MOLECULAR AND CELLULAR BIOLOGY, Sept. 1995, p. 4661–4671 Vol. 15, No. 9
0270-7306/95/$04.00+0

Characterization of the Five Replication Factor C Genes of *Saccharomyces cerevisiae*

GERHARD CULLMANN,[1] KAREN FIEN,[2]† RYUJI KOBAYASHI,[1] AND BRUCE STILLMAN[1]*

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724,[1] and Graduate Program in Molecular and Cellular Biology, State University of New York at Stony Brook, Stony Brook, New York 11794[2]

Received 27 February 1995/Returned for modification 28 April 1995/Accepted 26 May 1995

Replication factor C (RFC) is a five-subunit DNA polymerase accessory protein that functions as a structure-specific, DNA-dependent ATPase. The ATPase function of RFC is activated by proliferating cell nuclear antigen. RFC was originally purified from human cells on the basis of its requirement for simian virus 40 DNA replication in vitro. A functionally homologous protein complex from *Saccharomyces cerevisiae*, called ScRFC, has been identified. Here we report the cloning, by either peptide sequencing or by sequence similarity to the human cDNAs, of the *S. cerevisiae* genes *RFC1*, *RFC2*, *RFC3*, *RFC4*, and *RFC5*. The amino acid sequences are highly similar to the sequences of the homologous human RFC 140-, 37-, 36-, 40-, and 38-kDa subunits, respectively, and also show amino acid sequence similarity to functionally homologous proteins from *Escherichia coli* and the phage T4 replication apparatus. All five subunits show conserved regions characteristic of ATP/GTP-binding proteins and also have a significant degree of similarity among each other. We have identified eight segments of conserved amino acid sequences that define a family of related proteins. Despite their high degree of sequence similarity, all five *RFC* genes are essential for cell proliferation in *S. cerevisiae*. *RFC1* is identical to *CDC44*, a gene identified as a cell division cycle gene encoding a protein involved in DNA metabolism. *CDC44*/*RFC1* is known to interact genetically with the gene encoding proliferating cell nuclear antigen, confirming previous biochemical evidence of their functional interaction in DNA replication.

Replication factor C (RFC) is a multiprotein complex consisting of one large and four small subunits. The subunits of human RFC (hRFC) have apparent masses of 140, 40, 38, 37, and 36 kDa (23, 51). RFC has an associated ATPase activity that is stimulated by the binding of RFC to DNA and is further stimulated by proliferating cell nuclear antigen (PCNA) (54). RFC binds preferentially to the 3′ end of a DNA primer bound to a template DNA (22, 23, 54, 55). It is a structure-specific DNA-binding protein and acts as a primer recognition factor for DNA polymerases δ and ε (Pol δ and Pol ε) (5, 12, 22, 23, 42, 54–56). The large subunit of RFC contains a DNA-binding site, whereas at least one of the small subunits binds ATP (7, 55). PCNA recognizes and binds to the RFC-DNA complex in an ATP-dependent manner, and then DNA Pol δ or ε recognizes the RFC-PCNA complex bound to the primer-template prior to the start of DNA synthesis. In addition to stimulating polymerase loading, PCNA and possibly RFC also function as accessory proteins for Pol δ and ε by increasing their processivity (reviewed in reference 48).

RFC was shown to be essential for the simian virus 40 (SV40) in vitro DNA replication system (23, 53). It is responsible for a polymerase switch from DNA Pol α to Pol δ during initiation of leading-strand DNA replication at the SV40 origin and for the synthesis of Okazaki fragments during lagging-strand DNA synthesis (11, 52, 59). RFC binds to primers synthesized by Pol α-primase, blocks them for further elongation by this polymerase, and increases the affinity of Pol δ (or possibly Pol ε) for these primers (58). Studies with the SV40 DNA replication system, however, have not determined whether these same activities are required for replication of the eukaryotic genome and whether RFC has additional activities in vivo such as a role in DNA repair and recombination. To address these questions, we have turned to the genetically manipulatable organism *Saccharomyces cerevisiae*. In addition to being useful for determining the role of DNA replication factors in chromosomal replication, *S. cerevisiae* provides an opportunity to attempt reconstitution of DNA replication by using well-defined eukaryotic chromosomal origins of DNA replication.

A protein complex that is functionally homologous to hRFC from *S. cerevisiae* has been identified and called ScRFC (5, 12, 26, 61). ScRFC has activities similar to those of hRFC. It binds in a structure-specific manner to primer-template junctions and is also a DNA-stimulated ATPase that can be further stimulated by *S. cerevisiae* PCNA (ScPCNA) (5, 12). ScRFC was identified as an activity that stimulated *S. cerevisiae* Pol δ DNA synthesis by using a primed, circular single-stranded DNA template in the presence of ScPCNA and *S. cerevisiae* replication protein A (ScRPA) (12, 61). ScRFC does not cross-react with monoclonal antibodies that react with hRFC (3a), and hRFC will not cooperate with ScPCNA and ScRPA to stimulate *S. cerevisiae* Pol δ DNA synthesis (12).

ScRFC has a subunit composition similar to that of hRFC. It is a multisubunit protein with a large subunit with an apparent mass of 103 kDa and possibly four small subunits running as a doublet at 40/41 kDa and a single band at 36 kDa on sodium dodecyl sulfate (SDS)-polyacrylamide gels (12, 26). To characterize these protein subunits, we sequenced peptides derived from them and used this information to isolate the Sc*RFC* genes. Here we demonstrate that all five *RFC* genes are essential for cell proliferation and that the gene encoding the large subunit is identical to *CDC44*, a previously characterized gene encoding a cell division cycle protein (15). The genes encoding three of the small subunits also have recently been reported (24, 25, 34).

* Corresponding author. Mailing address: Cold Spring Harbor Laboratory, P.O. Box 100, Cold Spring Harbor, NY 11724. Phone: (516) 367-8383. Fax: (516) 367-8879.

† Present address: Department of Microbiology, Columbia University, New York, NY 10032.

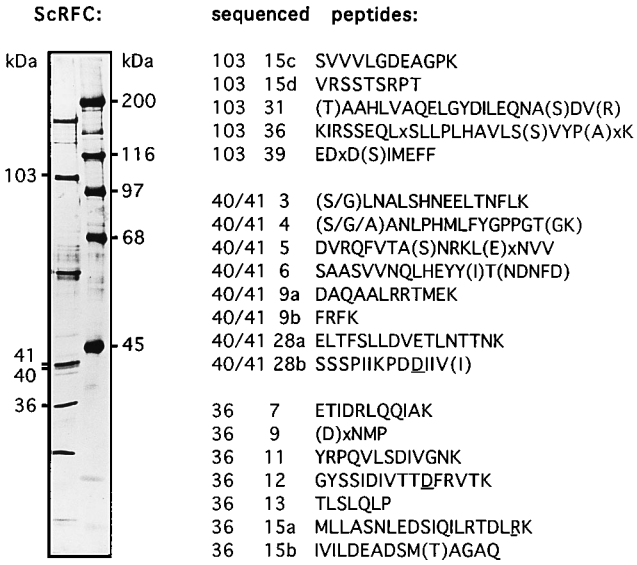


FIG. 1. Peptide sequences obtained from the ScRFC subunits. Shown is the result of SDS-PAGE of ScRFC (left lane; preparation used for peptide sequencing) next to molecular weight markers (right lane). The 103-kDa band, the 40/41-kDa doublet, and the 36-kDa band marked were used for peptide sequencing. The contaminating bands are not present in highly purified ScRFC (12). Listed are the amino acid sequences obtained from each band. Amino acids that were not unambiguously determined are in parentheses, and amino acids that do not correspond to the predicted ORFs are underlined. Amino acids that could not be determined are indicated by an x.

MATERIALS AND METHODS

Strains. The *S. cerevisiae* strains used were BJ926 (*MAT***a**/*MAT*α *TRP1*/*trp1 HIS1*/*his1 pep4-3*/*pep4-3 prb1-1122*/*prb1-1122 prc1-126*/*prc1-126 can1*/*can1 gal2*/*gal2*) (17), diploid W303 (*MAT***a**/MATα *ura3-1*/*ura3-1 his3-11,15*/*his3-11,15 trp1-1*/*trp1-1 leu2-3*/*leu2-3 ade2-1*/*ade2-1 can1-100*/*can1-100*), its derivative W937 (as W303, but *ura3-d1*/*ura3-d1* [43a]), YB10062 (as W303, but *RFC2*/*rfc2*Δ::*LEU2*), YB10063 (as W303, but *RFC3*/*rfc3*Δ::*HIS3*), YB10064 (as W937, but *RFC4*/*rfc4*Δ::*LEU2*), and YB10065 (as W937, but *RFC5*/*rfc5*Δ::*HIS3*).

ScRFC purification. ScRFC was purified as described by Fien and Stillman (12) from BJ926 cells, with the following modifications. The first phosphocellulose column was loaded with the crude lysate from BJ926 cells and then washed with 1 column volume of the same buffer containing 0.66 M NaCl. RFC and other proteins were eluted from the hydroxylapatite column with a 600-ml linear gradient from 0.1 to 0.6 M KPO_4 HAP buffer (12). The third step, a phosphocellulose column, was eliminated from the new purification scheme. Active fractions from the glycerol gradient were analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) and silver staining. The RFC fractions were concentrated on a 0.1-ml Q-Sepharose column and eluted in three 0.1-ml amounts of buffer A containing 0.2 M NaCl.

Isolation of lysyl-endopeptidase fragments and protein sequencing. Purified ScRFC was fractionated by electrophoresis in an SDS-polyacrylamide gel, and four bands corresponding to the ScRFC subunits (one band at 103 kDa, a doublet at 40/41 kDa, and one band at 36 kDa) were cut out from the gel. The polypeptides were digested with *Achromobacter* protease I (lysyl-endopeptidase; Wako) (1). Peptides were separated by reverse-phase high-pressure liquid chromatography (HPLC) on a Vydac C_{18} column and eluted with a linear gradient of 10 to 35% acetonitrile in water. Peptide fractions were sequenced directly on an Applied Biosystems model 475 automated sequencer with an on-line model 120A HPLC PTH analyzer. The sequences obtained are shown in Fig. 1.

Isolation of probes for the RFC genes. (i) *RFC1* probe. The amino acid sequence of peptide 103 31 (Fig. 1) was used to design four degenerate oligonucleotides: ScRFC1.4 (CGGAATTCGC HCAYCTDGTY GCHCARGA) and ScRFC1.5 (CGGAATTCGC HCAYTTRGTY GCHCARGA) from the N terminus of peptide 103 31 and ScRFC1.6 (CGGAATTCAC RTCDGADGCR TTYTGYTC) and ScRFC1.7 (CGGAATTCAC RTCRCTDGCR TTYTGYTC) from the C terminus (symbols for nucleotides according to the International Union of Biochemistry code; D = A, G, or T; H = A, C, or T; I = inosine; K = G or T; M = A or C; R = A or G; S = C or G; Y = C or T). The four primers were used in pairwise combination to generate PCR products of 75 bp, using *S. cerevisiae* genomic DNA as a template. The products were cloned in vector pBluescript KS+ (pKS+; Stratagene), using the *Eco*RI sites included in the primers, and sequenced.

