The EMBO Journal Vol.17 No.13 pp.3786–3795, 1998

DNA ligase I is recruited to sites of DNA replication by an interaction with proliferating cell nuclear antigen: identification of a common targeting mechanism for the assembly of replication factories

Alessandra Montecucco[1], Rossella Rossi[1,2], David S.Levin[3], Ronald Gary[4], Min S.Park[4], Teresa A.Motycka[3], Giovanni Ciarrocchi[1], Antonello Villa[5], Giuseppe Biamonti[1,6] and Alan E.Tomkinson[3]

[1]Istituto di Genetica Biochimica ed Evoluzionistica CNR, [2]Dipartimento di Genetica e Microbiologia 'A.Buzzati Traverso' Università di Pavia, via Abbiategrasso 207, 27100 Pavia, [5]Dipartimento di Farmacologia, CNR and B.Ceccarelli Centers and DIBIT Scientific Institute S.Raffaele, via Olgettina 60, Università di Milano, 20132 Milano, Italy, [3]Department of Molecular Medicine, Institute of Biotechnology, The University of Texas Health Science Center at San Antonio, San Antonio, TX 78245 and [4]Life Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

[6]Corresponding author
e-mail: biamonti@igbe.pv.cnr.it

In mammalian cells, DNA replication occurs at discrete nuclear sites termed replication factories. Here we demonstrate that DNA ligase I and the large subunit of replication factor C (RF-C p140) have a homologous sequence of ~20 amino acids at their N-termini that functions as a replication factory targeting sequence (RFTS). This motif consists of two boxes: box 1 contains the sequence IxxFF whereas box 2 is rich in positively charged residues. N-terminal fragments of DNA ligase I and the RF-C large subunit that contain the RFTS both interact with proliferating cell nuclear antigen (PCNA) *in vitro*. Moreover, the RFTS of DNA ligase I and of the RF-C large subunit is necessary and sufficient for the interaction with PCNA. Both subnuclear targeting and PCNA binding by the DNA ligase I RFTS are abolished by replacement of the adjacent phenylalanine residues within box 1. Since sequences similar to the RFTS/PCNA-binding motif have been identified in other DNA replication enzymes and in p21*$^{CIP1/WAF1}$*, we propose that, in addition to functioning as a DNA polymerase processivity factor, PCNA plays a central role in the recruitment and stable association of DNA replication proteins at replication factories.
Keywords: DNA ligase I/PCNA/replication foci/RF-C

Introduction

From bacteriophages to mammals, DNA replication is carried out by a complex multiprotein machinery. In mammalian cells, this process occurs at discrete sites on the nuclear matrix within a defined phase of the cell cycle. Progress in the study of the enzymology of mammalian DNA replication was greatly facilitated by the development of an *in vitro* assay in which bi-directional DNA replication is initiated at the SV40 replication origin in the presence of T antigen (Li and Kelly, 1984). Using this assay, mammalian replication proteins have been purified and the molecular mechanisms of DNA replication have been investigated in reconstitution experiments with replication protein A (RP-A), replication factor C (RF-C), proliferating cell nuclear antigen (PCNA), DNA polymerase δ, DNA polymerase α/primase, RNase H, FEN-1, DNA ligase I and SV40 T antigen (Waga and Stillman, 1994; Waga *et al.*, 1994). It has been suggested that after origin firing, synthesis of the leading strand is initiated by DNA polymerase α/primase and then extended by DNA polymerase δ (or ε), which is tethered to the DNA by an interaction with a PCNA trimer that encircles the template and acts as a sliding clamp. In contrast, synthesis of the lagging strand requires the repeated sequential action of two DNA polymerases to generate the Okazaki fragments that subsequently are joined together. After the initial RNA–DNA fragment is laid down by DNA polymerase α/primase, it is extended up to the 5′ phosphate of the adjacent Okazaki fragment by a complex of DNA polymerase δ (or ε) and PCNA. The involvement of PCNA in both leading- and lagging-strand DNA synthesis and its interactions with the clamp loader RF-C (Fotedar *et al.*, 1996; Cai *et al.*, 1997; Uhlmann *et al.*, 1997), DNA polymerase δ (Jonsson *et al.*, 1995), FEN-1 (Li *et al.*, 1995; Wu *et al.*, 1996) and DNA ligase I (Levin *et al.*, 1997) suggest that PCNA plays a central role in co-ordinating leading- and lagging-DNA synthesis at the replication fork.

To ensure that the genome is duplicated faithfully once and only once per cell cycle, DNA replication is regulated by a complex set of control mechanisms. Different regions of the genome, identified by their characteristic higher order chromatin structure, are duplicated in a precisely defined temporal order during S phase (Pardue and Gall, 1970). This appears to reflect the sequential firing of different origins of replication, but these origins and their regulatory mechanisms are as yet poorly understood. When proliferating cells are pulse-labelled with bromodeoxyuridine (BrdU), discrete sites of DNA replication, termed replication foci, are detectable within the nucleus by indirect immunofluorescence with anti-BrdU antibodies (Nakayasu and Berezney, 1989; Hozak *et al.*, 1994). The co-localization of DNA replication proteins and enzymes such as DNA polymerase α (Hozak *et al.*, 1993), PCNA (Bravo and MacDonald Bravo, 1987; Hozak *et al.*, 1993), DNA ligase I (Montecucco *et al.*, 1995) and RP-A (Cardoso *et al.*, 1993) with sites of BrdU incorporation has been demonstrated. In addition, other proteins such as cyclin A-dependent kinase (cdk) (Cardoso *et al.*, 1993) and DNA-(cytosine-5)-methyltransferase (MCMT) (Leonhardt *et al.*, 1992), which are involved in the regulation of DNA synthesis and methylation of newly synthesized DNA, respectively, also co-localize at the sites of DNA synthesis. Based on these kinds of observations, it has been proposed

that, in either late G_1 or early S phase, proteins involved in DNA replication are assembled at discrete sites termed replication factories (Hozak *et al.*, 1993), whose intranuclear location is maintained by binding to an underlying nuclear matrix.

A prediction of the model described above is that some, if not all, of the replication proteins should contain a sequence that allows them to be targeted to replication factories. Regions of MCMT (Leonhardt *et al.*, 1992) and DNA ligase I (Montecucco *et al.*, 1995; Cardoso *et al.*, 1997) that appear to mediate the recruitment of these enzymes to replication factories have been identified. The N-terminal region of DNA ligase I (residues 1–115) is necessary for subnuclear targeting (Montecucco *et al.*, 1995) but is not required for catalytic activity *in vitro* (Tomkinson *et al.*, 1990). Interestingly, this same fragment of DNA ligase I has been shown recently to bind to PCNA (Levin *et al.*, 1997). In this study, we have examined the relationship between the PCNA-binding and replication factory targeting functions that reside within the N-terminal 115 amino acids of DNA ligase I. We have identified a 20 amino acid sequence that is necessary and sufficient for interaction with PCNA *in vitro* and replication factory targeting *in vivo*. Since similar changes in this sequence disrupt both functions, it appears that the interaction of DNA ligase I with PCNA is required for the association of this enzyme with replication factories in S-phase cells. In addition, we have found that the N-terminus of the large subunit of RF-C, which is homologous to the 20 amino acid sequence of DNA ligase I, also functions as a replication factory targeting sequence (RFTS) and binds to PCNA.

Results

Identification and characterization of the DNA ligase I replication factory targeting sequence

DNA ligase I is composed of a catalytic C-terminal domain (residues 217–919) and an N-terminal domain (residues 1–216) that is not required for enzyme activity but is essential for *in vivo* function (Tomkinson *et al.*, 1990; Kodama *et al.*, 1991; Petrini *et al.*, 1995). In order to identify the protein determinants that are necessary and sufficient for the stable association of DNA ligase I with replication factories, a series of epitope-tagged deletion mutants were constructed as shown in Figure 1. After transient transfection of the recombinant plasmids into the human DNA ligase I-deficient fibroblast cell line 46BR.1G1 (Barnes *et al.*, 1992), the expression and subnuclear localization of the epitope-tagged polypeptides were evaluated by indirect immunofluorescence with the HUC1.1 monoclonal antibody (mAb) that is specific for the epitope tag. To determine whether the tagged proteins co-localized with sites of DNA synthesis, the transfected cells were incubated with the nucleotide analogue BrdU prior to fixation and co-stained with anti-BrdU antibodies. In a typical experiment, ~50% of the cells in the whole population were actively replicating their DNA as determined by BrdU incorporation. A similar proportion of cells expressing epitope-tagged polypeptides were in S phase. The behaviour of the epitope-tagged DNA ligase I polypeptides is summarized in Figure 1. Tagged proteins recruited to the replication factories co-localized with BrdU incorporation sites in at least 85% of the S-phase cells.

