A Novel Proteomic Approach Identifies New Interaction Partners for Proliferating Cell Nuclear Antigen

Laurence Meslet-Cladiere[1,x,*], Cedric Norais[1,x,*], Joëlle Kuhn[1], Julien Briffotaux[2], Jerry W. Sloostra[3], Elena Ferrari[4], Ulrich Hübscher[4], Didier Flament [2*] and Hannu Myllykallio [1*]

1 INSERM Avenir Group – Institut du Génétique et Microbiologie, CNRS UMR8621, U.Paris-Sud, 91405 Orsay Cedex, France
2 UMR 6197, Laboratoire de Microbiologie des Environnements Extrêmes, IFREMER Brest, 29280 Plouzane, France
3 Pepscan Systems, 8219 PH Lelystad, The Netherlands
4 Institute of Veterinary Biochemistry and Molecular Biology, University of Zurich-Irchel, Winterthurerstrasse 190, CH 8057 Zurich, Switzerland.

*Correspondence regarding SPR measurements should be addressed to D.F (didier.flament@ifremer.fr). All other correspondence should be addressed to H.M (hannu.myllykallio@igmors.u-psud.fr). **x**Two first authors contributed equally in this work.

Abstract:

During DNA replication and repair, many proteins bind to and dissociate in a highly specific and ordered manner from proliferating cell nuclear antigen (PCNA). We describe a combined approach of *in silico* searches at the genome level and combinatorial peptide synthesis to investigate the binding properties of hundreds of short PCNA-interacting peptides (PIP-peptides) to archaeal and eukaryal PCNAs. Biological relevance of our combined approach was demonstrated by identification an inactive complex of *Pyrococcus abyssi* ribonuclease HII with PCNA. Furthermore we show that PIP-peptides interact with PCNA largely in a sequence independent manner. Our experimental approach also identified many so far unidentified PCNA interacting peptides in a number of human proteins.

Keywords: PCNA; combinatorial peptide synthesis; ribonuclease HII

Abbreviations: PCNA, proliferating cell nuclear antigen; PIP-peptide, PCNA-interacting peptide; SPR, surface plasmon resonance

Proliferating cell nuclear antigen (replication clamp PCNA) was originally identified as an antigen present in the nuclei of dividing cells that reacted with an autoantibody in the sera of patients with lupus erythematosus [1]. Later studies demonstrated that PCNA is a trimeric processivity factor of eukaryotic replicative DNA polymerases [2]. The outer surface of the ring-shaped PCNA trimer is negatively charged whereas the central cavity accommodating DNA carries positive charges [3]. One PCNA monomer is composed of two domains that are linked by a domain-connecting loop. This loop is evolutionary conserved and plays a crucial role in mediating specific interactions of the replication clamp with its interaction partners in Okazaki fragment processing, DNA repair, translesion DNA synthesis, DNA methylation, chromatin remodeling and cell cycle regulation [4]. In order to be functionally active, different PCNA-interacting proteins must bind and dissociate in a highly specific and ordered manner. In view of this fact it is surprising that the identified peptide domain motifs of PCNA-interacting proteins (PIP-motifs) are quite similar and have overlapping binding sites with PCNA. For instance, the sequence motif Q-x(2)-(h)-x(2)-(a)-(a) (h, moderately hydrophobic residue; a, a highly hydrophobic residue) is known to mediate interactions of archaeal and human PCNA with many partners [5; 6]. As these motifs are short, poorly conserved and can be found in literally thousands of proteins, the possibilities to identify (new) PCNA interacting proteins based solely upon *in silico* sequence comparisons are limited, making functional screens necessary.

In this work we describe a combinatorial approach of *in silico* searches and combinatorial peptide synthesis to identify new PCNA interaction partners. In contrast to earlier screens that have used completely random peptide libraries, we have tested focused peptide libraries that were designed benefiting either from *P. abyssi* or human genome. As model for the proof of principle we used the presumed enzymatic machinery for the processing of Okazaki fragments in the hyperthermophilic archaeon *Pyrococcus abyssi* (*P. abyssi*). Our approach led to

identification of new interaction partners for archaeal PCNA and allows predicting hitherto unidentified interaction partners for human PCNA. Biological significance of our results was revealed through identification of an inactive complex of *P. abyssi* ribonuclease (RNase) HII with PCNA, indicating also a DNA repair function for RNase HII orthologs. We also show that the PIP-motif sustains a considerable amount of substitutions without having deleterious effect on the PCNA binding.

Results and Discussion

Screening of immobilized peptides with the robust archaeal *P. abyssi* PCNA

In order to experimentally identify PCNA-interacting peptides, we firstly screened a small library of 910 random X-C-X(8)-C-X peptides, immobilized onto a solid support via cysteine residues, with 2 μM of stable PCNA preparation from *P. abyssi*. *P. abyssi* PCNA carries a six histidine tag at the amino terminal end, allowing detection of the bound protein by a specific anti-His antibody. The background signal resulting from non-specific binding of *P. abyssi* PCNA corresponded to binding values smaller than 100 arbitrary units; no signal was observed when PCNA was omitted from the binding reaction. These analyses allowed the identification of 47 immobilized lead peptides with binding values higher than five hundred for *P. abyssi* PCNA (Figure 1A and Table S1). Manual alignment showed that some of these peptides (Figure 1B) possess the PIP-motif [6; 7]. However, many additional peptides with high binding values did not show obvious sequence similarity to the PCNA-interacting protein motif (see Table S1), suggesting that physico-chemical or structural properties of these peptides played an important role in binding. This idea was supported by the fact that the average predicted isoelectric point (pI) of 50 high affinity peptides was statistically significantly higher than the pI of non-interacting peptides (8.46 versus 6.31, see Figure 1A). This observation indicates that the positive charge of these peptides could contribute to the observed binding. It is of note that previously performed peptide screens to identify new PCNA-binding peptides failed to identify statistically significant correlation between the PCNA binding and positive charge [8; 9]. This likely reflects the fact that not enough PCNA binding peptides were identified in earlier peptide screens and/or that binding values were not recorded for peptides with low affinities towards PCNA.

We next synthesized several control peptides found in *P. abyssi* proteins (Figure 1C) that were immobilized on a solid support and screened as above with 1 μM *P. abyssi* PCNA.

Naturally occurring peptides with the highest binding values contained the conserved *P. abyssi* sequence motif Q-x(2)-L-x(2)-[WFT]-[LFT] (Figure 1C). We furthermore noted that this sequence motif is often flanked by positively charged residues (Figure 1C). Many "high affinity" peptides are present in proteins that are known to interact with the PCNA. However, we also found that peptides from *P. abyssi* ribonuclease H (RNase HII) and PAB2263, two proteins that previously have not been shown to interact with PCNA, contain a PIP-motif with a strong PCNA binding activity (Figure 1C).

The interaction domains of *P. abyssi* RNase HII, flap endonuclease 1 and DNA ligase with PCNA overlap

As a proof of principle, we next set out to further investigate how the above peptides when contained in archaeal proteins (their natural context) mediate interactions with PCNA. We studied the PCNA binding of the *P. abyssi* RNase HII, the structure-specific flap endonuclease 1 (Fen1) and the ATP-dependent DNA ligase using surface plasmon resonance (SPR). PCNA-interacting peptides located at the carboxyl termini of RNase HII and Fen1 have high binding values (Figure 1C), whereas the motif that was identified at the N-terminus of DNA ligase shows moderate affinity for PCNA. In this respect it is noteworthy that very recently an additional PIP-motif "QKSFF" in the central region of *P. furiosus* DNA ligase has been described [10], suggesting that the two motifs could participate in efficient binding. First, recombinant proteins for SPR measurements, carrying a six histidine tag at the N-terminus, were produced and purified to greater than 99% homogeneity using Ni-NTA agarose and gel filtration chromatography (Figure 2A). SPR measurements with immobilized homotrimeric PCNA were performed to analyze the physical interactions of the purified proteins with PCNA (Figures 2B-D). The results indicated that RNase HII, Fen1 and DNA ligase physically interact with *P. abyssi* PCNA at nanomolar concentrations (Figure 2B-D, trace 1). It is also of note that the high binding values observed for the individual peptides (Figure 1C and data not shown)

correlate with the K_D values in nanomolar range. These values were estimated using global fitting analysis that accounts for both association and dissociation phases as well as the maximal level of complex formation. It is interesting to note that the similar SPR data has been obtained for a PAB2263-PCNA complex of unknown function (manuscript in preparation).