(ii) *RFC2* and *RFC3* probes. The amino acid sequences of three peptides with significant similarity to the hRFC subunits (35) were used to design degenerate oligonucleotides (peptide 40/41 4 [ATP+; GGICCICCIG GIACIGGIAA RAC], peptide 36 15b [DEAD−; CATISIRTCI GCYTCRTCIA RDAT] and peptide 40/41 9a [ALRR−; TYTCIATIRY ICKICKIARI GC]; Fig. 1). Oligonucleotide ATP+ in combination with either DEAD− or ALRR− was used as a primer in PCR with *S. cerevisiae* genomic DNA. Several specific PCR products were obtained. The products were reamplified and cloned in the pGEMEX-derivative pDK101 (19), using their 3′ A overhangs. The predicted amino acid sequences of three clones showed strong similarity to the sequences of peptides from hRFC37 (27a/1 and 28/2 [*RFC2* probes]) and hRFC36 (29/1 [*RFC3* probe]).

(ii) *RFC4* probe. Primers DEAD+ and 15am (ATYTGRATDG ARTCYTC YAR RTT; derived from peptide 36 15a) were used for PCR with genomic *S. cerevisiae* DNA. They yielded a DNA fragment of 420 bp. Reamplification of this fragment with primers ALRR+ and 15am resulted in a fragment of 390 bp. The ALRR+/15am fragment was subcloned in pDK101 and sequenced. It showed strongest predicted amino acid sequence similarity to hRFC40 and was used as an *RFC4* probe.

(iv) *RFC5* probe. Primers derived from peptide 40/41 5 (5p [GAYGTIMGIC ARTTYGTIAC IGC]) and 40/41 28a (28A− [GTIGTRTTIA RIGTYTCIAC RTC]) were used in a PCR with genomic *S. cerevisiae* DNA and yielded a DNA fragment of 700 bp. This fragment could be reamplified with primers ALRR+ and 28A− and resulted in a 500-bp fragment. The ALRR+/28A− fragment was subcloned in pDK101 and sequenced. It showed strongest predicted amino acid sequence similarity to hRFC38 and was used as an *RFC5* probe.

Cloning of the *RFC* genes. The PCR-generated clones were used to screen an *S. cerevisiae* genomic DNA library in λ phage EMBL3A (gift of R. Young, Massachusetts Institute of Technology). A 2.1-kb *Bgl*II fragment from λ clone 2 and a 2.8-kb *Kpn*I-*Sal*I fragment of λ clone 11 hybridized with the *RFC1* probe and were subcloned in pKS+. A 3.8-kb *Eco*RI fragment of λ clone 2/2/1 hybridized with *RFC2* probe 27a/1, a 3.9-kb *Eco*RI fragment of λ clone 3/4/1 hybridized with *RFC3* probe 29/1, a 7.6-kb *Bam*HI fragment of λ clone 4/5/1 hybridized with the *RFC4* probe ALRR+/15am, and a 2.4-kb *Eco*RI and a 5.0-kb *Bam*HI-*Kpn*I fragment of λ clone 5/8/1 hybridized with the *RFC5* probe ALRR+/28A−. These fragments were subcloned in pBluescript SK+ (pSK+; Stratagene). The RFC genes were sequenced by primer walking, starting from the sequences known from the cloned PCR products.

Plasmid constructions. Plasmid pKSRFC1-2 contains the 5′ end of the *RFC1* open reading frame (ORF) on a 2.0-kb *Bgl*II fragment from λ clone 2 in pKS+ (Fig. 2). Plasmid pKSRFC1-11 contains the 3′ end of the *RFC1* ORF on a 2.8-kb *Kpn*I-*Sal*I fragment from λ clone 11 in pKS+. Plasmid pKSRFC1 was constructed by ligating the pKSRFC1-2 *Bgl*II fragment with pKSRFC1-11 digested with *Bam*HI and *Bgl*II. It contains the entire *RFC1* gene.

Plasmid pSKRFC2 contains the *RFC2* gene on a 3.8-kb *Eco*RI fragment from λ clone 2/2/1 in pSK+ Plasmid pSKRFC3 contains the *RFC3* gene on a 3.9-kb *Eco*RI fragment from λ clone 3/4/1a in pSK+. Plasmid pSKRFC4B contains the *RFC4* gene on a 7.6-kb *Bam*HI fragment in pSK+. Plasmid pSKRFC4H contains the *RFC4* gene on a 1.7-kb *Hin*dIII fragment in pSK+. Plasmid pSKRFC5BK contains the *RFC5* gene on a 5.0-kb *Bam*HI-*Kpn*I fragment in pSK+. Plasmid pSKRFC5E contains the *RFC5* gene on a 2.4-kb *Eco*RI fragment in pSK+. In plasmid pRFC2ko, the *Nde*I-*Hpa*I fragment of the *RFC2* ORF of pSKRFC2 was replaced with the *Sal*I fragment of *LEU2*. In plasmid pSKRFC3ko, the *Nco*I-*Nar*I fragment of the *RFC3* ORF was replaced with the *Bam*HI-*Cla*I fragment of *HIS3*. In plasmid pRFC4ko, the *Nco*I-*Nde*I fragment of the *RFC4* ORF of pSKRFC4H was replaced with the *Sal*I fragment of *LEU2*. In plasmid pSKRFC5ko, the *Mlu*I-*Bgl*II fragment of the *RFC5* ORF of pSKRFC5E was replaced with the *Bam*HI-*Cla*I fragment of *HIS3*.

***RFC* gene disruptions.** The inserts of plasmids pRFC2ko, pRFC3ko, pRFC4ko, and pRFC5ko were isolated by gel electrophoresis and transformed into diploid *S. cerevisiae* (pRFC2ko and pRFC3ko into yeast strain W303, and pRFC4ko and pRFC5ko into yeast strain W937). This resulted in the mutant diploid strains YB1062, YB1063, YB1064, and YB1065, respectively, in which one of the two *RFC* ORFs was replaced by the auxotrophic marker gene (Fig. 2). The disruption of one of the *RFC* alleles was confirmed by genomic Southern blot analysis and by PCR. The heterozygous diploids were then sporulated, and the tetrads were dissected (44).

Physical mapping. A set of overlapping *S. cerevisiae* genomic DNA clones in phage λ vectors and cosmids on nylon filters was provided by Linda Riles and Maynard Olson (Washington University, St. Louis, Mo.). Random hexamer-labeled probes for *RFC1* (*Acc*I-*Kpn*I fragment; Fig. 2), *RFC2* (*Nde*I-*Hpa*I fragment), *RFC3* (*Hin*dIII fragment), *RFC4* (*Eco*RI-*Mlu*I fragment), and *RFC5* (*Eco*RI and *Nde*I-*Sal*I fragments) were made and hybridized to the filters.

Nucleotide sequence accession numbers. The nucleotide sequences for the *RFC* genes have been submitted to GenBank under accession numbers U26027 (*RFC1*), U26028 (*RFC2*), U26029 (*RFC3*), U26030 (*RFC4*), and U26031 (*RFC5*).

RESULTS

Cloning of the Sc*RFC* genes. ScRFC was purified as described by Fien and Stillman (12), with slight modifications (see Materials and Methods). The individual protein subunits were separated by SDS-PAGE, and peptides from the 103-,

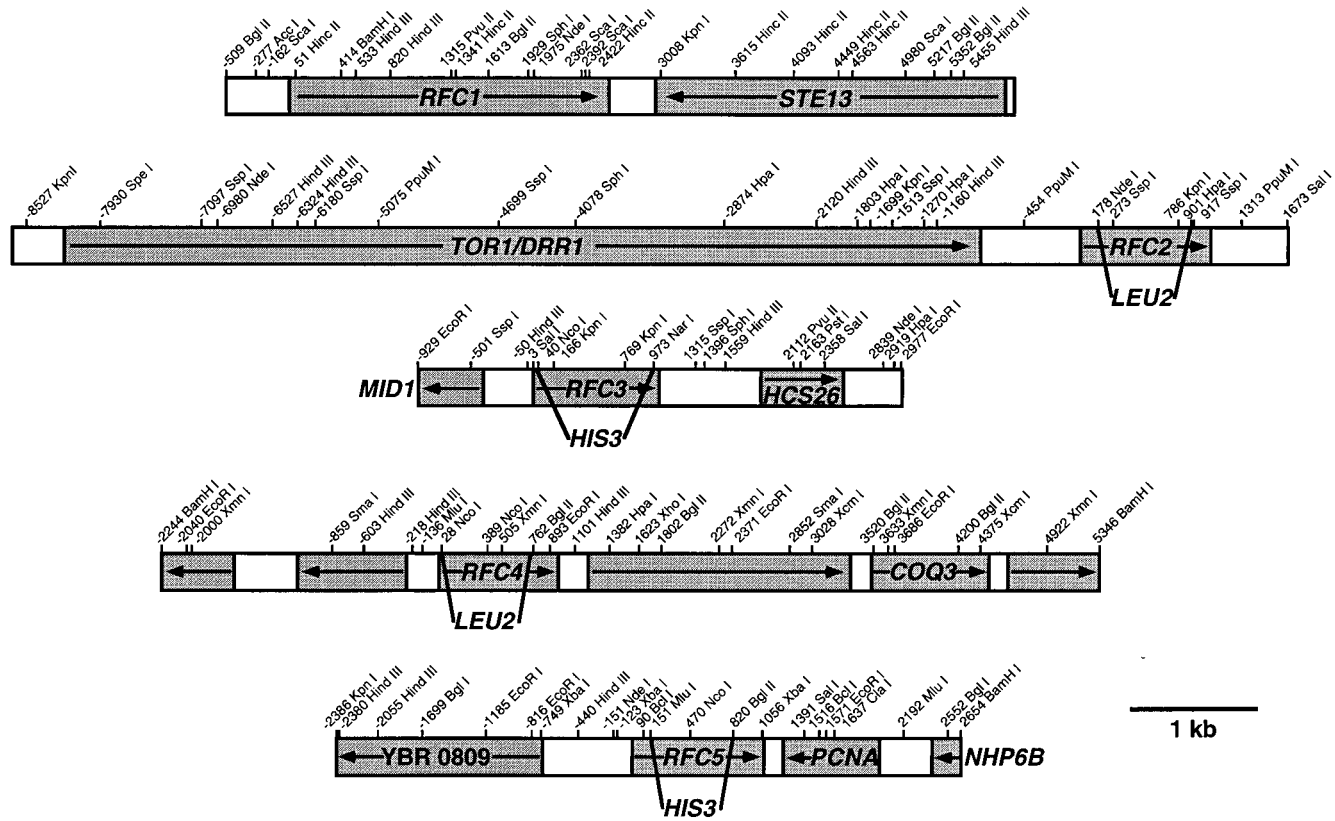


FIG. 2. Genomic organization of the five *S. cerevisiae RFC* genes. Shown are restriction maps of *RFC1* to *RFC5* and neighboring genes. The fragments replaced by auxotrophic markers in the disruptions are indicated. ORFs are shaded, and numbering of nucleotides is relative to the *RFC* start codons.