A nuclear localization signal (NLS), which is required for active transport into the nucleus through nuclear pores, has been identified in DNA ligase I at residues 119–131 (Montecucco *et al.*, 1995). Deletion of all the residues in the N-terminal domain that extend beyond the NLS (Δ132–216 mutant) had no effect on the recruitment of DNA ligase I catalytic domain to replication factories (Figure 2A). Thus, the N-terminal 131 residues of DNA ligase I are sufficient for both nuclear localization and the specific association of the tagged catalytic domain with replication factories. Subnuclear targeting but not nuclear localization of the tagged polypeptide to the factories was abolished by deletion of residues 2–6 (Δ2–6 mutant) (Figure 2B). In contrast, the polypeptides encoded by the deletion mutants L1–30/NLS, L1–20/NLS (data not shown) and L1–11/NLS (Figure 2C), which contained only the first 30, 20 or 11 residues and the NLS (residues 119–131) from the N-terminal domain, were both located in the nucleus and correctly targeted to the sites of DNA synthesis. These results suggest that the first 11 residues of DNA ligase I are essential for the recruitment of this enzyme to replication factories.

If the first 11 amino acids of DNA ligase I constitute an authentic RFTS, then this sequence should be necessary and sufficient to direct the recruitment of an unrelated polypeptide to sites of DNA replication. To test this, we constructed a plasmid in which the 11 amino acid putative RFTS was fused in-frame to the N-terminus of green fluorescent protein (GFP). Since GFP is small enough to enter the cell nucleus passively (Silver, 1991), we expected that the 11 residues would trigger GFP association with replication factories even in the absence of an NLS. Although this fusion protein was detectable in the nucleus of transfected cells, it was not recruited to sites of DNA synthesis (Figure 3A) and its subcellular distribution was indistinguishable from that of GFP alone (data not shown). However, association of GFP with replication factories was mediated efficiently by the first 20 residues of DNA ligase I as shown in Figure 3B. In S-phase cells, the polypeptide encoded by L(1–20)GFP was also found to co-localize with PCNA (data not shown), a previously identified component of replication factories (Bravo and MacDonald Bravo, 1987; Hozak *et al.*, 1993).

Residues 12–20 were not required for replication factory targeting of an epitope-tagged version of DNA ligase I, in which the first 11 amino acids, the NLS (residues 119–131) and the catalytic domain (residues 217–919) were fused (Figure 2C). In contrast, residues 12–20, in addition to the first 11 amino acids of DNA ligase I, were required to direct GFP to the sites of DNA synthesis (Figure 3A and B). Since residues 12–20 and the NLS of DNA ligase I both contain a high proportion of charged amino acids, we reasoned that the DNA ligase I NLS may be able to substitute for residues 12–20 if it is positioned adjacent to residues 1–11. To test this idea, we constructed a fusion protein with residues 1–11 of DNA ligase I followed immediately by the DNA ligase I NLS (residues 119–131) and GFP at the C-terminus. After transient transfection with this construct, a punctate nuclear pattern was observed in a subset of the transfected cells that was consistent with the association of the GFP fusion protein

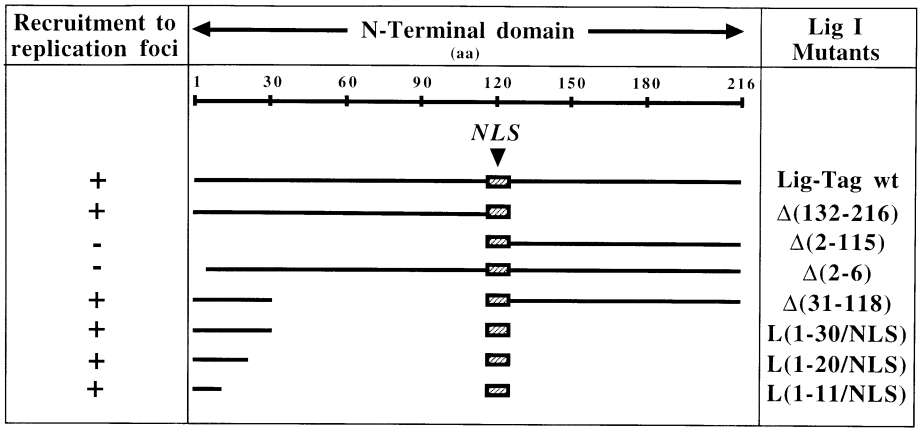

Fig. 1. Structure and subnuclear distribution of epitope-tagged versions of human DNA ligase I. Epitope-tagged full-length DNA ligase I (Lig-Tag wt) and versions of this polypeptide containing the indicated deletions were constructed as described in Materials and methods. All the deletions occur within the N-terminal domain (residues 1–216), while the catalytic domain (residues 217–919) is intact. The DNA ligase I NLS (residues 119–131), which is present in all the epitope-tagged polypeptides, is indicated by the arrowhead. The ability of the DNA ligase I polypeptides to associate with replication factories was assayed by indirect immunofluoresence microscopy and scored as + (targeting proficient) or – (targeting deficient).

with the replication factories (Figure 3C). To confirm that this was the case, the transfected cells were stained with a mAb to PCNA. The co-localization of GFP protein with PCNA in S-phase cells (Figure 3D) demonstrates that the L(1–11/NLS)GFP fusion protein is targeted to DNA replication factories (Bravo and MacDonald Bravo, 1987; Hozak *et al.*, 1993).

To directly address the role of DNA ligase I residues 12–20 in replication factory targeting, we constructed an epitope-tagged version of full-length DNA ligase I with just these residues deleted. This version of DNA ligase I entered the nucleus but, in contrast to epitope-tagged full-length DNA ligase I, did not associate with replication factories (Figure 3E). Thus, residues 12–20 are not required for nuclear localization but constitute an integral part of the DNA ligase I RFTS that cannot be substituted by other sequences rich in basic residues, such as the DNA ligase I NLS, unless the positively charged sequence is immediately adjacent to residues 1–11. Based on these results, we conclude that the DNA ligase I RFTS is situated between residues 1 and 20, and suggest that the RFTS can be subdivided into two distinct regions: box 1 (residues 1–11) that is rich in hydrophobic amino acid residues and box 2 (residues 12–20) that is rich in positively charged amino acid residues (Figure 4A).

In a recent study (Cardoso *et al.*, 1997) it was reported that targeting of DNA ligase I to replication factories is mediated by a bipartite sequence composed of residues 1–28 and 111–179. Since these authors identified this bipartite sequence by its ability to direct an unrelated polypeptide that normally is excluded from the nucleus, it is conceivable that residues 111–179, which contain the previously described DNA ligase I NLS (119–131) (Montecucco *et al.*, 1995), allow the fusion protein to enter the nucleus, whereas residues 1–28, which contain the RFTS defined in this study (1–20), direct the subnuclear localization of the fusion protein.

A homologous RFTS at the N-terminus of the large subunit of replication factor C

Having identified the DNA ligase I RFTS, we used this sequence to search the SwissProt data bank using the BLITZ program (Smith and Waterman, 1981). Although this search identified several proteins containing sequences exhibiting homology with box 1 of the DNA ligase I RFTS, only in the case of the large subunit of the RF-C was this sequence positioned at the N-terminus and flanked by a region rich in positively charged amino acids (Figure 4A). Next we sought to determine whether this motif, whose position and amino acid sequence are conserved among eukaryotic RF-C large subunits, functions as an RFTS. Plasmids expressing fusion proteins with either the first 11 (box 1 of the putative RFTS) or the first 24 residues (boxes 1 and 2 of the putative RFTS) of the human RF-C p140 subunit at the N-terminus of GFP were constructed. After transient transfection of the recombinant plasmids into the human fibroblast cell line 46BR.1G1, we observed that the GFP fusion protein with the intact putative RFTS was targeted efficiently to replication factories (Figure 4B) whereas the GFP fusion protein with only box 1 of the putative RFTS was not (Figure 4C). Interestingly, box 2 could be substituted for efficiently by PKKKRKV, the NLS of SV40 T antigen (Figure 4D), providing additional experimental evidence for the idea that the overall positive charge, not the primary sequence, of box 2 is critical for the function of the RFTS. These results are consistent with those obtained with the homologous RFTS of DNA ligase I and suggest that DNA ligase I and RF-C are targeted to replication factories by the same mechanism.