To exclude that the observed SPR data result from the oligomerisation of the proteins studied and/or from non-specific interactions under the experimental conditions used, we performed deletion analyses and competition experiments. RNase HII, with a 17 residue deletion at the C-terminus that removed the PIP motif, failed to interact with PCNA (Figure 2B, insert). The competition experiments were performed in the presence of a ten-fold excess of the peptide KRK**Q**LT**L**DN**F**LKR (bold residues correspond to conserved PIP motif residues). This peptide is a rationally designed composite peptide based on the RNase HII PIP-peptide (KRAQLTLDNFLK) where the alanine at the position 3 was changed to lysine that occurs in Fen1 and PAB2263 PIP-peptides in the corresponding positions. The binding value of this peptide is similar to those of Fen1 and PAB2263 (data not shown). The addition of this peptide substantially decreased the rate of complex formation and/or accelerated dissociation of RNase HII-, Fen1- and DNA ligase-PCNA complexes (Figure 2B-D, trace 3), whereas the non-binding control peptide (KEVKEEYKRFLEE) had no effect on complex formation (Figure 2B-D, trace 2). These results clearly indicated that the binding sites for all three proteins with PCNA do considerably overlap. To obtain further evidence for the physiological significance of the interactions detected *in vitro* between PCNA, RNase HII and Fen1, their complex formation in cell free extracts was analyzed (Figure 2E). Fen1 and RNase HII were precipitated together with PCNA from cell-free extracts, whereas the lack of detectable interaction between DNA ligase and PCNA could be explained by additional PCNA interaction partners, present in cell-free extracts, preventing the binding of DNA ligase to PCNA.

The specific physical interaction between *P. abyssi* PCNA and RNase HII measured by SPR suggested that PCNA might modulate the activity of RNase HII. We therefore examined the influence of PCNA on the enzymatic properties of *P. abyssi* RNase HII on heteroduplex substrate carrying embedded ribonucleotides, which likely corresponds to the physiologically relevant RNase HII substrate [11; 12]. Upon addition of PCNA the RNase HII activity on this substrate was strongly inhibited (Figure 2F). To test whether this inhibition results from a physical inhibitory interaction with PCNA, RNase HII activity was measured by changing the molar PCNA/RNase HII ratio in the reaction mixtures. These studies were performed with both wild-type (WT) RNase HII and its C-terminal deletion (ΔCter) derivative incapable of interacting with PCNA (Figure 2B). The results obtained revealed that inhibition of the ribonuclease activity of RNase HII by PCNA was dependent on the PIP-motif located at the C-terminus of the enzyme (Figure 2F). This inhibition could be explained by the fact that the substrate-binding domain of *Thermococcus kodakaraensis* RNase HII is located in a carboxy-terminus [11]. In control experiments we found that the PCNA preparation used was capable to activate *P. abyssi* Fen1 and DNA ligase (data not shown), indicating that PCNA was fully functional and properly folded.

Archaeal RNase H and Fen1 have a key role in DNA repair

To further define the physiologically relevant activities of archaeal RNase HII and Fen1, we inactivated the structural genes for RNase HII and Fen1 (Figure 3). For this we chose the halophile, *Haloferax volcanii*, since the necessary genetic tools have not been developed for *Pyrococcus* species. It is of note that we were unable to obtain a *H. volcanii* strain carrying a deletion of the structural gene for PCNA, suggesting that this processivity factor is essential in

this organism. While *P. abyssi* contains only one RNase H (family II encoded by *rnhB*), *H. volcanii* contains detectable orthologs for RNase HII and RNase HI [likely bacterial origin (encoded by *rnhA*)]. Mutant strains, lacking *fen1*, *rnhA*, *rnhB*, or different combinations of these, were constructed and confirmed as indicated in Materials and Methods section. On solid medium, the "wild type" and Δ*rnhB* Δ*rnhA* strains grew similarly whereas the growth of Δ*fen1* deletion strain was impaired. This visual observation was confirmed by measuring the doubling time for the parental and Δ*rnhB* Δ*rnhA* of 131±6 (n=5) and 108 (average of two measurements) minutes whereas Δ*fen1* strain had the doubling time of 189 minutes (average of two measurements). We also measured the incorporation rate of extracellular tritiated uracil into alkaline resistant form (DNA) using these different mutant strains (Table 2). These measurements suggested that archaeal Fen1 has a role in DNA replication, as earlier observed for Eukarya [13]. On the other hand, the *in vivo* role of RNase HII in DNA replication was not evident (Table 2). Considering that GC% of *H. volcanii* genome is relatively high, we can not exclude a possibility that *H. volcanii* RnhA origin participates in the processing of R-loops that could be found at some replication origins in this species [14; 15].

We also tested our deletion strains for possible differences regarding their sensitivity to UV radiation (Table 2, Figure 4). Our results revealed that the *fen1* deletion resulted in markedly enhanced UV sensitivity, indicating that archaeal Fen1 either participates in removal of photoproducts resulting from UV irradiation and/or indicate a role in DNA replication occurring downstream of unrepaired damage on lagging strand templates [13; 16]. Our results also indicated that both single RNase H mutant strains are moderately UV sensitive (Table 2), with an additive effect observed for the Δ*rnhA* Δ*rnhB* strain (Table 2 and Figure 3D). Thus, substrate specificity of archaeal RNase HII orthologs, inhibition of RNase HII activity by PCNA (Figure 2) and genetic studies (Figure 3) propose that major physiological role of archaeal RNase HII is not to participate in DNA replication but rather, possibly together with

Fen1 endonuclease [12], in repair of ribonucleotides misincorporated into chromosomal DNA. As translesion bypass polymerases capable of ribonucleotide misincorporation are often found in archaeal species exposed to UV radiation, for instance in *H. volcanii*, the *Sulfolobales*, and *Picrophilus* species [17], translesion synthesis through the UV damaged DNA, could be one mechanism that increases ribonucleotide misincorporation in chromosomal DNA.

Predicted interaction partners of human PCNA

The above experiments performed with archaeal proteins indicated that a combination of motif searches and screening with immobilized peptides can be used to identify new PCNA interacting partners. To expand this notion to different systems, we next synthesized 98 peptides carrying the Q-x(4,5)-[HDF]-[FY] motif occurring in human proteins. These peptides were subsequently screened with human PCNA carrying a C-terminal hexahistidine tag [7]. Figure 4A shows that, despite all 98 peptides screened contained the aforementioned sequence motif, the observed binding values varied between 80 and 1400 (left *y* axis). These experimental data are consistent with the earlier prediction that variable PIP-box sequences could result in a wide range of PCNA affinities [18]. The sliding average (in a window of 10 peptides) of the calculated pIs (Fig. 4A, right *y* axis), clearly indicates that high affinity peptides tend to be positively charged. Figure S1A shows an alignment of the peptides with observed binding values higher than 400. This list not only contains proteins known to interact with PCNA (bolded in Figure S1**A**), but also reveals several potential novel interaction partners for human PCNA. Obviously, more experiments are now required to address whether these proteins interact functionally and/or physically with human PCNA.

Since the interaction surface between the PCNA and its interaction partners is at least partially formed by the beta zipper peptide backbone [19; 20], formation of the complex interface could be largely independent of the protein sequence. To test this, we synthesized 60 different peptides where the bolded positions in the peptide KRK**Q**LT**L**DN**F**LKR were systematically

changed to all naturally occurring amino acids (Figure 4B) and performed a further functional screening. Our results revealed that not even the "fully conserved" glutamine residue is indispensable for binding in our experiments. On the other hand, the hydrophobic substitutions at the positions 7 and 10 were favored. These results indicate that the short PIP-motif sustains a high number of substitutions without having a drastic effect on the binding activity. These observations are in agreement with the notion that the beta zipper formation together with electrostatic and hydrophobic interactions, and not specific protein sequence, is important for binding. Similar results further indicating specific binding independent of peptide sequence were obtained after screening with 80 additional peptides based upon K**R**KQ**L**TLD**N**F**L**KR, where the underlined residues were changed in the same way as above (Figure S1B). With these peptides we also performed a screening with the QLGI125 mutant in the domain-connecting loop of human PCNA. This loop participates in specific interactions of PCNA with its interaction partners, and the mutant PCNA has a markedly lower affinity towards Fen1 than the wild type protein [7]. The signal observed with the QLGI125 mutant was consistently lower than observed for the wild type protein (Figure S1B), further indicating that the binding signal observed in our experiments is highly specific. The fact that binding by this mutant was not completely abolished in our experiments is in agreement with the proposed multi-factorial nature of this interaction.

Conclusions

We have developed an experimental approach that allows the rapid identification of peptide sequences with high affinity towards PCNA. Differently from the earlier peptide screens to identify PCNA binding peptides, we have benefited from the combined *P. abyssi* and human genome sequence information to synthesize a number of immobilized peptides carrying putative PIP-motifs. Usefulness and efficiency of this approach to identify novel PCNA

interaction partners were revealed by our finding that *P. abyssi* PCNA and RNase HII form an inactive complex. While proof of principle was obtained using robust archaeal proteins, our data indicate that this experimental screening method can be extrapolated to human systems. In particular, we have identified new peptide sequences in human proteins with high affinity towards PCNA (Figure S1). Moreover, we can predict that some of the 233 positively charged peptides (pI > 8.0) of human proteins carrying the modified PIP motif Q-X(2)-[LIMFV]-X(2)-[DHFYW]-[FYAQ] are likely to correspond to additional candidate PCNA interaction partners (Figure 4C). Our results also indicate that the PIP-motif sustains a considerable amount of substitutions, indicating that PCNA binding of the PIP-peptides is mediated through beta zipper formation at the interaction surface [19] and the positively charged residues (Figures 1 and 4). As this interaction is largely independent of the specific sequence, PCNA complexes are likely to persist even when the PIP-motif cannot be anymore recognized using computational approaches. Despite the power of genome comparison in the detection of linear interaction motifs being obvious [21; 22], proteomics approaches like the one described here are clearly needed to validate the observed predictions.

