Revealing Novel Telomere Proteins Using *in Vivo* Cross-linking, Tandem Affinity Purification, and Label-free Quantitative LC-FTICR-MS*[S]

Thalia Nittis‡, Lionel Guittat‡§, Richard D. LeDuc¶, Ben Dao‡, Julien P. Duxin‡, Henry Rohrs‖, R. Reid Townsend‡¶**, and Sheila A. Stewart‡¶‡‡

Telomeres are DNA-protein structures that protect chromosome ends from the actions of the DNA repair machinery. When telomeric integrity is compromised, genomic instability ensues. Considerable effort has focused on identification of telomere-binding proteins and elucidation of their functions. To date, protein identification has relied on classical immunoprecipitation and mass spectrometric approaches, primarily under conditions that favor isolation of proteins with strong or long lived interactions that are present at sufficient quantities to visualize by SDS-PAGE. To facilitate identification of low abundance and transiently associated telomere-binding proteins, we developed a novel approach that combines *in vivo* protein-protein cross-linking, tandem affinity purification, and stringent sequential endoprotease digestion. Peptides were identified by label-free comparative nano-LC-FTICR-MS. Here, we expressed an epitope-tagged telomere-binding protein and utilized a modified chromatin immunoprecipitation approach to cross-link associated proteins. The resulting immunoprecipitant contained telomeric DNA, establishing that this approach captures *bona fide* telomere binding complexes. To identify proteins present in the immunocaptured complexes, samples were reduced, alkylated, and digested with sequential endoprotease treatment. The resulting peptides were purified using a microscale porous graphite stationary phase and analyzed using nano-LC-FTICR-MS. Proteins enriched in cells expressing HA-FLAG-TIN2 were identified by label-free quantitative analysis of the FTICR mass spectra from different samples and ion trap tandem mass spectrometry followed by database searching. We identified all of the proteins that constitute the telomeric shelterin complex, thus validating the robustness of this approach. We also identified 62 novel telomere-binding proteins. These results demonstrate that DNA-bound protein complexes, including those present at low molar ratios, can be identified by this approach. The success of this approach will allow us to create a more complete understanding of telomere maintenance and have broad applicability. *Molecular & Cellular Proteomics 9: 1144–1156, 2010.*

Numerous redundant systems exist to maintain the genome and ensure proper segregation of genetic material upon cellular division. Elucidation of the molecular mechanisms that constitute these systems is an area of intense inquiry. In model systems, elegant genetic approaches have been used extensively to identify proteins and interrogate their role in these mechanisms. Unfortunately, mammalian systems are refractory to similar approaches, and thus protein identification has relied heavily on homology searches and mass spectrometry. For this reason, the development of isolation procedures and refined mass spectrometric approaches capable of identifying proteins within large protein complexes, including those present as transient interactors and in substoichiometric quantities, is an important area of research. Previous studies have successfully utilized quantitative proteomics with stable isotopic peptide labeling to identify specific components of cellular macromolecular complexes by affinity purification (1–6). More recently, high resolution mass spectrometry with label-free quantification has been shown to improve and extend quantitative proteomics toward comprehensive analysis of protein complexes (7).

Telomeres are DNA-protein structures located at the ends of linear eukaryotic chromosomes (see Fig. 1). The DNA portion of telomeres consists of a double-stranded region and a single-stranded 3′ overhang, both composed of repetitive non-coding G-rich sequences (TTAGGG). In addition to the DNA component, proteins bind the telomere and contribute to its stability. Six core proteins (TRF1, TRF2, POT1, TIN2, RAP1, and ACD/TPP1), collectively known as the shelterin (or telosome) complex, are constitutively present at the telomere (for reviews, see Refs. 8 and 9). Together, the telomeric DNA and shelterin complex maintain a "capped" or functional telomere that protects the end of the chromosome by distinguishing it from a *bona fide* double strand DNA break (10). When telomeres become uncapped or "dysfunctional," they no longer

carry out this protective function, rendering the chromosome ends susceptible to DNA repair enzymes. In the absence of functional checkpoints, uncapped telomeres can lead to end-to-end fusions that drive genomic instability, a hallmark of human cancer (11).

Recent work has revealed that in addition to the shelterin complex a growing list of proteins associate with the telomere and play essential roles in telomere maintenance (a subset of these proteins, colored in *gray*, is depicted in Fig. 1). Paradoxically, many of these proteins play roles in DNA repair and recombination. These proteins include the MRE11-Rad50-Nbs1 complex involved in recombinational repair (12); Ku70 and Ku80, which are members of the non-homologous end joining complex (13); the ERCC1/XPF nucleotide excision repair endonuclease (14); and the ataxia telagiectasia mutated (ATM) kinase (12, 15). Additional proteins have been found at the telomere in low stoichiometric ratios, including telomerase, which binds the telomere during S phase and adds telomeric repeats to the ends of the chromosomes (16, 17). The Werner helicase is also present at the telomeres during S phase where it plays an important role in lagging strand DNA replication (18). Despite the plethora of proteins known to bind to the telomere, many proteins that act in a transient manner and/or are present in substoichiometric quantities remain to be identified.

To identify novel telomere-binding proteins, we developed a method that involves chemical cross-linking of protein complexes in live cells to capture transient interactions followed by affinity purification of the cross-linked telomere complex with an epitope-tagged telomeric protein, TIN2. Using the affinity-captured protein preparations, we optimized cross-link reversal, sequential endoprotease digestion, and microscale solid phase peptide purification. The peptide pools were analyzed using nano-LC-FTICR-MS. Comparative quantitative analysis of affinity-purified proteins from cells overexpressing the epitope-tagged TIN2 and control cells was performed using the peptide ion currents at accurate m/z values from the aligned LC-MS chromatograms across multiple samples. The proteins were identified using tandem MS with spectral matching against protein databases. Using this approach, we identified the six members of the shelterin complex and other proteins previously reported to bind to the telomere. We also identified a novel group of candidate telomere-binding proteins that were significantly enriched in samples expressing epitope-tagged TIN2 (HA[1]-FLAG-TIN2) compared with non-expressing control cells. Importantly, the presence of telomeric DNA in our immunoprecipitants from cells expressing HA-FLAG-TIN2 but not in control cells demonstrates that it is possible to identify proteins bound to DNA by utilizing a protein-protein cross-linking reagent. This strategy will prove versatile for the identification of other proteins found in large protein complexes as well as bound to DNA.

EXPERIMENTAL PROCEDURES

Cell Culture and Generation of Vectors—293T cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (Sigma-Aldrich), 100 units/ml penicillin, and 100 units/ml streptomycin. The *TIN2* gene containing an HA and triple FLAG epitope tag was cloned into pShuttle-CMV vector (Clontech) between the NotI and HindIII sites and then reassorted into the AdEasy-1 vector, and large scale adenovirus stock was prepared according to the manufacturer's directions (Stratagene, La Jolla, CA). The HA-FLAG-tagged GFP construct was produced in an analogous manner.

Antibodies and Western Blot Analysis—Epitope-tagged TIN2 was detected with mouse monoclonal anti-FLAG M2 antibody (Sigma). TRF2 was detected with mouse monoclonal anti-TRF2 antibody (Upstate, Lake Placid, NY). Bound primary antibodies were detected with horseradish peroxidase-conjugated goat anti-mouse (Sigma) secondary antibodies, and proteins were visualized using ECL reagents.

Immunofluorescence—Cells transiently transfected with pShuttle-CMV HA-FLAG-TIN2 plasmid were fixed in fresh 4% paraformaldehyde, permeabilized in 0.2% Triton X-100, and rinsed with PBS, and following blocking in PBS containing 3% nonfat dry milk and 2% BSA, the cells were incubated with anti-FLAG primary antibody (1:1000) followed by goat anti-mouse antibody conjugated to Alexa 488 fluorophore (1:1000) (Molecular Probes, Eugene, OR). For fluorescence *in situ* hybridization (FISH) analysis, antibody-stained cells were fixed a second time with 4% paraformaldehyde and then incubated with 2× SSC and RNase A for 45 min at 37 °C. Coverslips were then dehydrated and hybridized with a Cy3-labeled peptide nucleic acid (PNA) probe $(CCCTAA)_{18}$ (Panagene, Daejeon, Korea) in 50% formamide. Coverslips were washed, mounted with Vectashield mounting medium (Vector Laboratories, Burlingame, CA) containing 2 μg/ml 4'6-diamidino-2-phenylindole (DAPI; Molecular Probes) to stain the DNA, and analyzed using a Nikon 90i microscope.