From the alignment of the DNA ligase I RFTS and the RFTS of the RF-C large subunit, it is apparent that the sequence IxxFF within box 1 is the most conserved sequence within the targeting motif (Figure 4A). Presumably the conservation of the hydrophobic residues reflects the critical roles that these amino acids play in the function of the RFTS. The inactivation of the DNA ligase I RFTS by the deletion of residues 2–6 (Figure 2) is consistent with the notion that the conserved isoleucine residue is essential for RFTS function, in particular since residues corresponding to amino acids 3 and 4 of DNA ligase I are not present in the RFTS of the RF-C large subunit (Figure 4A). To investigate the role of the adjacent phenylalanine residues in the RFTS, we constructed a

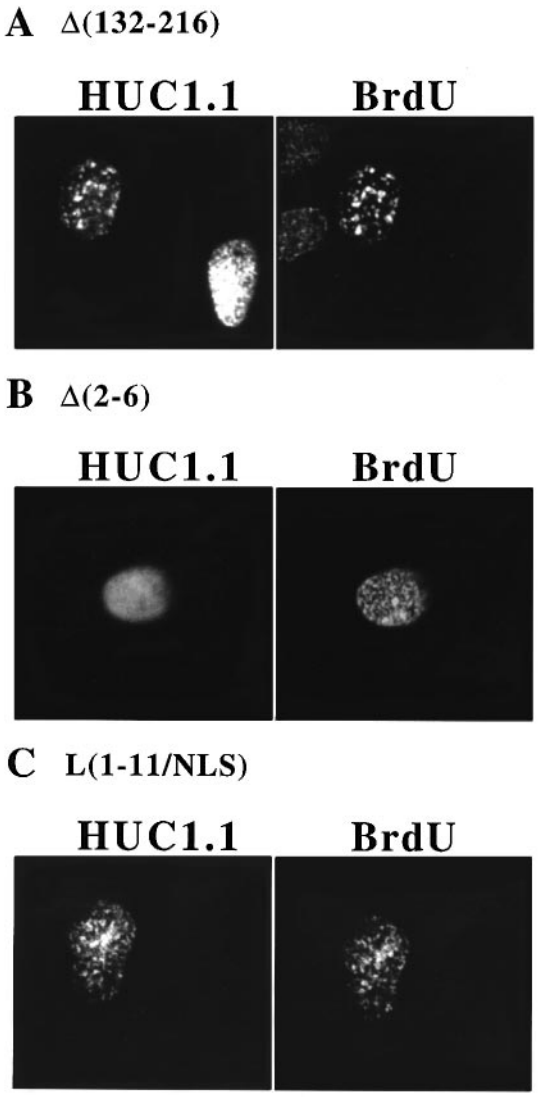

Fig. 2. Intracellular localization of epitope-tagged DNA ligase I polypeptides by indirect immunofluorescence. Human fibroblasts were transfected with plasmids encoding the following epitope-tagged versions of DNA ligase I that are described in Figure 1: (**A**) residues 132–216 deleted (Δ132–216); (**B**) residues 2–6 deleted (Δ2–6); (**C**) residues 1–11 plus the NLS [L(1–11/NLS)] and then pulse-labelled with BrdU prior to methanol fixation. Cells were co-stained with HUC1.1 and FITC-conjugated anti-BrdU mAbs. Epitope-tagged DNA ligase I polypeptides were detected with a rhodamine-conjugated sheep anti-mouse secondary antibody. Antigen–antibody complexes were visualized by confocal laser scanning microscopy. For a better visualization of the ability of the tagged proteins to co-localize with sites of BrdU incorporation, mid-late S-phase nuclei (characterized by a reduced number of large BrdU foci) are shown.

DNA ligase I mutant in which phenylalanines 8 and 9 were both substituted with glycine residues ($F/G_{8–9}$). Replacement of the conserved phenylalanines with glycine residues resulted in no co-localization of the tagged DNA ligase I L($F/G_{8–9}$) polypeptide with sites of DNA replication (Figure 4E), demonstrating that the phenylalanine residues are essential for the function of the RFTS.

The RFTS motif mediates the interaction of DNA ligase I with PCNA

Previously we have shown that DNA ligase I binds to PCNA and that this interaction with trimeric PCNA, which is topologically linked to duplex DNA, tethers DNA ligase I to the DNA molecule (Levin *et al.*, 1997). Since the binding to PCNA is mediated by residues within the N-terminal 118 amino acids of DNA ligase I (Levin *et al.*, 1997), we have examined the relationship between the replication factory targeting and PCNA binding functions that reside within the same region of DNA ligase I. Using a pull-down assay, we found that PCNA bound specifically to GST fusion proteins containing either residues 1–118 (Figure 5A, lanes 4 and 9) or 1–19 (Figure 5A, lanes 2 and 7) of DNA ligase I. Since similar quantities of PCNA bound to the glutathione beads with either of the GST–DNA ligase I fusion proteins as the ligand, it appears that residues 20–118 of DNA ligase I do not contribute to the formation of a stable interaction between PCNA and DNA ligase I.

To investigate the involvement of the phenylalanine residues at positions 8 and 9 of human DNA ligase I in the binding to PCNA, these residues were replaced by alanine residues in the GST–Lig $I_{1–118}$ and the GST–Lig $I_{1–19}$ fusion proteins. In both cases, PCNA binding was abolished by the amino acid substitutions (Figure 5A, lanes 3 and 5), indicating that the phenylalanine residues play a critical role in the protein–protein interaction. To confirm further that the N-terminal amino acids, in particular the phenylalanine residues at positions 8 and 9 of DNA ligase I, are required for the interaction with PCNA, we examined the effect of a peptide corresponding to the first 23 amino acids of DNA ligase I on the ability of PCNA to form a complex with DNA ligase I. As expected, pre-incubation of PCNA with this peptide inhibited co-immunoprecipitation of PCNA by a DNA ligase I antiserum in a concentration-dependent manner (Figure 5B). In contrast, an altered version of the peptide, in which the adjacent phenylalanine residues were replaced by alanine residues, had no such effect (Figure 5B).

A central region of the RF-C large subunit (residues 481–728) has been shown previously to form a stable complex with PCNA *in vitro* (Fotedar *et al.*, 1996). Homologous sequences that also appear to interact with PCNA are present in all the RF-C small subunits (Fotedar *et al.*, 1996; Cai *et al.*, 1997; Uhlmann *et al.*, 1997). However, these conserved regions do not exhibit homology with the RFTS identified in this study. Based on the results of the *in vitro* binding studies with DNA ligase I, we predicted that the RFTS of the RF-C large subunit (residues 1–24) would also interact with PCNA. In Figure 5C, we show that approximately equal amounts of PCNA were retained by glutathione beads with either GST–RF-$C_{1–130}$ or GST–$Lig_{1–118}$ as the ligand. Similar results were also obtained with GST–RF-$C_{1–16}$ as the ligand (data not shown). Since the RFTS of the RF-C large subunit interacts with PCNA, it appears that this polypeptide contains two distinct regions which independently can form a stable complex with PCNA.

Since it appears that the RFTS is composed of two boxes (Figure 4), we sought to determine whether the PCNA interaction also requires the presence of both boxes. The PCNA-binding properties of GST fusion proteins containing either box 1 (residues 1–11) or box 2 (residues 11–20) were examined. Although we were able to detect specific binding of PCNA to glutathione beads with GST–Lig $I_{1–11}$ as the ligand, a reduced amount of PCNA

