

# Telomeric Protein–DNA Point Contacts Identified by Photo-Cross-Linking Using 5-Bromodeoxyuridine<sup>†</sup>

Brian J. Hicke,<sup>†</sup> Michael C. Willis,<sup>§</sup> Tad H. Koch,<sup>§</sup> and Thomas R. Cech<sup>\*†§</sup>

Department of Molecular, Cellular, and Developmental Biology, Howard Hughes Medical Institute, and Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215

Received October 12, 1993; Revised Manuscript Received December 30, 1993\*

**ABSTRACT:** The *Oxytricha* telomere protein specifically recognizes single-stranded telomeric DNA, forming an extremely salt resistant and kinetically stable nucleoprotein complex. The absence of information on how this heterodimeric protein binds to DNA prompted this photo-cross-linking study. Multiple protein–DNA photo-cross-links are formed upon UV irradiation of *Oxytricha* telomeres reconstituted with a synthetic oligonucleotide terminating in 5′-T<sub>16</sub>T<sub>15</sub>T<sub>14</sub>T<sub>13</sub>G<sub>12</sub>G<sub>11</sub>G<sub>10</sub>G<sub>9</sub>T<sub>8</sub>T<sub>7</sub>T<sub>6</sub>T<sub>5</sub>G<sub>4</sub>G<sub>3</sub>G<sub>2</sub>G<sub>1</sub>-3′. Site-specific substitution of certain nucleotides with 5-bromodeoxyuridine (BrdU) greatly increased the photo-cross-linking yield, each substitution favoring a specific protein–DNA cross-link. For example, substitution of BrdU for T<sub>7</sub> resulted in 25% cross-linking of the bound DNA, a 10-fold increase over the unsubstituted DNA. Both subunits of the telomere protein cross-link to, and are therefore near, the DNA. Three point contacts within this nucleoprotein complex, involving the  $\alpha$  subunit, were established using BrdU substitution: Tyr239, Tyr142, and His292 cross-link to G<sub>3</sub>, T<sub>15</sub>, and T<sub>7</sub>, respectively. One photo-cross-link, Tyr239–G<sub>3</sub>, occurs amid a short acidic stretch of the  $\alpha$  subunit, counter to expectations for amino acids that approach the polyanionic DNA. The two remaining cross-links are to amino acids in hydrophobic regions of the primary polypeptide sequence, consistent with the hypothesis that hydrophobic interactions account for the salt resistance (>2 M NaCl) of this protein–DNA complex. These two photo-cross-links suggest that the telomere protein may bind telomeric single-stranded DNA by intercalation of aromatic residues into a nucleotide lattice.

Telomeres, the physical ends of eukaryotic chromosomes, consist of a canonical repetitive DNA and associated protein(s). Telomeres confer stability to chromosomes by forming a protective cap that deters exonucleolytic activities and prevents the end-to-end fusion that occurs at free ends generated by chromosome breakage. However, the telomere is not inert, but must dynamically interact with trans-acting factors to permit additional activities of telomeres: complete replication of the chromosomal terminus and possible participation in nuclear architecture. These features of telomeres have been reviewed (Zakian, 1989; Blackburn, 1990, 1991).

Telomeric DNA sequences from a variety of eukaryotes are all similar, consisting of tandem repeats of a short G-rich sequence and its complementary C-rich strand (Blackburn, 1991). For example, vertebrate telomeric DNA consists of repeats of T<sub>2</sub>AG<sub>3</sub> (Moyzis et al., 1988), and the unicellular parasite *Plasmodium* has TT(C/T)AGGG (Vernick & McCutchan, 1988), while telomeres of the ciliates *Tetrahymena* and *Oxytricha* have TTGGGG and TTTTGGGG, respectively (Blackburn & Gall, 1978; Klobutcher et al., 1981).

The *Oxytricha* macronuclear genome is highly enriched in telomeric DNA (Prescott, 1983), which has the following structure:

```
.....NNNNNNNGGGGTTTGGGGTTTGGGGTTTGGGGTTTGGGG-3'
.....NNNNNNNCCCCAAACCCCAACCC
```

The single-stranded 3′ protrusion seen in the *Oxytricha* telomeres above has also been identified in several other ciliates and a slime mold (Henderson & Blackburn, 1989) and recently in the yeast *Saccharomyces* (Wellinger et al., 1993) and may therefore be a general feature of eukaryotic telomeres. A ribonucleoprotein enzyme, telomerase, catalyzes the synthesis of the G-rich DNA strand (Yu et al., 1990), helping to address a long-appreciated problem in the maintenance of a linear genome: complete replication of the chromosomal terminus (Watson, 1972).

Telomeres packaged as protein–DNA complexes have been identified in several organisms (Blackburn & Chiou, 1981; Cheung et al., 1981; Gottschling & Cech, 1984; Berman et al., 1986; Wright et al., 1992). There are two classes of telomeric proteins, those that bind to internal, double-stranded telomeric DNA and those specific for the terminus. Terminus-binding proteins have been isolated and characterized in the ciliates *Oxytricha* (Gottschling & Zakian, 1986; Price & Cech, 1987) and *Euplotes* (Price, 1990; Price et al., 1992), and a candidate terminus-binding factor has been identified in the *Xenopus* egg (Cardenas et al., 1993). The ciliate terminus-binding proteins protect single-stranded telomeric DNA against nucleases and chemical modification by dimethyl sulfate, and therefore they provide the protective capping function of telomeres. The conservation of telomeric DNA sequence and the single-stranded 3′ overhang indicate that studies of the *Oxytricha* and *Euplotes* proteins may be

<sup>†</sup> This work was supported by NIH Grant GM 28039 to T.R.C. and by a grant of the Council for Tobacco Research, U.S.A., to T.H.K. B.J.H. was supported by an NSF predoctoral fellowship. T.R.C. is an investigator of the Howard Hughes Medical Institute and an American Cancer Society Professor.

\* Author to whom correspondence should be addressed.

<sup>†</sup> Department of Molecular, Cellular, and Developmental Biology and Howard Hughes Medical Institute.

<sup>§</sup> Department of Chemistry and Biochemistry.

\* Abstract published in *Advance ACS Abstracts*, February 15, 1994.

<sup>†</sup> Abbreviations used: O4 and O8, DNA oligonucleotides (T<sub>4</sub>G<sub>4</sub>)<sub>4</sub> and (T<sub>4</sub>G<sub>4</sub>)<sub>8</sub>, respectively; O2T, DNA oligonucleotide comprising 24 nt of nontelomeric DNA followed by (T<sub>4</sub>G<sub>4</sub>)<sub>2</sub>; O2BU3, O2BU7, and O2BU15, oligonucleotides identical to O2T except for substitutions of BrdU at nucleotides 3, 7, and 15, respectively, from the 3′ end; BrdU, 5-bromodeoxyuridine; PCR, polymerase chain reaction.

generally applicable to eukaryotic telomeres.

The *Oxytricha* telomere protein is an  $\alpha/\beta$  (56 kDa/41 kDa) heterodimer which binds to the single-stranded sequence TTTTGGGGTTTGGGG. The resulting nucleoprotein complex endures very high concentrations of salt, >2 M NaCl (Gottschling & Zakian, 1986; Price & Cech, 1987), and dissociates extremely slowly *in vitro* (Raghuraman & Cech, 1989; Fang et al., 1993). This long half-life implies that the complex is capable of enduring several rounds of the cell cycle intact, unless a mechanism exists for actively dissociating the complex. In addition, the telomere protein promotes formation of higher order structures *in vitro* (Raghuraman & Cech, 1989; Fang et al., 1993), potentially explaining the self-association of telomeres observed in *Oxytricha* (Lipps et al., 1982; Prescott, 1983). This unusual protein-DNA complex has been reconstituted *in vitro*, first using protein purified from *Oxytricha* (Raghuraman & Cech, 1989) and more recently using telomere protein purified from *Escherichia coli* expressing telomere protein genes (Gray et al., 1991). The *in vitro*-reconstituted complexes produce methylation protection patterns very similar to that observed in living *Oxytricha* (Price & Cech, 1987).

Ciliate telomeres thus represent an unusual class of protein-DNA interactions in their sequence specificity for single-stranded DNA, kinetic stability, salt stability, and ability to form higher order complexes. Therefore, understanding how this nucleoprotein complex forms and is regulated is of significant interest. Examination of the deduced protein sequences (Hicke et al., 1990; Gray et al., 1991) revealed no known DNA-binding motifs; although the  $\beta$  subunit contains a region of similarity to a histone H1 (Hicke et al., 1990), this region is dispensable for specific binding to telomeric DNA (Fang et al., 1993). Since the sequences provide no information on how this protein might bind specifically to single-stranded DNA, we have developed a photo-cross-linking assay to identify nucleotide-amino acid contacts and thereby infer regions of the protein-DNA interface.

We have employed 5-bromodeoxyuridine (BrdU<sup>1</sup>) and 308-nm excitation (Gott et al., 1991) to achieve high-yield photo-cross-linking which permits the correlation of a cross-linked amino acid with a specific site of BrdU incorporation in the DNA. BrdU is an attractive photo-cross-linking reagent because it is an excellent T analog, representing the replacement of a methyl group (2.0-Å radius) with bromine (1.95-Å radius). This substitution causes the UV absorption maximum to shift from 258 nm to 280 nm, permitting irradiation at wavelengths greater than 300 nm where the risk of photo-damage to nucleic acids and proteins is diminished. Studies using model compounds in solution indicate that, at 308 nm, 5-bromouracil cross-links to the oxidizable aromatic amino acids Tyr, Trp, and His, as well as to cystine (Dietz & Koch, 1987, 1989). The majority of UV cross-linking experiments which have identified cross-linkable amino acids in nucleoprotein complexes utilize 254-nm excitation, where significant absorption by nucleic acids and proteins occurs (cf. Paradiso et al. (1979), Barbier et al. (1984), Merrill et al. (1984), Shamoo et al. (1988), Mirzabekov et al. (1989), Catalano et al. (1990), Allen et al. (1991), Katouzian-Safadi et al. (1991), and Blatter et al. (1992)). The use of BrdU and irradiation at >300 nm maximizes cross-linking yield and minimizes photodamage.