Chromatin Immunoprecipitation (ChIP)—ChIP experiments were performed as described previously with some modifications (19). Briefly, 293T cells were cross-linked with 1% formaldehyde or 2 mM dithiobis(succinimidyl)propionate (DSP) for 1 h at RT, and the cross-linking was stopped by the addition of 0.125 M glycine. Cells cross-linked with DSP had been infected with HA-FLAG-TIN2 adenovirus. Cells were lysed and sonicated to obtain DNA fragments of 500–1000 base pairs. For each IP, 1 mg of protein lysate was used with 4 μg of the specified antibody (mouse IgG (Jackson ImmunoResearch Laboratories, West Grove, PA) used as control, mouse monoclonal anti-FLAG M2 antibody (Sigma), or mouse monoclonal anti-TRF2 antibody (Upstate)). The immunoprecipitated DNA was extracted and transferred to membranes Hybond-XL, GE Healthcare, Piscataway, NJ using a dot blot apparatus. Duplicate membranes were hybridized with a radioactive telomeric repeat probe or an Alu repeat probe. Dilutions of the lysates were also spotted nylon to the membranes to calculate the total amount of telomeric or Alu repeat DNA present in the IPs. The amount of telomeric DNA immunoprecipitated with each antibody was calculated as the percentage of the total or input DNA present in the IP.

Preparation of Cell Lysates—Approximately 1.25×10^9 293T cells were infected with TIN2 or GFP adenovirus or mock-infected. 48 h postinfection, cells were harvested, washed twice with PBS, and incubated for 5 min in 30 ml of PBS containing 0.055 mg/ml digitonin

---

[1] The abbreviations used are: HA, hemagglutinin; DSP, dithiobis(succinimidyl)propionate; ChIP, chromatin immunoprecipitation; IP, immunoprecipitation; TAP, tandem affinity purification; FISH, fluorescent *in situ* hybridization; MS1, parent scan mass spectra; MS2, tandem mass spectra; t_r, retention time; GFP, green fluorescent protein; DAPI, 4'6-diamidino-2-phenylindole; RT, room temperature; TCEP, tris(2-carboxyethyl)phosphine; LTQ, linear trap quadrupole; hnRNP, heterogeneous nuclear ribonucleoprotein; TERT, reverse transcriptase component of telomerase.

(Sigma) to permeabilize the cell membranes (20). After centrifugation, proteins were cross-linked with 2 mM DSP (Pierce) for 1 h at RT with rotation. A 500 mM DSP stock solution was prepared fresh in DMSO just prior to use, and 120 μl of this was dissolved in 30 ml of PBS, 20% DMSO at RT. The dissolved DSP solution was added directly to the cell pellet at RT and vortexed immediately to avoid precipitation. The cross-linking reaction was quenched by incubating the cells in 0.125 M glycine for 10 min (1.5 ml of a 2.5 M stock added directly to the mixture). Cells were washed once with 30 ml of PBS, 20% DMSO and twice more with 30 ml of PBS. Cells were then resuspended in 15 ml of lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, pH 8.0, 0.1% SDS, 1% Triton X-100, and freshly added protease inhibitors: 1 μg/ml aprotinin, 1 μg/ml leupeptin, 1 μg/ml pepstatin, and 1 mM phenylmethylsulfonyl fluoride (all from Sigma)). Cells were lysed on ice for 1 h followed by repeated sonication (six times for 30 s with 30-s rests on ice in between using a Misonix Sonicator 3000, D.A.I Scientific Equipment, Mundelein, IL) to obtain average chromatin fragments of 500–1000 base pairs. The cell lysates were clarified by centrifugation, and protein content was determined by the Bradford assay (Bio-Rad).

Immunoprecipitation—120 mg of total protein was used per sample (control and HA-FLAG-TIN2-infected) for the large scale immunoprecipitation experiment, or 10 mg was used for the HA-FLAG-GFP-infected cells. The clarified lysates were diluted with lysis buffer to a concentration of 2.4 mg/ml and precleared by incubating with 0.5 ml of mouse IgG-agarose (Sigma A0919) on a rotator at 4 °C for 2 h. After centrifugation, the lysates were added to 1 ml of anti-FLAG M2 affinity gel (Sigma) and incubated on a rotator overnight at 4 °C. Lysates were poured into columns (Poly Prep, 0.8 × 4 cm; Bio-Rad) and allowed to drain, and the beads were washed with 15 ml of wash buffer (100 mM KCl, 20 mM Tris-HCl, pH 8.0, 5 mM MgCl$_2$, 0.2 mM EDTA, 10% glycerol, 0.1% Tween, 10 mM 2-mercaptoethanol added fresh) (21). The bottoms of the columns were capped, and 1.5 ml of wash buffer containing 2 mg/ml HPLC-purified FLAG peptide (Tufts Core Facility, Boston, MA) was added. The tops of the columns were capped, and proteins were eluted off the beads by rotating the columns at RT for 3.5 h. After the eluate was allowed to drain from the column, the elution step was repeated, and the eluates were combined. The eluates were next applied to 200 μl of EZview Red anti-HA affinity gel (Sigma E6779) and rotated at 4 °C overnight. After pelleting by centrifugation at maximum speed and removing the supernatant, the beads were washed three times with 1 ml of wash buffer. Protein complexes were eluted with 300 μl of wash buffer containing 1 mg/ml HPLC-purified HA peptide (Tufts Core Facility, Boston, MA). The elution step was repeated, and eluates were combined. Each elution step was performed at RT for 3.5 h with rotation. The final eluates were forced through Filter-in-a-tips (Glygen Corp., Columbia, MD) to remove any remaining beads. Rapigest SF reagent (Waters) was added to a final concentration of 0.02%, and tris(2-carboxyethyl)phosphine (TCEP; Sigma) was added to a final concentration of 50 mM. The pH was adjusted to 8.0 with KOH, and the samples were incubated for 1 h at RT to reverse the cross-links. Samples were next placed in dialysis cassettes (10,000 molecular weight cutoff; Pierce) and dialyzed overnight at 4 °C against 1 liter of PBS. Samples were concentrated to a volume of ~200 μl using Pierce's Slide-A-Lyzer concentrating solution and stored at −20 °C.

Protein Digestion and Mass Spectrometry—All reagents were made on the day of sample preparation. Solutions were prepared in H$_2$O in 1.5-ml microcentrifuge tubes unless otherwise indicated. Enzyme solutions were rehydrated in the vendor vials. Bovine serum albumin (100 ng) was first added to each sample to serve as an internal digestion control. Samples were precipitated using the 2-D protein clean-up kit (GE Healthcare) according to the manufacturer's instructions. The pellet was dissolved in 40 μl of 9 M urea and

aliquoted into two 1.5-ml microcentrifuge tubes. Samples (20 μl in 9 M urea) were reduced with 5 mM TCEP (2.2 μl of 50 mM stock) at pH 8.0 for 30 min at RT and alkylated with 10 mM iodoacetamide (1 μl of 250 mM stock; Bio-Rad) in the dark at RT for 30 min. TCEP and iodoacetamide were quenched by adding 5 mM DTT (1 μl of 125 mM stock; Bio-Rad) and incubating at RT for 10 min. The reduced and alkylated proteins were digested with 1 μg of endoproteinase Lys-C (2 μl of a 0.5 μg/μl stock; Roche Applied Science) overnight at 37 °C (22). Samples were diluted with 63.8 μl of H$_2$O to reduce the concentration of urea to 2 M and digested with 4 μg of trypsin (10 μl of a 0.4 μg/μl stock; Sigma) overnight at 37 °C. The pH was adjusted to 8.3 with 0.75 M NH$_4$OH. Trypsin (20 μg) was dissolved in 50 μl of 1 mM triethylammonium bicarbonate and 1 μl of 0.075 M NH$_4$OH, pH 8.0. Peptides were acidified with 5.5 μl of formic acid (Sigma) and extracted six times with 10–200-μl NuTip porous graphite carbon wedge tips (Glygen Corp.) according to the manufacturer's directions. Peptides were eluted into a 1.5-ml autosampler vial with 60% acetonitrile (Burdick & Jackson, Muskegon, MI) in 0.1% formic acid. The peptide digests were first evaluated for quality and detergent contaminants using MALDI-TOF/TOF (23) followed by nano-LC-FTICR-MS analysis. For MALDI-TOF/TOF analysis, the peptide sample (0.5 μl) was mixed with an equal volume of MALDI matrix solution (Agilent Technologies, Santa Clara, CA) prior to spotting. For nano-LC-FTICR-MS analysis, the peptide sample was dried and immediately dissolved in 10 μl of aqueous acetonitrile/formic acid (1%/1%).