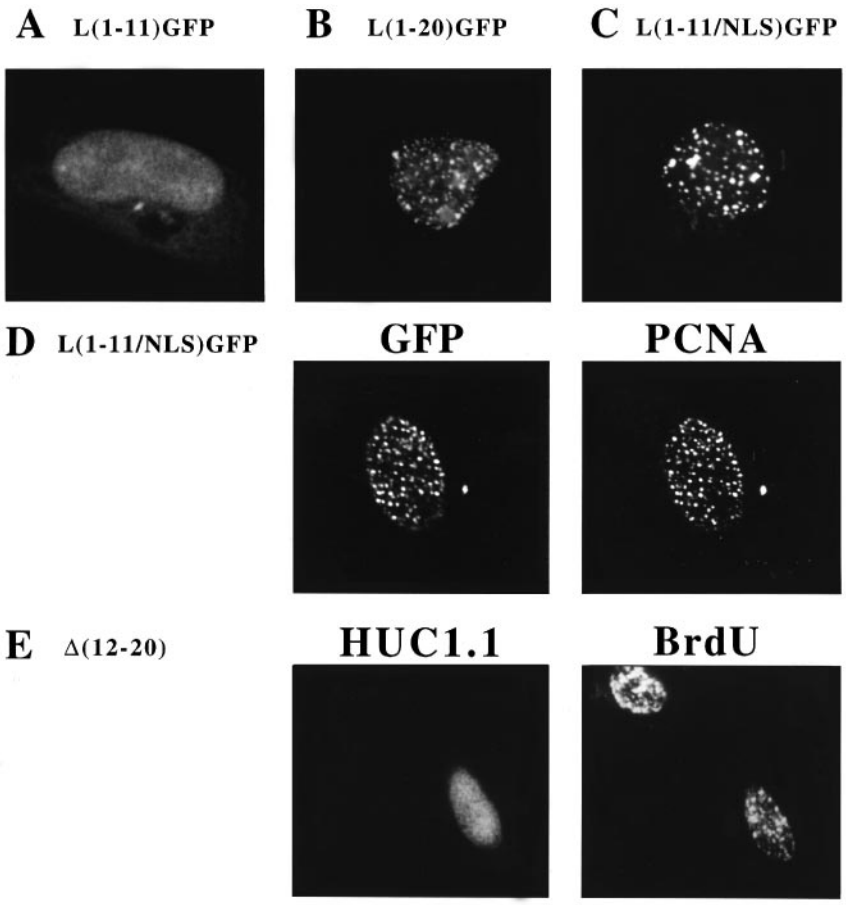

Fig. 3. Identification of the DNA ligase I RFTS by immunofluorescence. Human fibroblasts were transfected with plasmids encoding the following fusion proteins: (**A**) residues 1–11 of DNA ligase I fused to GFP [L(1–11)GFP]; (**B**) residues 1–20 of DNA ligase I fused to GFP [L(1–20)GFP]; (**C**) residues 1–11 and the NLS of DNA ligase I fused to GFP [L(1–11/NLS)] and fixed 48 h later with 4% paraformaldehyde. The autofluorescent signal of GFP was visualized by confocal laser scanning microscopy. About 30% of cells expressing either L(1–20)GFP or L(1–11/NLS) showed a punctate nuclear pattern as exemplified in (B) and (C). The images shown in (A), (B) and (C) are equivalent exposures. (**D**) Cells transfected with the plasmid encoding residues 1–11 and the NLS of DNA ligase I fused to GFP [L(1–11/NLS)] were fixed and stained with anti-PCNA (PC10) mAb. PCNA was detected with a rhodamine-conjugated sheep anti-mouse Ig secondary antibody. The staining pattern of PCNA and the autofluorescent signal of GFP in the same cell are shown. Cells in S phase were identified by their punctate staining pattern after incubation with the PCNA antibody. In 80% of S-phase cells that were also expressing the GFP fusion protein, the distribution of PCNA and fusion protein was coincident. The absence of such a relationship in the remaining fraction of cells appeared to be due to the high level of expression of the fusion protein. (**E**) Human fibroblasts were transfected with a plasmid encoding an epitope-tagged version of DNA ligase I, in which residues 12–20 were deleted (Δ12–20), and 48 h later pulse-labelled with BrdU prior to methanol fixation. Cells were co-stained with HUC1.1 and FITC-conjugated anti-BrdU mAb. The epitope-tagged DNA ligase I polypeptide was detected with a rhodamine-conjugated sheep anti-mouse secondary antibody. Antigen–antibody complexes were visualized by confocal laser scanning microscopy.

(20–50%) bound to these beads compared with glutathione beads with GST–Lig $I_{1–118}$ as the ligand (data not shown). No specific binding of PCNA to glutathione beads with GST–Lig $I_{11–20}$ as the ligand was observed under the same reaction conditions. Thus, it appears that both box 1 and box 2 of the RFTS contribute to the interaction of DNA ligase I with PCNA. We suggest that the conserved hydrophobic amino acids within box 1 are essential for specific binding whereas the amino acids in box 2 stabilize or increase the affinity of the interaction. In summary, we have demonstrated that subnuclear targeting and PCNA-binding functions both reside at the N-termini of both DNA ligase I and the large subunit of RF-C. Since similar amino acid changes within this region inactivate both activities, we conclude that the stable association of DNA ligase I and presumably the RF-C complex with replication factories in S-phase cells occurs via an interaction with PCNA.

Discussion

The molecular mechanisms by which DNA replication proteins are assembled at discrete sites on the nuclear matrix, so-called replication factories (Hozak *et al.*, 1993), are not well understood. By constructing a series of deletion mutants, we have shown that the first 20 amino acids of the non-catalytic N-terminal domain of DNA ligase I are necessary for the association of this enzyme with replication factories in S-phase cells. To provide further evidence that this 20 amino acid sequence functions as an authentic RFTS, we demonstrated that its addition to an unrelated polypeptide, GFP, was sufficient to cause the recruitment of the fusion protein to the sites of ongoing DNA replication.

A sequence with homology to the DNA ligase I RFTS was detected at the N-terminus of the large subunit of RF-C, another essential DNA replication enzyme. Since addition of the N-terminal 24 residues of the human RF-C

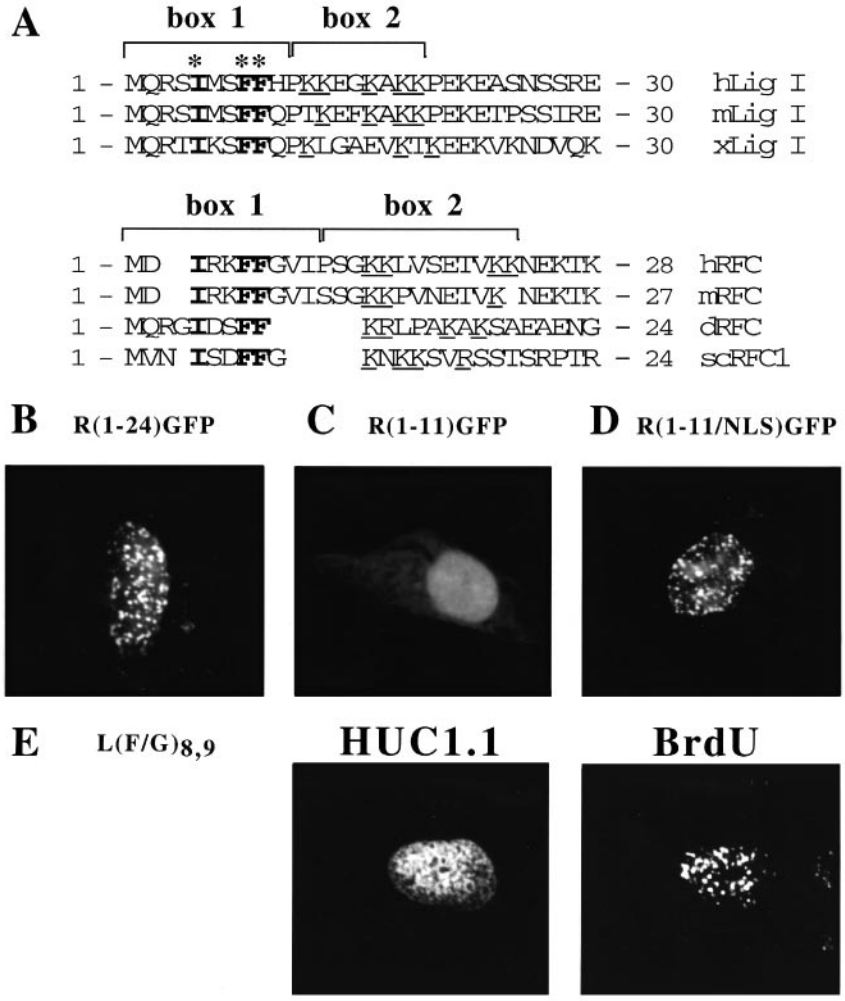

Fig. 4. Identification of an RFTS in the large subunit of RF-C: analysis of RFTS function by amino acid substitution. (**A**) Alignment of sequences at the N-terminus of DNA ligase I from vertebrates and at the N-terminus of the RF-C large subunit from eukaryotes. The positions of box 1 and box 2 within the RFTS are outlined. Conserved hydrophobic amino acids (in bold) are also indicated. Basic residues in box 2 are underlined. h, human; m, mouse; x, *Xenopus laevis*; d, *Drosophila melanogaster*; sc, *Saccharomyces cerevisiae*. Human fibroblasts were transfected with plasmids encoding the following fusion proteins; (**B**) residues 1–24 of human p140 RF-C subunit fused to GFP [R(1–24)GFP]; (**C**) residues 1–11 of human p140 RF-C subunit fused to GFP [R(1–11)GFP]; (**D**) residues 1–11 of human p140 RF-C subunit and the SV40 NLS fused to GFP [R(1–11/NLS)] and fixed 48 h later with 4% paraformaldehyde. The autofluorescent signal of GFP was detected by confocal laser scanning microscopy. About 30% of cells expressing either R(1–24)GFP or R(1–11/NLS) showed a punctate nuclear pattern as exemplified in (B) and (D). The images shown in (B), (C) and (D) are equivalent exposures. (**E**) Human fibroblasts were transfected with a plasmid encoding an epitope-tagged version of DNA ligase I, $L(F/G_{8-9})$, in which the phenylalanine residues at positions 8 and 9 were replaced by glycine residues, and 48 h later pulse-labelled with BrdU prior to methanol fixation. Cells were co-stained with HUC1.1 and FITC-conjugated anti-BrdU mAb. The epitope-tagged DNA ligase I polypeptide was detected with a rhodamine-conjugated sheep anti-mouse secondary antibody. Antigen–antibody complexes were visualized by confocal laser scanning microscopy.

large subunit to GFP also resulted in the association of the fusion protein with sites of replicative DNA synthesis, it appears that this sequence may constitute an RFTS for RF-C. The RFTS motif has been highly conserved in replicative DNA ligases of vertebrates and in the RF-C large subunit encoded by different eukaryotes, indicating the biological importance of this subnuclear targeting function. This is the first example of the identification of a short conserved motif within essential DNA replication enzymes that mediates the recruitment of these polypeptides to replication foci.

The RFTS of DNA ligase I and of RF-C appears to consist of two regions or boxes of ~10 residues in length. The first box contains the conserved sequence motif IxxFF which is essential for RFTS function. Within the second box there is no obvious conservation of amino acid sequence but this box does contain a high proportion of basic amino acids, implying that the net positive charge of this region may be functionally important. This conclusion is supported by the observation that other sequences rich in positively charged amino acids such as the NLS of DNA ligase I or of SV40 T antigen can substitute effectively for box 2 of the RFTS. However, box 2 is not required for nuclear localization but is essential for subnuclear targeting, indicating that these residues are an integral part of the RFTS.