We present a method for isolating and sequencing peptides cross-linked to a nucleic acid of defined length. We identify four amino acid residues of the *Oxytricha* telomere protein that cross-link to telomeric DNA:  $\alpha$ Tyr142,  $\alpha$ Tyr239,  $\alpha$ His292, and  $\beta$ Tyr134. Three of these residues represent

regiospecific cross-links, permitting the physical correlation of these amino acids with their cognate cross-linking nucleotides within the telomere. These photo-cross-linking studies establish point contacts in the telomeric nucleoprotein complex and test models for the assembly of this complex.

## EXPERIMENTAL PROCEDURES

**DNAs.** Oligonucleotides were synthesized on an Applied Biosystems 380B synthesizer. The unmodified oligonucleotide, O2T, had the sequence 5'-AAGACGACATCGCTCAGC-CAGACATTTTGGGGTTTGGGG-3'. 5-BrdU phosphoramidites were obtained from Cruachem or Glen Research and used according to manufacturer's instructions. Deprotected oligonucleotides were dried under vacuum and separated in 12% polyacrylamide/7 M urea gels. The DNA was cut from the gel by visualizing with UV shadowing for the minimal time (<5 s) to avoid inactivation of the BrdU chromophore. The gel slice was crushed and the DNA eluted by slow shaking for 12–18 h at room temperature in 10 mM Tris, pH 7.5, 0.1 mM EDTA, and 250 mM NaCl. The sample was filtered by centrifugation through a paper disk (Quick-Sep column, Isolab, Akron, OH) and precipitated by addition of 2.5 vol of 100% ethanol and storage at 4 °C overnight.

**Telomere Protein Isolation.** *E. coli* BL21(DE3) containing plasmids bearing telomere protein genes were grown, harvested, and lysed according to Fang et al. (1993). The cleared supernatant from a 750-mL culture was loaded onto a 5-mL S-Sepharose Econo-Pak cartridge (BioRad), and the cartridge was washed in 20 mM MOPS pH 7.5, 0.1 mM EDTA, 200 mM NaCl ( $\alpha$ ) or 400 mM NaCl ( $\beta$ ). Protein was eluted by stepping [NaCl] to 300 mM ( $\alpha$ ) or 500 mM ( $\beta$ ), dialyzed against MNG 20/20/20 (20 mM MOPS pH 7.5, 20 mM NaCl, 20% glycerol) or spin dialyzed (Centricon 30 filter, Amicon) into MNG 20/20/20, and stored at -20 °C. Protein prepared in this fashion was analyzed by SDS-PAGE, and its purity and concentration were estimated by comparison to dilution series of protein purified as described (Gray et al., 1991).

**Analytical Irradiations.** Telomere protein subunits (0.5–3  $\mu$ M) were incubated with 5' <sup>32</sup>P- or <sup>33</sup>P-labeled telomeric DNAs (0.1–2  $\mu$ M) in 20 mM Tris (pH 7.5) and 20–300 mM NaCl for 10 min at 25 °C before irradiation. A Lambda Physik EMG-101 excimer laser was operated at 308 nm by charging with 60 mbar of xenon and 80 mbar of 5% hydrogen chloride in helium and filled to 2500 mbar with helium. The laser output was passed through a spherical quartz lens of 10-cm focal length and into protein-DNA samples contained within 6- or 8-mm-diameter Pyrex tubing at ambient temperature (20 °C). Samples were placed at varying distances from the focal point to account for variance in laser output. Sample volumes of 10–100  $\mu$ L were irradiated in the range of 10–40 mJ pulse<sup>-1</sup> cm<sup>-2</sup>, using a repetition rate of 10 Hz.

Irradiated samples were heated to 80 °C for 2 min in Laemmli sample buffer and electrophoresed in 10% polyacrylamide/SDS gels made as described (Harlow & Lane, 1988) or using a Phast System (Pharmacia LKB), fixed in 30% methanol/10% acetic acid, dried, and exposed to film or quantitated using a PhosphorImager (Molecular Dynamics).

Alternatively, a 312-nm transilluminator (Model FBTIV-816, Fisher Biotech) was used at maximum power to irradiate samples (typically 10  $\mu$ L) contained in inverted 0.5-mL polypropylene tubes with caps removed. The transilluminator was kept in a fume hood during irradiation to aid heat dispersal.

**Preparative Irradiations and Cross-Link Identification.** Telomere protein (50 nmol at 3  $\mu$ M) was incubated with 100 nmol of BrdU-substituted DNA, and 2-mL aliquots were

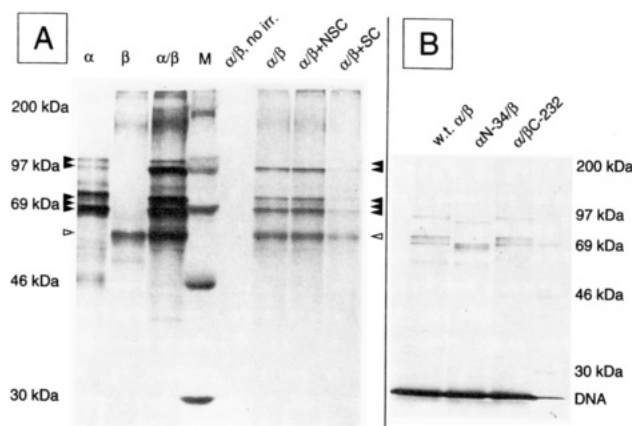
irradiated for 5 min each, with 308-nm emission incident upon the sample at 10 mJ pulse<sup>-1</sup> cm<sup>-2</sup>. A sample was analyzed by SDS-PAGE and autoradiography, and the remainder was denatured by the addition of urea to 5 M and heating to 65 °C. DNAs and DNA-protein cross-links were precipitated for 1.5–12 h by the addition of CTAB (cetyltrimethylammonium bromide, Sigma) to 1 mM and (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> to 15 mM and cooling to 0 °C (Mirzabekov et al., 1989). This step reduced total protein from roughly 5 mg to 0.5 mg. Precipitate was collected by centrifugation (10000g for 30 min), washed with 70% ethanol to solubilize CTAB, recentrifuged, and then resuspended and separated by SDS-PAGE. Cross-linked complexes were cut from an 8% polyacrylamide gel (0.75 mm × 15 cm × 20 cm) using an autoradiogram as a template, eluted by crushing the gel and soaking for 1–5 days at 37 °C in 0.5 M NH<sub>4</sub>OAc, 0.1% SDS, 0.5 mM EDTA, and precipitated with 2.5 vol of ethanol, 12–18 h at –20 °C. After centrifugation and washing in 70% ethanol, the precipitate was resuspended in 100 mL of 0.1 M Tris pH 8.3, 0.1% SDS, and proteolyzed by the addition of 2 µg of freshly resuspended sequencing-grade trypsin (Boehringer Mannheim) for 2–5 h at 25 °C. The BrdU3 cross-link resisted this treatment and was effectively digested by the addition of CaCl<sub>2</sub> to 10 mM and 2 µg of chymotrypsin (Boehringer Mannheim). After complete digestion, as monitored by SDS-PAGE and autoradiography, samples were precipitated with 0.3 M NaOAc/70% ethanol, –20 °C overnight, and then subjected to electrophoresis on a 12% polyacrylamide/7 M urea/1× TBE gel designed to separate nucleic acids (Sambrook et al., 1989). After gel separation, samples were electrophoretically blotted onto a polyvinylidene difluoride (PVDF) membrane (Immobilon, Millipore, or Trans-Blot, Bio-Rad) at 200 mA for 20 min using 1× TBE in a Bio-Rad Trans Blotter. The PVDF membrane was washed for 1 min in H<sub>2</sub>O and dried at 25 °C, and a piece of membrane bearing the cross-linked peptide(s) was cut out using an autoradiograph as a template. Yield of cross-link was estimated by Cerenkov counting of the dried filter fragment. The immobilized peptide was sequenced by automated Edman degradation, performed on an Applied Biosystems 470A sequencer using manufacturer's methods and protocols (Clive Slaughter, Howard Hughes Medical Institute, University of Texas, Southwestern).

**Analytical Cross-Linking of Mutant and Truncated Forms of Protein.** Extracts of *E. coli* expressing mutant and truncated polypeptides were prepared and incubated with telomeric DNA as described (Fang et al., 1993), and cross-linking was performed as described above.

**Site-Directed Mutagenesis of Telomere Protein Genes.** DNA fragments containing site-directed mutations were generated by adapting methodology as described (Higuchi, 1990). In Higuchi's method, both "left" and "right" PCR were carried out using mutagenic primers. The mutagenic PCR performed in this study created a mutation in only the "left" PCR, and the overlap between the middle primers was limited to the nonmutagenic portion of the primer, 5' of the desired mutation site. The introduction of each mutation was accompanied by the creation of a new restriction site to aid in identifying plasmids containing mutations. To minimize the size of PCR-amplified DNA introduced into the expression vectors, mutation-containing PCR fragments were cut with restriction endonucleases and inserted as restriction fragment cassettes into expression vectors p41A and p56A [vectors described in Gray et al. (1991)].