Materials and Methods

Bacterial Strains

Escherichia coli SURE [*e14– (McrA–) Δ(mcrCB-hsdSMR-mrr)171 endA1 supE44 thi-1 gyrA96*

relA1 lac recB recJ sbcC umuC::Tn5 (Kan^r) *uvrC* [F' *proAB lacI^{q}Z ΔM15 Tn10* (Tet^r)]]

(Stratagene) was used as a host strain for maintaining plasmids. For protein expression, *cis*-

repressed pQE-80L (Qiagen) derivatives were transformed into *E. coli* BL21 CodonPlus [*E.*

coli B F– *ompT hsdS*(r$_B$– m$_B$–) *dcm+* Tet^r *gal λ* (DE3) *endA* Hte [*argU ileY leuW* Cam^r]]

(Stratagene) containing extra copies for arginine, isoleucine and leucine tRNAs.

Expression and purification of recombinant proteins

Genes encoding *P. abyssi* PCNA (PAB1465), RNase HII (PAB0352), Fen1 (PAB1877)

and DNA ligase (PAB2002) were amplified using 30 cycles of PCR with CsCl gradient

purified chromosomal DNA as template (primer sequences will be made available upon

request). pQE-80L (Qiagen) expression vector containing Isopropyl-β-D-thiogalactopyranoside

(IPTG)-inducible bacteriophage T5 promoter was used to express all proteins. *P. abyssi* PCNA

was cloned using *BamH*I and *Sal*I restriction sites (all primer sequences are indicated in Table

2); all other proteins were cloned into *BamH*I and *Pst*I sites of the pQE-80L expression vector.

The resulting expression constructs encode full length proteins with a six-histidine tag at their

amino-termini. A construct encoding truncated version of RNase HII lacking the 17 carboxy-

terminal residues and carrying PCNA interaction motif was produced in a similar manner.

Expression constructs were transformed into *E. coli* BL21-CodonPlus-RIL strain (Stratagene)

using solid LB medium containing 100 µg/mL ampicillin. Protein expression was induced in

exponential phase cultures by adding 0.5 mM IPTG. After 2 h of induction, the cells were

collected by centrifugation. Cell lysis was performed in buffer A (30 mM Hepes pH 8, 300 mM

NaCl) with several freeze-thaw cycles, followed by brief sonication to decrease viscosity of supernatants. Cellular debris was eliminated by centrifugation. All chromatographic procedures were performed with an ÄKTA FPLC system at 10 °C (Amersham Biosciences). Tagged proteins were purified on immobilized Ni-NTA beads, according to the manufacturer's instructions (Qiagen), followed by purification on an S-200 gel filtration column (Amersham). All protein samples were analyzed for purity and integrity using 13% SDS-polyacrylamide gel electrophoresis and by MALDI-TOF analyses (Innova Proteomics, France). Recombinant human wildtype PCNA as well as its QLGI125 mutant were purified to over 99% homogeneity as previously described [7; 23].

Synthesis of peptides and Pepscan screening

The overlapping synthetic peptides were synthesized and screened using credit-card format mini-PEPSCAN cards (455-well-plate with 3 µl wells) as described previously [24]. The binding of antibodies to each peptide was tested in a PEPSCAN-based enzyme-linked immuno assay (ELISA). The 455-well creditcard-format polypropylene cards, containing the covalently linked peptides, were incubated with sample, for example 1 µg/ml PCNA diluted in a PBS solution which contains 5% horse-serum (v/v) and 5% ovalbumin (w/v) and 1% Tween 80 (4°C, overnight). After washing, the peptides were incubated with an anti-HIS tag antibody (dilution 1-1000, Novagen) (1 h, 25°C) and subsequently with a rabbit anti-mouse antibody peroxidase (dilution 1/1000, Dako) (1 h, 25°C). After washing the peroxidase substrate, 2,2'-azino-di-3-ethylbenzthiazoline sulfonate (ABTS) and 2 µl/ml 3% H_2O_2 were added. After 1 h color development of the ELISA was quantified with a CCD-camera and an image processing system.

Interaction studies

Surface plasmon resonance studies were performed in a BIAcore X apparatus (BIAcore, Uppsala Sweden), the interactions were monitored at 25 °C at a flow rate of 30 µl/min in HBS-

P buffer (10 mM Hepes, pH 7.4; 150 mM NaCl and 0.005% (v/v) P20). PCNA was amine-coupled on the surface of CM5 sensor chips (BIAcore) in 10 mM sodium acetate buffer, pH 4.5. The level of immobilisation of PCNA was of about 150 resonance units (RU).

P. abyssi GE5 (strain Orsay) cell free extracts were prepared from anaerobically grown 2 liter cultures using a gas-lift bioreactor, which was prepared as previously described [25]. Cell suspensions were harvested in exponential phase (approximately 2.0 x10^8 cells/ml) and centrifuged at 5000 x g for 45 min. Cells were suspended in 50 mM Tris-HCl, pH 8; 0.5 mM dithiothreitol, 150 mM NaCl and lysed by sonication using a Vibracell sonifier (15 pulses of 30 sec, 375 W, 40% amplitude). After centrifugation at 10 000 x g for 15 min, the supernatant was used as the cell free extract. 20 μg of His-tagged recombinant proteins were incubated with 1.6 mg of Co^{2+} magnetic micro beads (Dynabeads® TALON™, Dynal Biotech) in 1 ml of binding and washing buffer (50 mM Na-phosphate, pH 8; 300 mM NaCl and 0.01% Tween 20) for 20 min at 25 °C. Proteins bound to the beads were washed three times in the same buffer. The beads were then incubated with 12.5 mg of whole-cell extract for 2 h at 4°C. After five washing steps using the above mentioned buffer, proteins bound to the target proteins were eluted in 1 M Tris, pH 8.8; 2% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 8% (v/v) sucrose and 0.1% (w/v) bromophenol blue (SDS sample buffer). The eluted samples were separated in an SDS 12% polyacrylamide gel and transferred to a nitrocellulose membrane. After membranes were blocked with 5% milk in TBST buffer (10 mM Tris-HCl, pH 7.5; 150 mM NaCl and 0.05% (v/v) Tween 20), polyclonal anti-PCNA antibody (raised in rabbits by Eurogentec) was incubated for 1 h at 25°C and then washed in TBST. Immunocomplexes were detected using an enhanced chemiluminescence procedure according to the manufacturer's recommendations (Amersham).

RNase HII assays

Activity assays using RNase HII were performed essentially as described [26]. A substrate consisting of oligonucleotides O8 and O9 (Table 1) was obtained by mixing a deoxyoligonucleotide carrying an embedded ribonucleotide in a 1:1 molar ratio with the complementary strand in 20 mM Tris-HCl, pH 8,0 and 150 mM NaCl, followed by an incubation at 75°C for 15 min. After this incubation step, oligonucleotides were cooled slowly at room temperature and conserved at –20°C until use.

Where indicated, one of oligonucleotides used in the hybridizations was labelled either on the 5'-end with [γ−32P] ATP by using T4 Polynucleotide kinase (Promega) or, alternatively, the 3'-end was marked using the Terminal Desoxynucleotide Transferase (Promega) and [α−32P] ddATP (Amersham). Free label was removed from oligonucleotide solutions using Sephacryl G25 MicroSpin columns (Amersham). Reaction products and substrates were separated under denaturing conditions. Polyacrylamide gels were dried on Whatman filter paper and reactions were quantified using Storm system.

Gene deletion in *Haloferax volcanii*

Gene deletion in *H. volcanii* strain H53 (pHV2⁻ Δ*pyrE2* Δ*trpA*) and segregation of mutant alleles were carried out using the optimized pop-in/pop-out system [27; 28]. A non-selective enriched medium Hv-YPC or a selective Hv-Ca medium were used for cell growth. Where appropriate, selective medium was supplemented with uracil and tryptophane (50µg/ml each) or, alternatively, using a combination of 5-fluoroorotic acid (50µg/ml) and uracil (10µg/ml) with or without tryptophane (50µg/ml). Plasmids used for *H. volcanii* gene inactivation carried approximately 500 bp 5' upstream and 3' downstream the ORF sequences (oligonucleotides used are indicated in Table 1). PCR analyses were used to confirm that the constructed strains corresponded to the expected deletion strains (Figure 3A-C). Flanking regions were amplified using PCR and cloned into pTA131 containing a *pyrE2* selection marker using *Eco*RI and *Xba*I

restriction sites, yielding plasmids pCN7 (carrying the Δ*rnhB* allele) and pCN6 (Δ*fen1*). Non-completed genome sequence of *H. volcanii* was used for primer design (www.tigr.org). *trpA* selection marker from pTA298 [27] was inserted at the *Bam*HI site of pHM306 (Δ*pcna::trpA*), pYGH40 (Δ*rnhA::trpA*) and pCN9 (Δ*rnhB::trpA*) that was created between two inserted fragments. The non-replicative plasmids obtained were reisolated from *E. coli* SCS110 [*dam*⁻/*dcm*⁻ (Stratagene)] and non-methylated DNA was used to transform *H. volcanii* using polyethylene glycol 600 as described [29].