To begin to understand the molecular mechanisms by which the RFTS of DNA ligase I and of RF-C mediate subnuclear targeting, we next sought to identify the protein or proteins that interact with this RFTS motif. Previously we have shown that PCNA interacts with the N-terminal 118 amino acids of DNA ligase I (Levin *et al.*, 1997). In

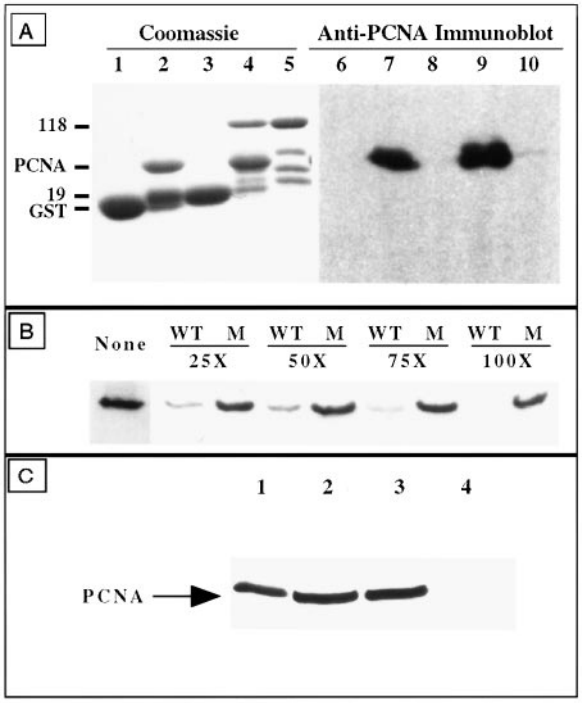

Fig. 5. The PCNA-binding activity of N-terminal fragments of DNA ligase I and the RF-C large subunit: analysis of PCNA binding by amino acid substitution. (**A**) Glutathione beads were incubated with cleared lysates from bacteria expressing GST (lane 1), a GST fusion protein containing residues 1–19 of human DNA ligase I (lane 2), GST–Lig $I_{1–19}$ fusion with an F8A/F9A double amino acid substitution (lane 3), GST fusion protein containing residues 1–118 of human DNA ligase I (lane 4), or GST–Lig $I_{1–118}$ fusion with an F8A/F9A double amino acid substitution (lane 5). An equal aliquot of lysate from bacteria expressing human PCNA was added to each assay. After washing, protein complexes were denatured and separated by SDS–PAGE. Polypeptides were detected by staining with Coomassie Blue (lanes 1–5) or by immunoblotting with PCNA antibody (lanes 6–10). The positions of PCNA (PCNA), GST (GST), GST–Lig $I_{1–118}$ (118) and GST–Lig $I_{1–19}$ (19) are indicated. The additional bands in lanes 4 and 5 are probably due to degradation of the N-terminal fragment of DNA ligase I which is susceptible to proteolysis (Tomkinson *et al.*, 1990). (**B**) PCNA (35 pmol) was pre-incubated with a peptide corresponding to the first 23 residues of human DNA ligase I (WT) or a version of this polypeptide in which the phenylalanine residues at positions 8 and 9 were replaced by alanine residues (M). The molar ratio of the peptide to trimeric PCNA is indicated. The PCNA–peptide mixtures were incubated with DNA ligase I that was linked to protein A–agarose beads by a DNA ligase I antibody. PCNA that bound to the DNA ligase I beads was detected and quantitated by immunoblotting. (**C**) Glutathione beads with either GST–Lig $I_{1–118}$ (lane 2), GST–$RFC_{1–130}$ (lane 3) or GST (lane 4) as the ligand were incubated with PCNA (0.5 μg) as described in Materials and methods. PCNA retained by the beads was detected by immunoblotting. Lane 1, an aliquot of the input PCNA (50 ng).

this study, this mapping has been refined and we have found that the N-terminal 19 amino acids of human DNA ligase I mediate the interaction with PCNA. Thus, replication factory targeting and PCNA binding are mediated by the same short N-terminal fragment of DNA ligase I. Furthermore, both of these activities are inactivated by similar amino acid substitutions and deletions, strongly suggesting that they are not separable functions. The simplest interpretation of these results is that DNA ligase I becomes stably associated with replication factories in S-phase cells by binding to PCNA.

Previous studies have identified a central domain of

```
141-      KRRQTSMTDFYHSKRRLIFSKRKP    -164  p21
162-        TRQTTTTSHFAKGPAKRKP       -174  MCMT
328-LSKSRQGSTQGRLDDFFKVIGSLSSAKR      -355  FEN1
981- KQLDAQQTQLRIDSFFRLAQQEKEDAKR     -1009 XPG
  1-          MDIRKFFGVIPSGKKLVSEIVKKN-24   RFC
  1-        MQRSIMSFFHPKKEGKAKKP      -20   LigI
```

Fig. 6. Alignment of the RFTS of DNA ligase I and of the large subunit of RF-C with regions of $p21^{CIP1/WAF1}$, MCMT, FEN-1 and XPG that interact with PCNA. Amino acid sequences of the proteins shown were obtained from public databases. Conserved hydrophobic residues are indicated in bold. Positively charged residues in the regions corresponding to box 2 of the RFTS are underlined.

RF-C p140 that binds to PCNA (Fotedar *et al.*, 1996; Cai *et al.*, 1997; Uhlmann *et al.*, 1997). Interestingly, this domain, which is conserved in all the smaller subunits of RF-C, does not bind to the same region of PCNA that is bound by the DNA replication inhibitor, $p21^{CIP1/WAF1}$ (Cai *et al.*, 1997; Uhlmann *et al.*, 1997). Here we have shown that RF-C p140 has a second PCNA-interacting region which is located at the N-terminus of this polypeptide and which probably binds to the interdomain connector loop of PCNA (see below). Although the N-terminal region of RF-C p140 is dispensable for the catalytic activities of the RF-C complex (Uhlmann *et al.*, 1997), the conservation of amino acid sequence at the N-terminus of eukaryotic RF-C large subunits implies that this region is functionally important. The existence of a further PCNA-binding domain in the N-terminal region of p140 could explain why an N-terminal fragment of this subunit, that is not required for the catalytic activity of RF-C complex *in vitro* (Cai *et al.*, 1997; Uhlmann *et al.*, 1997), has a concentration-dependent inhibitory effect in an *in vitro* DNA replication assay that can be partially alleviated by increasing the levels of RF-C and PCNA (Rhéaume *et al.*, 1997). We suggest that PCNA binding by the N-terminus of RF-C p140 recruits RF-C to replication factories whereas the central domain of RF-C p140 and similar regions in the small subunits of RF-C interact with PCNA when this clamp is loaded onto and unloaded from the DNA template.

During DNA replication, PCNA also interacts with DNA polymerase δ (Jonsson *et al.*, 1995), DNA ligase I (Levin *et al.*, 1997), FEN-1 (Li *et al.*, 1995; Wu *et al.*, 1996) and possibly DNA polymerase ε (Maga and Hubscher, 1995). Furthermore, XPG, a nuclease involved in nucleotide excision repair (Gary *et al.*, 1997), MCMT (Chuang *et al.*, 1997), which methylates newly synthesized DNA, and $p21^{CIP1/WAF1}$, an inhibitor of cdk kinase and DNA replication (Flores-Rozas *et al.*, 1994), also bind specifically to PCNA. The regions of FEN-1, XPG, MCMT and $p21^{CIP1/WAF1}$ that interact with PCNA have been identified (Krishna *et al.*, 1994; Li *et al.*, 1995; Chuang *et al.*, 1997; Gary *et al.*, 1997), and all share homology with the RFTS motif of DNA ligase I and RF-C (Figure 6). Recently, it has been shown that the PCNA-binding region of MCMT is also responsible for the recruitment of this enzyme to replication factories (Chuang *et al.*, 1997). Thus, it appears that the RFTS/PCNA-binding motif does not have to be positioned at the N-terminus as it is in DNA ligase I and RF-C. Peptide mapping studies (Warbrick *et al.*, 1995) and the subsequent elucidation of the crystal structure of the PCNA-binding peptide of $p21^{CIP1/WAF1}$ complexed with PCNA identified the interdomain connector loop of PCNA as the binding site

for p21$^{CIP1/WAF1}$ (Gulbis *et al.*, 1996). This implies that all of the other DNA replication proteins that contain a similar PCNA-binding motif bind to the same or at least overlapping regions of PCNA. Consistent with this idea, p21$^{CIP1/WAF1}$ or the PCNA-binding peptide of p21$^{CIP1/WAF1}$ inhibits the binding of FEN-1, XPG, MCMT and DNA ligase I to PCNA (Chen *et al.*, 1996; Chuang *et al.*, 1997; Gary *et al.*, 1997; Levin *et al.*, 1997). We propose that PCNA binding is a common mechanism employed by many different proteins involved either in DNA replication or in its regulation to ensure their stable association with replication factories that assemble in late G_1 or early S phase. Given the number of proteins that interact with PCNA, a prediction of this model for the assembly of replication factories is that PCNA will be present in significant molar excess compared with other DNA replication proteins.