## RESULTS

**Formation and Characterization of UV Cross-Links.** Telomeric protein–DNA complexes were formed using ra-



**FIGURE 1:** Cross-linked products separated by SDS/polyacrylamide gel electrophoresis. Laser irradiations were at 308 nm, 37 mJ/pulse/cm<sup>2</sup> at 10 Hz for 360 s. (A) Products of irradiation of telomeric protein subunits complexed with telomeric DNA. Lanes 1–3: 2 µM  $\alpha$  and/or 6 µM  $\beta$  were incubated with 50 nM 5' <sup>33</sup>P-labeled O2T in 20 mM Tris pH 7.5, 20 mM NaCl. O2T is an oligonucleotide containing (T<sub>4</sub>G<sub>4</sub>)<sub>2</sub> at its 3' end (see Experimental Procedures). All samples contained 10 µM NSC, a nonspecific competitor DNA unrelated to telomeres, except  $\alpha/\beta$  + SC, which contained a specific competitor (O2T at 5 µM). Lane M: <sup>14</sup>C protein markers. Open arrowhead: cross-link to the  $\beta$  subunit. Filled arrowheads: cross-links to  $\alpha$ . Arrows designate cross-linked species in the  $\alpha/\beta$  lanes; those in the  $\alpha$  lane have somewhat different migration. (B) Irradiation products using extracts from *E. coli* expressing telomere protein subunits. W.t.: wild-type.  $\alpha$ N34: truncation of first 33 amino acids of  $\alpha$ .  $\beta$ C232: truncation at residue 232 resulting in loss of the remaining C-terminal residues. Both truncated subunits bind to DNA and form an apparently normal ternary complex (Fang et al., 1993).

diolabeled telomeric DNA and irradiated using either 308-nm monochromatic light or 312-nm transilluminator light. SDS-PAGE/autoradiography revealed four major and several minor species representing cross-links between the telomere protein and the DNA when the  $\alpha$ – $\beta$ –DNA complex was irradiated (Figure 1). The cross-linked forms can be grouped as an upper doublet migrating with the 97-kDa marker (the lower of which is one of the most prominent species), a triplet migrating near the 69-kDa marker, and a prominent singlet migrating at 58 kDa. For comparison, the un-cross-linked telomere protein subunits (not shown here) migrate at 41 kDa ( $\beta$ ) and 56 kDa ( $\alpha$ ).

The specificity of cross-linking was established by competition experiments: excess unlabeled telomeric DNA reduced cross-linking, while excess nonspecific single-stranded oligonucleotide had no effect (Figure 1A). Additionally, the cross-linking patterns of the recombinant protein and the protein purified from its native source, *Oxytricha nova*, were similar (data not shown).

When purified  $\alpha$  was incubated alone with telomeric DNA and irradiated, a subset of the  $\alpha$ – $\beta$ –DNA cross-linking pattern was generated, but bands with slightly different electrophoretic mobility were present as well (see the bands near the 69-kDa marker, Figure 1A). When  $\beta$  alone was incubated with telomeric DNA, a band comigrating with the fastest band in the ternary complex ( $\alpha$ – $\beta$ –DNA) appeared after irradiation. The cross-linking patterns of the two individual subunits did not add to give that of the ternary complex, indicating that the interactions involving  $\alpha$  are altered in the transition from the  $\alpha$ –DNA to the  $\alpha$ – $\beta$ –DNA complex.

**Assignment of Cross-Linked Forms to the  $\alpha$  or  $\beta$  Subunit.** To assign cross-linked species formed when irradiating the ternary complex, we utilized truncated forms of each subunit (Figure 1B). These truncated subunits bind DNA almost as tightly as does wild-type telomere protein and produce the same methylation protection pattern observed for wild-type

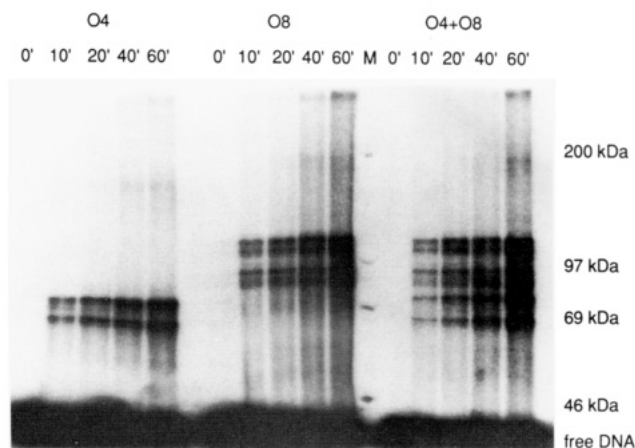


FIGURE 2: Assignment of one DNA molecule per cross-linked form. Protein purified from *O. nova* (Gottschling & Zakian, 1986; Price & Cech, 1987) was incubated with O4 [(T<sub>4</sub>G<sub>4</sub>)<sub>4</sub>] or O8 [(T<sub>4</sub>G<sub>4</sub>)<sub>8</sub>] or an equimolar mixture of O4 and O8. A time course of irradiation by a 312-nm transilluminator was performed and analyzed by SDS/8% PAGE and autoradiography. In this experiment, the protein preparation contained a proteolyzed, 26-kDa form of  $\beta$ . Thus, the  $\beta$  cross-linked form is not detected because it would coelectrophorese with the free DNA in this 8% gel. The gel pattern is also less distinct due to the use of <sup>32</sup>P labeling, whereas <sup>33</sup>P gives improved resolution (cf. Figure 4). Thus, the bands of the upper doublet and middle triplet are not resolved.

protein (Fang et al., 1993). In this experiment, cross-linking was performed using extracts made from *E. coli* expressing either  $\alpha$  or  $\beta$ , and the DNA substrate was a 40-mer with two repeats of *Oxytricha* telomeric DNA as the 3' end (O2T). Telomere protein in these extracts cross-links to O2T, giving the same pattern as seen with purified protein (compare Figure 1A to 1B). Figure 1B shows that all bands except the 58-kDa band shift to a faster mobility with a truncation of  $\alpha$ . The 58-kDa band, barely detectable in this autoradiogram, shifts to a faster mobility, or perhaps is lost, upon truncation of  $\beta$  (in Figure 1A, purified protein was used, which gave a higher yield of this band). Thus, in the ternary complex cross-linking assay, the upper two groups of bands represent cross-links to  $\alpha$ , while the 58-kDa band represents a cross-link to  $\beta$ .

**Electrophoretic Mobility of Cross-Linked Forms.** The cross-linking pattern revealed by SDS-PAGE was unusually complex, prompting us to ask whether multiple DNAs or polypeptides were involved. In 8% polyacrylamide gels, the slowest migrating cross-linked form had an apparent molecular weight of 80 kDa (Figure 2, O4 lanes). It is unlikely that any combination of multimeric polypeptides cross-linked to DNA could migrate this rapidly, since the minimum molecular weight of such a complex,  $\beta$ - $\beta$ -DNA, is 94 kDa. In addition, irradiation time courses indicated that each cross-link formed at a similar relative rate; if a double cross-link were formed, it would be expected to appear at a significantly lower relative rate.

To address the possibility that the slowly migrating cross-linked forms contain more than one DNA, we used O4, O8, and an equimolar mixture of O4/O8 in cross-linking assays. Figure 2 shows that the additional 32 nucleotides present in O8 significantly retarded the mobility of cross-linked forms. No new bands appeared in the O4/O8 mixture, as would be generated if a cross-linked form contained both O4 and O8; thus, each cross-linked species contains only one DNA molecule. We conclude that each band in the autoradiogram represents a single cross-link between DNA and protein, and that the slow electrophoretic mobility of the upper doublet is likely due to a structural feature of the protein-DNA cross-link which retards its progress through the gel.

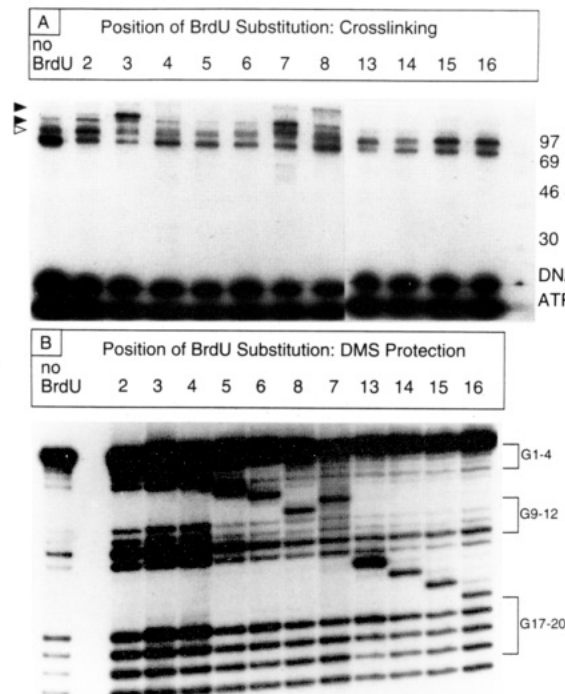


FIGURE 3: Methylation protection footprint on BrdU-substituted DNAs. In each DNA, BrdU was substituted by automated DNA synthesis for one nucleotide in O2T. Position 1 is the 3' G. (A) Irradiation of BrdU-substituted nucleoprotein complexes, followed by SDS-PAGE/autoradiography. Open arrowhead: cross-link to the  $\beta$  subunit. Filled arrowheads: cross-links to  $\alpha$ . (B) BrdU-substituted nucleoprotein complexes analyzed by methylation protection as described (Fang et al., 1993). Not shown is the methylation of protein-free DNA, which produces a uniform ladder of cleavage at Gs; cf. Gray et al. (1991).