40/41-, and 36-kDa bands of ScRFC were used to obtain amino acid sequences from these subunits (Fig. 1). Degenerate oligonucleotides were designed on the basis of these peptide sequences and used to amplify probes by PCR with *S. cerevisiae* genomic DNA as a template. The resulting PCR products were subcloned and sequenced to verify identities with peptide sequences and similarities to the hRFC subunits (3, 35). They were then used to screen an *S. cerevisiae* genomic DNA library. Phage λ clones were obtained for all of the subunits. Figure 2 shows the genomic organization of the genes *RFC1*, *RFC2*, *RFC3*, *RFC4*, and *RFC5*. *RFC1* is located next to *STE13* (accession number U08230), the *S. cerevisiae* dipeptidyl-aminopeptidase A. *RFC2* is downstream of the phosphatidylinositol kinase gene *DRR1* (6) or *TOR1* (14). Downstream of *RFC3* lies the *HCS26* gene (36), encoding a G_1 cyclin. *RFC4* is framed by two unknown ORFs. The one downstream of *RFC4* is similar in predicted amino acid sequence to DNA helicases. *RFC5* is also known as YBR0810, which was sequenced as part of the *S. cerevisiae* genome sequencing project (accession number X78993). YBR0809 is located upstream of *RFC5*, and the *POL30* gene corresponding to ScPCNA (YBR0811) is located downstream of *RFC5* (Fig. 2).

The Sc*RFC* genes were then subcloned and sequenced. The results are shown in Fig. 3, along with the predicted translation products. The amino acid sequences obtained by direct amino acid sequencing of peptides and the binding sites of primers used for PCR are underlined. Peptide sequences were obtained for *RFC1*, *RFC2*, *RFC4*, and *RFC5*; no peptide sequences were obtained for *RFC3*. All peptide sequences obtained could be accounted for in the predicted ORFs. The genes encoding ScRFC2, ScRFC3, and ScRFC4 have been isolated independently (24, 25, 34) on the basis of the similarity in sequence to the hRFC clones, and hence only the peptide sequences of these proteins are shown in Fig. 3.

RFC1 contains one major ORF encoding 861 amino acids, and the gene was found to be identical to a recently cloned cell division cycle gene called *CDC44* (15). The predicted molecular weight is 94.9 kDa, and the protein is rich in lysine (K) residues. The calculated pI is 10.0. The predicted *RFC2* ORF product of 353 amino acids has a calculated molecular mass of 39.7 kDa and a pI of 8.6. The *RFC3* ORF product has a length of 340 amino acids. The predicted molecular weight is 38.2 kDa, and the calculated pI is 6.4. *RFC4* has an ORF encoding 323 amino acids; the predicted molecular weight is 36.2 kDa, and the calculated pI is 9.8. The *RFC5* ORF product is 354 amino acids in length; the molecular weight is predicted 39.9 kDa, with a calculated pI of 8.2.

Physical mapping. Probes for all five *RFC* genes were hybridized to filters containing an ordered *S. cerevisiae* genomic library provided by L. Riles and M. Olson (see Materials and Methods). The *RFC1* probe detected phage λ clones 3556 and 4182. The gene was mapped to the right arm of chromosome XV, next to *STE13*. The *RFC2* probe detected clone 3881, and the gene was mapped to the right arm of chromosome X, between *CDC8* and *CDC11*. The *RFC3* probe hybridized with clone 5398, which is located on the left arm of chromosome XIV, between *SUF6* and *MET2*. The *RFC4* probe detected phage λ clones 5929 and 6006. The gene was localized to the left arm of chromosome XV between *SUF1* and *adh1*. Neither the *RFC5* probe nor the probe for the gene encoding PCNA, which is located next to *RFC5*, detected any clones in this genomic representation, which is 96% complete. Nevertheless,

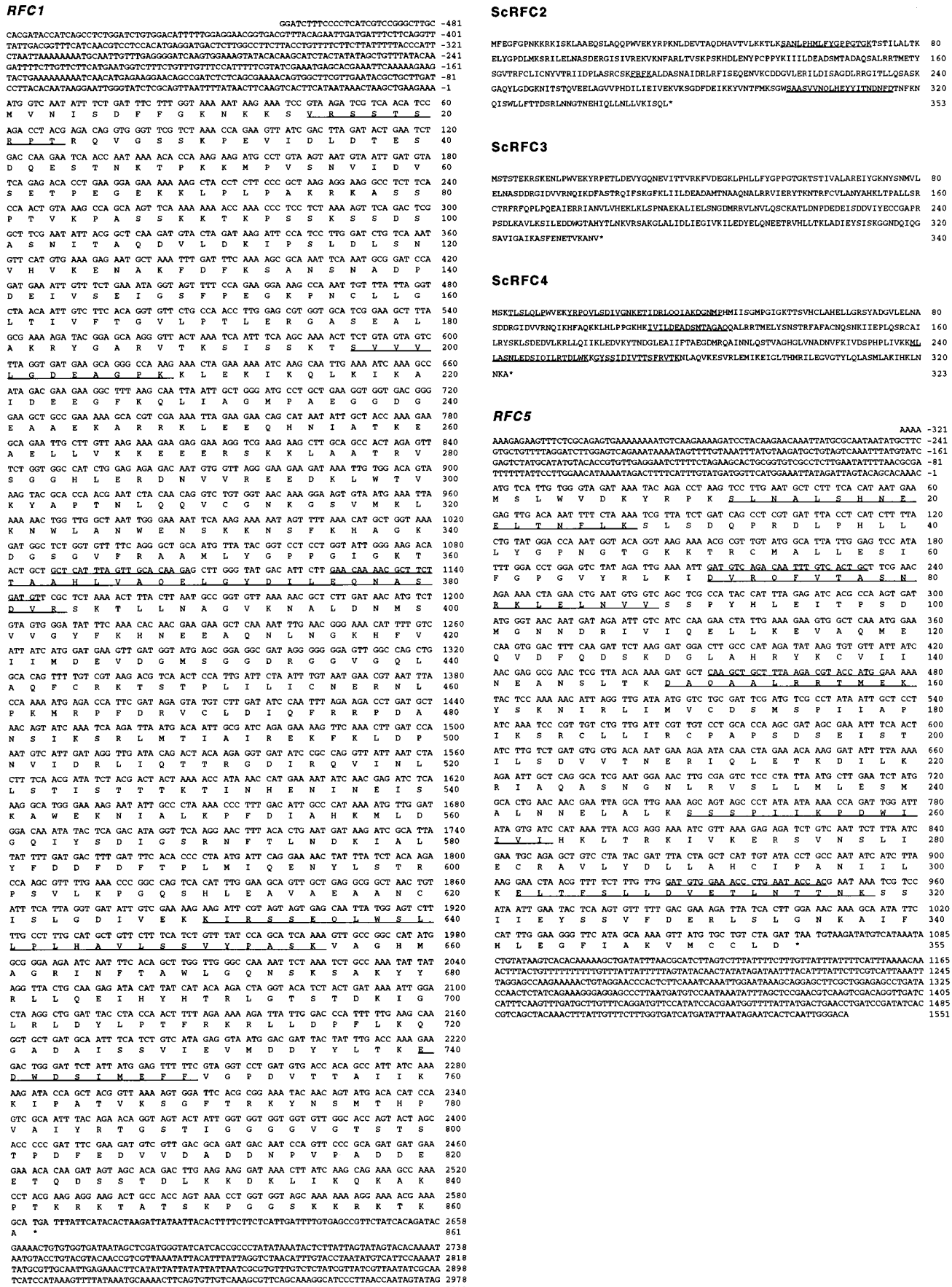


FIG. 3. DNA sequences of *RFC1* and *RFC5* and predicted translation products of *RFC1* to *RFC5*. The DNA primers used for amplification of RFC probes and the sequenced peptides are underlined.

TABLE 1. Disruption of the *RFC* genes shows they are essential

Disrupted gene	No. with indicated dissection pattern (viable[a]:nonviable)			Total no. of tetrads
	2:2	1:3	0:4	
RFC2	15	5		20
RFC3	7	3		10
RFC4	19	9	2	30
RFC5	32	7	1	40

[a] All viable spores tested negative for the auxotrophic marker used for the disruption and thus contained an intact copy of the *RFC* gene.

RFC5 could be mapped to chromosome II, since this chromosome was sequenced as part of the *S. cerevisiae* genome project (accession number X78993). It is located on the right arm, 38 kb proximal from *CMD1*.

RFC genes are required for *S. cerevisiae* viability. To determine whether all of the ScRFC subunits perform an essential function, the coding regions of one of the alleles of the *RFC2* to *RFC5* genes were replaced by auxotrophic markers in diploid yeast strains (see Materials and Methods). The *RFC1* gene, which is identical to *CDC44*, was shown recently to be essential (15), and hence we did not disrupt this gene. Sporulation and tetrad dissection predominantly yielded a 2:2 segregation pattern, as shown in Table 1. All of the viable spores exhibited a phenotype negative for the relevant auxotrophic marker. If the diploid strains carried a plasmid with the respective *RFC* gene prior to sporulation and dissection, the segregation pattern was switched to 3:1 or 4:0 (data not shown). This finding confirmed that the lethality was due to the disruption of the respective *RFC* gene. Cells with the deleted *RFC* gene typically arrested as microcolonies at the 2- to 32-cell stage. Most of the cells were budded, which is consistent with an arrest in S phase.

Comparison of *S. cerevisiae* and human RFC subunits. Figure 4 shows a comparison between the amino acid sequences of the hRFC subunits and the ScRFC subunits. The amino acid identity among all the subunits is high (24 to 37%), but amino acid comparisons indicate that there are homologous pairs of human and yeast subunits with distinctly higher amino acid identity and similarity (Fig. 4 and 5). The large *S. cerevisiae* subunit, ScRFC1, is somewhat smaller than the human subunit, hRFC140. This variation in size is mostly due to differences in the N terminus. The pI of this protein is very basic in both species. Traditionally, the small human subunits have been ordered according to their apparent sizes in SDS-PAGE, and this ordering does not coincide with their calculated molecular masses. The ScRFC subunits are numbered according to decreasing calculated molecular mass, except for ScRFC5. This causes a slightly different order of the homologous subunits. ScRFC2 and hRFC37 have virtually identical calculated masses and basic pIs. ScRFC3 is slightly shorter than hRFC36 and has a slightly acidic pI, compared with the neutral pI of hRFC36. ScRFC4 and hRFC40, despite being the most homologous pair, show the biggest difference in mass and pI. ScRFC5 and hRFC38 are the least homologous small subunits but have similar masses and pIs.

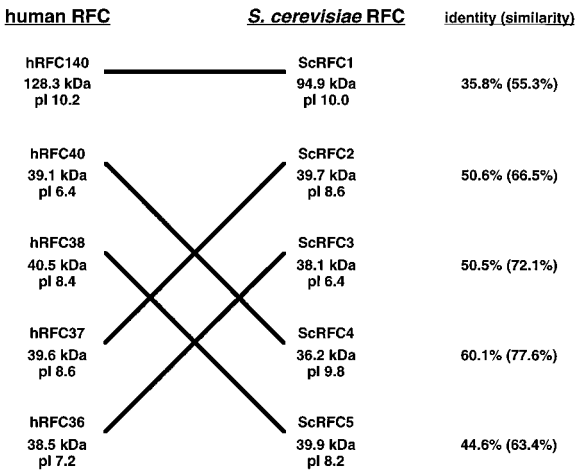


FIG. 4. Comparison between the hRFC and ScRFC subunits. Indicated are calculated masses and pIs as well as the levels of amino acid identity and similarity between the homologous subunit pairs connected by bars. Differences in nomenclature lead to a somewhat different order of the subunits in human and *S. cerevisiae* proteins.