If PCNA binding can result in recruitment to replication factories, it seems paradoxical that XPG, which is required for nucleotide excision repair but is not known to function in DNA replication, should interact with PCNA via a region of protein that resembles an RFTS. Although the PCNA-binding regions of the DNA structure-specific endonucleases FEN-1 and XPG are homologous, the binding of PCNA to FEN-1 and to XPG differs significantly when these interactions are examined as a function of salt concentration (Gary *et al.*, 1997). Thus, XPG may be able to interact with PCNA at the sites of nucleotide excision repair but not with PCNA at the sites of replicative DNA synthesis. In fact, a region of XPG protein that is distinct from the PCNA-binding motif is responsible for the dynamic redistribution of this protein which occurs after DNA damage by UV light (Park *et al.*, 1996; Gary *et al.*, 1997). This suggests that the recruitment of XPG to sites of DNA damage occurs by a different mechanism that does not involve PCNA.

In summary, we have identified a conserved RFTS motif within DNA ligase I and the large subunit of RF-C. Furthermore, we have shown that these polypeptides interact with PCNA via their RFTS motifs, indicating that DNA ligase I and probably the RF-C complex are recruited to replication factories by binding to PCNA. Since this RFTS/PCNA-binding motif is present in several other proteins involved in DNA replication, we suggest that PCNA binding may play a major role in the recruitment of DNA replication proteins to replication factories. Interestingly, the cell-cycle inhibitor, p21$^{CIP1/WAF1}$, binds to PCNA via the same motif. Therefore, it is possible that p21$^{CIP1/WAF1}$ does not inhibit DNA replication directly by binding to the PCNA molecules functioning at the replication fork but instead binds to PCNA within replication factories and causes the release of DNA replication proteins from these structures.

Materials and methods

Construction of tagged DNA ligase I mutants

All the Lig I mutants were obtained by PCR-mediated mutagenesis of the pLig I-Tag plasmid (Montecucco *et al.*, 1995) that encodes DNA ligase I with a C-terminal muscular actin epitope which is recognized by the HUC1.1 mAb (ICN, USA). The mutagenesis procedure entails the replacement of the 1438 bp *Eco*RI fragment of pLig I-Tag with a mutated version amplified from human DNA ligase I cDNA (ATTC #65857; Barnes *et al.*, 1990) by PCR. The *Eco*RI fragment (nucleotides 105–1543) contains the Kozak consensus sequence, the ATG (nucleotide 121) and codes for the N-terminal domain (216 residues), which contains the NLS. The plasmid construct, pΔ(2–6), which encodes a polypeptide with the indicated N-terminal deletion, was obtained by using the primer pair (S2,A1) (Table I). The remaining plasmid constructs, which contain either internal deletions or substitutions within the DNA ligase I open reading frame (ORF), were produced by a two-step PCR procedure that has been described previously (Montecucco *et al.*, 1995). Briefly, human DNA ligase I cDNA was the template in two amplification reactions carried out with two different pairs of primers. One primer of each pair was complementary to a cDNA sequence outside the 1438 bp *Eco*RI fragment. In all cases, S1 and A1 were the external primers. The remaining primers, which were located within the *Eco*RI fragment, were complementary to each other at their 5′ ends. Since the mutation to be introduced was located within the complementary region, the two resulting products covered the entire *Eco*RI fragment but only overlapped in a short region containing the mutation. The amplification products were gel-purified, mixed and further amplified using the two external primers S1 and A1. The plasmid construct pΔ(132–216) was produced by using primer pairs (S1,A3/S4,A1) and DNA ligase I cDNA as the template. To generate the plasmids pL(1–30/NLS), pL(1–20/NLS) and pL(1–11/NLS), we used the following pairs of primers: (S1,A2/S3,A1), (S1,A4/S5,A1) and (S1,A5/S6,A1), respectively, and pΔ(132–216) as the template. To produce pΔ(12–20), we used primers (S1,A12/S11,A1) and DNA ligase I cDNA as the template. The double substituted clone pL(F/$G_{8–9}$) was obtained by amplifying human DNA ligase I cDNA with primers (S1,A6/S7,A1). All amplification reactions were carried out with *Pwo* DNA polymerase (Boehringer Mannheim, Germany). Plasmids were analysed by restriction mapping, and the presence of mutation was verified by DNA sequencing (Thermo Sequenase kit, Amersham, UK). All the oligonucleotides were purchased from Tib.Mol.-Biol. (Genova, Italy). Expression of an appropriately sized fusion protein was verified by immunoblotting analysis of extracts from transfected cells with HUC1.1-1 mAb.

Construction of GFP fusions

Human DNA ligase I and RF-C p140 cDNA sequences were subcloned in-frame at the N-terminus of the GFP ORF in a modified pEGFP-N1 vector (Clontech, USA), in which the ATG and the Kozak consensus sequence were deleted to avoid undesirable translation products starting at the first methionine of the GFP protein. The plasmid construct pL(1–11/NLS)GFP fusion was obtained by subcloning a fragment amplified from the pL(1–11/NLS) plasmid with the primer pair (S1,A7) into the *Eco*RI–*Bam*HI sites of the vector. The pL(1–20)GFP fusion was obtained by subcloning a fragment amplified from DNA ligase I cDNA with the primer pair (S1,A8) into the *Eco*RI–*Kpn*I sites of the vector. To produce pL(1–11)GFP, the complementary oligonucleotides S8 and A9 were annealed and the resultant duplex cloned into the *Eco*RI–*Bam*HI sites of the vector. To obtain plasmids pR(1–11)GFP and pR(1–24)GFP, complementary oligonucleotides (S9,A10) and (S10,A11) were annealed and the resultant duplexes cloned into the *Eco*RI–*Kpn*I sites of the vector. To obtain pR(1–11/NLS)GFP, the complementary oligonucleotides (S12,A13) were annealed and the resultant duplex was cloned into the *Kpn*I–*Bam*HI sites of pR(1–11)GFP.

Cell culture, transfection and immunofluorescence

The SV40-transformed human fibroblast cell line, 46BR.1G1 (ECACC, UK) was grown in complete Dulbecco's modified Eagle's medium (DMEM; Sigma) with 15% fetal calf serum (FCS). Transfections, cell fixation and indirect immunofluorescence microscopy were performed as described previously (Weighardt *et al.*, 1995). Epitope-tagged proteins were detected with the anti-muscular actin mAb, HUC1.1 (ICN, USA) and a rhodamine-conjugated sheep anti-mouse IgG F(ab′)$_2$ fragment (Boehringer Mannheim, Germany). To detect epitope-tagged polypeptides and sites of DNA synthesis simultaneously, the cells were grown in 50 μM BrdU (Sigma) for 1 h immediately prior to methanol fixation (Montecucco *et al.*, 1995). After sequential incubation with HUC1.1 mAb and anti-mouse secondary antibody, antigen–antibody complexes were fixed in place by immersion in ice-cold methanol for 5 min. The coverslips were treated with 2 M HCl at 37°C for 1 h to denature DNA and then neutralized with 0.1 M borate buffer (pH 8.5). After washing, cells were incubated for 60 min at room temperature with the fluorescein isothiocyanate (FITC)-conjugated anti-BrdU mAb (Boehringer Mannheim, Germany). To detect GFP fluorescence, the cells were fixed in 4% paraformaldehyde (Sigma) for 10 min at room temperature. PCNA was detected in paraformaldehyde-fixed cells after permeabilization with cold methanol (5 min), by indirect immunofluorescence with the