**Substitution of T Residues in O4 with BrdU.** BrdU is a T analog known to enhance UV cross-linking efficiency (Weintraub, 1974). Because the *Oxytricha* telomere protein contacts several T residues in O4 (Raghuraman & Cech, 1989), we synthesized several derivatives of O4, each containing a single, site-specific, T-to-BrdU substitution. In addition, each of three G residues at the 3' end was altered to BrdU.

DNAs brominated at positions 2–8 and 13–16, where position 1 is the 3' end, were assayed for photo-cross-linking to telomere protein (Figure 3A) using a 312-nm transilluminator. BrdU-independent cross-linking is represented by the "no BrdU" lane (Figure 3A). In this experiment the  $\beta$  cross-link was unusually strong in the "no BrdU" lane; its intensity is more typically represented in the remaining lanes. Enhancement of cross-links to  $\alpha$  are seen at positions 3, 7, 15, and 16. Positions 9–12 represent Gs which were not substituted in this experiment.

To verify that the T-to-BrdU substitutions do not alter protein-DNA contacts and to examine the effect of heterologous G-to-BrdU substitutions on protein binding at the 3' end, BrdU-substituted complexes were assayed using DMS methylation protection (Figure 3B). The normal  $\alpha$ - $\beta$ -DNA methylation protection pattern (Price & Cech, 1987; Raghuraman & Cech, 1989; Gray et al., 1991; Fang et al., 1993) was reproduced faithfully in all T-to-BrdU substituted DNAs, indicating that protein binding at G tracts is not disrupted. [DNA cleavage at the position of bromination is evident by examination of the gel pattern; this cleavage prevents assessment of protection at positions 2, 3, and 4 for the DNAs substituted with BrdU at these positions. While it is evident that Gs in the 3' G tract are protected from methylation in the G-BrdU substituted DNAs (Figure 3B, lanes 2–4, compare



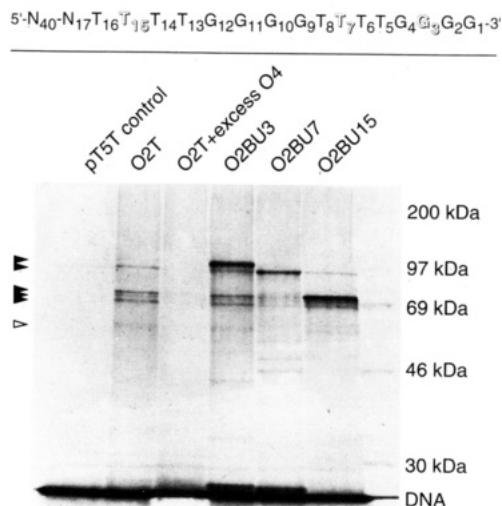


FIGURE 4: UV cross-linking of BrdU-substituted DNAs. Upper panel: sequence of telomeric DNA O2T. Three nucleotide positions substituted with BrdU in this experiment are outlined. Lower panel: extracts of *E. coli* expressing  $\alpha$  or  $\beta$  were mixed at protein excess with 2  $\mu$ M  $^{33}$ P-labeled DNAs, irradiated at 308 nm, and analyzed by SDS-PAGE and autoradiography. pT5T control: an extract of *E. coli* containing the pT5T expression vector without an inserted gene. O2BU3, O2BU7, and O2BU15 are identical to O2T except that BrdU is substituted at position 3, 7, or 15, respectively. Markers are  $^{14}$ C-labeled polypeptides. Lane 3 contains 200  $\mu$ M unlabeled O4, and the remaining lanes contain 200  $\mu$ M NSC, a nontelomeric oligonucleotide competitor. Irradiation of all samples was at 308 nm, 33 mJ pulse $^{-1}$  cm $^{-2}$  at 10 Hz, for 120 s, except for O2BU7, which was irradiated for 30 s. The higher relative rate of this cross-link's formation, and thus higher relative rate of degradation which causes smearing of the gel pattern, forces irradiation of O2BU7 for the shorter time. Open arrowhead: cross-link to the  $\beta$  subunit. Filled arrowheads: cross-links to  $\alpha$ .

G1-4 to G17-20), the missing information at the brominated positions prevents us from distinguishing binary ( $\alpha$ -DNA) from ternary ( $\alpha$ - $\beta$ -DNA) complex footprints for the DNAs substituted at positions 2, 3, and 4.]

Photo-cross-linking with DNA substituted at positions 3, 7, and 15, which consistently demonstrated enhanced cross-linking relative to the unsubstituted DNA, was analyzed on higher resolution gels (Figure 4).

All three enhanced cross-linked forms appear to comigrate with bands which are seen with the unsubstituted DNA: bromination at position 3 or 7 resulted in cross-linked forms which comigrated with the two bands of the upper doublet, while bromination at position 15 resulted in a cross-linked form which comigrated with the slowest migrating band of the middle triplet (Figure 4). The presence of BrdU-independent cross-linked forms in the BrdU-substituted DNA provided an internal control, showing that protein-DNA contacts required for cross-linking to both  $\alpha$  and  $\beta$  are maintained in these complexes despite the BrdU substitution for G<sub>3</sub> (Figure 4). Note that irradiation was performed at concentrations 10<sup>3</sup>-fold higher than the  $K_d$  of 1 nM<sup>2</sup> (Fang et al., 1993). Thus, effects of the G-to-BrdU substitutions which had less than a 10<sup>3</sup>-fold effect on the  $K_d$  would have gone undetected.

**Laser versus Transilluminator Irradiation.** The 308-nm-emitting XeCl excimer laser provides a powerful source of monochromatic light which has been used previously in protein-RNA cross-linking (Gott et al., 1991). To optimize cross-linking for preparative work, we performed irradiation time courses. An example (Figure 5) compares the unsubstituted DNA, O2T, to a BrdU-substituted DNA, O2BU7. Maximal cross-linking of the BrdU7 band occurs at 2 min, and photodegradation becomes evident after this time point

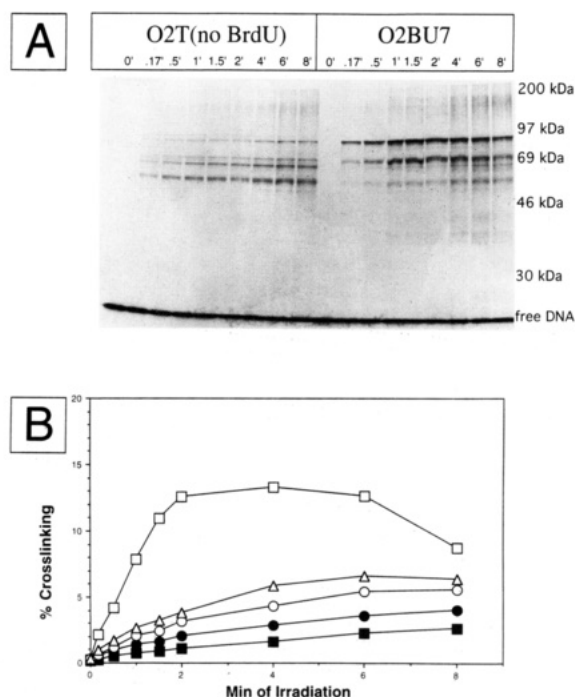


FIGURE 5: BrdU enhances photo-cross-linking yield. (A) SDS-polyacrylamide gel showing a time course of irradiation comparing BrdU-substituted with unsubstituted DNA. Partially purified  $\alpha$  and  $\beta$  were incubated at 3 and 6  $\mu$ M, respectively, with 1  $\mu$ M  $^{33}$ P-labeled O2T or O2BU7 in 10 mM Tris pH 7.5, 20 mM NaCl, and 7  $\mu$ M NSC, a nontelomeric oligonucleotide. Irradiation was at 308 nm, 33 mJ pulse $^{-1}$  cm $^{-2}$  at 10 Hz. (B) Data from the upper panel quantified using a PhosphorImager (Molecular Dynamics). 100% cross-linking = radioactivity in the free DNA band at time 0. Only the BrdU-enhanced band running just below 97 kDa was quantified in the O2BU7 time course (open squares). The remaining data come from the O2T time course: closed squares, the uppermost strong band at 97 kDa; closed circles, the upper band of the middle triplet centered around 69 kDa; open circles, the lower band of the middle triplet; open triangles, the  $\beta$  cross-link at 58 kDa.

(Figure 5A). In contrast, the percent cross-linking of the BrdU-independent bands continues to increase through 8 min. The yield (2 min) of the cross-linked species migrating slightly faster than 97 kDa was increased more than 10 times by BrdU substitution (compare open and closed squares in Figure 5B). Percent cross-linking for the BrdU-independent bands ranged from 2% to 7%, while the enhanced cross-linked species generated upon BrdU substitution at positions 3, 7, and 15 were obtained with yields of 25, 13, and 16%, respectively. An additional enhanced photo-cross-link at 69 kDa caused by bromination at position 7 is evident in Figure 5A but absent in Figure 4, lane O2BU7. This additional enhanced band was seen only when the protein was present in molar excess.

While optimal laser power and time required to reach maximal cross-linking varied for each of the three cross-links, we found that the total cross-linking yield depends on total photon input rather than on light intensity (data not shown).