Computational methods

Scan prosite (http://www.expasy.org/tools/scanprosite/) with variations of PIP-motifs was used to scan *P. abyssi* or *Homo sapiens* protein sequence(s) from Swiss-Prot and TrEMBL databases. For human proteins, description filter "DNA" was used to limit the number of the hits. The sequence manipulation suite hosted at http://www.bioinformatics.vg/sms/protein_iep.html was used to predict isoelectric points for the different peptides. Statistical tests were performed using OriginPro 7.5 software package (OriginLab Corporation).

References

1.   Miyachi, K., Fritzler, M. J. & Tan, E. M. (1978). Autoantibody to a nuclear antigen in proliferating cells. *J Immunol* **121**, 2228-34.
2.   Prelich, G., Tan, C. K., Kostura, M., Mathews, M. B., So, A. G., Downey, K. M. & Stillman, B. (1987). Functional identity of proliferating cell nuclear antigen and a DNA polymerase-delta auxiliary protein. *Nature* **326**, 517-20.
3.   Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M. & Kuriyan, J. (1996). Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell* **87**, 297-306.
4.   Maga, G. & Hubscher, U. (2003). Proliferating cell nuclear antigen (PCNA): a dancer with many partners. *J Cell Sci* **116**, 3051-60.
5.   Duggin, I. G. & Bell, S. D. (2006). The chromosome replication machinery of the archaeon Sulfolobus solfataricus. *J Biol Chem* **281**, 15029-32.
6.   Warbrick, E. (1998). PCNA binding through a conserved motif. *Bioessays* **20**, 195-9.
7.   Jonsson, Z. O., Hindges, R. & Hubscher, U. (1998). Regulation of DNA replication and repair proteins through interaction with the front side of proliferating cell nuclear antigen. *Embo J* **17**, 2412-25.
8.   Xu, H., Zhang, P., Liu, L. & Lee, M. Y. (2001). A novel PCNA-binding motif identified by the panning of a random peptide display library. *Biochemistry* **40**, 4512-20.
9.   Warbrick, E. (2006). A functional analysis of PCNA-binding peptides derived from protein sequence, interaction screening and rational design. *Oncogene* **25**, 2850-9.
10.   Kiyonari, S., Takayama, K., Nishida, H. & Ishino, Y. (2006). Identification of a novel binding motif in pyrococcus furiosus DNA ligase for the functional interaction with proliferating cell nuclear antigen. *J Biol Chem*.
11.   Haruki, M., Tsunaka, Y., Morikawa, M. & Kanaya, S. (2002). Cleavage of a DNA-RNA-DNA/DNA chimeric substrate containing a single ribonucleotide at the DNA-RNA junction with prokaryotic RNases HII. *FEBS Lett* **531**, 204-8.
12.   Rydberg, B. & Game, J. (2002). Excision of misincorporated ribonucleotides in DNA by RNase H (type 2) and FEN-1 in cell-free extracts. *Proc Natl Acad Sci U S A* **99**, 16654-9.
13.   Liu, Y., Kao, H.-I. & Bambara, R. A. (2004). FLAP ENDONUCLEASE 1: A Central Component of DNA Metabolism. *Annual Review of Biochemistry* **73**, 589-615.
14.   Sandler, S. J. (2005). Requirements for replication restart proteins during constitutive stable DNA replication in Escherichia coli K-12. *Genetics* **169**, 1799-806.
15.   Yu, K., Roy, D., Huang, F. T. & Lieber, M. R. (2006). Detection and structural analysis of R-loops. *Methods Enzymol* **409**, 316-29.
16.   Rossi, M. L., Purohit, V., Brandt, P. D. & Bambara, R. A. (2006). Lagging strand replication proteins in genome stability and DNA repair. *Chem Rev* **106**, 453-73.
17.   Kelman, Z. & White, M. F. (2005). Archaeal DNA replication and repair. *Curr Opin Microbiol* **8**, 669-76.
18.   Bruning, J. B. & Shamoo, Y. (2004). Structural and thermodynamic analysis of human PCNA with peptides derived from DNA polymerase-delta p66 subunit and flap endonuclease-1. *Structure* **12**, 2209-19.
19.   Chapados, B. R., Hosfield, D. J., Han, S., Qiu, J., Yelent, B., Shen, B. & Tainer, J. A. (2004). Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. *Cell* **116**, 39-50.
20.   Vijayakumar, S., Chapados, B. R., Schmidt, K. H., Kolodner, R. D., Tainer, J. A. & Tomkinson, A. E. (2007). The C-terminal domain of yeast PCNA is required for

physical and functional interactions with Cdc9 DNA ligase. *Nucleic Acids Res* **35**, 1624-37.

21. Neduva, V. & Russell, R. B. (2005). Linear motifs: evolutionary interaction switches. *FEBS Lett* **579**, 3342-5.

22. Neduva, V. & Russell, R. B. (2006). Peptides mediating interaction networks: new leads at last. *Curr Opin Biotechnol* **17**, 465-71.

23. Schurtenberger, P., Egelhaaf, S. U., Hindges, R., Maga, G., Jonsson, Z. O., May, R. P., Glatter, O. & Hubscher, U. (1998). The solution structure of functionally active human proliferating cell nuclear antigen determined by small-angle neutron scattering. *J Mol Biol* **275**, 123-32.

24. Slootstra, J. W., Puijk, W. C., Ligtvoet, G. J., Langeveld, J. P. & Meloen, R. H. (1996). Structural aspects of antibody-antigen interaction revealed through small random peptide libraries. *Mol Divers* **1**, 87-96.

25. Sharp, R. J. & Raven, N. D. H. (1997). Isolation and growth of hyperthermophiles. In *Applied Microbial Physiology: A practical approach* (Rhodes, P. M. & Stanbury, P. F., eds.), pp. 23-51. IRL Press, Oxford University press.

26. Chapados, B. R., Chai, Q., Hosfield, D. J., Qiu, J., Shen, B. & Tainer, J. A. (2001). Structural biochemistry of a type 2 RNase H: RNA primer recognition and removal during DNA replication. *J Mol Biol* **307**, 541-56.

27. Allers, T., Ngo, H.-P., Mevarech, M. & Lloyd, R. G. (2004). Development of Additional Selectable Markers for the Halophilic Archaeon Haloferax volcanii Based on the leuB and trpA Genes. *Appl. Environ. Microbiol.* **70**, 943-953.

28. Bitan-Banin, G., Ortenberg, R. & Mevarech, M. (2003). Development of a Gene Knockout System for the Halophilic Archaeon Haloferax volcanii by Use of the pyrE Gene. *J. Bacteriol.* **185**, 772-778.

29. Cline, S. W., Schalkwyk, L. C. & Doolittle, W. F. (1989). Transformation of the archaebacterium Halobacterium volcanii with genomic DNA. *J Bacteriol* **171**, 4987-91.

30. Myllykallio, H., Lopez, P., Lopez-Garcia, P., Heilig, R., Saurin, W., Zivanovic, Y., Philippe, H. & Forterre, P. (2000). Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. *Science* **288**, 2212-5.

Table 1: Oligonucleotides used in this study.

Usage		Oligonucleotides	Sequence 5' to 3'
Recombinant proteins:			
PCNA		PCNA BglII	ccatagatctATGCCATTCGAGATAGTC
		PCNA SalI	ccatgtccgagtcgacTTACTCCTCAACCCTGGG
Fen1		Fen1 BamHI	ccatggatccATGGGAGTTCCTATAGGGG
		Fen1 PstI	ccatctgcagTTAGGGTTTCTTTTTGATGAACCAGC
RnaseHII		Rnase HII BamHI	ccatggatccATGAAGGTTGCCGGAGCAGATG
		Rnase HII PstI	ccatctgcagTTAATTGCGAAACCGCTTAAG
Ligase		Lig1 BamHI	ccatggatccATGAGGTACATAGAGCTGGCC
		Lig1 PstI	ccatctgcagTTACACCTTTCCCTTCATCCTC
Gene deletion:			
pcna deletion		Hvo1 PCNA5	gatcgaattcGTCGCGGGCGTCGCCGCGCCA
		Hvo1 PCNA3	gatcggatccATATGATGTTCGCACGCTCTC
		Hvo2 PCNA5	gatcggtccGAGCCGTCCCACTCGTTTCTC
		Hvo2 PCNA3	gatctctagaGGGTCGAAGTCGAGCGGGCCT
fen1 deletion		OCN39	gctctagagCAGTTCGGCCATGGAGTTCG
		OCN40	cgggatcccgAGTGACCGGCGTCGATTTCG
		OCN41	cgggatcccGGCGCGCTCGCATCGAGACA
		OCN42	ggaattccAGCATTCCGTAGGTTGCCGAG
fen1 control	α	OCN72	GTCGTAGACGGGTGCGAAG
	β	OCN93	GCGCGAAGGAGTTGAACC
	γ	OCN144	GGTGTCGTTTGACGAGGTG
	δ	OCN145	GAGCTGTCGGAGCGTGTAG
	ε	OCN92	CGAACCAGCCCGTGTAAA
rnhA deletion		OCN138	ggaattcCCTCCTTCCGACA
		OCN141	gctctagACTCGAGCGACTCCGGTAA
		YGH19	gttttatccacgggatccAGTTGGACCTCGACAACCTC
		OCN140	acgcacataacggatccAGAACTCCCCGAAGACGTG
rnhA control	α	YGH32	GATTTCCTTGCGCTCGAC
	β	YGH33	CCGGTCTCCTCGTGAATC
	γ	OCN150	CGGGAAAGTCGTCGTACAG
	δ	OCN151	CCTCGTTGGCTAGGGAGTC
	ε	YGH34	AGCACGTCTTCGGGGAGT
rnhB deletion		OCN43	ccgctcgagCGACATGGTCGAGCGCGTG
		OCN44	cctcaagcttACCTGGTCGGGTCTTGGGGTG
		OCN45	gctcaagcttGGCGTCGGGGAACCGGCGG
		OCN46	cgcggatccCGGCCTCGACGATTCTCGAAG
rnhB control	α	OCN74	CGTGGTCGAAATCATGGAG
	β	OCN94	ACCGCAACCTCTAATTTTCC
	γ	OCN152	CGAGACGGACATGAACCTG
	δ	OCN153	ACGTACTCCCGCAGGAACT
	ε	OCN95	GTAATCGTCTCCGCGATTC
RNase HII Substrate (Bold A refers to a ribonucleotide):			
		O8	CCGTCCC**A**CCGTGCC
		O9	GGCACGGTGGGACGG