Table I. PCR primers

S-1	5′-C$_{95}$AGGAAGGGAGAATTCTGACGC$_{116}$
S-2	5′-G$_{105}$AATTCTGACGCCAAC$_{120}$-A$_{136}$TG TCA TTT TTC CAC CCC$_{153}$
S-3	5′-C$_{191}$A TCC AAT AGC AGC AGA GAG$_{210}$-C$_{475}$CG AAG CGT CGC ACA GCT CG$_{494}$
S-4	5′-G$_{494}$G AAG CAG CTC CCG AAA CGG$_{513}$-C$_{769}$AG ACC AAG CCT CCC CGC AG$_{788}$
S-5	5′-G$_{162}$ GGT AAA GCA AAG AAG CCT$_{180}$-C$_{475}$CG AAG CGT CGC ACA GCT CG$_{494}$
S-6	5′-C$_{135}$ ATG TCA ATT TTC CAC CCC$_{153}$-C$_{475}$CG AAG CGT CGC ACA GCT CG$_{494}$
S-7	5′-C$_{135}$ ATG TCA GGT GGC CAC CCC AAG AAA GAG GG$_{164}$
S-8	5′-AATTCTGACGCCAAC A$_{121}$TG CAG CGA AGT ATC ATG TCA TTT TTC CAC CC$_{152}$G
S-9	5′-AATTCG$_{201}$CTGCG ATG GAC ATT CGG AAA TTC TTT GGA GTA ATA CCA$_{239}$ CCGGTAC
S-10	5′-AATTCG$_{201}$CTGCG ATG GAC ATT CGG AAA TTC TTT GGA GTA ATA CCA AGT GGA AAG AAA CTT GTA AGT GAA ACA GTA AAG AAG AAT$_{278}$ CCGGTAC
S-11	5′-T$_{144}$TTC CAC CCC$_{153}$-G$_{181}$AG AAG GAG GCA TCC AAT AG$_{200}$
S-12	5′-CAA AAA AGA AGA GAA AGG TAC GG
A-1	5′-G$_{1556}$C TGG TGG GAA TTC TTG GCC C$_{1536}$
A-2	5′-C$_{494}$G AGC TGT GCG ACG CTT CGG$_{475}$-C$_{210}$TC TCT GCT GCT ATT GGA TG$_{191}$
A-3	5′-C$_{788}$T GCG GGG AGG CTT GGT CTG$_{769}$-C$_{513}$CG TTT CGG GAG CTG CTT CC$_{494}$
A-4	5′-C$_{494}$G AGC TGT GCG ACG CTT CGG$_{475}$-A$_{180}$GG CTT CTT TGC TTT ACC C$_{62}$
A-5	5′-C$_{494}$G AGC TGT GCG ACG CTT CGG$_{475}$-G$_{153}$GG GTG GAA AAA TGA CAT G$_{135}$
A-6	5′-G$_{153}$ GGG TGG CCA CCT GAC ATG ATA CTT CGC TG$_{123}$
A-7	5′-GGGGATC C$_{512}$TG AAT GGT CCG TTT CGG GAG CTG CTT C$_{495}$
A-8	5′-GGGGGTACCGG A$_{180}$GG CTT CTT TGC TTT ACC C$_{162}$
A-9	5′-GATCCG$_{152}$G GTG GAA AAA TGA CAT GAT ACT TCG CTG CAT$_{121}$ GTTGGCGTCAG
A-10	5′-CGGT$_{239}$GG TAT TAC TCC AAA GAA TTT CCG AAT GTC CAT CGCAGC$_{201}$G
A-11	5′-CGG A$_{278}$TT CTT CTT TAC TGT TTC ACT TAC AAG TTT CTT TCC ACT TGG TAT TAC TCC AAA GAA TTT CCG AAT GTC CAT CGCAGC$_{201}$G
A-12	5′-C$_{190}$ CTC CTT CTC181-G$_{153}$GG GTG GAA AAA TGA CAT G$_{135}$
A-13	5′-GAT CCC GTA CCT TTC TCT TCT TTT TTG GTA C

PC10 mAb (Santa Cruz Biotechnology, USA). DNA was stained with 0.2 μg/ml 4,6′-diamidino-2-phenylindole (DAPI; Sigma). Optical sections were obtained using the Bio-Rad MRC-1024 confocal microscope (Bio-Rad). Micrographs were taken using a Focus Imagecorder Plus (Focus Graphics Inc., USA) on Kodak Tmax 100 film.

Bacterial expression plasmids

A fragment of human DNA ligase I cDNA encoding residues 1–118 of human DNA ligase I (Levin *et al.*, 1997) was subcloned in-frame into a pGEX vector (Pharmacia) to generate a recombinant plasmid encoding a GST–Lig I$_{1–118}$ fusion protein. An F8A/F9A derivative of this plasmid, in which phenylalanine residues 8 and 9 were replaced with alanines, was made using the QuickChange Mutagenesis procedure (Stratagene, La Jolla, CA) and mutagenic primer pairs 5′-CAGCGAAGTATCATGT-CTGCAGCCCACCCCAAGAAAGAGGGTAAAG-3′ and 5′- CTTTA-CCCTCTTTCTTGGGGTGGGCTGCAGACATGATACTTCGCTG-3′. Mutagenesis created a *Pst*I site to facilitate screening.

An expression plasmid for the production of a GST fusion protein containing residues 1–19 of DNA ligase I was made by ligating annealed synthetic oligonucleotides into an *Eco*RI–*Xho*I-linearized pGEX-4T-1 vector (Pharmacia Biotech, Piscataway, NJ). The oligonucleotides corresponded to each strand of the first 19 codons of the human DNA ligase I cDNA sequence, flanked at the 5′ side by an *Eco*RI site for cloning and a *Hin*dIII site to facilitate screening, and flanked at the 3′ side by a TAA stop codon followed by a *Xho*I site for cloning. A second plasmid directing the expression of an F8A/F9A mutant form of the GST–Lig I$_{1–19}$ fusion protein was created similarly, except that codons 8 and 9 in the annealed oligo pair were GCA and GCA. Recombinant plasmids expressing GST–Lig I$_{1–11}$ and GST–Lig I$_{11–20}$ fusion proteins were constructed by essentially the same methodology. A fragment of cDNA encoding the N-terminal 130 amino acids of the RF-C large subunit was amplified from a human testis cDNA library (Clontech) by PCR and subcloned in-frame into the vector pGSTag (Ron and Dressler, 1992) to generate the recombinant plasmid, pGST-RFC$_{1–130}$. The sequence of the subcloned fragment was verified by dideoxy sequencing.

Fusion protein binding assay

The *Escherichia coli* strain BL21(DE3) was used as the host to express human PCNA and the GST fusion proteins. Expression was induced with 0.8 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Cells were lysed in wash buffer (50 mM Tris–HCl, 150 mM NaCl pH 7.4) supplemented with 2 mM EDTA and 0.2 mg/ml lysozyme at a ratio of 1 ml per 30 ml of culture. The cell lysates were clarified by centrifugation at 16 000 *g*, and 240 μl of GST fusion protein cell lysate and 200 μl of PCNA cell lysate were mixed with 100 μl of 40% glutathione–agarose beads (Sigma, St Louis, MO). The reaction mixtures were incubated with gentle rocking for 2 h at 4°C. After collection by centrifugation, the beads were washed five times with 0.8 ml of wash buffer. Protein complexes bound to the beads were denatured by heating to 100°C with 80 μl of 2× Laemmli sample buffer (Laemmli, 1970). After separation by SDS–PAGE, polypeptides were detected by Coomassie Blue staining or transferred to nitrocellulose membranes prior to detection by immunoblotting with the PCNA PC10 mAb from Santa Cruz Biotechnology (Levin *et al.*, 1997).

GST–Lig I$_{1–118}$, GST–Lig I$_{1–11}$, GST–Lig I$_{11–20,}$ GST and human PCNA were purified as described previously (Levin *et al.*, 1997). Glutathione beads with 2 μg of GST–Lig I$_{1–11}$, GST–Lig I$_{11–20}$, GST–RFC$_{1–130}$ or GST as the ligand were mixed with 0.5 μg of purified PCNA in binding buffer (50 mM HEPES–KOH pH 7.5, 100 mM NaCl, 1 mM dithiothreitol, 0.1 mM EDTA, 1% NP-40, 10% glycerol) at 25°C for 20 min. After washing extensively with binding buffer, proteins bound to the beads were denatured by heating to 90°C with 2× Laemmli sample buffer (Laemmli, 1970). After separation by SDS–PAGE, PCNA was detected by immunoblotting with the PCNA PC10 mAb from Santa Cruz Biotechnology (Levin *et al.*, 1997).

Peptide competition assay

A peptide corresponding to the first 23 amino acids of human DNA ligase I and a version of this peptide in which the phenylanine residues at positions 8 and 9 were replaced by alanine residues, were synthesized by the peptide synthesis facility at The University of Texas Health Science Center at San Antonio, TX. Recombinant DNA ligase I (35 pmol) purified from baculovirus-infected insect cells was incubated with (5 μl) rabbit polyclonal antiserum specific for DNA ligase I in 200 μl of binding buffer for 1 h at 4°C (Levin *et al.*, 1997). Protein A–agarose beads (20 μl) were added and the incubation was continued for 1 h. After collection by centrifugation, the agarose beads were washed extensively with binding buffer. Purified human PCNA (35 pmol) was pre-incubated with either a wild type or mutant DNA ligase I peptide for 10 min at 4°C in 200 μl of binding buffer with 5% bovine serum

albumin (BSA). The PCNA–peptide solution was then added to the protein A–DNA ligase I agarose beads and mixed for 20 min at 4°C. After collection by centrifugation, the agarose beads were washed extensively with binding buffer. PCNA bound to the protein A–agarose beads was detected by immunoblotting. In assays in which either DNA ligase I or the DNA ligase I antiserum was omitted, PCNA did not bind to the protein A–agarose beads. Addition of the DNA ligase I peptides had no effect on the amount of DNA ligase I bound to the protein A–agarose beads (data not shown).

Acknowledgements

We thank Dr Giovanni Maga for helpful discussion. This work was supported by grants from the Associazione Italiana per la Ricerca sul Cancro (AIRC), the PF 'Ciclo Cellulare e Apoptosi' CNR, the San Antonio Cancer Institute (to A.E.T.) and the Department of Health and Human Services (GM47251 to A.E.T. and CA71630 to M.S.P.). R.R. was supported by a fellowship from the Italian PhD program. D.S.L. was supported by the Training Program in the Molecular Basis of Breast Cancer (DAMD1794-J-4147) from the Department of the Army. T.M. was supported by a summer fellowship from the Institute of Biotechnology.