With the transilluminator as the radiation source, 30–60 min of irradiation was required to reach maximal yields (data not shown), which was about 50–100-fold slower than with the excimer laser at 40 mJ/pulse/cm<sup>2</sup>. Maximum yield with the transilluminator was roughly half that obtained with the laser, while the qualitative cross-linking pattern was identical. In conclusion, laser and transilluminator irradiation provided similar qualitative results, with laser excitation enabling a much faster approach to a 2-fold-higher cross-linking yield [see Gott et al. (1991) for further comparison of transilluminator vs laser excitation of an RNA-protein complex].

**Sequences Identifying Cross-Linked Amino Acids.** Because BrdU substitution at sites 3, 7, and 15 enhanced cross-linking

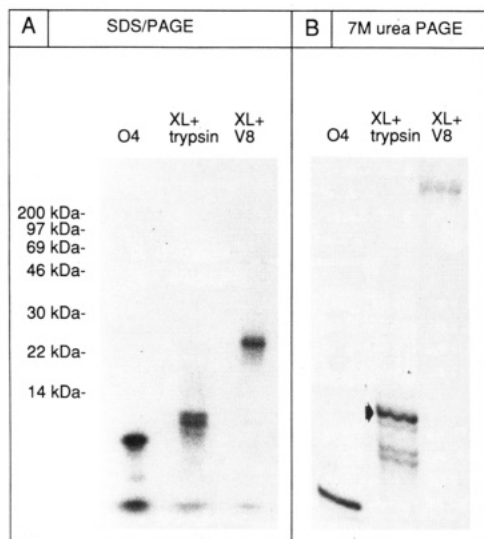


FIGURE 6: Electrophoretic purification of cross-linked peptide. (A) Analysis of cross-linked peptides by SDS-15% polyacrylamide gel electrophoresis. Gel-purified O2BU7-cross-linked form was proteolyzed by trypsin (XL + trypsin) or V8 protease (XL + V8) and electrophoresed next to O4, the un-cross-linked DNA marker. The fastest migrating band in this gel represents a radiolabeled product which has been hydrolyzed from the DNA during incubation. (B) The same samples from panel A electrophoresed through a preparative 7 M urea/12% polyacrylamide gel. Arrow indicates the band transferred for sequence analysis.

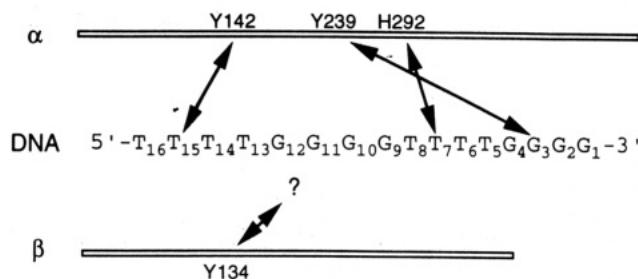


FIGURE 7: Physical map of cross-links in the  $\alpha$ - $\beta$ -O2T DNA complex identified by the sequencing.

efficiency and provided reference sites on the DNA, these three DNAs became candidates for preparative scale cross-linking to identify the corresponding amino acid residues. Peptides derived from cross-links at these three positions were isolated and sequenced (Figure 6; see supplementary material). The cross-linked amino acids were identified as  $\alpha$ Tyr239 (G<sub>3</sub>),  $\alpha$ His292 (T<sub>7</sub>), and  $\alpha$ Tyr142 (T<sub>15</sub>), all on the  $\alpha$  subunit (Figure 7). The cross-linked amino acid was inferred through its absence: the covalently attached DNA alters the amino acid such that it is not detected, and a blank appears in the sequence. In one case, BrdU3, two cycles of Edman degradation were missing a residue,  $\alpha$ His234 and  $\alpha$ Tyr239. We believe that the blank at  $\alpha$ His234 was artifactual, due to adsorption of histidine to the HPLC column used in sequencing, and have confirmed  $\alpha$ Tyr239 as the cross-linked residue by site-directed mutagenesis (see below). In addition, a peptide was isolated from the  $\beta$  cross-link band which represents a cross-link to  $\beta$ Tyr134; because the  $\beta$  cross-link formed independent of BrdU substitution (Figure 3A), the cross-linking partner or partners on O2T DNA remain unidentified (Figure 7). Thus, the complex pattern generated by UV cross-linking has revealed four points of contact, or at least close association, between protein and DNA in the telomeric complex. The  $\alpha$  subunit places its residues  $\alpha$ Tyr239,  $\alpha$ His292, and  $\alpha$ Tyr142 in close proximity to nucleotides BrdU3, BrdU7, and BrdU15, respectively, while the  $\beta$  subunit places its  $\beta$ Tyr134 near an unknown nucleotide(s).

An alignment of four  $\alpha$  subunit sequences from hypotrichous ciliates shows that two of the three  $\alpha$  subunit cross-linking residues are conserved (Figure 8). The third,  $\alpha$ His292, is not indicated as conserved in this figure, but His residues are present one residue away. Secondary structure prediction algorithms (Gibrat et al., 1987) predict that  $\alpha$ His292 is found in a  $\beta$  turn (data not shown), which could likely accommodate a slight shifting to place  $\alpha$ His292 in a geometry similar to that of nearby His residues in the other two polypeptides.

**Primary Sequence Features of Cross-Linked Peptides.** A hydrophobicity profile (Figure 9) indicates that the BrdU15 cross-link,  $\alpha$ Tyr142, occurs amid one of the most hydrophobic regions of this DNA-binding domain, while the BrdU7 cross-link,  $\alpha$ His292, is also in a hydrophobic stretch, close to the transition from a hydrophilic stretch (Figure 8). Both cross-links occur at a pair of aromatic residues near serine residues. The position 7 cross-link is directly flanked by two serines, and the aromatic pair near BrdU15 has three serines on its C-terminal side. BrdU3, very near the 3' end, cross-links to Tyr239, which is found in a hydrophilic/acidic stretch within which four of eight residues are either Glu or Asp. The  $\beta$  cross-link,  $\beta$ Tyr134, occurs in a region of moderate hydrophobicity [sequence in Hicke et al. (1990)].

**A Sequence in  $\beta$  Which Resembles the RNA Recognition Motif.** Downstream from the  $\beta$ Tyr134 cross-linking residue, visual inspection revealed a block of amino acids with similarity to the RNA recognition motif (RRM), which is conserved in a class of RNA-binding proteins (Bandziulis et al., 1989). RNP-1, an octamer sequence contained within the RRM, represents the most conserved region of the RRM and has the consensus sequence K/R-G-F/Y-G/A-F-V-X-F/Y, where X represents any amino acid. Four of seven possible matches to the RNP-1 octamer are found in residues 196–203 of  $\beta$ : K-G-D-E-F-S-D-F. This finding led to the hypothesis that the telomere protein  $\beta$  subunit binds DNA using a binding motif similar to the RRM.

A test of this hypothesis was performed by mutating the two phenylalanines in the octamer sequence to valine, a double mutation which destroys RNA binding by the U1-A protein (Lutz-Freyermuth et al., 1990). This hydrophobic pair represents the most highly conserved residues in the RRM. The doubly mutant  $\beta$  subunit retains the ability to bind telomeric DNA as determined by a cross-linking experiment (data not shown). In addition, the doubly mutant  $\beta$  subunit participates with  $\alpha$  to generate the wild-type methylation protection pattern, indicating its competence to interact with  $\alpha$  (described below). It must be noted that these experiments were performed using crude *E. coli* extracts at roughly 1  $\mu$ M telomere protein, far above the  $K_d$  of 1 nM<sup>2</sup>. Nevertheless, we can conclude that the sequence under investigation is not essential for DNA binding.

**Site-Directed Mutagenesis of Cross-Linking Amino Acids.** To confirm that peptide sequencing had correctly identified the cross-linked amino acids, we mutated each candidate to Ala, expressed the mutant polypeptides in *E. coli*, and tested the mutants for cross-linking. Figure 10 shows that in each case the mutation to Ala knocks out the BrdU-enhanced cross-linking. The mutation His234Ala did not abolish cross-linking, indicating that the absence of His234 from the cross-linked peptide's sequence was not due to photo-cross-linking. Interestingly, a faint band is present at the position of the formerly BrdU-enhanced band in all three  $\alpha$  mutants, where we had expected to see cross-linking completely abolished by the mutation. Mutation of the amino acid involved in the  $\beta$  cross-link,  $\beta$ Tyr134 to Ala, did not abolish cross-linking (data not