Nano-LC-FTICR-MS—Analysis was performed using a hybrid linear quadrupole ion trap Fourier transform-ion cyclotron resonance mass spectrometer (LTQ-FT, Thermo-Fisher, San Jose, CA). The nanoflow HPLC system (Nano LC-1D, Eksigent, Dublin, CA) was interfaced to the LTQ-FT with a nanospray source (PicoView PV550, New Objective, Woburn, MA). Sample injection was performed with an autosampler (AS1, Eksigent). Reverse phase C$_{18}$ columns (75 × 10 μm; PicoFrit, New Objective) were used for gradient separation. Both the aqueous phase (LC-MS water, Riedel-de Haen) and organic phase (LC-MS acetonitrile, Riedel-de Haen) were modified with 0.1% formic acid (Sigma-Aldrich). 5-μl samples were loaded at 1 μl/min from a 10-μl loop. After an initial aqueous wash at 260 nl/min, the organic phase for the analytical gradient was increased at 0.6–1.2%/min up to 70% organic also at 260 nl/min. The nanospray source was operated between 1.8 and 2.3 kV with sheath gas, and the spray was visually optimized with ~20% organic flow at 260 nl/min. The capillary temperature was 240 °C. The LTQ-FT was operated in both the data-dependent and targeted analysis (*i.e.* parent ion inclusion list) modes. Full MS scans were acquired at 100,000 resolving power (m/z 421.75) with a target value of 1,000,000. The ion trap MS^n target was 20,000. For data-dependent scans, the six most intense ions were selected for wideband collisional activation and detection in the ion trap (parent threshold, 1000; isolation width, 2.0 Da; normalized collision energy, 35; activation Q, 0.250; activation time, 30 ms). Dynamic exclusion was used to expand selection. For targeted acquisitions, comparative analysis of MS data was used to select parent ions for an inclusion list. Earlier studies had shown that optimum performance was achieved for low abundance species when the MS^n trap times and thresholds were optimized for sample concentration. In these experiments, the time was 500 ms, and the intensity minimum was 500. All other parameters remained the same other than only selecting parent masses from an inclusion list.

Data Analysis—The MS1 and MS2 data from the LTQ-FT mass spectrometer (ThermoFisher, San Diego, CA) were acquired in the profile mode. To perform quantitative label-free analysis, the nano-LC-FTICR-MS data from separate LC analyses of the peptide mixtures from control and cells overexpressing HA-FLAG-TIN2 were analyzed using Rosetta Elucidator software (version 3.0; Rosetta Biosoftware, Seattle, WA) (24). The "raw" files were imported for feature

FIG. 1. **Fluorescent *in situ* hybridization reveals presence of telomeres at termini of human chromosomes.** *Top panel*, representative metaphase spread from human cells. FISH analysis reveals the presence of telomeres (*red*) and centromeres (*green*), and chromosomal DNA (*blue*) was detected by DAPI staining. *Bottom panel*, schematic drawing of a telomere loop (*T-Loop*) showing the shelterin core complex (TRF1, TRF2, POT1, TIN2, RAP1, and TPP1) as well as a subset of known telomere-binding proteins (in *gray*). *Question marks* indicate that more telomere-binding proteins remain to be identified. *WRN*, Werner, *BLM*, Bloom, and *XPF*, xeroderma pigmentosum type F.
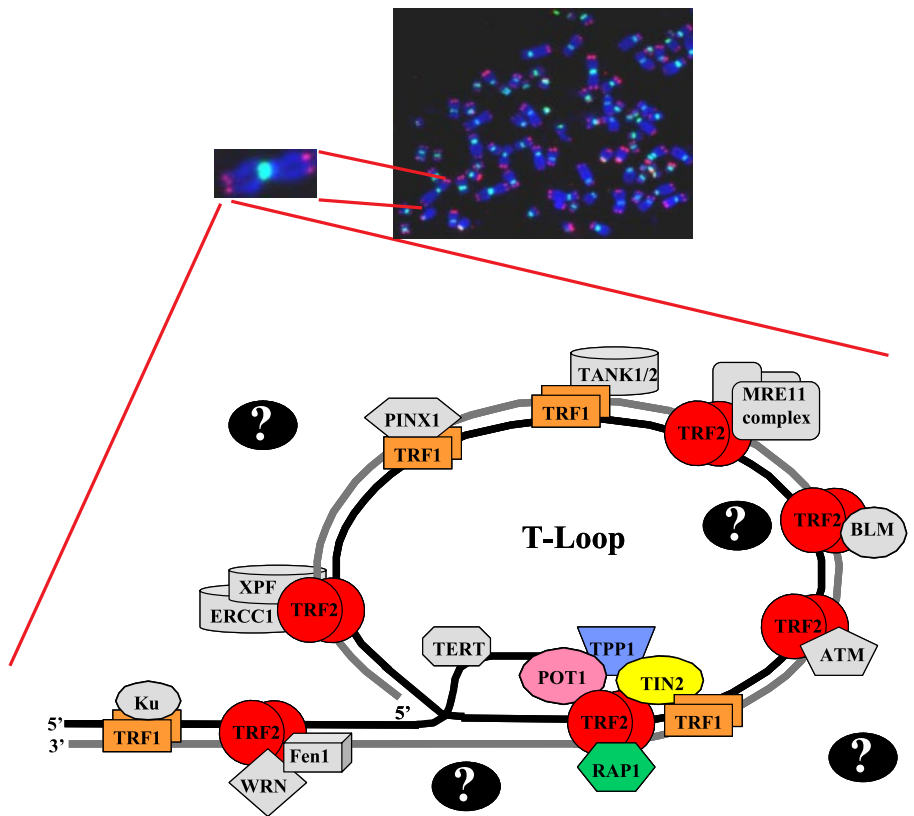
retention time alignment, definition, and volume determination within the selected LC-MS time windows. The "PeakTeller" algorithm in the software performed background subtraction and smoothing in both the retention time and *m/z* dimensions using scores of 0 and 0.5, respectively. The "adaptive alignment" option was selected, and the following additional parameters were used during the alignment process: instrument mass accuracy, 10 ppm; "expected retention time shift," 2 min; and "noise removal strength" for retention time and $m/z = 1$ for both. The peak width time was set at >0.1 min. Intensity scaling was based on the mean intensity of all quality features (as defined above) and was performed after a 10% outlier trim to correct for variations in the total ion current between individual LC-MS analyses.

For analysis of the tandem spectra from spectral acquisitions in the ion trap (MS2), the raw files were processed using MASCOT Distiller (Matrix Science, Oxford, UK) with the following settings: 1) MS processing: 200 data points per Da; no aggregation method; maximum charge state, 5+; minimum number of peaks, 1; 2) MS/MS processing: 200 data points per Da; time domain aggregation method enabled; minimum number of peaks, 10; precursor charge and *m/z*, "try to redetermine from the survey scan (tolerance, 2.5 Da)"; charge defaults, 2+/3+; maximum charge state, 2+; 3) time domain parameters: minimum precursor mass, 700; maximum precursor mass, 16,000; precursor *m/z* tolerance for grouping, 0.1; maximum number of intermediate scans, 5; minimum number of scans in a group, 1; peak picking: maximum iterations, 500; correlation threshold, 0.90; minimum signal to noise, 3; minimum peak *m/z*, 50; maximum peak *m/z*, 100,000; minimum peak width, 0.001; maximum peak width, 2; expected peak width, 0.01. The files from the MASCOT Distiller output (mgf) for each individual LC-MS analysis were concatenated and searched against the National Center for Biotechnology Information non-redundant (NCBInr) database (downloaded July 8, 2008) with a human taxonomy filter. All peptide identifications were done

using MASCOT version 2.2.04 with the following parameters: enzyme, trypsin; MS tolerance, 10 ppm; MS/MS tolerance, 0.8 Da with a fixed carbamidomethylation of Cys residues and the following variable modifications: methionine oxidation, pyro-Glu (N terminus), and deamidation (Gln and Asn residues); maximum missed cleavages, 5; charge states, 1+, 2+, and 3+. This helped reduce errors from mass inference calculations to determine the ^{12}C isotopic signal from the parent isotopic cluster. The proteins with single tandem spectra with MASCOT ion scores <40 were manually interpreted and annotated as shown in supplemental Table 1.