References

Barnes,D.E., Johnston,L.H., Kodama,K., Tomkinson,A.E., Lasko,D.D. and Lindahl,T. (1990) Human DNA ligase I cDNA: cloning and functional expression in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA*, **87**, 6679–6683.

Barnes,D.E., Tomkinson,A.E., Lehman,A.R., Webster,A.D.B. and Lindahl,T. (1992) Mutations in the DNA ligase I gene in an individual with immunodeficiencies and cellular hypersensitivity to DNA-damaging agents. *Cell*, **69**, 495–504.

Bravo,R. and MacDonald Bravo,H. (1987) Existence of two populations of cyclin/proliferating cell nuclear antigen during the cell cycle: association with DNA replication sites. *J. Cell Biol.*, **105**, 1549–1554.

Cai,J., Gibbs,E., Uhlmann,F., Phillips,B., Yao,N., O'Donnel,M. and Hurwitz,J. (1997) A complex consisting of human replication factor C p40, p37 and p36 subunits is a DNA-dependent ATPase and an intermediate in the assembly of the holoenzyme. *J. Biol. Chem.*, **272**, 18974–18981.

Cardoso,M.C., Leonhardt,H. and Nadal-Ginard,B. (1993) Reversal of terminal differentiation and control of DNA replication: cyclin A and cdk2 specifically localize at subnuclear sites of DNA replication. *Cell*, **74**, 979–992.

Cardoso,M.C., Joseph,C., Rahn,H.-P., Reusch,R., Nadal-Ginard,B. and Leonhardt,H. (1997) Mapping and use of a sequence that targets DNA ligase I to sites of DNA replication *in vivo*. *J. Cell Biol.*, **139**, 579–587.

Chen,J., Chen,S., Saha,P. and Dutta,A. (1996) $p21^{Cip1/Waf1}$disrupts the recruitment of human Fen1 by proliferating-cell nuclear antigen into the DNA replication complex. *Proc. Natl Acad. Sci. USA*, **93**, 11597–11602.

Chuang,L.S.-H., Ian,H.-I., Koh,T.-W., Ng,H.-H., Xu,G. and Li,B.F.L. (1997) Human DNA-(cytosine-5) methyltransferase–PCNA complex as a target for $p21^{WAF1}$. *Science*, **277**, 1996–2000.

Flores-Rozas,H., Kelman,Z., Dean,F.B., Pan,Z.-Q., Harper,J.W., Elledge,S.J., O'Donnell,M. and Hurwitz,J. (1994) Cdk-interacting protein 1 directly binds with proliferating cell nuclear antigen and inhibits DNA replication catalyzed by the DNA polymerase δ holoenzyme. *Proc. Natl Acad. Sci. USA*, **91**, 8655–8659.

Fotedar,R. *et al*. (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.

Gary,R., Ludwig,D.L., Cornelius,H.L., MacInnes,M.A. and Park,M.S. (1997) The DNA repair endonuclease XPG binds to proliferating cell nuclear antigen (PCNA) and shares sequence elements with the PCNA-binding regions of FEN-1 and cyclin-dependent kinase inhibitor p21. *J. Biol. Chem.*, **272**, 24522–24529.

Gulbis,J.M., Kelman,Z., Hurwitz,J., O'Donnell,M. and Kuriyan,J. (1996) Structure of the C-terminal region of $p21^{WAF1/CIP1}$ complexed with human PCNA. *Cell*, **87**, 297–306.

Hozak,P., Hassan,A.B., Jackson,D.A. and Cook,P.R. (1993) Visualization of replication factories attached to a nucleoskeleton. *Cell*, **73**, 361–373.

Hozak,P., Jackson,D.A. and Cook,P.R. (1994) Replication factories and nuclear bodies: the ultrastructural characterization of replication sites during the cell cycle. *Cell*, **107**, 2191–2202.

Jonsson,Z.O., Podust,V.N., Podust,L.M. and Hubsher,U. (1995) Tyrosine 114 is essential for the trimeric structure and the functional activities of human proliferating cell nuclear antigen. *EMBO J.*, **14**, 5745–5751.

Kodama,K., Barnes,D.E. and Lindahl,T. (1991) *In vitro* mutagenesis and functional expression in *E.coli* of a cDNA encoding the catalytic domain of human DNA ligase I. *Nucleic Acids Res.*, **19**, 6093–6099.

Krishna,T.S.R., Kong,X.P., Gary,S., Burgers,P.M. and Kuryan,J. (1994) Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell*, **79**, 1233–1243.

Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.

Leonhardt,H., Page,A.W., Weier,H.U. and Bestor,T.H. (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell*, **71**, 865–873.

Levin,D.S., Bai,W., Yao,N., O'Donnell,M. and Tomkinson,A.E. (1997) An interaction between DNA ligase I and proliferating cell nuclear antigen: implication for Okazaki fragment synthesis and joining. *Proc. Natl Acad. Sci. USA*, **94**, 12863–12868.

Li,J.J. and Kelly,T.J. (1984) Simian virus 40 DNA replication *in vitro*. *Proc. Natl Acad. Sci. USA*, **81**, 6973–6977.

Li,X., Li,J., Harrington,J., Lieber,M.R. and Burgers,P.M. (1995) Lagging strand DNA synthesis at the eukaryotic replication fork involves binding and stimulation of FEN-1 by PCNA. *J. Biol. Chem.*, **270**, 22109–22112.

Maga,G. and Hubscher,U. (1995) DNA polymerase ε interacts with proliferating cell nuclear antigen in primer recognition and elongation. *Biochemistry*, **34**, 891–901.

Montecucco,A., Savini,E., Weighardt,F., Rossi,R., Ciarrocchi,G., Villa,A. and Biamonti,G. (1995) The N-terminal domain of human DNA ligase I contains the nuclear localisation signal and directs the enzyme to sites of DNA replication. *EMBO J.*, **14**, 5379–5386.

Nakayasu,H. and Berezney,R. (1989) Mapping replication sites in the eukaryotic cell nucleus. *J. Cell Biol.*, **108**, 1–11.

Pardue,M.L. and Gall,J.G. (1970) Chromosomal localization of mouse satellite DNA. *Science*, **168**, 1356–1358.

Park,M.S., Knauf,J.A., Pendergrass,S.H., Coulon,C.H., Strniste,G.F., Marrone,B.L. and MacInnes,M.A. (1996) Ultraviolet-induced movement of the human DNA repair protein, xeroderma pigmentosum type G, in the nucleus. *Proc. Natl Acad. Sci. USA*, **93**, 8368–8373.

Petrini,H.J., Xiao,Y. and Weaver,D.T. (1995) DNA ligase I mediates essential functions in mammalian cells. *Mol. Cell. Biol.*, **15**, 4303–4308.

Rhéaume,E., Cohen,L.Y., Uhlmann,F., Lazure,C., Alam,A., Hurwitz,J., Sékaly,R.-P. and Denis,F. (1997) The large subunit of replication factor C is a substrate for capsase-3 *in vitro* and is cleaved by a capsase-3-like protease during Fas-mediated apoptosis. *EMBO J.*, **16**, 6346–6354.

Ron,D. and Dressler,H. (1992) pGSTag—a versatile bacterial expression plasmid for enzymatic labeling of recombinant proteins. *BioTechniques*, **13**, 866–888.

Silver,P.A. (1991) How proteins enter the nucleus. *Cell*, **64**, 489–497.

Smith,T.F. and Waterman,M.S. (1981) Identification of common molecular sequences. *J. Mol. Biol.*, **147**, 195–197.

Tomkinson,A.E., Lasko,D.D., Daly,G. and Lindahl,T. (1990) Mammalian DNA ligases. Catalytic domain and size of DNA ligase I. *J. Biol. Chem.*, **265**, 12611–12617.

Uhlmann,F., Cai,J., Gibbs,E., O'Donnell,M. and Hurwitz,J. (1997) Deletion analysis of the large subunit p140 in human replication factor C reveals regions required for complex formation and replication activities. *J. Biol. Chem.*, **272**, 10058–10064.

Waga,S. and Stillman,B. (1994) Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication *in vitro*. *Nature*, **369**, 207–212.

Waga,S., Bauer,G. and Stillman,B. (1994) Reconstitution of complete SV40 DNA replication with purified replication factor. *J. Biol. Chem.*, **269**, 10923–10934.

Warbrick,E., Lane,D.P., Glover,D.M. and Cox,L.S. (1995) A small peptide inhibitor of DNA replication defines the site of interaction between the cyclin-dependent kinase inhibitor p21WAF1 and proliferating cell nuclear antigen. *Curr. Biol.*, **5**, 275–282.

Weighardt,F., Biamonti,G. and Riva,S. (1995) Nucleo-cytoplasmic distribution of human hnRNP proteins: a search for the targeting domains in hnRNP A1. *J. Cell Sci.*, **108**, 545–555.

Wu,X., Li,J., Li,X., Hsieh,C.-L., Burgers,P.M.J. and Lieber,M.R. (1996) Processing of branched DNA intermediates by complex of human FEN-1 and PCNA. *Nucleic Acids Res.*, **24**, 2036–2043.

Received March 16, 1998; revised and accepted April 30, 1998