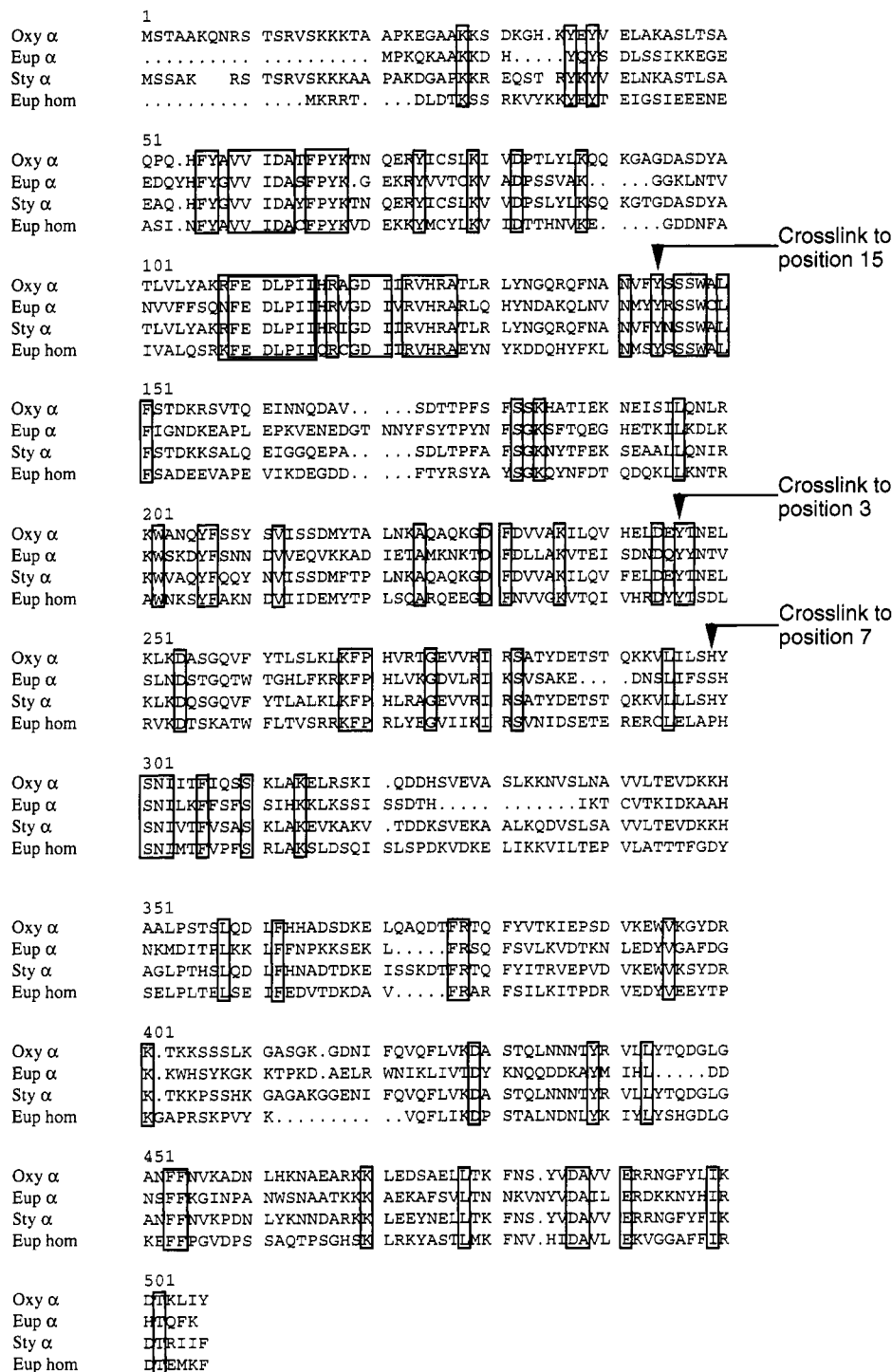


FIGURE 8: Alignment of four  $\alpha$  subunit sequences from ciliated protozoa (Wang et al., 1992). Cross-linking amino acids of the *Oxytricha*  $\alpha$  subunit are indicated, and completely conserved amino acids are boxed. Eup: *Euplotes*  $\alpha$  subunit. Sty, *Stylonychia*  $\alpha$  subunit. Oxy: *Oxytricha*  $\alpha$  subunit. Eup hom: a homolog of the *Euplotes*  $\alpha$  subunit protein whose function is yet undetermined.

shown), nor did the mutation inhibit  $\beta$ 's ability to interact with  $\alpha$  in forming the ternary complex (described below).

$\alpha$ Tyr239 photo-cross-links to the DNA at the 3' G-tract, and  $\alpha$ His234 could be nearby the DNA due to its proximity to  $\alpha$ Tyr239 in the primary sequence. We therefore tested the ability of mutants  $\alpha$ Tyr239Ala and  $\alpha$ His234Ala to protect the 3' end of telomeric DNA in a methylation protection assay. No alteration of the wild-type footprint is seen (Figure 11), indicating that contacts to N7 of guanines are not disrupted by these mutations. We also tested two mutant forms of the  $\beta$  subunit for their ability to participate in complex formation: a mutant which disrupts the RRM-like sequence, and a mutant which alters a photo-cross-linking residue,

$\beta$ Tyr134Ala. Neither  $\beta$  mutant displayed an altered methylation protection pattern in the presence of  $\alpha$ .

## DISCUSSION

We have characterized photo-cross-linking in *Oxytricha* telomeres reconstituted from the  $\alpha$  and  $\beta$  telomeric polypeptides and an oligonucleotide of telomeric DNA sequence. Regiospecific photo-cross-links formed by BrdU substitution identify three nucleotide-amino acid point contacts within this complex, providing information on the assembly of the chromosomal terminus into a nucleoprotein complex.

**Specificity of Cross-Linking.** The photo-cross-links detected here in reconstituted *Oxytricha* telomeres accurately

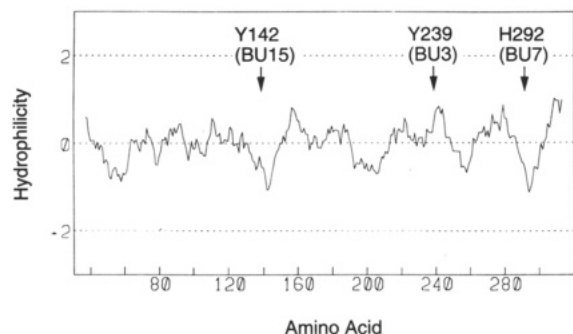


FIGURE 9: Hopp-Wood hydrophilicity analysis of amino acids 30–320 of the  $\alpha$  subunit. The positions of the three cross-links are indicated. Residues 30–320 represent a domain capable of binding to telomeric DNA (Fang et al., 1993).

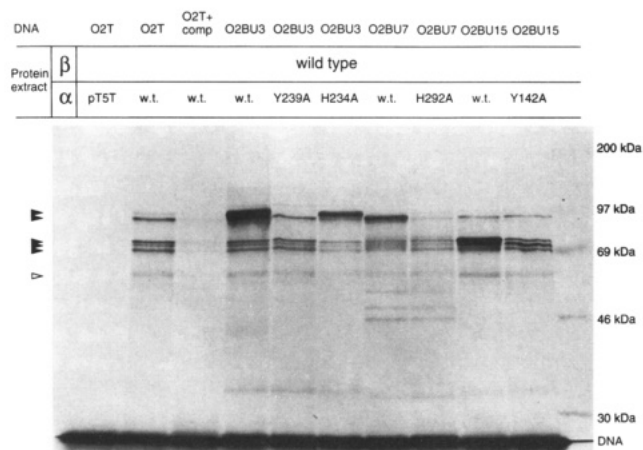


FIGURE 10: Cross-linking of mutant telomere protein subunits to BrdU-substituted DNAs. O2T + comp: 200  $\mu$ M unlabeled O4 is added as a specific competitor. Protein: crude extracts were used as described in the caption below Figure 4. pT5T: extract of *E. coli* containing the expression vector without an inserted telomere protein gene. Mutant nomenclature: Y239A is a mutation of  $\alpha$ Tyr 239 to Ala. W.t.: wild-type protein. Open arrowhead: cross-link to the  $\beta$  subunit. Filled arrowheads: cross-links to  $\alpha$ .

report protein–DNA interactions occurring at telomeres *in vivo*, as evidenced by three criteria: First, protein–DNA interactions detected at telomeres *in vivo* using a DMS methylation protection assay (Gray et al., 1991) are faithfully reproduced using the recombinant telomere protein and the oligodeoxynucleotide d(T<sub>4</sub>G<sub>4</sub>)<sub>4</sub>. Second, the specificity of cross-linking was established by competition experiments: excess unlabeled telomeric DNA reduces cross-linking, while excess nonspecific single-stranded oligonucleotide has no effect (Figure 1). Third, the cross-linking patterns of the recombinant protein and the protein purified from its native source, *O. nova*, are similar (Figure 2 and additional data not shown).

**Complexity of Cross-Linking.** The UV cross-linking assay developed here revealed six distinct cross-links between the *Oxytricha* telomere protein and its cognate DNA. Five cross-links were to  $\alpha$ , and one was to  $\beta$  (Figure 1). The appearance of multiple cross-links was unexpected and indicates that this protein binds to single-stranded DNA in a manner that places several of its aromatic side chains near the DNA in positions favorable for cross-linking. The five  $\alpha$  cross-linked forms have a wide range of electrophoretic mobilities, with the apparent molecular weights dependent on gel composition. Cross-linking is known to alter electrophoretic mobility of proteins (Steele & Nielsen, 1978) and nucleic acids. Mixing experiments with different lengths of telomeric DNA, along with comparison of the rates of cross-link formation, showed that each band represents a single cross-linking event.

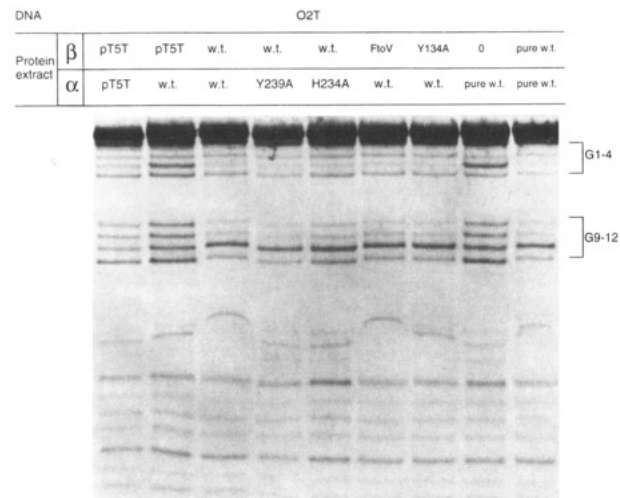


FIGURE 11: DMS methylation protection by mutant proteins. Crude extracts of telomere protein, described in the caption below Figure 4, were incubated with 5' <sup>33</sup>P-labeled O2T which in this experiment has a biotin replacing the A next to the 5' end (Fang et al., 1993). DMS was added to methylate DNA, and the DNA was isolated by using streptavidin-conjugated magnetic beads and analyzed according to described methods (Fang et al., 1993). Mutants are as described in the caption below Figure 10, except for two  $\beta$  mutations: FtoV represents a double mutation of Phe200 and Phe203, each to Val, and Y134A is a Tyr to Ala mutation. Pure w.t.: purified wild-type polypeptides.