Table 2: Genotypes and phenotypes of various *Haloferax volcanii* strains.

Name	Genotype	Incorporation rate[1]	% of survival[2] Ultraviolet light (112 J/m^2)
H53	Wild type[3]	100%	0.25
CN7	Δ*rnhB*	130%	0.035
CN8	Δ*fen1*	70%	0.000125
CN9	Δ*fen1*/Δ*rnhB*	65%	0.0025
CN10	Δ*rnhA*	75%	0.15
CN11	Δ*rnhA*/Δ*rnhB*	108%	0.005
CN16	Δ*fen1*/Δ*rnhA*	68%	0.0003
CN17	Δ*fen1*/ Δ*rnhA*/Δ*rnhB*	33%	0.006

[1] Relative incorporation rates of tritiated uracil (marked at the positions 5 and 6 of pyrimidine ring) into alkaline resistant form (chromosomal DNA) was measured during early exponential growth phase using TCA precipitation as described 30. Myllykallio, H., Lopez, P., Lopez-Garcia, P., Heilig, R., Saurin, W., Zivanovic, Y., Philippe, H. & Forterre, P. (2000). Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon. *Science* **288**, 2212-5. 100 % corresponds to incorporation of 0.5 fmol of tritiated uracil into alkaline resistant form in one hour by the strain H53. Cell numbers were normalized using OD_{600} values prior all labelling experiments.

[2] Percentage of survival was measured using a 90 s exposure with a dose of 1.24 J/m^2/s. Survival curves for H53, CN8 and CN11 strains are shown in Figure 3.

[3] *Haloferax volcanii* strain H53 is referred to as "wild type" through out this work.

Figure legends:

Figure 1: **Screening of immobilized peptides with the robust archaeal *P. abyssi* replication clamp. (A)** *Screening of 910 random peptides with P. abyssi PCNA.* Binding is indicated in arbitrarily chosen fluorescence units. Peptides were sorted according to the observed binding values. Predicted pI values for 50 high "affinity" peptides were statistically higher than for 300 peptides with the lowest binding values. Synthetic peptides were synthesized and screened using credit-card format mini-PEPSCAN cards (455-well-plate with 3 µl wells) as described in the Materials and Methods section. **(B)** *Lead peptides isolated from a random library with high affinity for P. abyssi PCNA.* Hydrophobic residues in the vicinity of the conserved glutamine (in blue) are indicated in bold and positively charged residues are shown in red. **(C)** *Screening of candidate P. abyssi PCNA interacting peptides.* Sequences were identified using ScanProsite in the proteins encoded by the *P. abyssi* genome sequence. Relative binding values in arbitrary units are indicated on the right. The concentration of *P. abyssi* PCNA used for binding studies was 1 µg/ml. Several control negative control peptides with their binding values are shown at the bottom of Figure 1C. ND, not detected.

Figure 2: The interaction domains of *P. abyssi* RNase HII, Fen1 and DNA ligase with PCNA overlap (A) *The proteins used in this study:* 1 µg of each purified *P. abyssi* or human DNA replication proteins were separated on a 13% SDS-PAGE and stained by Coomassie blue. RNase HIIΔCt refers to the carboxy-terminal deletion derivative of *P. abyssi* RNase HII (see Material and methods). **(B-D)** *Experiments on PCNA chips.* Ten pmoles of RNase HII **(B),** Fen1 **(C)** or DNA ligase **(D)** were injected at time 0 over the PCNA chip (trace **1**; panels B, C and D). All data are indicated in RU units as a function of time. When indicated, 100 pmoles of non-interacting peptide control (trace 2) or a control peptide KRK**Q**LT**L**DN**F**LKR were included during injection (trace 3). As an additional control an interacting peptide alone over

the PCNA chip was injected (trace 4). The background resulting from buffer injection was always subtracted. The insert in **B** shows that the C-terminal PIP-domain of the *P. abyssi* RNase HII is necessary and sufficient for binding to the PCNA chip. + Cter, wild type *P. abyssi* RNase HII; ΔCter, *P. abyssi* RNase HII lacking 17 residues from the amino terminus. **(E)** *A pull down experiment on magnetic beads.* 20 μg of histidine-tagged Fen1, RNaseHII or DNA ligase (from left to right) bound to magnetic beads, and a negative control that consists of beads alone, were incubated with 12.5 mg of *P. abyssi* cellular extract. Control lanes contained 7 μg of cell free extract and 15 ng of PCNA. Pull-down assay, SDS-polyacrylamide gel electrophoresis and immunoblot using anti-PCNA antibodies were performed as described in Material and Methods. **(F)** *RNase H activities in dependence of PCNA.* Activities of RNase HII and RNase HIIΔCter (1 pmol) were tested on a DNA substrate radioactively labelled at 5'-terminus containing a single embedded ribonucleotide (1 pmol *per* 20 μl reaction) in the absence and the presence of *P. abyssi* PCNA.

Figure 3: Deletion of *fen1*, *rnhA* and *rnhB* genes in *Haloferax volcanii.* (A) *Schematics indicating the annealing positions of the primers (arrows) used to confirm the deletions for each gene.* Homology sequences used for the pop-in/pop-out constructs are shown by shaded boxes, the strait line corresponds to the chromosome and the undulated line to the remains of plasmidic DNA after the pop-out. **(B)** *The expected sizes of the PCR products obtained with the different primer sets are shown for each deletion.* Table also indicated genotypes for the strains used in this work **(C):** *Agarose gel electrophoresis confirming gene deletions.* The PCR products obtained using either the wild type H53 or the triple mutant CN17 DNA. Each product presents the expected size indicating the loss of *fen1*, *rnhA* and *rnhB* genes in the strain CN17. Similar results were obtained for the other strains indicated in Table 2. **(D)** *Fraction of surviving H. volcanii cells on solid enriched media as a function of UV exposure is shown.* One

second exposure corresponds to 1.24 J/m^2. The strains used were H53 (wild type), CN11 (ΔrnhAΔrnhB), and CN8 (Δfen1).

Figure 4: **Predicted interaction partners of human PCNA (A)** *Screening of 98 naturally occurring peptides present in human proteins annotated in Swiss-Prot database containing the sequence motif x(3)-Q-x(4,5)-[HDF]-[YF]-x(3).* The screening was performed with human PCNA (1 μg/ml). The average pI value for 25 top binders was 8.40 (n=25) and for the non-binders 5.87 (n=25). Two sample independent t-test revealed that the difference of population means was significantly different (p-value=0.00113) **(B)** *Positional scan of the underlined residues in the peptide KRKQLTLDNFLKK.* These residues correspond to the conserved residues of the archaeal and eukaryal PIP-motif. Green columns correspond to "wild type" peptides whereas peptides carrying substitution are indicated in blue. Binding values where normalized using the values obtained for "wild type" peptide. **(C)** *Frequency distribution of pI values for 779 human peptides with the predicted PIP-motif.* Note that the data shown in Figures 1C and 3A already provide experimental support suggesting that some positively charged peptides in this data set could be used to identify novel interaction partners for the human PCNA.

Figure

A.

Average pI

$\bar{x} = 8.46 \pm 0.55$
$n = 50$

$\bar{x} = 6.31 \pm 1.90$
$n = 300$

Binding (arbitrary units)

910 random peptides
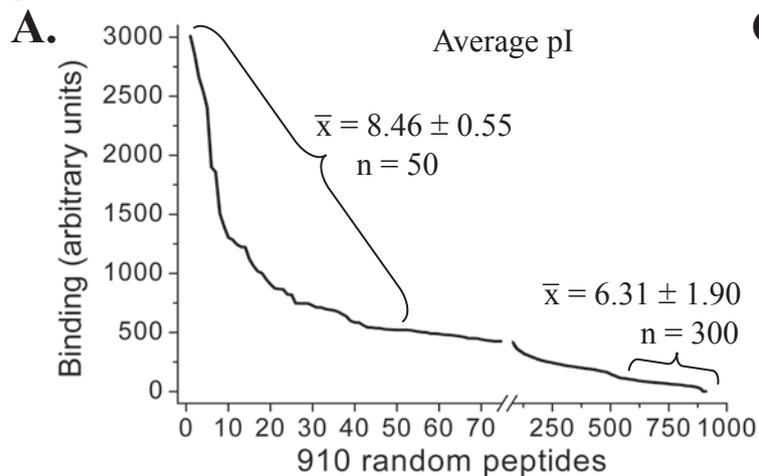
B.