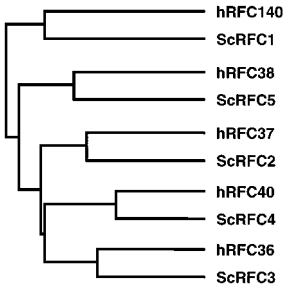
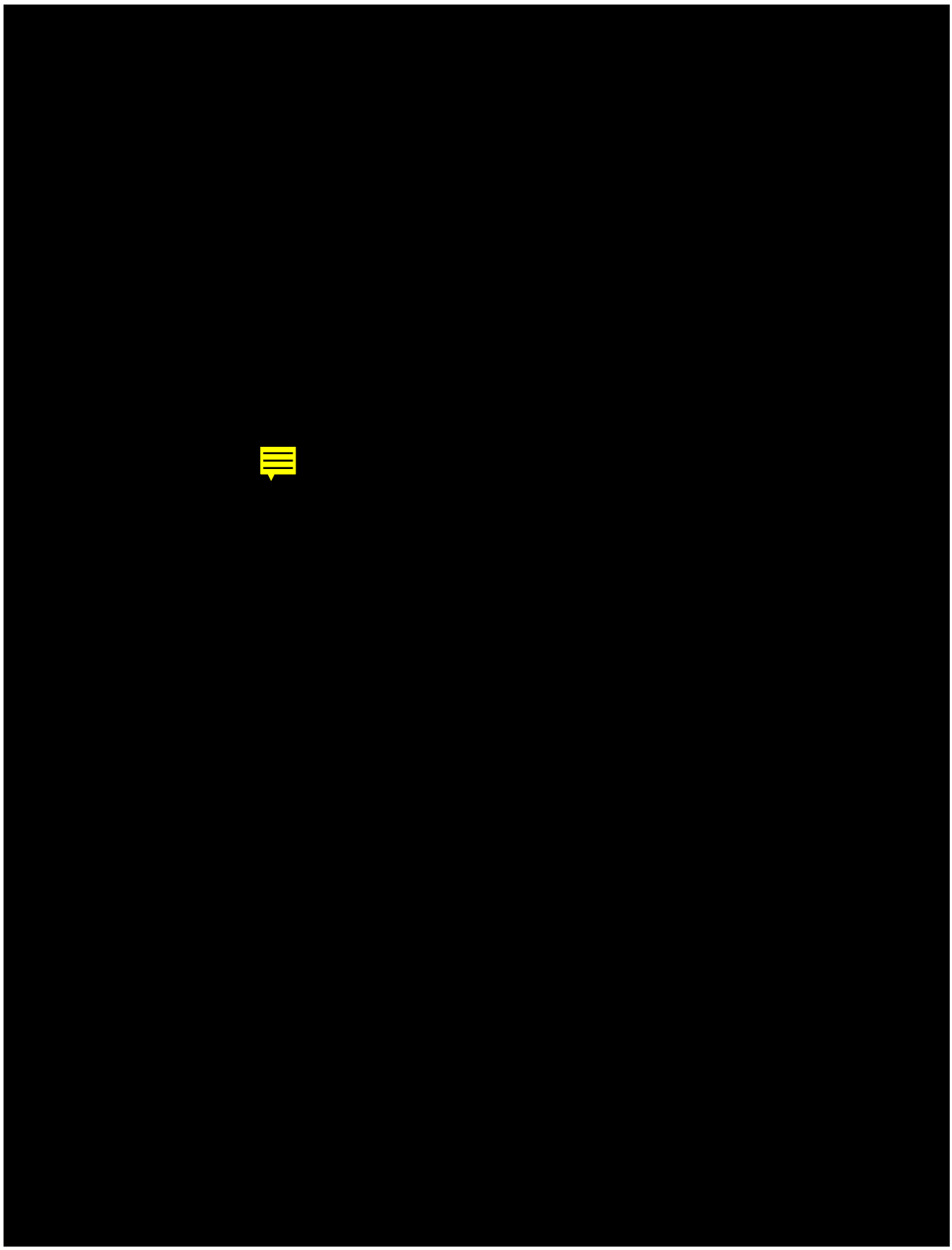


FIG. 5. Evolutionary tree for the hRFC and ScRFC subunits. Although the amino acid identity among all subunits is high (24 to 37%), there are homologous pairs with distinctly higher amino acid identity and similarity.

Comparison of amino acid sequences of RFC subunits. Figure 6 shows an alignment of all RFC subunits from *S. cerevisiae* and humans (summarized in Fig. 8). They were compared with the functionally homologous protein gp44 from phage T4 (47) and the prokaryotic DNA polymerase III holoenzyme subunits *Escherichia coli* γ/τ (13, 60) and δ' (10) and *Bacillus subtilis* DnaH (37), an *E. coli* γ/τ homolog (γ consists of the same amino acid sequence as τ but ends as a result of a translational frameshift at amino acid 431).

In addition to these functionally related proteins, a search of the sequence databases revealed sequence similarity to a predicted protein sequence from the *CHL12* gene (18) of *S. cerevisiae*. There is a significant similarity between the *CHL12* gene product and the RFC subunits, ranging from 20 to 25% amino acid identity and up to 50% similarity. The *CHL12* ORF predicts an 84.3-kDa protein with a pI of 8.5. RFC boxes II to V, VII, and VIII are conserved, but CHL12 is not known to be a component of ScRFC. *CHL12* was isolated as a chromosome loss mutation (18) and is identical to *CTF18* (46). Mutants in *CHL12* are incapable of stable maintenance of circular and linear artificial chromosomes. Mitotic recombination frequency and sensitivity to UV and γ irradiation are normal, suggesting that DNA repair is not affected by *CHL12* mutations.

There is an overall similarity between the prokaryotic and eukaryotic proteins, as has been noted previously for the human sequences (7, 8, 35) and some of the *S. cerevisiae* sequences (24, 25, 34). The small RFC subunits and the bacterial and phage T4 polypeptides align with the central part of the large RFC subunits from different species. We named the similar regions RFC boxes I to VIII, numbered from the N terminus toward the C terminus (Fig. 6; see also 8). The most obvious feature of all the sequences is a conserved ATP/GTP-



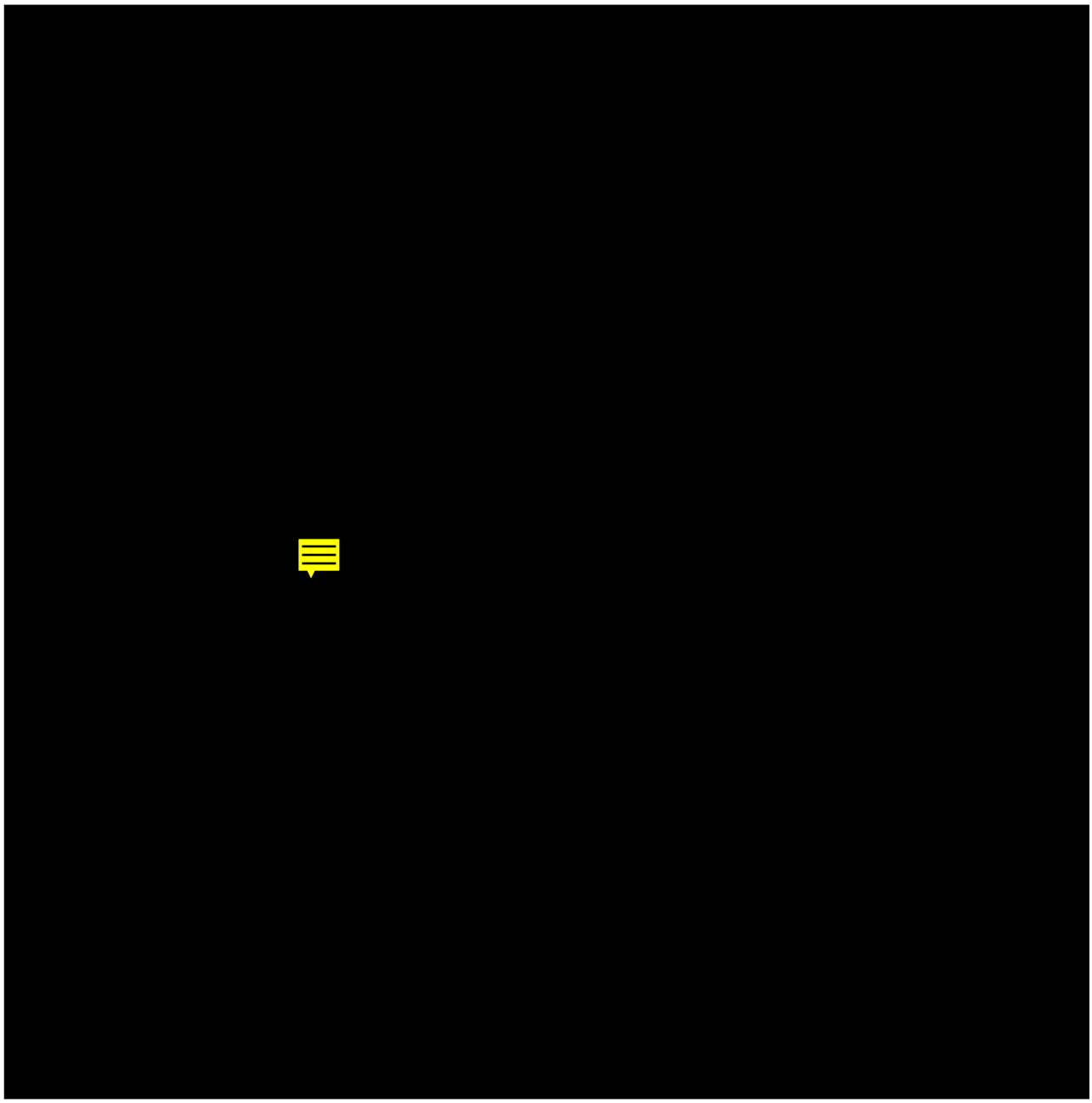


FIG. 6. Alignment of the amino acid sequences of all RFC subunits from *S. cerevisiae* (this report and references 24, 25, and 34) and human (3, 7, 8, 15a, 35) proteins in comparison with sequences of the functionally homologous proteins gp44 from bacteriophage T4 (T4gp44) (47), *E. coli* γ/τ (EcTau) (13, 60), *E. coli* δ′ (EcDelta′) (10), and *B. subtilis* DnaH (BsdnaH) (37) and the sequence of the product of a chromosome loss mutation from *S. cerevisiae*, *CHL12* (CHL12) (18). Similar amino acids are assigned similar colors so that conserved regions become obvious. Indicated are RFC box I (ligase homology) and boxes II to VIII, which are conserved in all RFC subunits. Box VIa is conserved in the large RFC subunits, and box VIb is conserved in the other proteins. Amino acids affected by mutations in the *CDC44* gene, resulting in a cold-sensitive phenotype, are boxed in black in the ScRFC1 protein sequence.

binding region. It consists of several motifs in the N-terminal half of the small RFC subunits and the equivalent region of the large subunit.