**Protein–DNA Cross-Linking with BrdU-Substituted DNA.** The site-specific incorporation of the T analog 5-bromodeoxyuridine (BrdU) into telomeric DNAs facilitated the initially daunting task of identifying individual nucleotide–amino acid cross-links within this complex photo-cross-linking pattern. The high yield of photo-cross-linking (up to 25%) that could be achieved in 30–120 s supports the hypothesis that these cross-links are representative of the majority of telomeric complexes, and not a small fraction of aberrant complexes.

The techniques described here are advantageous not only in producing high-yield photo-cross-linking but also in their adaptability to commonly available laboratory equipment. Peptide isolation requires only a single step of gel electrophoresis after proteolysis, and the gel is then blotted onto a PVDF membrane and sequenced directly. The photo-cross-linking methodology described here can be applied using a 308-nm monochromatic light source or a 312-nm transilluminator. We have recently found that 5-iododeoxyuracil (IdU) substitution produces extremely efficient (70%) photo-cross-linking at 308 or 312 nm, representing a further improvement over the yields with BrdU (Willis et al., 1993).

**Cross-Linked Amino Acid Residues.** BrdU substitution at positions 3, 7, and 15 of the telomeric oligonucleotide resulted in photo-cross-linking to  $\alpha$ Tyr239,  $\alpha$ His292, and  $\alpha$ Tyr142, respectively, of the *O. nova* telomere protein. To confirm the identification of these residues as photo-cross-linking partners, each was altered to Ala. These mutations did not destroy binding activity, but each mutation eliminated BrdU-enhanced photo-cross-linking, strongly supporting their identification as the cross-linked residues. The formation of the normal pattern of BrdU-independent photoproducts served as a control for correct formation of a *bona fide* telomeric complex containing the mutant polypeptides. Furthermore, when tested for their ability to protect the 3' terminal Gs from methylation, the mutants  $\alpha$ Tyr239Ala and  $\alpha$ His234Ala produced a wild-type protection pattern, indicating that the protein–DNA complex was not globally disrupted by these mutations.

Though mutation destroyed BrdU-enhanced cross-linking in all three cases, in two cases a relatively inefficiently photo-



cross-linked form remained which had the same electrophoretic migration. In both cases,  $\alpha$ His292Ala and  $\alpha$ Tyr142Ala, an aromatic residue is a nearest neighbor in the polypeptide primary sequence. Possibly the photo-cross-links revealed upon mutation represent adduct formation between BrdU and the neighboring aromatic side chain which is now in a cross-linkable conformation due to the mutation. Alternatively, the photo-cross-linked form revealed by mutation could represent an unrelated photoadduct, not dependent on BrdU, which simply comigrates in polyacrylamide gels.

While the ability to photo-cross-link is an indicator of physical proximity and cross-linkable geometry of amino acid and nucleotide, such proximity does not necessarily imply that the identified amino acid is forming a critical contact (Allen et al., 1991). The mutation  $\alpha$ Tyr239Ala alters a cross-linking amino acid, yet no change in the DMS protection of G<sub>3</sub> could be detected using this mutant. Since DMS protection monitors only the N7 of guanine, more extensive studies are necessary to determine the role, if any, that these cross-linkable amino acids play in complex formation.

**DNA-Binding Regions of the Telomere Protein  $\alpha$  Subunit.** Since BrdU7 and BrdU15 occupy analogous positions in neighboring T tracts of telomeric DNA (cf. Figure 7), we examined the character of the residues spanning these cross-links to see if any similarities existed in peptides cross-linked to positions 7 and 15. Both cross-linked amino acids,  $\alpha$ His292 and  $\alpha$ Tyr142, are part of an aromatic pair which is flanked by hydrophilic but uncharged side chains: Ser-His292-Tyr-Ser-Asn and Asn-Val-Phe-Tyr142-Ser-Ser-Ser. The primary sequence in this region cannot provide information about how the binding pocket is shaped. Nevertheless, it is possible to speculate that the arrangement of two aromatic residues together, flanked by hydrophilic but uncharged residues, places  $\alpha$ Tyr142 and  $\alpha$ His292 in favorable cross-linking environments, perhaps through intercalation of the pairs of aromatic side chains into a nucleotide base stacking lattice. An additional example of this arrangement is found just C-terminal to the BrdU7-cross-linked  $\alpha$ His292, where Phe is flanked by uncharged hydrophilic residues (Figure 8).

The three  $\alpha$  amino acids which cross-link to DNA,  $\alpha$ Tyr142,  $\alpha$ Tyr239, and  $\alpha$ His292, fall in conserved stretches of the protein (Figure 8). This conservation supports the hypothesis that cross-linking residues (or at least residues near the cross-links) are not simply in proximity to DNA, but are important for DNA recognition. Furthermore, the formation of multiple photo-cross-links implicates quite a number of aromatic side chains as participants in this nucleoprotein complex. It has been noted that a disproportionate number of the conserved amino acids in the  $\alpha$  subunit are aromatic residues (Wang et al., 1992). In fact, of 88 completely conserved residues in the alignment of the four  $\alpha$  sequences, 24, or nearly one-third of the conserved residues, are aromatic. There are 17 additional cases in which an aromatic residue is always present or is present in three of the four sequences, but the identity of the aromatic residue varies. Thus aromatic residues appear to play a large role in this protein's function.

NMR studies indicate that nonspecific single-stranded DNA binding proteins of bacteriophages T4 and fd contain aromatic residues which interact with the DNA bases (O'Connor & Coleman, 1983; Prigodich et al., 1984). Prigodich et al. (1984) proposed that nonspecific single-stranded DNA binding proteins in general may employ aromatic side chains by intercalating them into the base stacking lattice of a polynucleotide. Though proteins for which such intercalation has been hypothesized bind to DNA nonspecifically, there may

be similarities between them and telomere protein in the way DNA is bound.

A hydrophilicity profile indicates that cross-linking to T tracts, at BrdU7 and BrdU15, occurs amid the most hydrophobic regions of  $\alpha$ 's DNA binding domain. One explanation for the stability of this nucleoprotein complex in high salt (>2 M NaCl) is the formation of hydrophobic interactions. Such hydrophobic interactions could be at protein-protein or at protein-DNA interfaces. The existence of cross-links to predicted hydrophobic regions of  $\alpha$  indicates that salt-stabilizing hydrophobic interactions may in fact be present at protein-DNA interfaces in telomeres. The character of the  $\alpha$ Tyr239-BrdU3 cross-link, which represents a cross-link near the extreme terminus of an *Oxytricha* chromosome, appears different from the two cross-links to T tracts.  $\alpha$ Tyr239 is found in the midst of an acidic block in which four of eight residues are Glu or Asp. The high concentration of negative charges here would not be expected to approach DNA; perhaps salt bridges are formed with positively charged regions of  $\alpha$  or  $\beta$  which neutralize the negative charge density.

The four cross-linking residues identified here all fall within regions of telomere protein required for DNA binding, as defined by a truncation analysis of each subunit (Fang et al., 1993). Thus there is good agreement between the photo-cross-linking and molecular genetic analyses of DNA binding regions. The *Euplotes* 50-kDa protein can be proteolytically truncated to produce a DNA-binding fragment (Price et al., 1992) similar to the domain identified for the *Oxytricha*  $\alpha$  subunit.

**Photo-Cross-Linking to BrdU-Substituted vs Unsubstituted DNA.** Three of the six gel bands representing cross-linking were intensified by BrdU substitutions of G<sub>3</sub>, T<sub>7</sub>, and T<sub>15</sub>. This observation does not establish that the coelectrophoresing, BrdU-independent cross-linked forms represent cross-links between the same amino acid and nucleotide in the unmodified DNA. However, such an identity is consistent with model photochemical studies (Shetlar et al., 1984). Using irradiation at 254 nm, polyG reacted at greatest yield with Phe, Arg, Cys, Tyr, and Lys, and polyT with Cys, Phe, Tyr, Lys, Asn, Arg, and His. The three photo-cross-links identified in this study fall within the highest reactivity combinations. Further, the photoproduct from reaction of thymine with *N*-acetyltyrosine [*N*-acetyl-4-hydroxy-3-(5,6-dihydrothymine-5-yl)phenylalanine, Shaw et al. (1992)] is similar to the photoproduct from reaction of 5-bromouracil with *N*-acetyltyrosine *N*-ethylamide (Dietz & Koch, 1987). Taken together, these studies suggest that BrdU substitution may simply enhance formation of a cross-link that can be formed by the unsubstituted DNA.

**The  $\beta$  Subunit Photo-Cross-Link.** Mutation of the amino acid thought to be responsible for the  $\beta$  subunit-DNA cross-link,  $\beta$ Tyr134, did not abolish  $\beta$  cross-linking.  $\beta$ Tyr134 is, therefore, not the only  $\beta$  residue which can cross-link to the DNA. Because the  $\beta$  cross-link is a diffuse band in a gel, it is possible that other cross-links are formed which contribute to this band. Since the role of  $\beta$ Tyr134 in cross-linking has not been confirmed by independent means, the placement of  $\beta$ Tyr134 near the telomeric DNA must remain hypothetical at this time.