RESULTS

Expression and Purification of TIN2 Complexes—TIN2 is critical for the assembly of the six-protein shelterin complex and forms a bridge between the TRF1 and TRF2 subcomplexes and TPP1, which interacts with the telomeric single strand-binding protein POT1 (Fig. 1, *lower panel*) (25, 26). To isolate new telomere-interacting proteins, we epitope-tagged (5′ FLAG-HA) and expressed TIN2. To confirm that the modified TIN2 protein localized to telomeres, we carried out immunofluorescence. Utilizing an anti-FLAG antibody that specifically recognized the ectopically expressed TIN2 protein, we observed the presence of distinct dots that colocalized with the telomeric DNA within the nucleus, indicative of telomere binding (Fig. 2). We next isolated TIN2-containing complexes using a modified ChIP approach as detailed in Fig. 3.

To facilitate isolation of TIN2-interacting proteins, including those bound transiently or weakly to the telomere, we treated
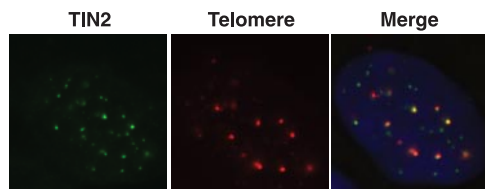
FIG. 2. **Immunofluorescence localization of HA-FLAG-TIN2.** Cells transiently expressing HA-FLAG-TIN2 were fixed on slides, permeabilized, and examined following treatment with an antibody against FLAG and FISH analysis for telomere sequences as described under "Experimental Procedures." DNA was stained with DAPI (*blue*). The *left panel* depicts HA-FLAG-tagged TIN2 (*green*), the *middle panel* depicts telomere staining (*red*) in the same cell, and the *right panel* is the merged image showing colocalization within the nucleus as expected.
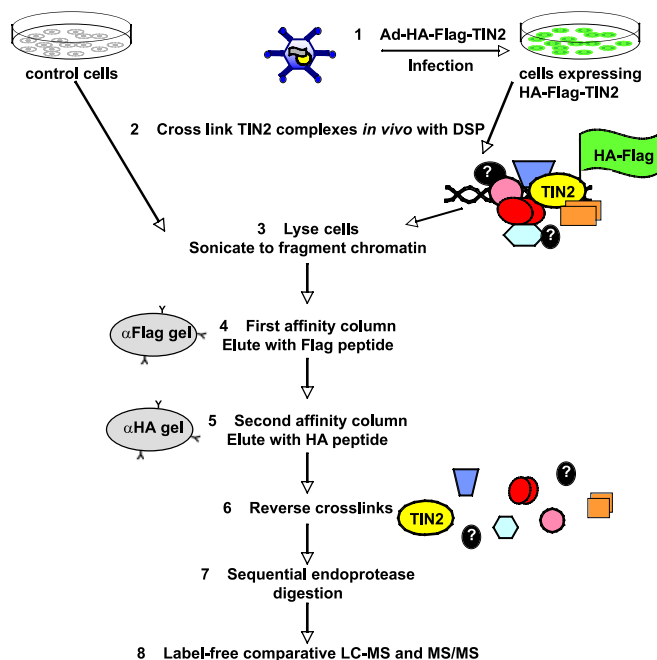

FIG. 3. **Flow chart of experimental procedure.** Control 293T cells or 293T cells infected with HA-FLAG-TIN2 or HA-FLAG-GFP (not shown) adenovirus were treated with DSP to cross-link protein complexes, lysed, and sonicated to fragment the chromatin. Cleared lysates were applied to FLAG and HA affinity columns sequentially to isolate TIN2-containing protein complexes. Cross-links were reversed, and eluted proteins were digested sequentially with two endoproteases and analyzed by label-free comparative LC-MS and MS/MS.

live cells with a cross-linking reagent prior to cell lysis and sonication (Fig. 3, *step 2*). Previous work demonstrated that novel proteins could be identified by mass spectrometry following formaldehyde cross-linking and immunoprecipitation of Ras-containing complexes (27). Therefore, we hypothesized that a similar approach would yield novel telomere-binding proteins. We first optimized cross-linking conditions (Fig. 4A). Because we wished to capture TIN2-containing complexes that were present at the telomere, we utilized 1% formaldehyde to cross-link protein complexes to telomeric
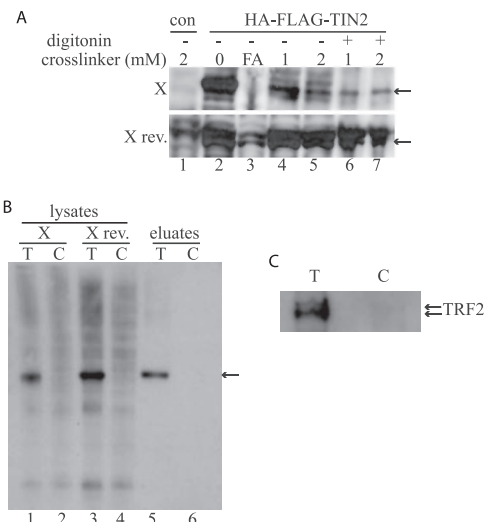

FIG. 4. **Immunoblot analysis of TIN2 and TRF2 from cross-linked lysates.** *A*, control 293T cells (*C*) and 293T cells expressing HA-FLAG-TIN2 (*T*) were treated with cross-linkers (DSP at 0, 1, or 2 mM and formaldehyde (*FA*) at 1%) for 1 h at RT. Some cells were pretreated with digitonin for 5 min prior to cross-linking. Cells were washed and lysed, and equivalent amounts of protein lysate were loaded onto protein gels and subjected to Western blot analysis with anti-FLAG antibody. In the *top panel*, the cross-links were not reversed (*X*), whereas in the *bottom panel* they were reduced by boiling for 30 min in the presence of 10% 2-mercaptoethanol (*X rev.*). The *arrows* point to HA-FLAG-TIN2. *B*, immunoblot showing the successful purification of HA-FLAG-TIN2 from cross-linked 293T cells. *Lanes 1* and *2* contain total protein lysates (25 μg) in which the cross-links were not reversed (*X*), whereas *lanes 3* and *4* contain samples in which the cross-links were reversed (*X rev.*). *Lanes 5* and *6* contain the final eluates (1/60) after the two-step purification. *C*, one-quarter of the final eluate was concentrated, the cross-links were reversed, and the sample was examined for the presence of TRF2 by immunoblot analysis. *Arrows* point to TRF2, present as a doublet only in the sample prepared from HA-FLAG-TIN2-expressing cells.