LEAD PEPTIDES

FCQFKYMKMSCA
QCKFFKDIGKCD
YCNGGQTKFTCR
VCKVIQDGKGCF
QCQFQMPTRWCY
KCMQRHAWVKCE
YCQTWNSDRTCK
QCKTYEPLQRCG
RCQYKENIRPCL
RCQGTSSPGNCR
QCSPKFVQASCH
NCVLDKKFKDCL

C.

PROTEIN	PEPTIDE	VALUE
Known interactions		
RF-C	KGKQVTLFD**FI**KK	2608
Fen1	AGKQSTLES**WF**IK	1260
PolB	KTKQVGLGA**WL**KF	384
DNA ligase	QLYQKLEKT**TM**KL	151
Predicted PCNA interacting motifs		
RnaseHII	KRAQLTLDN**FL**KR	2578
PAB2263	RSKQKTLDF**FT**P–	1884
Negative control peptides		
	QLNQITIMS**FL**GV	39
	RRGQISIEF**LF**LI	13
	LKAQISIDF**LI**AL	ND
	NFFVTTIIL**LF**IL	16
	RSLVASLII**FI**LI	15
	SLGVRSLAL**FL**RI	11
	TTGVKGLDE**LL**GG	ND
	NKRVLGIIS**LL**VL	ND
	TTGVKGLDE**LL**GG	ND
	KEVKEEYKR**FL**EE	ND

Figure

Figure

A

Wild type

α β γ δ ε
gene of interest

Deletion mutant

α β ε
*trp*A

or

α β ε
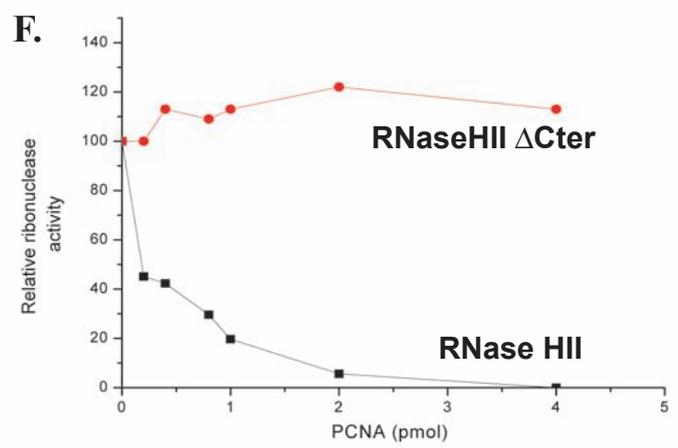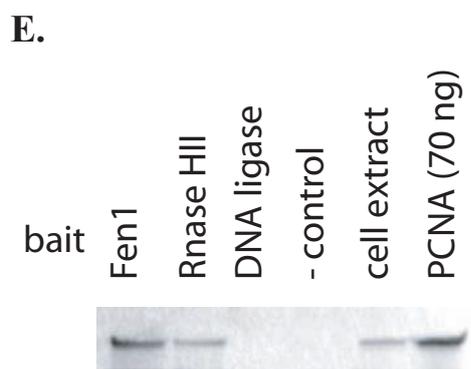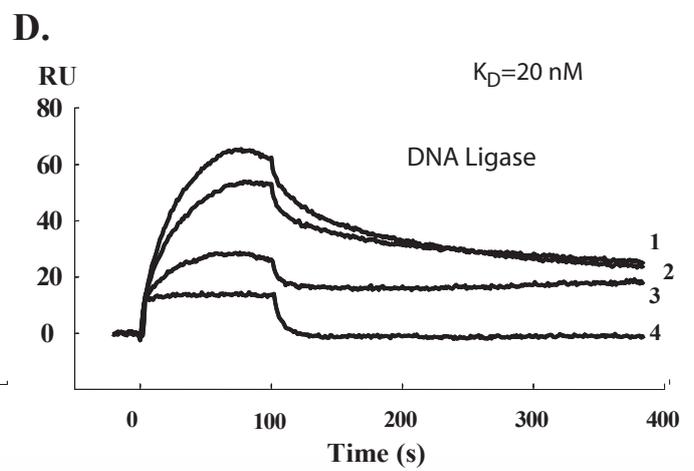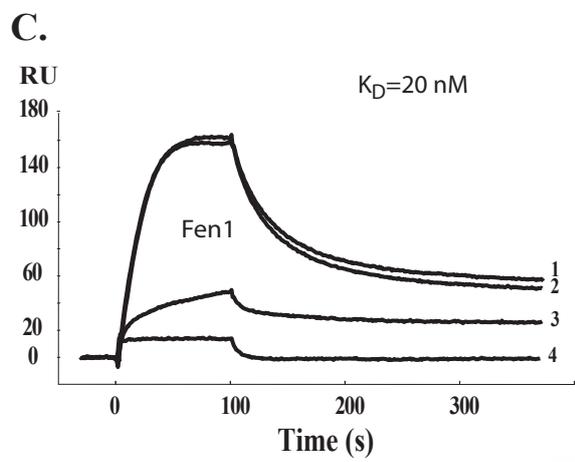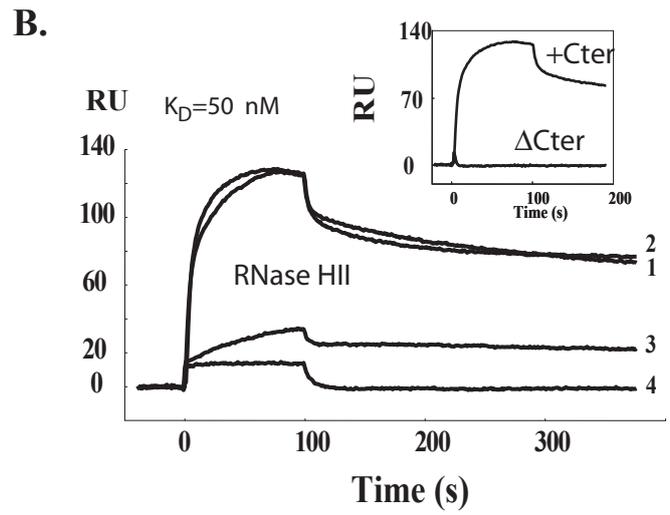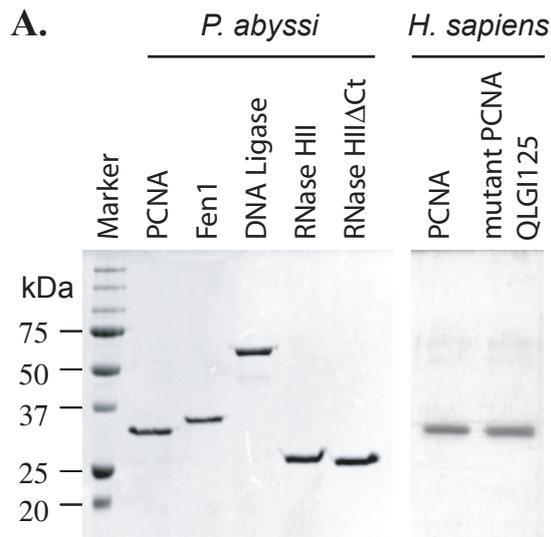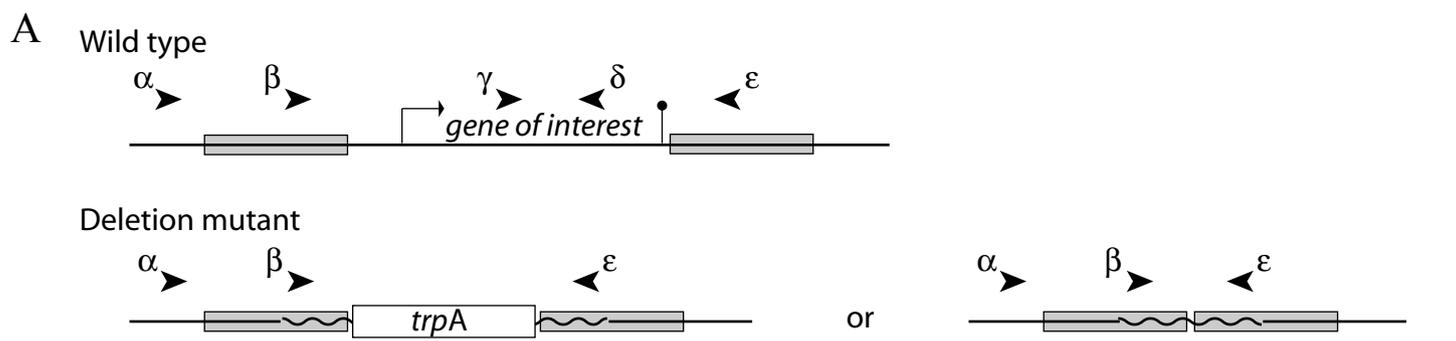
B