**Ternary Complex Formation.** The existence of a cross-link between  $\alpha$ Tyr239 and BrdU-substituted G<sub>3</sub> allows evaluation of models of ternary complex ( $\alpha$ - $\beta$ -DNA) formation. When  $\beta$  is withheld,  $\alpha$  protects only the terminal G tract against methylation by DMS, forming an  $\alpha$  pattern (Gray et al., 1991). The addition of  $\beta$  translates the  $\alpha$  pattern to the next G tract inward, and a new pattern, the  $\alpha/\beta$  pattern, is seen at the 3' G tract. The two simplest models for ternary

complex formation are that  $\beta$  contacts the 3' tract with  $\alpha$  sliding inward, or that  $\beta$  contacts only  $\alpha$  and alters its conformation (Gray et al., 1991). Since we now know, by photo-cross-linking, that  $\alpha$  is near the 3' end and that  $\beta$  can be cross-linked to the DNA, a new model can be envisioned in which  $\beta$  contacts telomeric DNA as well as  $\alpha$ . When  $\beta$  binds,  $\alpha$  changes its protein-DNA contacts, now contacting the 3' G tract and both T tracts.

In addition to testing models for telomere assembly, these photo-cross-links map point contacts in the nucleoprotein complex. Analysis of these point contacts supports the hypothesis that hydrophobic interactions are important contributors to complex formation. In addition, the formation of multiple photo-cross-links provides support for the hypothesis that the telomere protein utilizes multiple aromatic residues in single-stranded DNA binding.

## ACKNOWLEDGMENT

We are grateful to Clive Slaughter (HHMI, U.T. Southwestern) for peptide sequencing and helpful suggestions. Guowei Fang and John Gray were instrumental in providing helpful discussions and materials. Steve Schultz and James Ruggles designed the protein purification protocol currently in use.

## SUPPLEMENTARY MATERIAL AVAILABLE

One figure showing amino acid sequencing data of peptides cross-linked to BrdU-substituted O2T DNA (1 page). Ordering information is given on any current masthead page.

## REFERENCES

- Allen, T. D., Wick, K. L., & Matthews, K. S. (1991) *J. Biol. Chem.* 266, 6113-6119.
- Bandziulis, R. J., Swanson, M. S., & Dreyfuss, G. (1989) *Genes Dev.* 3, 431-435.
- Barbier, B., Charlier, M., & Maurizot, J. C. (1984) *Biochemistry* 23, 2933-2939.
- Berman, J., Tachibana, C. Y., & Tye, B.-K. (1986) *Proc. Natl. Acad. Sci. U.S.A.* 83, 3713-3717.
- Blackburn, E. H. (1990) *Science* 249, 489-490.
- Blackburn, E. H. (1991) *Nature* 350, 569-573.
- Blackburn, E. H., & Gall, J. G. (1978) *J. Mol. Biol.* 120, 33-53.
- Blackburn, E. H., & Chiou, S.-S. (1981) *Proc. Natl. Acad. Sci. U.S.A.* 78, 2263-2267.
- Blatter, E. E., Ebright, Y. W., & Ebright, R. H. (1992) *Nature* 359, 650-652.
- Cardenas, M. E., Bianchi, A., & de Lange, T. (1993) *Genes Dev.* 7, 883-894.
- Catalano, C. E., Allen, D. J., & Benkovic, S. J. (1990) *Biochemistry* 29, 3612-3621.
- Cheung, M. K., Drivas, D. T., Littau, V. C., & Johnson, E. M. (1981) *J. Cell Biol.* 91, 309-314.
- Dietz, T. M., & Koch, T. H. (1987) *Photochem. Photobiol.* 46, 971-978.
- Dietz, T. M., & Koch, T. H. (1989) *Photochem. Photobiol.* 49, 121-129.
- Fang, G., Gray, J. T., & Cech, T. R. (1993) *Genes Dev.* 7, 870-882.
- Gibrat, J. F., Garnier, J., & Robson, B. (1987) *J. Mol. Biol.* 3, 425-443.
- Gott, J. M., Willis, M. C., Koch, T. H., & Uhlenbeck, O. C. (1991) *Biochemistry* 30, 6290-6295.
- Gottschling, D. E., & Cech, T. R. (1984) *Cell* 38, 501-510.
- Gottschling, D. E., & Zakian, V. A. (1986) *Cell* 47, 195-205.
- Gray, J. T. (1993) Ph.D. Thesis, University of Colorado, Boulder.
- Gray, J. T., Celander, D. W., Price, C. M., & Cech, T. R. (1991) *Cell* 67, 807-814.
- Harlow, E., & Lane, D. (1988) *Antibodies. A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Henderson, E. R., & Blackburn, E. H. (1989) *Mol. Cell. Biol.* 9, 345-348.
- Hicke, B. J., Celander, D. W., MacDonald, G. H., Price, C. M., & Cech, T. R. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 1481-1485.
- Higuchi, R. (1990) Recombinant PCR, in *PCR Protocols*, pp 177-183, Academic Press, San Diego.
- Hoffman, D. W., Query, C. C., Golden, B. C., White, S. W., & Keene, J. D. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 88, 2495-2499.
- Katouzian-Safadi, M., Laine, B., Chartier, F., Cremet, J. Y., Belaiche, D., Sautiere, P., & Charlier, M. (1991) *Nucleic Acids Res.* 19, 4937-4941.
- Kenan, D. J., Query, C. C., & Keene, J. D. (1991) *TIBS* 16, 214-220.
- Klobutcher, L. A., Swanton, M. T., Donini, P., & Prescott, D. M. (1981) *Proc. Natl. Acad. Sci. U.S.A.* 78, 3015-3019.
- Lipps, H. J., Gruissem, W., & Prescott, D. M. (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79, 2495-2499.
- Lutz-Freyermuth, C., Query, C. C., & Keene, J. D. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 6393-6397.
- Merrill, B. M., Williams, K. R., Chase, J. W., & Konigsberg, W. H. (1984) *J. Biol. Chem.* 259, 10850-10856.
- Mirzabekov, A. D., Bavykin, S. G., Belyavsky, A. V., Karpov, V. L., Preobrazhenskaya, O. V., Shick, V. V., & Ebralidse, K. K. (1989) *Methods Enzymol.* 170, 386-408.
- Moyzis, R. K., Buckingham, J. M., Cram, L. S., Dani, M., Deaven, L. L., Jones, M. D., Meyne, J., Ratliff, R. L., & Wu, J. R. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 6622-6626.
- Nagai, K., Oubridge, C., Jessen, T. H., Li, J., & Evans, P. R. (1990) *Nature* 348, 515-520.
- O'Connor, T. P., & Coleman, J. E. (1983) *Biochemistry* 22, 3375-3381.
- Paradiso, P. R., Nakashima, Y., & Konigsberg, W. (1979) *J. Biol. Chem.* 254, 4739-4744.
- Prescott, D. M. (1983) *Mod. Cell Biol.* 2, 329-352.
- Price, C. M. (1990) *Mol. Cell. Biol.* 10, 3421-3431.
- Price, C. M., & Cech, T. R. (1987) *Genes Dev.* 1, 783-793.
- Price, C. M., Skopp, R., Krueger, J., & Williams, D. (1992) *Biochemistry* 31, 10835-10843.
- Prigodich, R. V., Casa-Finet, J., Williams, K. R., Konigsberg, W., & Coleman, J. E. (1984) *Biochemistry* 23, 522-529.
- Raghuraman, M. K., & Cech, T. R. (1989) *Cell* 59, 719-728.
- Raghuraman, M. K., Dunn, C. J., Hicke, B. J., & Cech, T. R. (1989) *Nucleic Acids Res.* 17, 4235-4253.
- Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Shamoo, Y., Williams, K. R., & Konigsberg, W. H. (1988) *Proteins* 4, 1-6.
- Shaw, A. A., Falick, A. M., & Shetlar, M. D. (1992) *Biochemistry* 31, 10976-10983.
- Shetlar, M. D., Hom, K., Carbonne, J., Moy, D., Steady, E., & Watanabe, M. (1984) *Photochem. Photobiol.* 39, 135-140.
- Steele, J. C. H. J., & Nielsen, T. B. (1978) *Anal. Biochem.* 84, 218-224.
- Vernick, K. D., & McCutchan, T. F. (1988) *Mol. Biochem. Parasitol.* 28, 85-94.
- Wang, W., Skopp, M., Scofield, R., & Price, C. (1992) *Nucleic Acids Res.* 20, 6621-6629.
- Watson, J. D. (1972) *Nature, New Biol.* 239, 197-201.
- Weintraub, H. (1974) *Cold Spring Harbor Symp. Quant. Biol.* 38, 247-256.
- Wellinger, R. J., Wolf, A. J., & Zakian, V. A. (1993) *Cell* 72, 51-60.
- Willis, M. C., Hicke, B. J., Uhlenbeck, O. C., Cech, T. R., & Koch, T. (1993) *Science* 262, 1255-1257.
- Wright, J. H., Gottschling, D. E., & Zakian, V. A. (1992) *Genes Dev.* 6, 197-210.
- Yu, G. L., Bradley, J. D., Attardi, L. D., & Blackburn, E. H. (1990) *Nature* 344, 126-132.
- Zakian, V. A. (1989) *Annu. Rev. Genet.* 23, 579-604.