The most conserved motif is within box III and is the phosphate-binding loop (P loop) with the consensus sequence GxxxxGK(S/T). This loop usually contains additional glycines and prolines, and we know from analyses of $p21^{ras}$, for which the tertiary structure has been solved, that it is involved in the binding of the phosphate groups of the nucleotide (39, 40). In the case of the RFC subunits, for which we named this domain RFC box III, it has the consensus sequence phUUuyGPPGt GKT(S/T)t (where U stands for a bulky aliphatic residue such as I, L, V, or M).

The second-most-conserved domain is RFC box V, with the consensus sequence (F/H/Y)kUUUUDE(V/A)D for the RFC subunits. It bears similarity to the DEAD-box proteins, a family of putative RNA helicases which also have P loops and are ATPases (27, 41). There is no further similarity between the RFC subunits and the DEAD-box proteins, and no helicase activity could be found for RFC.

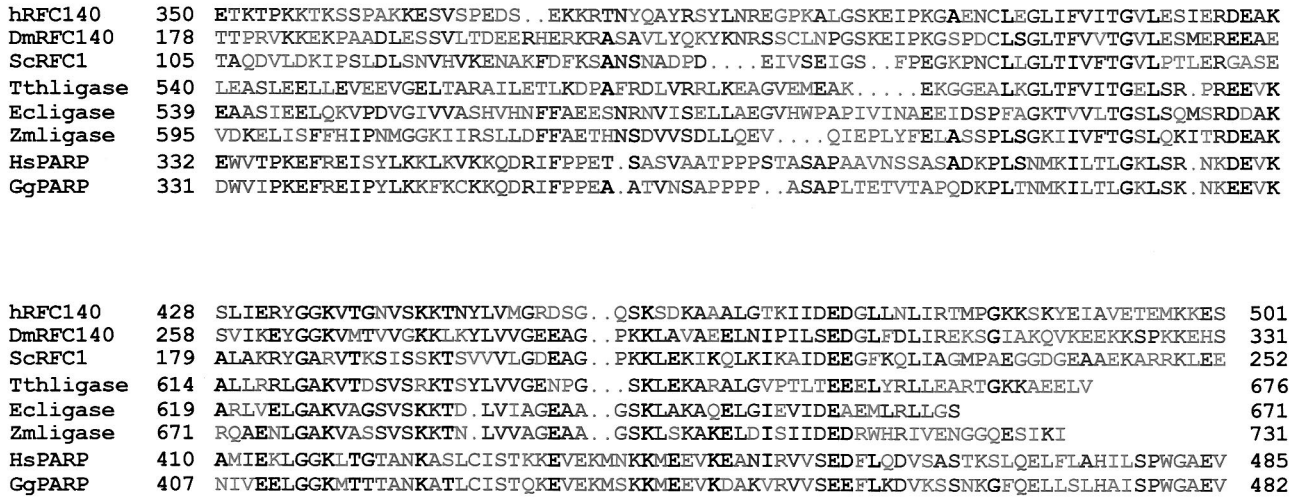


FIG. 7. RFC box I (ligase homology) compared with prokaryotic DNA ligases from *Thermus thermophilus* (Tthligase) (21), *E. coli* (Ecligase) (accession number M24278), and *Zymomonas mobilis* (Zmligase) (accession number Z11910) and PARPs from humans (HsPARP) (57) and chickens (GgPARP) (16). Shading of the amino acids is according to conservation in one, two, or three of the protein families.

The ATP/GTP-binding region also includes RFC boxes II, IV, and VI, which are not present in other ATP/GTP-binding proteins but are quite unique to RFC and the related proteins. RFC box II shows a high degree of similarity among the RFC subunits and a conserved RP dipeptide in the related proteins. The consensus sequence for the RFC subunits is (L/P)WV(E/D)KYrPxxU.

RFC box IV is only weakly conserved in the prokaryotic proteins, but it is found in T4 gp44 and the CHL12 protein with the consensus sequence LEUNaSD.

RFC box VI is different in the small and the large RFC subunits. RFC box VIa, present in the large subunits, has the consensus sequence gMaGneDRGGUqeL and is not conserved in other proteins. RFC box VIb, present in the small subunits, is somewhat conserved in the prokaryotic accessory proteins and in T4 gp44. The consensus sequence among the small RFC subunits is s(M/L)TxxAQxALRRtmE.

RFC box VII, SRC, is conserved within the small subunits, the prokaryotic accessory proteins, and T4 gp44, but only the Cys is present in the large RFC subunits and CHL12. Between box VII and box VIII, we can align single amino acids present in most of the proteins, but they do not cluster as a conserved box. RFC box VIII has the consensus sequence gdURxx(L/I)xxlq, and mutations in the codons for G and D have been shown to cause a cold-sensitive phenotype in *cdc44* mutants.

Box I is present only in the large RFC subunits (3, 4, 29). It consists of about 90 amino acids, and similar boxes can be found in all three known prokaryotic DNA ligases and, to a lesser extent, in all known poly(ADP-ribose) polymerases (PARPs) (Fig. 7). This region has been designated the ligase homology domain and has been recognized previously in the mammalian RFC (3, 4, 29).

DISCUSSION

The DNA replication apparatus is conserved in function, structure, and amino acid sequence from yeasts to mammals among all eukaryotes investigated so far. Most of the factors also seem to be functionally conserved in bacteria like *E. coli* and bacteriophages such as T4 (48). Some of these DNA replication proteins, such as the DNA polymerases, even show amino acid similarity (2), but this is not always the case (e.g., PCNA and *E. coli* β subunit).

RFC is functionally similar to the bacteriophage T4 accessory protein complex gp44-gp62 and to the *E. coli* γ/δ complex (35, 48). Each of these proteins binds to primer-template structures and loads the respective donut-shaped sliding DNA clamp, gp45 for T4, β for *E. coli*, and PCNA for the eukaryotes, onto the DNA template. Besides their similar functions, there also exists considerable amino acid sequence similarity between subunits of these protein complexes (3, 4, 24, 25, 29, 34, 35) (Fig. 6). Nevertheless, the subunit compositions of the respective complexes are quite different. T4 gp44-gp62 seems to be composed of four gp44 subunits and one gp62 subunit (62), with gp62 having no similarity to any RFC subunit. The *E. coli* proteins can be found in two different complexes. The τ homodimer has a DNA-dependent ATPase activity (50) and might work as a bridge between two Pol III core DNA polymerases (49). γ binds ATP in a similar fashion, which is expected since γ consists of the N-terminal three-fourths of the τ subunit, including the ATPase consensus domain. But γ becomes an ATPase only as part of the γ complex with δ, δ′, Ψ, and χ (38). This ATPase activity is, like τ, stimulated by DNA but can be further stimulated by the β clamp. ATPase activity stimulated by DNA can be reproduced by γδ, γδ′, or γδδ′ complexes, but only the activity of complexes containing γ and δ can be further stimulated by the β clamp (38).

We used peptide sequences obtained by sequencing the purified ScRFC subunits and the sequence similarity of hRFC to T4 gp44 and *E. coli* γ/τ and δ′ (35) to design degenerate oligonucleotides. PCR yielded probes for all five *RFC* genes, which were then used to clone *RFC1* to *RFC5* from a genomic *S. cerevisiae* library. All peptide sequences could be accounted for by the amino acid sequences of *RFC1*, *RFC2*, *RFC4*, and *RFC5* (Fig. 3). No peptide sequences were obtained from *RFC3*. However, *RFC2* (34), *RFC3* (25), and *RFC4* (24) were cloned recently by others, and Li and Burgers (25) have shown that a polyclonal antiserum directed against bacterially expressed ScRFC3 reacts with a band of about 40 kDa in an immunoblot with biochemically isolated ScRFC. These data and our peptide sequencing data indicate that ScRFC, like hRFC, is a protein complex of five different subunits.

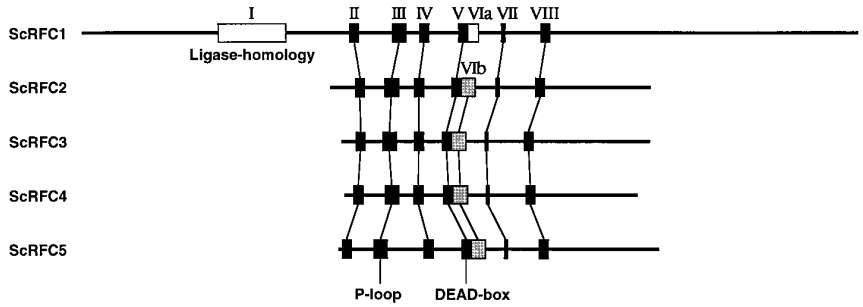


FIG. 8. Summary of the ScRFC subunits. The small subunits align to the middle part of the large subunit. There are eight conserved RFC boxes numbered consecutively from N terminus to C terminus. Box I is the DNA ligase homology domain, and boxes II to VIII contain within them an ATP-binding region.

The *RFC1* gene is identical to the *CDC44* gene. Cold-sensitive *cdc44* mutations isolated by Moir et al. (32) and Howell et al. (15) show a cell division cycle arrest phenotype and arrest as large-budded cells with the nucleus at the neck between mother and daughter cells. The phenotypes observed in *cdc44* mutants, while complicated, are consistent with a role for this protein in DNA metabolism in cells (15) and thus consistent with previous biochemical characterization of both human and yeast RFCs. The amino acid changes in *cdc44* mutants are all due to defects in or near the ATP/GTPase consensus domain in RFC boxes V and VI or in box VIII, which is also conserved in all five RFC subunits (Fig. 6; summarized in Fig. 8). These mutations are likely to affect the ATPase activity of ScRFC1 (since GTP is not essential for RFC activity and substitutes only poorly for ATP [23], we assume that those sites bind primarily ATP). RFC most likely acts as a protein topoisomerase in that it opens up the donut-shaped PCNA trimer and loads it onto partially duplex DNA in an energy-consuming step. The *RFC1* mutations can be suppressed by point mutations in the *POL30* gene, which encodes PCNA (31). Because of their positions in the protein, the amino acid changes are likely to affect the interaction between the PCNA monomers within the trimer structure in a way that compensates for the reduced activity of the RFC complex (20). It is not clear from these data whether the interaction between ScRFC1 and PCNA is direct or exerted through the small RFC subunits, but the data show that this interaction is essential in vivo.

Some investigators cloned RFC1 from human (28, 29), mouse (4, 29), and *Drosophila melanogaster* (accession number 17340) cells in attempts to screen expression libraries with double-stranded oligonucleotides as a probe for sequence-specific DNA-binding proteins. In none of these cases, however, was DNA sequence specificity of the RFC large subunit demonstrated. It is most likely that the RFC large subunit bound to nicks or single-stranded–double-stranded DNA junctions present in the probe. Nevertheless, these results clearly show that the RFC large subunit has a DNA-binding activity by itself. This was demonstrated directly (55), and the DNA-binding activity was mapped to a region containing the ligase homology domain (4).