	Strain		**H53**	**CN7**	**CN8**	**CN9**	**CNY10**	**CNY11**	**CN16**	**CN17**
	Genotype		pHV2-Δ*pyrE2* Δ*trpA*	H53 Δ*rnhB*	H53 Δ*fen1*	H53 Δ*fen1* Δ*rnhB* ::*trpA*	H53 Δ*rnhA* ::*trpA*	H53 Δ*rnhB* Δ*rnhA* ::*trpA*	H53 Δ*fen1* Δ*rnhA* ::*trpA*	H53 Δ*fen1* Δ*rnhB* Δ*rnhA* ::*trpA*
Gene	Oligonucleotides	position								
fen1	OCN72 - OCN92	α − ε	1792		782	782			782	782
	OCN93 - OCN92	β − ε	1114		104	104			104	104
	OCN144-OCN145	γ − δ	529		o	o			o	o
rnhA	YGH32 - YGH34	α − ε	1387				1555	1555	1555	1555
	YGH33 - YGH34	β − ε	900				1068	1068	1068	1068
	OCN150-OCN151	γ − δ	443				o	o	o	o
rnhB	OCN74 - OCN95	α − ε	1335	693		1658		693		693
	OCN94 - OCN95	β − ε	758	116		1081		116		116
	OCN152-OCN153	γ − δ	333	o		o		o		o

C

fen1   *rnhA*   *rnhB*   locus
α − ε   β − ε   γ − δ   α − ε   β − ε   γ − δ   α − ε   β − ε   γ − δ   primers
H53 CN17 H53 CN17 H53 CN17   H53 CN17 H53 CN17 H53 CN17   H53 CN17 H53 CN17 H53 CN17   DNA

- 10 kb
- 3 kb
- 2 kb
- 1.5 kb
- 1 kb
- 750 b
- 500 b
- 250 b
- 0 b

D

% of survival vs UV dose (seconds)

H53 (WT) t_d=131 min

CN11 (Δ*rnhA*Δ*rnhB*) t_d=108 min

t_d=189 min

CN8 (Δ*fen1*)
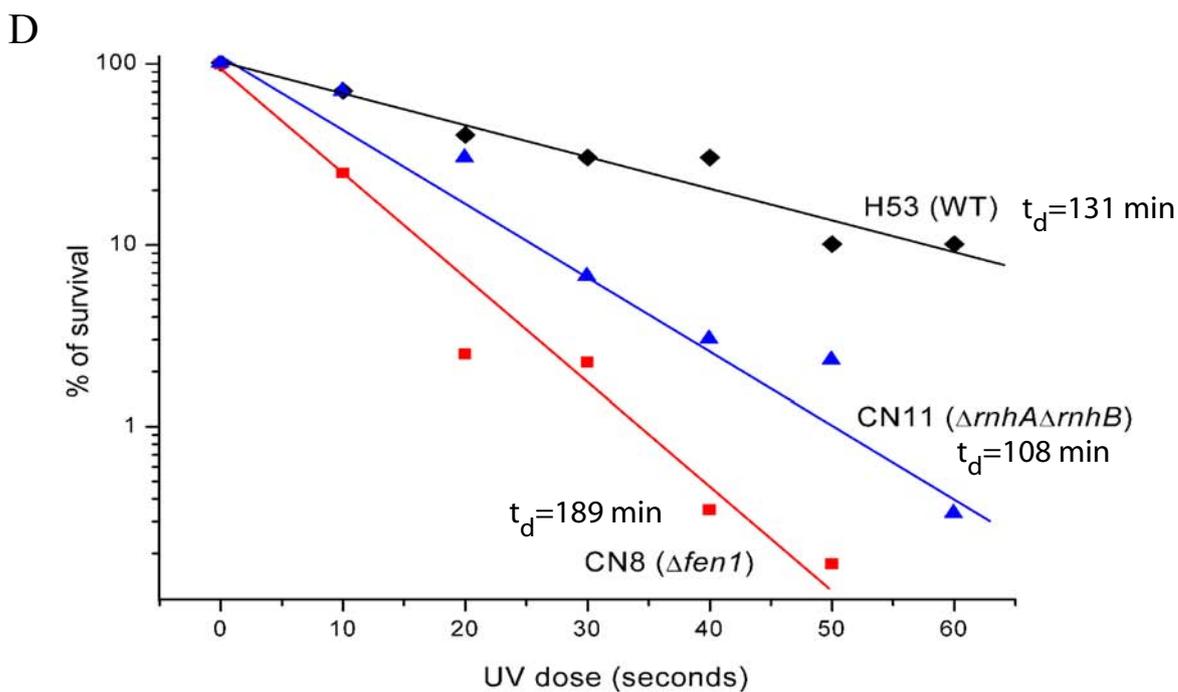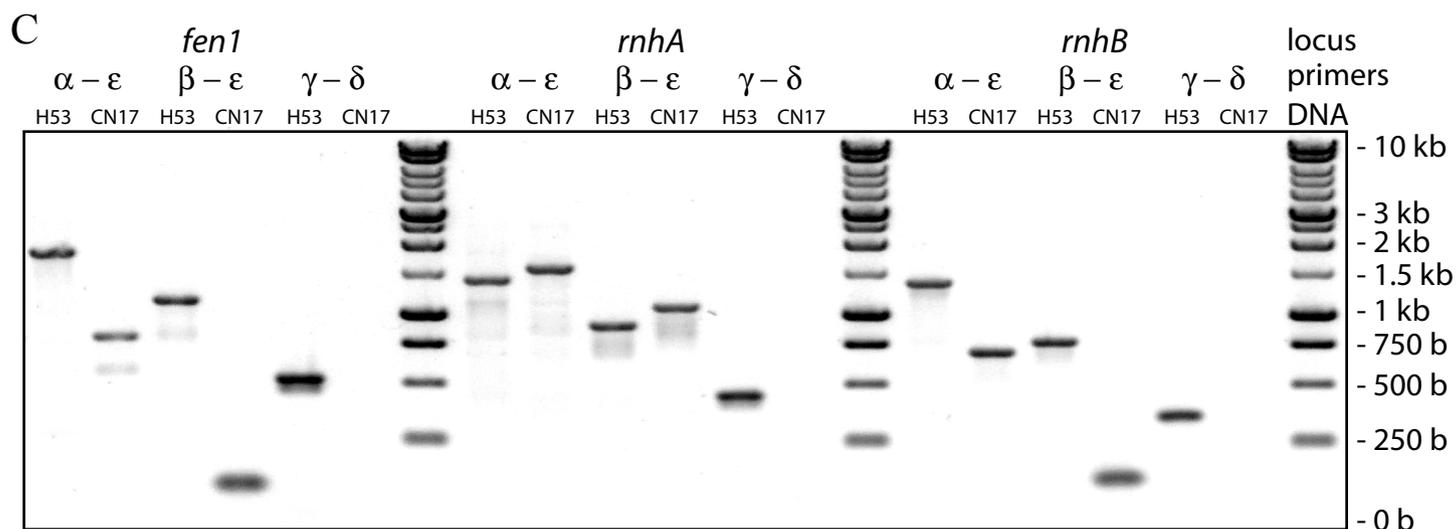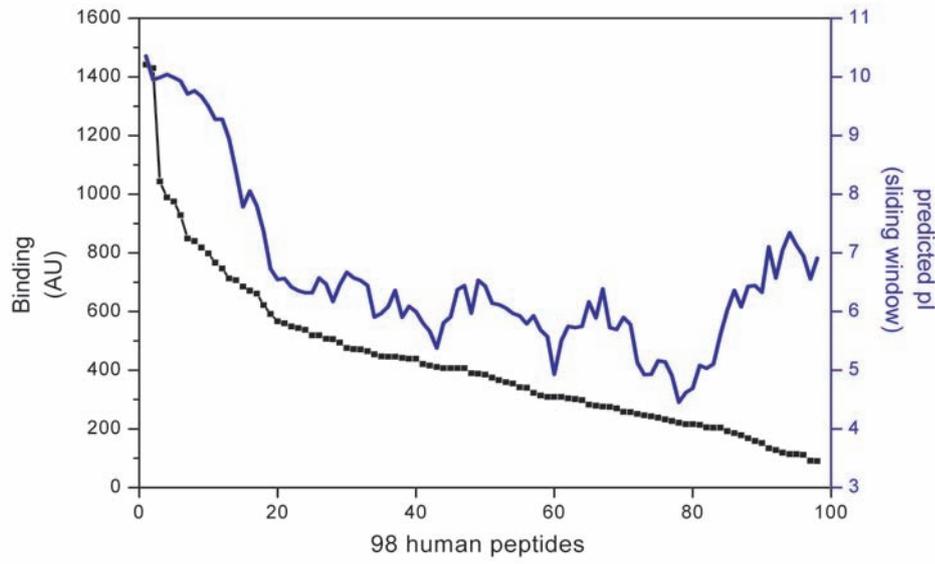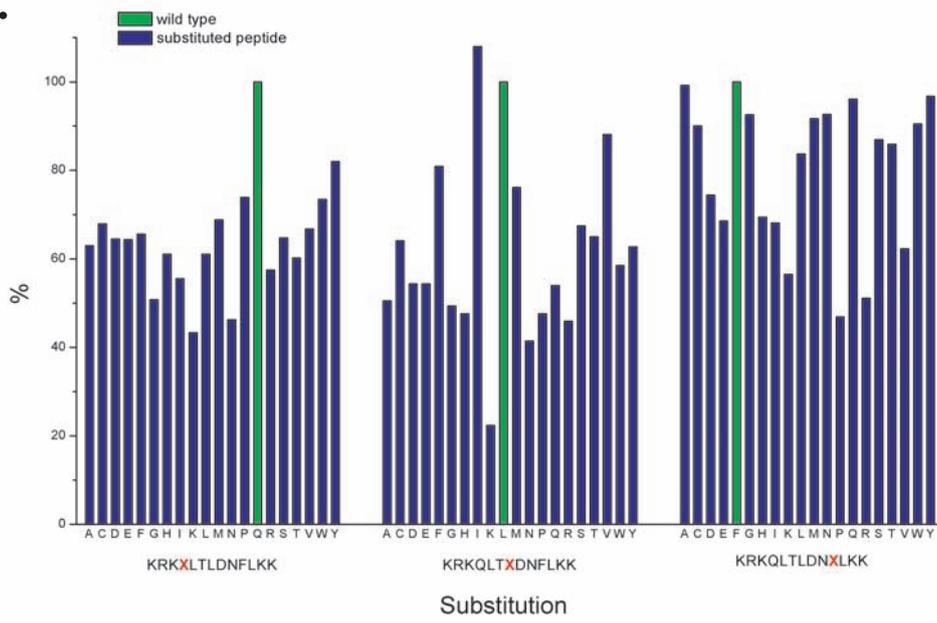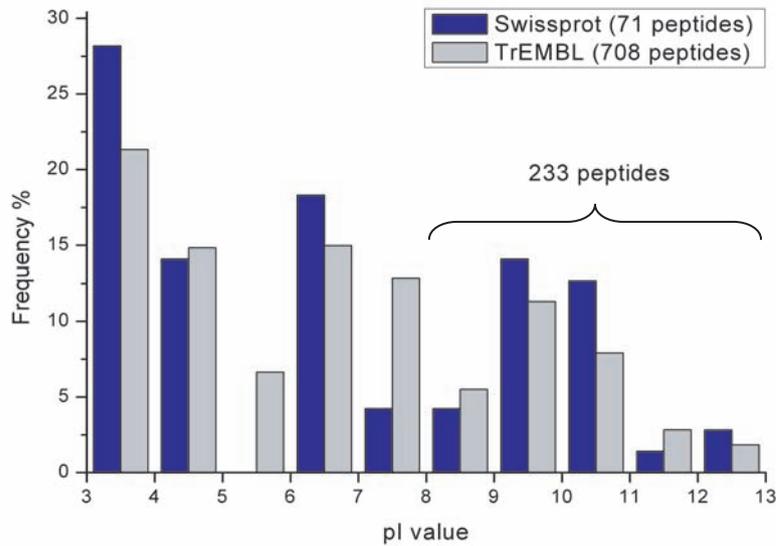
**Figure**