DNA. Indeed, ChIP has been used to demonstrate telomeric association for numerous proteins (19, 28). Formaldehyde treatment resulted in effective cross-linking of the epitope-tagged TIN2 as evidenced by the appearance of large TIN2-containing complexes that do not migrate through SDS-polyacrylamide gels (Fig. 4A, *top panel*, compare *lanes 1–3*). We next reversed the formaldehyde cross-links in our immunoprecipitants by boiling the samples in the presence of 10% 2-mercaptoethanol, and the resulting mixtures were analyzed by Western blot analysis. Western blot analysis revealed that both the reversal of the formaldehyde cross-links and immunoprecipitation of formaldehyde-cross-linked HA-FLAG-TIN2 with the FLAG antibody were inefficient (Fig. 4A, *top* and *bottom panels*, compare *lanes 1–3*). The observed low yields of cross-linking reversal and the described chemical complexity of amino acid side chain modifications following formaldehyde treatment (29) indicated that another cross-linking reagent would be more advantageous for quantitative MS analysis and protein identification using tandem mass spectrometry.
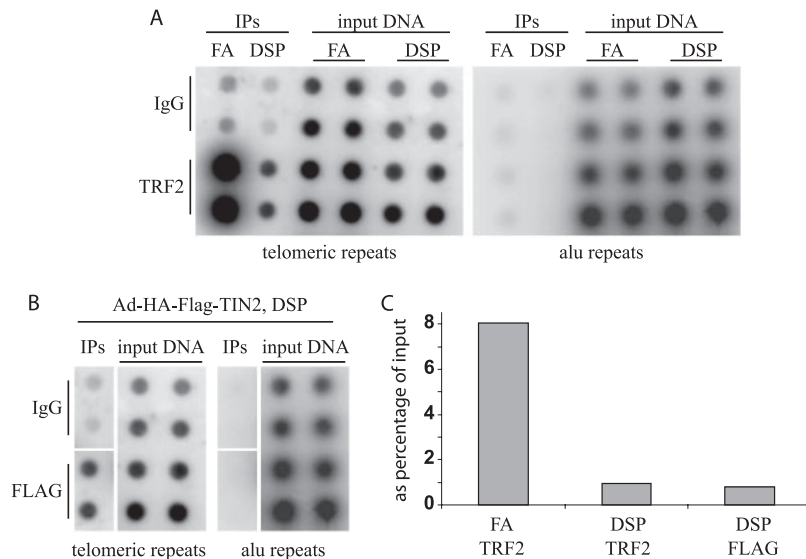
FIG. 5. **Telomere-specific localization of HA-FLAG-TIN2.** *A*, ChIP experiments were carried out with lysates from 293T cells. Cells were cross-linked *in vivo* with either formaldehyde (*FA*) or DSP. Protein-DNA complexes were immunoprecipitated with IgG from mouse (negative control) or anti-TRF2 antibodies. Each IP was carried out in duplicate. Immunoprecipitated DNA was extracted from each sample and spotted onto nylon membranes using a dot blot apparatus. Duplicate dot blots were probed with either telomeric or Alu repeat sequences. Each blot also contains DNA dilutions (1:20 and 1:50) from input DNA that was not immunoprecipitated to allow for the calculation of the amount of telomeric or Alu repeat DNA immunoprecipitated with each antibody. *B*, ChIP experiments were carried out with lysates from 293T cells infected with HA-FLAG-TIN2 adenovirus and cross-linked with DSP. Protein-DNA complexes were immunoprecipitated with IgG from mouse or anti-FLAG antibodies. Each IP was carried out in duplicate, and the extracted DNA was analyzed as described above. Each blot also contains DNA dilutions (1:20 and 1:50) from input DNA that was not immunoprecipitated. *C*, quantification of the telomeric signal from *A* and *B*. The amount of telomeric DNA immunoprecipitated with the TRF2 or FLAG antibodies is expressed as a percentage of the total telomeric DNA used in each IP.

DSP is a bifunctional, thiol-cleavable protein-protein cross-linker (30). DSP principally reacts with primary amines on the side chains of lysine residues as well as the N termini of polypeptides (31, 32). Given the restricted reactivity of DSP, we postulated that it would allow isolation and identification of telomere-binding proteins. As shown in Fig. 4A (*lanes 4–7*), DSP was an effective, reversible cross-linking reagent for TIN2. A concentration of 2 mM DSP produced maximal cross-linked protein (Fig. 4A, *lanes 4–7*, compare 1 and 2 mM DSP concentrations). At concentrations higher than 2 mM, DSP precipitated, and cross-linking was inefficient (data not shown). To facilitate entry of DSP into the nucleus and further increase cross-linking efficiency of TIN2-containing complexes, we examined the effect of treating cells with digitonin prior to addition of DSP. We found that treatment with digitonin (20) prior to DSP addition increased the amount of cross-linked protein (Fig. 4A, *lanes 6* and *7*). However, as expected, the amount of total protein cross-linked with DSP was less than that cross-linked with formaldehyde (Fig. 4A, *top panel*, compare *lanes 4–7* with *lane 3*). Despite the reduced efficiency of protein cross-linking observed upon DSP treatment, ChIP experiments (Fig. 5) showed that telomeric DNA was specifically immunoprecipitated with both the anti-TRF2 and anti-FLAG antibodies (Fig. 5, *A* and *B*, *left panels*, and *C*). Thus, despite the fact that DSP is not known to cross-link proteins to DNA directly, telomeric DNA was immu-

noprecipitated with TIN2 (Fig. 5*B*) albeit to a lesser degree (Fig. 5*C*) than when formaldehyde was used. As expected, immunoprecipitation with both the TRF2 and HA-FLAG antibodies was specific for telomeric DNA (note the lack of signal from each IP on the Alu blot; Fig. 5, *A* and *B*, *right panels*). Furthermore, the anti-FLAG antibody immunoprecipitated telomeric protein/DNA only from 293T cells expressing the tagged TIN2 construct, demonstrating the specificity of the antibody. Importantly, the isolation of telomeric DNA indicates that we successfully isolated proteins complexed to telomeric DNA, indicating that our approach is capable of identifying protein complexes bound to DNA.

Affinity purification of protein complexes often results in the isolation of nonspecific proteins that complicates the identification of components that constitute the macromolecular complexes *in situ*. One approach that has successfully reduced the purification of nonspecific proteins is the utilization of tandem affinity purification (TAP) (21, 33). We chose to utilize such an approach by adding both FLAG and HA epitopes to the 5′ terminus of TIN2, which allowed us to carry out a two-step TAP procedure (Fig. 3, *steps 4* and *5*). Furthermore, we precleared the cell lysates with IgG-agarose prior to binding to the FLAG and HA affinity columns. SYPRO Ruby-stained gels (data not shown) indicated that both the preclearing step with IgG-agarose and the two-step purification procedure significantly reduced nonspecifically bound proteins (data not shown).
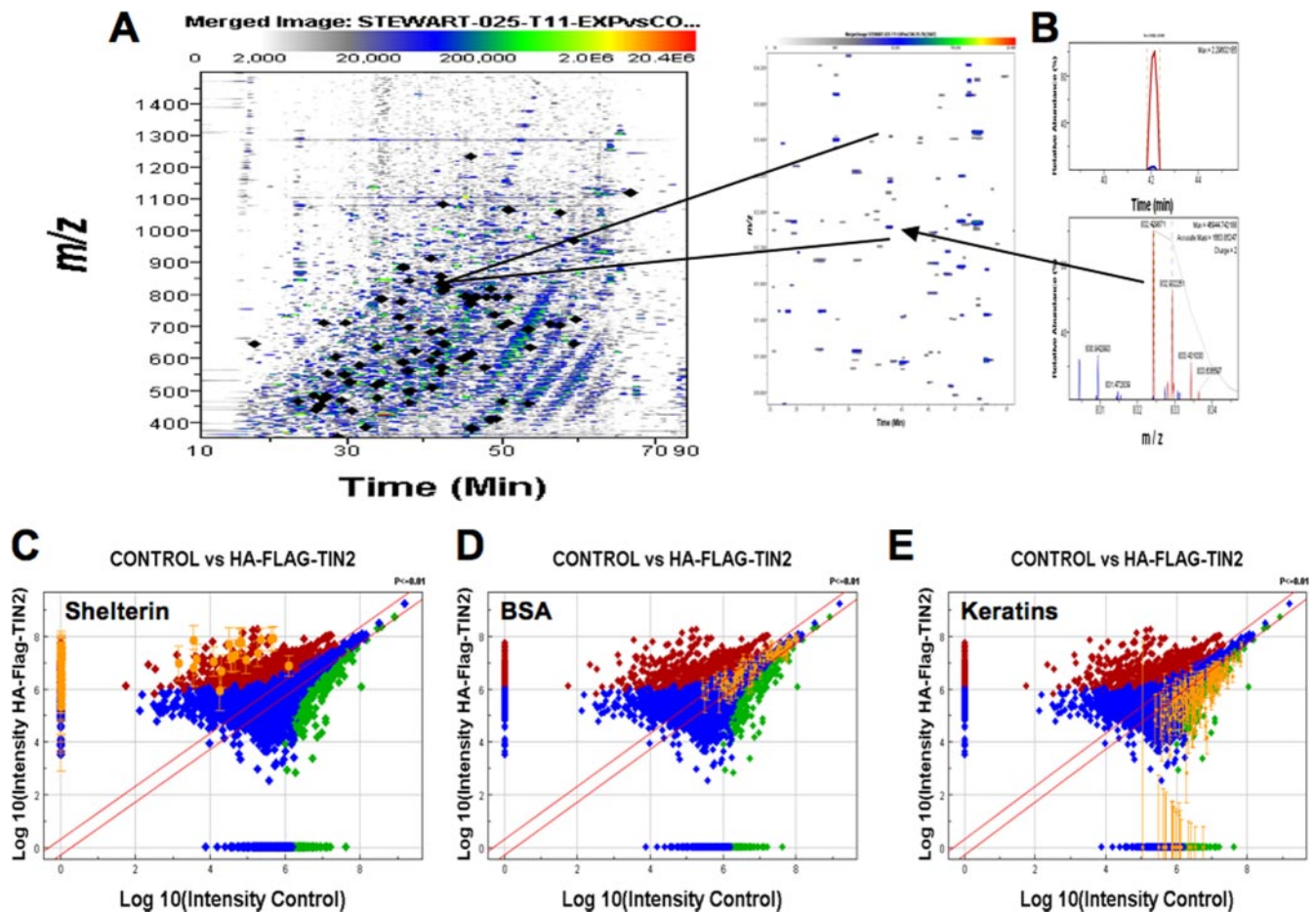
FIG. 6. **Label-free quantitative nano-LC-FT-MS analysis of affinity-purified shelterin peptides.** *A*, signal intensity map (*m/z* 400–1500; retention time window, 12–90 min) of a sequential endoprotease digest (Lys-C and trypsin) of tandem affinity-purified proteins from control and HA-FLAG-TIN2-expressing 293T cells. The locations of the MS2 sequenced shelterin peptides are indicated by *filled diamonds*. *B*, the expanded view of the features representing the selected ion chromatogram (*m/z* 832.430; retention time, 41.38–42.38 min) and isotopic cluster from the $[M + 2H]^{2+}$ ion of the peptide from TRF1 that was identified by tandem MS. *C*, *D*, and *E* are log/linear intensity ratio plots of the concatenated monoisotopic peak of the peptides that were prepared from affinity-isolated proteins (control *versus* HA-FLAG-TIN2) and analyzed by LC-FTICR-MS. The *red*, *green*, and *blue* data points show the peptide isotope groups that were significantly ($p < 0.001$) increased, decreased, or not changed, respectively, using a software (Rosetta Elucidator™) error model. *B*, *C*, and *D* highlight the peptides from the internal standard (spiked BSA), shelterin, and keratin peptides, respectively. The *diagonal lines* demarcate the significance boundaries ($p < 0.01$) from the peak error model in the Rosetta software.