RFC box I shows similarity to a region in prokaryotic DNA ligases and in procyclic acidic repetitive proteins (PARPs) from eukaryotes (Fig. 7). All three protein groups bind to primer-template or nick structures in DNA, and box I is the only conserved region in the fragment mapped by Burbelo et al. (4) as the DNA-binding domain of the large RFC subunit. The DNA-binding domains of the ligases have has not been mapped. In the PARPs, the region of similarity is contained in the automodification domain and not in the known DNA-binding domain (9). More work is necessary to characterize the DNA binding by these proteins.

Because of the high similarity between the five ScRFC subunits, it was surprising that each one is essential for cell proliferation, and we therefore assume that each one has a unique function in DNA replication, repair, or recombination. This could be accounted for by significant differences in the ATPase domains of the five subunits. In ScRFC5 and the homolog hRFC38, for example, we can see divergences from the consensus sequence of RFC boxes III (P loop), IV, and V (DEAD box) that affect residues that are important for the GTPase activity of $p21^{ras}$ (39, 40). Homologous RFC subunits also show good conservation outside the ATP/GTPase domain, where there is little similarity among the five subunits. This finding indicates that these regions (e.g., the part C terminal of RFC box VIII) are important for subunit-specific functions.

Biochemical data so far provide little clue as to what the specific functions of the four small RFC subunits might be. It is not clear whether all of the RFC subunits actually have an ATPase activity. An ATPase activity that is stimulated most efficiently by single-stranded DNA or primed single-stranded DNA was shown for ScRFC3 (25). There are also reports about ATP cross-linking to ScRFC2 (34), to a ScRFC 40-kDa band (26), and to one of the small hRFC subunits that runs as 41 kDa (55). But even if not all subunits have an ATPase activity, the question remains of why these domains (II to VIII) are conserved in all subunits. We can speculate that some subunits might have regulatory functions. We also do not know whether all of the RFC subunits associate in a single complex or whether several complexes with different subunit composition, e.g., the large subunit with different small subunits, may act in different processes such as DNA replication and repair. PCNA has been demonstrated to play a role in both DNA replication (43) and DNA repair (30, 33, 45). Although RFC has a clear role in DNA replication, its role in DNA repair remains to be determined.

ACKNOWLEDGMENTS

We thank J. Hurwitz for communication of the unpublished sequence from the hRFC38 cDNA.

This research was supported by grants AI20460 and CA13106 from the National Institutes of Health. G.C. was supported by a joint postdoctoral fellowship from the Swiss National Science Foundation and the Deutsche Forschungsgemeinschaft.

REFERENCES

1. **Bell, S. P., R. Kobayashi, and B. Stillman.** 1993. Yeast origin recognition complex functions in transcription silencing and DNA replication. Science **262:**1844–1870.

2. **Blanco, L., A. Bernad, M. A. Blasco, and M. Salas.** 1991. A general structure for DNA-dependent DNA polymerases. Gene **100:**27–38.
3. **Bunz, F., R. Kobayashi, and B. Stillman.** 1993. cDNAs encoding the large subunit of human replication factor C. Proc. Natl. Acad. Sci. USA **90:**11014–11018.

3a. **Bunz, F., and B. Stillman.** Unpublished results.

4. **Burbelo, P. D., A. Utani, Z.-Q. Pan, and Y. Yamada.** 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.
5. **Burgers, P. M. J.** 1991. *Saccharomyces cerevisiae* replication factor C. II. Formation and activity of complexes with the proliferating cell nuclear antigen and with DNA polymerases δ and ε. J. Biol. Chem. **266:**22698–22706.
6. **Cafferkey, R., P. R. Young, M. McLaughlin, D. J. Bergsma, Y. Koltin, G. M. Sathe, L. Faucette, W.-K. Eng, R. K. Johnson, and G. P. Livi.** 1993. Dominant missense mutations in a novel yeast protein related to mammalian phosphatidylinositol 3-kinase and VPS34 abrogate rapamycin cytotoxicity. Mol. Cell. Biol. **13:**6012–6023.
7. **Chen, M., Z.-Q. Pan, and J. Hurwitz.** 1992. Sequence and expression in *Escherichia coli* of the 40-kDa subunit of activator 1 (replication factor C) of HeLa cells. Proc. Natl. Acad. Sci. USA **89:**2516–2520.
8. **Chen, M., Z.-Q. Pan, and J. Hurwitz.** 1992. Studies of the cloned 37-kDa subunit of activator 1 (replication factor C) of HeLa cells. Proc. Natl. Acad. Sci. USA **89:**5211–5215.
9. **de Murcia, G., and J. M. de Murcia.** 1994. Poly(ADP-ribose) polymerase: a molecular nick-sensor. Trends Biochem. Sci. **19:**172–176.
10. **Dong, Z., R. Onrust, M. Skangalis, and M. O'Donnell.** 1993. DNA polymerase III accessory proteins. J. Biol. Chem. **268:**11758–11765.
11. **Eki, T., T. Matsumoto, Y. Murakami, and J. Hurwitz.** 1992. The replication of DNA containing the simian virus 40 origin by the monopolymerase and dipolymerase systems. J. Biol. Chem. **267:**7284–7294.
12. **Fien, K., and B. Stillman.** 1992. Identification of replication factor C from *Saccharomyces cerevisiae*: a component of the leading-strand DNA replication complex. Mol. Cell. Biol. **12:**155–163.
13. **Flower, A. M., and C. S. McHenry.** 1986. The adjacent *dnaZ* and *dnaX* genes of *Escherichia coli* are contained within one continuous open reading frame. Nucleic Acids Res. **14:**8091–8101.
14. **Helliwell, S. B., P. Wagner, J. Kunz, M. Deuter-Reinhard, R. Henriquez, and M. N. Hall.** 1994. *TOR1* and *TOR2* are structurally and functionally similar but not identical phosphatidylinositol kinase homologs in yeast. Mol. Biol. Cell **5:**105–118.
15. **Howell, E. A., M. A. McAlear, D. Rose, and C. Holm.** 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.

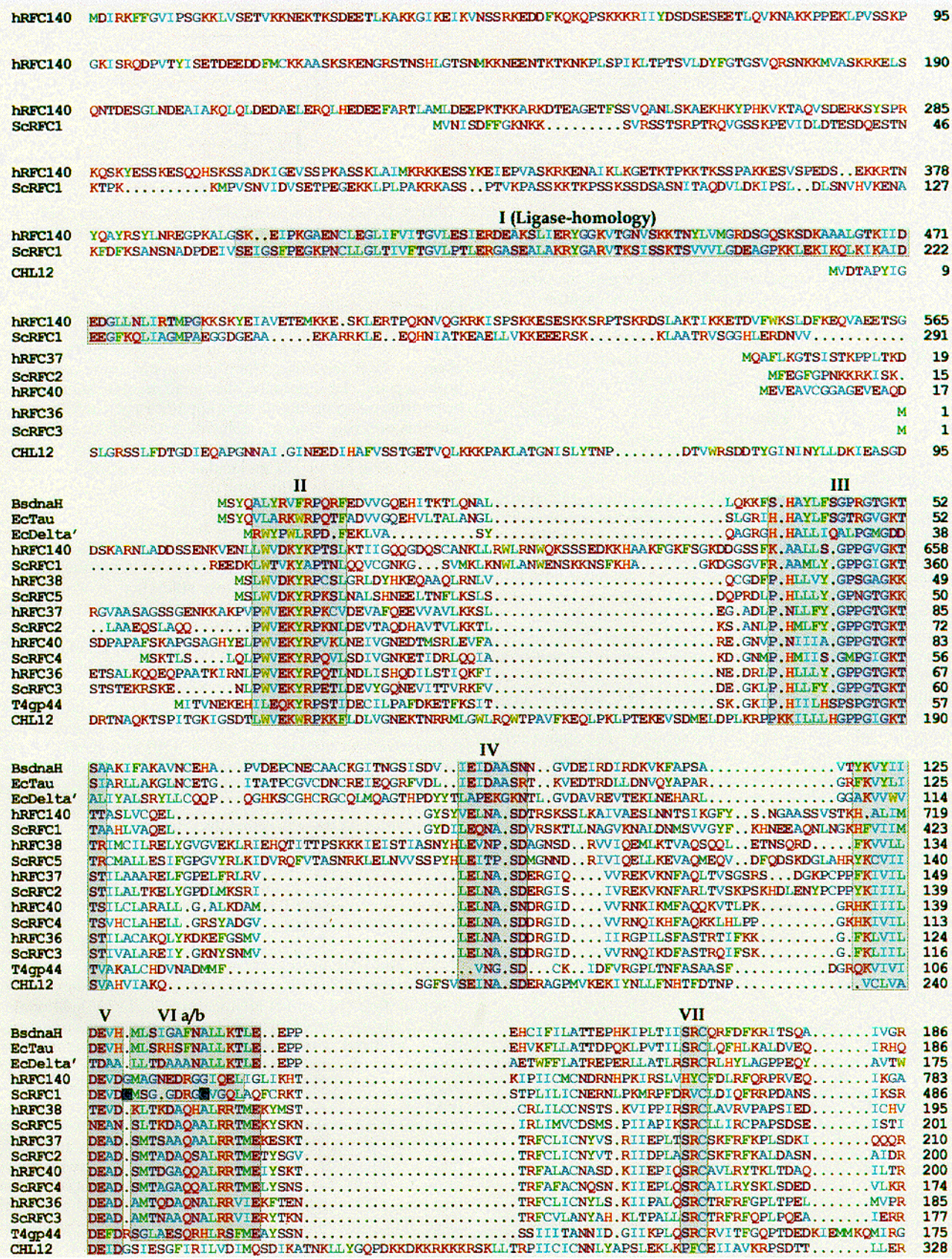
15a. **Hurwitz, J.** Personal communication.