Table S1

Random peptides with the binding values higher than 500. The alignment of the peptides shown in the bold type is shown in Figure 1 B.

Peptide sequence	Binding value
FCQFKYMKMSCA	3008
VCNRVGPIRKCD	2854
QCKFFKDIGKCD	2660
KCMKQSQNLQCW	2549
RCWQKMMGVRCD	2404
YCNGGQTKFTCR	1900
VCKVIQDGKGCF	1858
SCPGATKHLSCY	1507
QCQFQMPTRWCY	1397
YCQTWNSDRTCK	1305
QCKTYEPLQRCG	1285
MCKEVRISKDCA	1246
FCTKMGSIQACA	1225
RCQYKENIRPCL	1223
NCVLDKKFKDCL	1122
LCGAKKKTESCN	1062
MCYKSNGGRMCI	1018
RCQGTSSPGNCR	1002
SCRMLQHLMACY	950
QCSPKFVQASCH	908
QCRFMAQLRSCV	874
RCARKTAEAMCT	868
KCMQRHAWVKCE	865
LCWKNKILEMCS	823
ICTKKMHIMLCD	819
KCSNSLNQFDCL	747
WCQTMWDKQRCV	747
VCLRDITKMACM	746
RCDKQTAPWTCM	746
TCNQYQRILRCG	727
WCKSINNMRICA	712
WCTPPNKFTHCL	712
RCIRWKANSDCM	699
SCHKWENVIRCR	693
LCRAAQAIPRCV	687
GCASLRHGGQCS	676
RCPPSSFRMSCV	656
KCPWKAEHRQCV	637
DCMLGLAYRPCK	598
KCFHKEFVGYCR	585
RCDMGTRAHTCH	584
QCPAMMFYPNCK	561
QCQVHKRYTGCG	543
YCLQSALNIKCS	542
PCKFYREKHICD	537
KCRGTMDHDFCK	536

Figure S1A

The list of peptides with the motif Q-x(4,5)-[HFD][FY] showing top a binding signal higher than 400 is shown. Known interaction partners of the human PCNA are indicated in the bold.

```
DNMT1_HUMAN    STRQTTITSHFAKG  ⎫ > 1000 arbitrary units
MSH6_HUMAN     MSRQSTLYSFFPKS  ⎭ Q-x(4,5)-[HF][F]
TOP2B_HUMAN    SKCQLS-EKFFKAA
GCF_HUMAN      SSNQALNCKFYKSM
CHD1_HUMAN     QAKQLQTRADYLIK  ⎫
BLM_HUMAN      TTNQQRVKDFFKNA  ⎬ > 800 arbitrary units
DNJA1_HUMAN    DPNQER-RRHYNGE  ⎪ Q-x(4,5)-[HFD][FY]
CHD2_HUMAN     QGKQLQTRADYLLK  ⎪
MUTYH_HUMAN    RMGQQVLDNFFRSH  ⎭
RA54B_HUMAN    SMSQLKQWKHFSGD  ⎫
TDT_HUMAN      GKHQLVVRRDYSDS  ⎪
PRKDC_HUMAN    LYFQvLKSKDFVQV  ⎪
RECQ4_HUMAN    RSRQGSVP-DYGQR  ⎪
DNL3_HUMAN     MAEQRFCV-DYAKR  ⎬ > 600 arbitrary units
DPOL_HUMAN     TTQQAIGLKHYSDF  ⎪ Q-x(4,5)-[HFD][FY]
DPOA_HUMAN     DPTQFR-VHHYHKD  ⎪
DNS2B_HUMAN    QSAQGQKFLHFAKS  ⎪
TATD2_HUMAN    KEFQGC-ISDFCDP  ⎭
DPOE1_HUMAN    TICQRE-NSFYVDT  ⎫
GRLF1_HUMAN    CTDQLGLEQDFEQK  ⎪
DPOD3_HUMAN    ANRQVSITGFFQRK  ⎪
DPOM_HUMAN     TLSQLQGLPHFGEH  ⎪
DNSL2_HUMAN    LKPQSATVHDFQEE  ⎪
RFA1_HUMAN     NLGQGDKPDYFSS-  ⎪
DPOQ_HUMAN     PNIQNV-PRDFEIK  ⎪
PRKDC_HUMAN    GSSQGS-VIHFSHG  ⎪
FBX18_HUMAN    QGSQRCIPEFFLAG  ⎪
MSH6_HUMAN     KEAQKG-GHFYSAK  ⎪
DHX9_HUMAN     QVPQFI-LDDFIQN  ⎪
KBF2_HUMAN     AVPQLLHMPDFEGL  ⎪
ZN148_HUMAN    EVLQTL-LDHYSHK  ⎪
RECQ5_HUMAN    KDSQSI-ARFFCRR  ⎬ > 400 arbitrary units
CHD4_HUMAN     QDDQSDNQSDYSVA  ⎪
DNJC1_HUMAN    HLIQDA-GQFYAKY  ⎪
PRKDC_HUMAN    NKLQYFMEQFYGII  ⎪
ERCC5_HUMAN    NQKQALQEEFFHNP  ⎪
RA51B_HUMAN    SRHQILTCQDFLCL  ⎪
REV1_HUMAN     PNLQAV-PYDFHAY  ⎪
DNAS1_HUMAN    QLAQAI-SDHYPVE  ⎪
RECQ1_HUMAN    FLIQQYLKEDYSFT  ⎪
MSH3_HUMAN     HAFQAV-TEFYAKD  ⎪
DNM3A_HUMAN    SIKQGK-DQHFPVF  ⎪
DHX9_HUMAN     KDAQSNAARDFVNY  ⎪
DPOG1_HUMAN    WVVQSS-AVDYLHL  ⎪
RAD50_HUMAN    NYIQDG-KDDYKKQ  ⎪
TOP1M_HUMAN    DGCQHVVEFDFLGK  ⎪
TOP3B_HUMAN    LEKQTN-PPDYLTE  ⎭
```