Our initial work indicated that DSP cross-linking was sufficient to isolate telomere-bound protein complexes (Fig. 5). Therefore, having optimized cross-linking conditions and the TAP procedure, we next carried out large scale affinity purifications of HA-FLAG-TIN2 from DSP-cross-linked 293T cell lysates. Fig. 4*B* shows the isolation of tagged TIN2 from 293T cells. In addition, Western blot analysis demonstrated that the known TIN2-interacting partner (TRF2) co-immunoprecipitated from cells ectopically expressing tagged TIN2 (*T lane*) *versus* control cells (*C lane*) (Fig. 4*C*). Having established that our method captured a known TIN2-interacting protein with the bait protein, we proceeded to identify proteins that associated with TIN2 after *in situ* cross-linking and tandem affinity purification. We used nano-LC-FTICR-MS to analyze the peptide pools after denaturation, disulfide reduction, alkylation, sequential endoprotease digestion, and peptide solid phase extraction of the immunocaptured proteins from cells expressing HA-FLAG-TIN2. To identify proteins that nonspecifically bound to the affinity columns under the TAP conditions as well as proteins that copurified with a TAP-tagged GFP protein, we identified the proteins isolated from control 293T cells and cells expressing HA-FLAG-GFP.

Identification of Shelterin Proteins by Nano-LC-FTICR-MS—To determine whether the cross-linking, affinity capture procedure resulted in enhanced isolation of proteins in the telomere core complex (Fig. 1), we compared the ion currents of peptides from affinity-isolated proteins that were from cells expressing the HA-FLAG-TIN2. Fig. 6*A* shows a combined (control and HA-FLAG-TIN2-expressing cells), high resolution

two-dimensional plot of the aligned MS1 peptide ion chromatograms from the affinity-captured proteins. A total of 11,697 features were detected that were concatenated by the described software into 6808 isotope groups (charge state 2+, 3+, 4+, or 5+). Fig. 6*B* details the isotope cluster of the peptide from the telomeric repeat-binding protein factor (TERF1), QSAVTESSGTVSLLR. The details of the MS1 and MS2 data are given in supplemental Table 1. The peptide sequence was deduced from the MS2 spectra and assigned to the MS1 intensity using the accurate mass measurement ($\pm$10 ppm) and retention time ($\pm$4 min) tolerances as described under "Experimental Procedures." The selected ion chromatograms (t_r = 41.82–42.38 min) and the isotope signals from the $[M + H]^{2+}$ ion (m/z = 832.430) are shown in Fig. 6*B*. The ion current of this peptide from the HA-FLAG-TIN2-expressing cells (Fig. 6*B*, *red trace*) was 32-fold greater than that detected from the control cells (Fig. 6*B*, *blue trace*) (see supplemental Table 1 for intensity values), indicating a significant enrichment of this TRF1 peptide in the HA-FLAG-TIN2 sample. Table I shows that in addition to the ectopically expressed TIN2 the five endogenous shelterin proteins (TRF1, TRF2, RAP1, POT1, and TPP1) were identified using MS2 spectra and database searching. The location of the sequenced peptides from the shelterin proteins is indicated by the *black diamonds* in Fig. 6*A*. The log/linear ratio plot of the peptide intensities that were derived from the control and HA-FLAG-TIN2-expressing cells is shown in Fig. 6*C*. The peptide signals that were identified for the shelterin proteins are highlighted in *orange* with many of the peptides not detected in the control sample as shown by the cluster of highlighted points along the *y axis*. A total of 420 unique peptides were sequenced from the six shelterin proteins (supplemental Table 1).

We also assessed the overall variability of the peptide preparation steps from reduction/alkylation, endoprotease digestion, sample preparation, and nano-LC-FTICR-MS analysis by adding BSA after the cross-linking was reversed (Fig. 3, *step 6*) and quantifying the MS1 ion currents from the BSA peptides in the control and HA-FLAG-TIN2 sample. From the spiked BSA (100 ng), we observed 87% sequence coverage from peptides deduced from the MS2 data as detailed in supplemental Table 1. Fig. 6*D* shows the intensity ratio plot with the BSA peptide intensities highlighted and with ratios within the *lines* that demarcate differences with a *p* value of <0.01. We also determined that there was no significant difference at the protein level in the mean of the z score normalized intensities of the BSA peptides (*n* = 96). The coefficient of variation for the paired intensities was 23.25% from the 96 BSA peptides that were identified from MS2 data, similar to errors that have been reported in other comparative, label-free LC-MS studies (for a review, see Ref. 34). The peptides that were increased in the control samples were determined to be primarily from human keratin. These peptides are highlighted in Fig. 6*E*, and the mass spectrometry data are given in supplemental Table 1.

Table I summarizes the proteins that were identified in both the control and HA-FLAG-TIN2 samples. Some of the proteins identified in the HA-FLAG-TIN2 samples were previously associated with telomeres, thus further validating our approach. For example, nucleolin was shown to bind *in vitro* to both single-stranded and duplex telomeric DNA (35, 36). Similarly, the heterogeneous nuclear ribonucleoprotein (hnRNP) family of proteins has widely diverse biological functions, including pre-mRNA processing, transcriptional regulation, recombination, and a telomere maintenance function (37–41). The 14-3-3 protein was shown to bind the human reverse transcriptase component of telomerase (TERT) and is thought to enhance nuclear localization of TERT (42). Heat shock protein 90 (HSP90) was shown to directly interact with TERT (43, 44). In addition to known telomere-binding proteins, we also identified several candidate TIN2- and/or telomere-binding proteins. A subset of these proteins includes DEAD/DEAH box helicases postulated to play roles in DNA replication (encoded by the *DDX17*, *DDX3X*, and *DHX9* genes). We identified proteins involved in chromatin structure and remodeling, such as core histone proteins (encoded by the *HIST1H4A*, *HIST1H2BL*, *HIST2H2AA3*, and *HIST1H1A* genes) and the histone-binding proteins encoded by the *RBBP4* and *RBBP7* genes. Interestingly, we also identified a number of proteins involved in the transmission of genetic information. Given recent reports demonstrating that mammalian telomeres are transcribed, these proteins may prove critical to telomeric transcription (45). These include proteins that play roles in transcription (MATR3), mRNA processing (hnRNPs (A/B, A1, A2B1, C, F, G, H1, K, L, M, R, and U), small nuclear ribonucleoproteins (D1 and D3), NPM1, SERBP1, and TAR DNA-binding protein), and translation (ribosomal proteins (23A, 28, 29, 20, 14, 4X, and 9) and elongation factors EEF1A1 and EEF2). Finally, we isolated a number of proteins with no ascribed functions. The role of these proteins in telomere biology will require further experimentation. Importantly, none of these proteins were found associated with GFP (data not shown), underscoring the specificity of our approach.