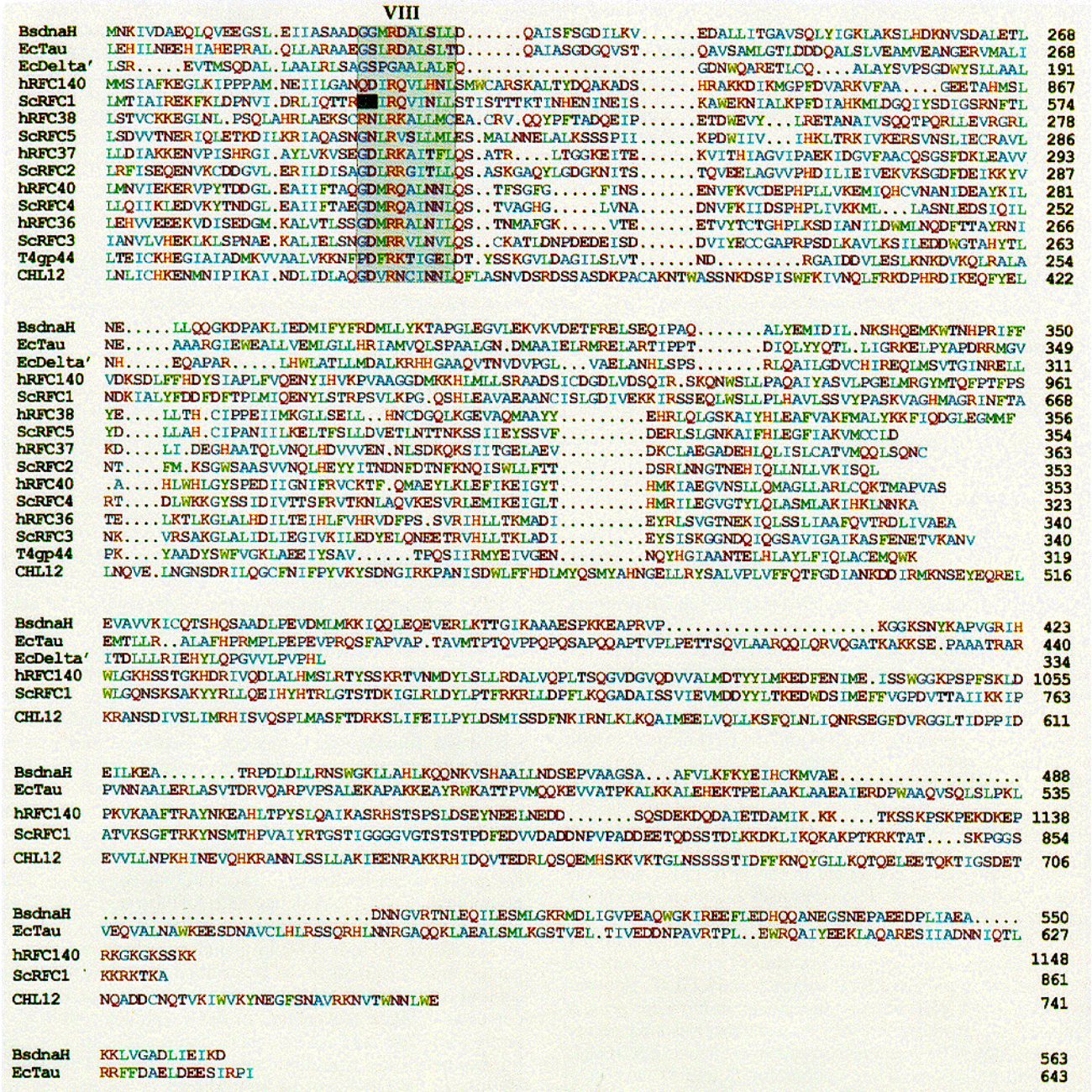
16. **Ittel, M.-E., J.-M. Garnier, J.-M. Jeltsch, and C. P. Niedergang.** 1991. Chicken poly(ADP-ribose) synthetase: complete deduced amino acid sequence and comparison with mammalian enzyme sequences. Gene **102:**157–164.
17. **Jones, W. E.** 1991. Tackling the protease problem in *Saccharomyces cerevisiae*. Methods Enzymol. **194:**428–453.
18. **Kouprina, N., A. Tsouladze, M. Koryabin, P. Hieter, F. Spencer, and V. Larionov.** 1993. Identification and genetic mapping of *CHL* genes controlling mitotic chromosome transmission in yeast. Yeast **9:**11–19.
19. **Kovalic, D., J.-H. Kwak, and B. Weisblum.** 1991. General method for direct cloning of DNA fragments generated by the polymerase chain reaction. Nucleic Acids Res. **19:**4560.
20. **Krishna, T. S. R., X.-P. Kong, S. Gary, P. M. Burgers, and J. Kuriyan.** 1994. Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. Cell **79:**1233–1243.
21. **Lauer, G., E. A. Rudd, D. L. McKay, A. Ally, D. Ally, and K. Backman.** 1991. Cloning, nucleotide sequence, and engineered expression of *Thermus thermophilus* DNA ligase, a homolog of *Escherichia coli* DNA ligase. J. Bacteriol. **173:**5047–5053.
22. **Lee, S.-H., and J. Hurwitz.** 1990. Mechanism of elongation of primed DNA by DNA polymerase δ, proliferating cell nuclear antigen, and activator I. Proc. Natl. Acad. Sci. USA **87:**5672–5676.
23. **Lee, S.-H., A. D. Kwong, Z.-Q. Pan, and J. Hurwitz.** 1991. Studies on the activator 1 protein complex, an accessory factor for proliferating cell nuclear antigen-dependent DNA polymerase δ J. Biol. Chem. **266:**594–602.
24. **Li, X., and P. M. J. Burgers.** 1994. Cloning and characterization of the essential *Saccharomyces cerevisiae RFC4* gene encoding the 37-kDa subunit of replication factor C. J. Biol. Chem. **269:**21880–21884.
25. **Li, X., and P. M. J. Burgers.** 1994. Molecular cloning and expression of the *Saccharomyces cerevisiae RFC3* gene, an essential component of replication factor C. Proc. Natl. Acad. Sci. USA **91:**868–872.
26. **Li, X., B. L. Yoder, and P. M. J. Burgers.** 1992. A *Saccharomyces cerevisiae* DNA helicase associated with replication factor C. J. Biol. Chem. **267:**25321–25327.
27. **Linder, P., P. F. Lasko, M. Ashburner, P. Leroy, P. J. Nielsen, K. Nishi, J. Schnier, and P. P. Slonimski.** 1989. Birth of the D-E-A-D box. Nature (London) **337:**121–122.
28. **Lu, Y., and A. T. Riegel.** 1994. The human DNA-binding protein, PO-GA, is homologous to the large subunit of mouse replication factor C: regulation by alternate 3′ processing of mRNA. Gene **145:**261–265.
29. **Luckow, B., F. Bunz, B. Stillman, P. Lichter, and G. Schütz.** 1994. Cloning, expression, and chromosomal localization of the 140-kilodalton subunit of replication factor C from mice and humans. Mol. Cell. Biol. **14:**1626–1634.
30. **Matsumoto, Y., K. Kim, and D. F. Bogenhagen.** 1994. Proliferating cell nuclear antigen-dependent abasic site repair in *Xenopus laevis* oocytes: an alternative pathway of base excision DNA repair. Mol. Cell. Biol. **14:**6187–6197.
31. **McAlear, M. A., E. A. Howell, K. K. Espenshade, and C. Holm.** 1994. Proliferating cell nuclear antigen (*pol30*) mutations suppress *cdc44* mutations and identify potential regions of interaction between the two encoded proteins. Mol. Cell. Biol. **14:**4390–4397.
32. **Moir, D., S. E. Stewart, B. C. Osmond, and D. Botstein.** 1982. Cold-sensitive cell-division-cycle mutants of yeast: isolation, properties, and pseudoreversion studies. Genetics **100:**547–563.
33. **Nichols, A. F., and A. Sancar.** 1992. Purification of PCNA as a nucleotide excision repair protein. Nucleic Acids Res. **20:**3559–3564.
34. **Noskov, V., S. Maki, Y. Kawasaki, S.-H. Leem, B.-I. Ono, H. Araki, Y. Pavlov, and A. Sugino.** 1994. The *RFC2* gene encoding a subunit of replication factor C of *Saccharomyces cerevisiae*. Nucleic Acids Res. **22:**1527–1535.
35. **O'Donnell, M., R. Onrust, F. B. Dean, M. Chen, and J. Hurwitz.** 1993. Homology in accessory proteins of replicative polymerase-*E. coli* to humans. Nucleic Acids Res. **21:**1–3.
36. **Ogas, J., B. J. Andrews, and I. Herskowitz.** 1991. Transcriptional activation of *CLN1*, *CLN2*, and a putative new G1 cyclin (*HCS26*) by SWI4, a positive regulator of G1-specific transcription. Cell **66:**1015–1026.
37. **Ogasawara, N., S. Nakai, and H. Yoshikawa.** 1994. Systematic sequencing of the 180 kilobase region of the *Bacillus subtilis* chromosome containing the replication origin. DNA Res. **1:**1–14.
38. **Onrust, R., P. T. Stuckenberg, and M. O'Donnell.** 1991. Analysis of the ATPase subassembly which initiates processive DNA synthesis by DNA polymerase III. J. Biol. Chem. **266:**21681–21686.
39. **Pai, E. F., W. Kabsch, U. Krengel, K. C. Holmes, J. John, and A. Wittighofer.** 1989. Structure of the guanine-nucleotide-binding domain of the Ha-*ras* oncogene product p21 in the triphosphate conformation. Nature (London) **341:**209–214.
40. **Pai, E. F., U. Krengel, G. A. Petsko, R. S. Goody, W. Kabsch, and A. Wittighofer.** 1990. Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. EMBO J. **9:**2351–2359.
41. **Pause, A., and N. Sonenberg.** 1992. Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. EMBO J. **11:**2643–2654.
42. **Podust, V. N., A. Georgaki, B. Strack, and U. Hübscher.** 1992. Calf thymus RF-C as an essential component for DNA polymerase δ and ε holoenzymes function. Nucleic Acids Res. **20:**4159–4165.
43. **Prelich, G., M. Kostura, D. R. Marshak, M. B. Mathews, and B. Stillman.** 1987. The cell-cycle regulated proliferating cell nuclear antigen is required for SV40 DNA replication *in vitro*. Nature (London) **326:**471–475.

43a. **Seufert, W., and B. Futcher.** Unpublished results.

44. **Sherman, F., and J. Hicks.** 1991. Micromanipulation and dissection of asci. Methods Enzymol. **194:**21–37.
45. **Shivji, M. K. K., M. K. Kenny, and R. D. Wood.** 1992. Proliferating cell nuclear antigen is required for DNA excision repair. Cell **69:**367–374.
46. **Spencer, F., S. L. Gerring, C. Connelly, and P. Hieter.** 1990. Mitotic chromosome transmission fidelity mutants in *Saccharomyces cerevisiae*. Genetics **124:**237–249.
47. **Spicer, E. K., N. G. Nossal, and K. R. Williams.** 1984. Bacteriophage T4 gene 44 DNA polymerase accessory protein: sequence of gene 44 and its protein product. J. Biol. Chem. **259:**15425–15432.
48. **Stillman, B.** 1994. Smart machines at the DNA replication fork. Cell **78:**725–728.
49. **Stukenberg, P. T., J. Turner, and M. O'Donnell.** 1994. An explanation for lagging strand replication: polymerase hopping among DNA sliding clamps. Cell **78:**877–887.
50. **Tsuchihashi, Z., and A. Kornberg.** 1989. ATP interactions of the τ and γ subunits of DNA polymerase III holoenzyme of *Escherichia coli*. J. Biol. Chem. **264:**17790–17795.
51. **Tsurimoto, T., M. P. Fairman, and B. Stillman.** 1989. Simian virus 40 DNA replication in vitro: identification of multiple stages of initiation. Mol. Cell. Biol. **9:**3839–3849.
52. **Tsurimoto, T., T. Melendy, and B. Stillman.** 1990. Sequential initiation of lagging and leading strand synthesis by two DNA polymerase complexes at the SV40 DNA replication origin. Nature (London) **346:**534–539.
53. **Tsurimoto, T., and B. Stillman.** 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.
54. **Tsurimoto, T., and B. Stillman.** 1990. Functions of replication factor C and proliferating cell nuclear antigen: functional similarity of DNA polymerase accessory proteins from human cells and bacteriophage T4. Proc. Natl. Acad.

