, 1069–1076
90. Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., et al. (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**, 1058–1073
91. Li, J. J., and Desjaies, R. J. (1993) Exercising self-restraint: Discouraging illicit acts of S and M in eukaryotes. *Cell* **74**, 223–226
92. Weinert, T. A. (1992) Dual cell cycle checkpoints sensitive to chromosome replication and DNA damage in the budding yeast *Saccharomyces cerevisiae*. *Radiat. Res.* **132**, 141–143
93. Doherty, A. J., Serpell, L. C., and Ponting, C. P. (1996) The helix-hairpin-helix DNA-binding motif: a structural basis for non-sequence-specific recognition of DNA. *Nucleic Acids Res.* **24**, 2488–2497
94. Hagemeier, C., Bannister, A. J., Cook, A., and Kouzarides, T. (1993) The activation domain of transcription factor PU.1 binds the retinoblastoma (RB) protein and the transcription factor TFIID in vitro: RB shows sequence similarity to TFIID and FIIB. *Proc. Natl. Acad. Sci. USA* **90**, 1580–1584
95. Gibson, T. J., Thompson, J. D., Blocker, A., and Kouzarides, T. (1994) Evidence for a protein domain superfamily shared by the cyclins, TFIIB, and RB/p107. *Nucleic Acids Res.* **22**, 946–952

*Received for publication August 9, 1996.*
*Accepted for publication October 28, 1996.*