DISCUSSION

In this study, we describe a novel in strategy to identify and characterize proteins that comprise the central telomeric complex and other associated proteins. A combination of *in vivo* cross-linking, tandem affinity purification, and label-free, quantitative, high resolution mass spectrometry was used to identify 62 proteins that were enriched in lysates from cells expressing a HA-FLAG-TIN2 protein construct. Using this approach, we identified the core telomere-binding proteins that constitute the shelterin complex (TRF1, TRF2, TIN2, RAP1, POT1, and TPP1). Although classical biochemical approaches have successfully identified telomere-binding proteins (14, 28, 46–48), we aimed to develop a method in which telomere-interacting proteins from low microgram quantities

TABLE I

Proteins identified from 293T cells after tandem affinity purification using HA-FLAG-TIN2 epitope tag

Proteins with peptide ion currents that increased in samples from cells overexpressing the HA-FLAG-TIN2 protein are shown.

Protein number	Gi codes	Gene symbols[a]	Protein descriptors
1	6581060	TINF2/TIN2	TERF1 (TRF1)-interacting nuclear factor 2
2	52627149	TERF2IP/RAP1	Telomeric repeat binding factor 2, interacting protein
3	16877291	ACD/TPP1	Adrenocortical dysplasia homolog
4	5032169	TERF2/TRF2	Telomeric repeat binding factor 2
5	7022355	POT1	Protection of telomeres
6	2507149	TERF1/TRF1	Telomeric repeat binding factor (NIMA-interacting)
7	14141152	HNRPM	Heterogeneous nuclear ribonucleoprotein M
8	133254	HNRPA1	Heterogeneous nuclear ribonucleoprotein A1
9	306875	HNRPC	Heterogeneous nuclear ribonucleoprotein C
10	5031753	HNRPH1	Heterogeneous nuclear ribonucleoprotein H1
11	13384620	HNRPK	Heterogeneous nuclear ribonucleoprotein K
12	32358	HNRPU	Heterogeneous nuclear ribonucleoprotein U
13	4504447	HNRPA2B1	Heterogeneous nuclear ribonucleoprotein A2/B1
14	337451	HNRPAB	Heterogeneous nuclear ribonucleoprotein A/B
15	109124618	HNRPL	Heterogeneous nuclear ribonucleoprotein L
16	5031755	HNRPR	Heterogeneous nuclear ribonucleoprotein R
17	4826760	HNRPF	Heterogeneous nuclear ribonucleoprotein F
18	542850	HNRPG	Heterogeneous nuclear ribonucleoprotein G
19	5902102	SNRPD1	Small nuclear ribonucleoprotein D1
20	4759160	SNRPD3	Small nuclear ribonucleoprotein D3
21	52114	ELAVL1	HuR RNA-binding protein
22	678271	TARDBP	TAR DNA-binding protein
23	4504259	HIST1H2BL	Histone cluster 1, H2b1
24	4504301	HIST1H4A	Histone cluster 1, H4a
25	4885373	HIST1H1A	Histone cluster 1, H1a
26	4504251	HIST2H2AA3	Histone cluster 2, H2aa3
27	189306	NCL	Nucleolin
28	21626466	MATR3	Matrin 3
29	825671	NPM1	Nucleophosmin 1
30	5803225	YWHAE	Tyr 3-monooxygenase/Trp 5-monooxygenase ε
31	2580550	DDX3X	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3
32	100913206	DHX9	DEAH (Asp-Glu-Ala-Asp) box polypeptide 9
33	38201710	DDX17	DEAD (Asp-Glu-Ala-Asp) box polypeptide 17
34	66346685	SERBP1	SERPINE1 mRNA-binding protein 1
35	4826998	SFPQ	Splicing factor proline/glutamine-rich
36	47604944	SCYL2	SCY1-like 2
37	52632381	MAGED1	NRAGE
38	62896605	EEF1A1	Eukaryotic translation elongation factor 1 α 1
39	4503483	EEF2	Eukaryotic translation elongation factor 2
40	404015	RPL23A	Ribosomal protein L23a
41	550023	RPS9	Ribosomal protein S9
42	337930	RPS4X	Ribosomal protein S4, X-linked
43	793843	RPL29	Ribosomal protein L29
44	506697	RPS20	Ribosomal protein S20
45	5032051	RPS14	Ribosomal protein S14
46	2088834	RPSL28	Ribosomal protein L28
47	4505995	PPM1B	Protein phosphatase 1B
48	5032027	RBBP4	Retinoblastoma-binding protein 4
49	4506439	RBBP7	Retinoblastoma-binding protein 7
50	62898329	SEC61B	Protein translocation complex
51	348239	NONO	Non-POU domain-containing, octamer-binding
52	5031741	DNAJA2	DnaJ (Hsp40) homolog, subfamily A, member 2
53	219588	DNAJA1	DnaJ (Hsp40) homolog, subfamily A, member 1
54	7706326	SF2B14	Splicing factor 3B, 14-kDa subunit
55	4758302	ERH	Enhancer of rudimentary homolog
56	4885487	MLF2	Myeloid leukemia factor 2
57	5729877	HSPA1A	Heat shock 70-kDa protein 8
58	4529892	HSPA1B	Heat shock 70-kDa protein 1B

TABLE I—*continued*

Protein number	Gi codes	Gene symbols[a]	Protein descriptors
59	31542947	HSPD1	Heat shock 60-kDa protein 1 (chaperonin)
60	9082289	HSP90AB1	Heat shock protein 90-kDa α, class B, member 1
61	61656603	HSP90AA1	Heat shock protein 90-kDa α, class A, member 1
62	292059	HSPA9	Heat shock 70-kDa protein 9
63	35505	PKM2	Pyruvate kinase isozymes M1/M2
64	16877071	ATP5C1	ATP synthase
65	41872631	FASN	Fatty-acid synthase
66	13491174	MARCKSL1	MARCKS-like 1
67	2627129	UBB	Ubiquitin
68	5031877	LMNB1	Lamin B1
69	4557553	EMD	Emerin
70	62122917	FLG2	Filaggrin family member 2
71	416917	DSG1	Desmoglein 1
72	6005854	PHB2	Prohibitin-2
73	119354192	KRT1	Keratin 1
74	547754	KRT2	Keratin 2
75	908801	KRT82	Keratin type II
76	18999435	KRT5	Keratin 5
77	55956899	KRT9	Keratin 9
78	40354192	KRT10	Keratin 10
79	12803709	KRT14	Keratin 14
80	1195531	KRT16	Keratin 16
81	4557701	KRT17	Keratin 17
82	51921	DCD	Dermcidin
83	57546919	HRNR	Hornerin
84	68563515	KPRP	Keratinocyte protein-rich protein
85	190688	S100A7	S100-A7
86	16924319	ACTG1	Actin, γ1
87	10835838	1FH5-L	Immunoglobulin, chain L
88	425518	AAB28159	Immunoglobulin, heavy chain
89	24158782	1KFA-L	Immunoglobulin, chain L
90	7106439	TUBB	Tubulin-β
91	34740335	TUBA1B	Tubulin-α1b
92	47115317	VIM	Vimentin

[a] Symbols from Human Genome Organisation (HUGO) Gene Nomenclature Database (51). NIMA, never in mitosis gene a, HuR, Hu antigen R, TAR, transactivation-responsive region, NRAGE, neutrophin receptor MAGE, homologue, POU, Pit-1, Oct1, UNC-86, MARCKS, macrophage myristoylated slsnin-tich C kinase substrate.

of highly complex protein mixtures could be identified. In addition, we sought to identify proteins that weakly and/or transiently interacted with the telomere that would have been absent from previous studies. To facilitate these studies, we created an epitope-tagged TIN2 construct that allowed purification of protein complexes using two-tag TAP technology and used the chemical cross-linker DSP to capture telomeric complexes in their *in vivo* state. Indeed, we found that the TAP method reduced the isolation of nonspecifically bound proteins that tend to increase when cross-linkers are used (data not shown and Ref. 49). We optimized a sequential endoprotease digestion and peptide solid phase extraction method for reproducibility and peptide recovery to use label-free quantitative LC-MS to determine the peptides that were significantly enriched by the cross-linking/tandem affinity protocol. Finally, we used comparative label-free LC-FTCIR-MS to identify peptides that were quantitatively enriched in the samples purified by TAP from epitope-tagged TIN2-expressing cells, bypassing the need for comparative SDS-

PAGE. This approach increased the sensitivity of protein identification and led to the discovery of putative telomere-associated proteins.