Sci. USA **87:**1023–1027.
55. **Tsurimoto, T., and B. Stillman.** 1991. Replication factors required for SV40 DNA replication *in vitro*. I. DNA structure specific recognition of a primer-template junction by eukaryotic DNA polymerases and their accessory factors. J. Biol. Chem. **266:**1950–1960.
56. **Tsurimoto, T., and B. Stillman.** 1991. Replication factors required for SV40 DNA replication *in vitro*. II. Switching of DNA polymerase α and δ during initiation of leading and lagging strand synthesis. J. Biol. Chem. **266:**1961–1968.
57. **Uchida, K., T. Morita, T. Sato, T. Ogura, R. Yamashita, S. Noguchi, H. Suzuki, H. Nyunoya, M. Miwa, and T. Sugimura.** 1987. Nucleotide sequence of a full-length cDNA for human fibroblast Poly(ADP-ribose) polymerase. Biochem. Biophys. Res. Commun. **148:**617–622.
58. **Waga, S., and B. Stillman.** 1994. Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication *in vitro*. Nature (London) **369:**207–212.
59. **Weinberg, D. H., K. L. Collins, P. Simancek, A. Russo, M. S. Old, D. M. Virshup, and T. J. Kelly.** 1990. Reconstitution of simian virus 40 DNA replication with purified proteins. Proc. Natl. Acad. Sci. USA **87:**8692–8696.
60. **Yin, K. C., A. Blinkowa, and J. R. Walker.** 1986. *E. coli* dnaZX replication gene. Nucleic Acids Res. **14:**6541–6549.
61. **Yoder, B. L., and P. M. J. Burgers.** 1991. *Saccharomyces cerevisiae* replication factor C I. Purification and characterization of its ATPase activity. J. Biol. Chem. **266:**22689–22697.
62. **Young, M. C., M. K. Reddy, and P. H. von Hippel.** 1992. Structure and function of the bacteriophage T4 DNA polymerase holoenzyme. Biochemistry **31:**8675–8690.


I (Ligase-homology)
II
III
IV
V
VI a/b
VII


VIII
BsdnaH MNKIVDAEQLQVEEGSL.EIIASAADGGMRDALSLLD.....QAISFSGDILKV.....EDALLITGAVSQLYIGKLAKSLHDKNVSDALETL 268
EcTau LEHIILNEEHIAHEPRAL.QLLARAAEGSLRDALSLTD.....QAIASGDGQVST.....QAVSAMLGTLDDDQALSLVEAMVEANGERVMALI 268
EcDelta' LSR.....EVTMSQDAL.LAALRLSAGSPGAALALFQ.....GDNWQARETLCQ....ALAYSVPSGDWYSLLAAL 191
hRFC140 MMSIAFKEGLKIPPPAM.NEIILGANQDIRQVLHNLSMWCARSKALTYDQAKADS.....HRAKKDIKMGPFDVARKVFAA....GEETAHMSL 867
ScRFC1 LMTIAIREKFKLDPNVI.DRLIQTTR[redacted]IRQVINLLSTISTTTKTINHENINEIS.....KAWEKNIALKPFDIAHKMLDGQIYSDIGSRNFTL 574
hRFC38 LSTVCKKEGLNL.PSQLAHRLAEKSCRNLRKALLMCEA.CRV.QQYPFTADQEIP.....ETDWEVY...LRETANAIVSQQTPQRLLEVRGRL 278
ScRFC5 LSDVVTNERIQLETKDILKRIAQASNGNLRVSLLMLES.MALNNELALKSSSPII.....KPDWIIV...IHKLTRKIVKERSVNSLIECRAVL 286
hRFC37 LLDIAKKENVPISHRGI.AYLVKVSEGDLRKAITFLQS.ATR....LTGGKEITE.....KVITHIAGVIPAEKIDGVFAACQSGSFDKLEAVV 293
ScRFC2 LRFISEQENVKCDDGVL.ERILDISAGDLRRGITLLQS.ASKGAQYLGDGKNITS.....TQVEELAGVVPHDILIEIVEKVKSGDFDEIKKYV 287
hRFC40 LMNVIEKERVPYTDDGL.EAIIFTAQGDMRQAINNLQS..TFSGFG.....FINS.....ENVFKVCDEPHPLLVKEMIQHCVNANIDEAYKIL 281
ScRFC4 LLQIIKLEDVKYTNDGL.EAIIFTAEGDMRQAINNLQS..TVAGHG.....LVNA.....DNVFKIIDSPHPLIVKKML...LASNLEDSIQIL 252
hRFC36 LEHVVEEEKVDISEDGM.KALVTLSSGDMRRAINILQS..TNMAFGK.....VTE.....ETVYTCTGHPLKSDIANILDWMLNQDFTTAYRNI 266
ScRFC3 IANVLVHEKLKLSPNAE.KALIELSNGDMRRVLNVLQS..CKATLDNPDEDEISD.....DVIYECCGAPRPSDLKAVLKSILEDDWGTAHYTL 263
T4gp44 LTEICKHEGIAIADMKVVAALVKKNFPDFRKTIGELDT.YSSKGVLDAGILSLVT.....ND.....RGAIDDVLESLKNKDVKQLRALA 254
CHL12 LNLICHKENMNIPIKAI.NDLIDLAQGDVRNCINNLQFLASNVDSRDSSASDKPACAKNTWASSNKDSPISWFKIVNQLFRKDPHRDIKEQFYEL 422
BsdnaH NE.....LLQQGKDPAKLIEDMIFYFRDMLLYKTAPGLEGVLEKVKVDETFRELSEQIPAQ.....ALYEMIDIL.NKSHQEMKWTNHPRIFF 350
EcTau NE.....AAARGIEWEALLVEMLGLLHRIAMVQLSPAALGN.DMAAIELRMRELARTIPPT.....DIQLYYQTL.LIGRKELPYAPDRRMGV 349
EcDelta' NH.....EQAPAR.....LHWLATLLMDALKRHHGAAQVTNVDVPGL...VAELANHLSPS.....RLQAILGDVCHIREQLMSVTGINRELL 311
hRFC140 VDKSDLFFHDYSIAPLFVQENYIHVKPVAAGGDMKKHIMLLSRAADSICDGDLVDSQIR.SKQNWSLLPAQAIYASVLPGELMRGYMTQFPTFPS 961
ScRFC1 NDKIALYFDDFDFTPLMIQENYLSTRPSVLKPG.QSHIEAVAEAANCISLGDIVEKKIRSSEQLWSLLPLHAVLSSVYPASKVAGHMAGRINFTA 668
hRFC38 YE....LLTH.CIPPEIIMKGLLSELL..HNCDGQLKGEVAQMAAYY.....EHRLQLGSKAIYHLEAFVAKFMALYKKFIQDGLEGMMF 356
ScRFC5 YD....LLAH.CIPANIILKELTFSLLDVETLNTTNKSSIIEYSSVF.....DERLSLGNKAIFHLEGFIAKVMCCID 354
hRFC37 KD....LI.DEGHAATQLVNQLHDVVVEN.NLSDKQKSIITGELAEV.....DKCLAEGADEHLQLISLCATVMQQLSQNC 363
ScRFC2 NT....FM.KSGWSAASVVNQLHEYYITNDNFDTNFKNQISWLLFTT.....DSRLNNGTNEHIQLLNLLVKISQL 353
hRFC40 .A....HLWHLGYSPEDIIGNIFRVCKTF.QMAEYLKLEFIKEIGYT.....HMKIAEGVNSLLQMAGLLARLCQKTMAPVAS 353
ScRFC4 RT....DLWKKGYSSIDIVTTSFRVTKNLAQVKESVRLEMIKEIGLT.....HMRILEGVGTYLQLASMLAKIHKLNNKA 323
hRFC36 TE....LKTLKGLALHDILTEIHLFVHRVDFPS.SVRIHLLTKMADI.....EYRLSVGTNEKIQLSSLIAAFQVTRDLIVAEA 340
ScRFC3 NK....VRSAKGLALIDLIEGIVKILEDYELQNEETRVHLLTKLADI.....EYSISKGGNDQIQGSAVIGAIKASFENETVKANV 340
T4gp44 PK....YAADYSWFVGKLAEEIYSAV.....TPQSIIRMYEIVGEN.....NQYHGIAANTELHLAYLFIQLACEMQWK 319
CHL12 LNQVE.LNGNSDRILQGCFNIFPYVKYSDNGIRKPANISDWLFFHDLMYQSMYAHNGELLRYSALVPLVFFQTFGDIANKDDIRMKNSEYEQREL 516
BsdnaH EVAVVKICQTSHQSAADLPEVDMLMKKIQQLEQEVERLKTTGIKAAAESPKKEAPRVP.....KGGKSNYKAPVGRIH 423
EcTau EMTLLR..ALAFHPRMPLPEPEVPRQSFAPVAP.TAVMTPTQVPPQPQSAPQQAPTVPLPETTSQVLAARQQLQRVQGATKAKKSE.PAAATRAR 440
EcDelta' ITDLLLRIEHYLQPGVVLPVPHL 334
hRFC140 WLGKHSSTGKHDRIVQDLALHMSLRTYSSKRTVNMDYLSLLRDALVQPLTSQGVDGVQDVVALMDTYYLMKEDFENIME.ISSWGGKPSPFSKLD 1055
ScRFC1 WLGQNSKSAKYYRLLQEIHYHTRLGTSTDKIGLRLDYLPTFRKRLLDPFLKQGADAISSVIEVMDDYYLTKEDWDSIMEFFVGPDVTTAIIKKIP 763
CHL12 KRANSDIVSLIMRHISVQSPLMASFTDRKSLIFEILPYLDSMISSDFNKIRNLKLKQAIMEELVQLLKSFQLNLIQNRSEGFDVRGGLTIDPPID 611
BsdnaH EILKEA.....TRPDLDLLRNSWGKLLAHLKQQNKVSHAALLNDSEPVAAGSA...AFVLKFKYEIHCKMVAE..... 488
EcTau PVNNAALERLASVTDRVQARPVPSALEKAPAKKEAYRWKATTPVMQQKEVVATPKALKKALEHEKTPELAAKLAAEAIERDPWAAQVSQLSLPKL 535
hRFC140 PKVKAAFTRAYNKEAHLTPYSLQAIKASRHSTSPSLDSEYNEELNEDD.....SQSDEKDQDAIETDAMIK.KK....TKSSKPSKPEKDKEP 1138
ScRFC1 ATVKSGFTRKYNSMTHPVAIYRTGSTIGGGGVGTSTSTPDFEDVVDADDNPVPADDEETQDSSTDLKKDKLIKQKAKPTKRKTAT....SKPGGS 854
CHL12 EVVLLNPKHINEVQHKRANNLSSLLAKIEENRAKKRHIDQVTEDRLQSQEMHSKKVKTGLNSSSSTIDFFKNQYGLLKQTQELEETQKTIGSDET 706
BsdnaH .....DNNGVRTNLEQILESMLGKRMDLIGVPEAQWGKIREEFLEDHQQANEGSNEPAEEDPLIAEA..... 550
EcTau VEQVALNAWKEESDNAVCLHLRSSQRHLNNRGAQQKLAEALSMLKGSTVEL.TIVEDDNPAVRTPL..EWRQAIYEEKLAQARESIIADNNIQTL 627
hRFC140 RKGKGKSSKK 1148
ScRFC1 KKRKTKA 861
CHL12 NQADDCNQTVKIWVKYNEGFSNAVRKNVTWNNLWE 741
BsdnaH KKLVGADLIEIKD 563
EcTau RRFFDAELDEESIRPI 643