To facilitate protein identification within our complex peptide mixtures, we used LC-FTICR-MS to quantify tryptic peptides that were prepared from the tandem affinity-captured proteins. The peptide pools from samples that were isolated from cells expressing the HA-FLAG-TIN2 construct and control cells were subjected to separate LC-MS analyses. The raw MS data from multiple LC-MS analyses were imported directly into software that aligned, identified, normalized, and quantified the ion currents from individual peptides that were measured with high resolution. A total of $\sim$11,000 signal features were detected that could be concatenated into $\sim$6000 isotope groups comprising the detected charge states of the individual peptides (from 2+ to 5+). Using this analytical approach and the described software, we readily identified peptides enriched in the experimental samples and contaminants that were increased in

the control samples and assessed the reproducibility of peptide digestion and preparation. Importantly, our approach was able to 1) identify the shelterin complex using 10–40-fold less cell lysate than was used in previous experiments, demonstrating that our approach is capable of identifying numerous proteins within a complex mixture regardless of their stoichiometries, and 2) identify proteins known to be present in higher order TIN2 complexes but that do not directly interact with TIN2, demonstrating that our protein cross-linking was effective at capturing important protein interactors.

Isolation of the shelterin components demonstrates that our approach is able to isolate large protein complexes. In addition, isolation of telomeric DNA indicated that this approach resulted in the isolation of telomeric bound protein complexes. In addition to the known TIN2-interacting proteins, we isolated numerous proteins in samples prepared from cells expressing HA-FLAG-TIN2 but not from control cells that have not been described to interact with the telomeric complex. One concern that could be raised with our approach is that it resulted in nonspecific cross-linking to proteins in the neighborhood of HA-FLAG-TIN2. We find this possibility unlikely for three reasons. First, peptides that represent the shelterin complex were enriched in our samples and were the most predominant peptides present in the samples. This included POT1, which does not directly interact with TIN2 but instead interacts through TPP1. If the cross-linking were random, we would not expect to see POT1 peptides so highly enriched. Second, the shelterin proteins were not observed in lysates from HA-FLAG-GFP-expressing cells. This indicated that the shelterin proteins do not associate with a nonspecific HA-FLAG-tagged protein such as GFP. Third, if treatment with DSP led to significant nonspecific cross-linking, we would expect to see proteins that are present in high molar ratios such as nuclear membrane proteins. Although we did observe both ribosomal and hnRNP proteins in our TIN2 pulldowns, we did not observe these proteins in our HA-FLAG-GFP pulldowns, suggesting that their interaction with TIN2 was specific and not due to their high expression within the cell. Finally, the TIN2-binding protein TRF2 traffics through the nucleolus (50), and in several instances we isolated nucleolar proteins. However, we did not isolate a significant number of nucleolar proteins, arguing that the proteins that were isolated specifically interact with TIN2 possibly through TRF2. Interestingly, most of the novel proteins identified are involved in the transmission of genetic information, including replication, transcription, mRNA processing, translation, and chromatin structure/remodeling, processes that are linked to telomere biology. Even the process of transcription has been recently linked to telomere biology (45) in a study that showed mammalian telomeres are transcribed into telomeric repeat-containing RNA (TERRA). Prior to this report, telomeres were considered to be transcriptionally silent. Thus, it is likely that a number of the proteins identified in our study will prove to be biologically relevant to telomere biology.

Our approach successfully identified many telomere-binding proteins, but a number of known telomere-binding proteins were not identified in our experiments. Although it is not clear why we failed to identify these proteins, it is possible that 1) their identification required larger amounts of cell lysate or 2) the choice of DSP as a cross-linking reagent might preclude isolation and/or identification of certain proteins. Indeed, cross-linkers add permanent modifications to proteins resulting in alterations in peptide mass, making the detection of some peptides impossible. In addition, the use of a protein cross-linker could alter antibody binding sites, resulting in reduced protein recovery following immunoprecipitation. Alternatively, the method of isolation may impact the proteins identified. Indeed, a recent study utilized sequence-specific nucleic acids to isolate telomere-binding proteins from formaldehyde-cross-linked lysates (referred to as proteomics of isolated chromatin (PICh)) (51). Using this approach, all of the components of the shelterin complex were identified as well as additional proteins previously reported to bind the telomere. Novel telomere-binding proteins were also identified, and interestingly there was no overlap with our list of novel telomere-interacting proteins. This finding likely reflects differences in the isolation procedures and cross-linking reagent and underscores the need for multiple, complementary approaches to identify components of large multiprotein complexes such as those found at the telomere.

‡‡ A Sidney Kimmel Scholar. To whom correspondence may be addressed: Dept. of Cell Biology and Physiology, Washington University School of Medicine, 660 S. Euclid Ave., St. Louis, MO 63110-1093. Tel.: 314-362-7437; Fax: 314-362-7463; E-mail: Sheila.stewart@wustl.edu.

REFERENCES

1. Ranish, J. A., Yi, E. C., Leslie, D. M., Purvine, S. O., Goodlett, D. R., Eng, J., and Aebersold, R. (2003) The study of macromolecular complexes by quantitative proteomics. *Nat. Genet.* **33,** 349–355

2. Blagoev, B., Kratchmarova, I., Ong, S. E., Nielsen, M., Foster, L. J., and Mann, M. (2003) A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol.* **21,** 315–318

3. Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A., and Pappin, D. J. (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **3,** 1154–1169

4. Schmidt, A., Kellermann, J., and Lottspeich, F. (2005) A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* **5,** 4–15

5. Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **1,** 376–386

6. Himeda, C. L., Ranish, J. A., Angello, J. C., Maire, P., Aebersold, R., and Hauschka, S. D. (2004) Quantitative proteomic identification of six4 as the trex-binding factor in the muscle creatine kinase enhancer. *Mol. Cell. Biol.* **24,** 2132–2143

7. Rinner, O., Mueller, L. N., Hubálek, M., Müller, M., Gstaiger, M., and Aebersold, R. (2007) An integrated mass spectrometric and computational framework for the analysis of protein interaction networks. *Nat. Biotechnol.* **25,** 345–352

8. de Lange, T. (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev.* **19,** 2100–2110

9. Liu, D., O'Connor, M. S., Qin, J., and Songyang, Z. (2004) Telosome, a mammalian telomere-associated complex formed by multiple telomeric proteins. *J. Biol. Chem.* **279,** 51338–51342

10. Blackburn, E. H. (2001) Switching and signaling at the telomere. *Cell* **106,** 661–673

11. Ferreira, M. G., Miller, K. M., and Cooper, J. P. (2004) Indecent exposure: when telomeres become uncapped. *Mol. Cell* **13,** 7–18

12. Takai, H., Smogorzewska, A., and de Lange, T. (2003) DNA damage foci at dysfunctional telomeres. *Curr. Biol.* **13,** 1549–1556

13. Hsu, H. L., Gilley, D., Blackburn, E. H., and Chen, D. J. (1999) Ku is associated with the telomere in mammals. *Proc. Natl. Acad. Sci. U.S.A.* **96,** 12454–12458

14. Zhu, X. D., Niedernhofer, L., Kuster, B., Mann, M., Hoeijmakers, J. H., and de Lange, T. (2003) ERCC1/XPF removes the 3′ overhang from uncapped telomeres and represses formation of telomeric DNA-containing double minute chromosomes. *Mol. Cell* **12,** 1489–1498

15. Karlseder, J., Hoke, K., Mirzoeva, O. K., Bakkenist, C., Kastan, M. B., Petrini, J. H., and de Lange, T. (2004) The telomeric protein TRF2 binds the ATM kinase and can inhibit the ATM-dependent DNA damage response. *PLoS Biol.* **2,** E240

16. Wright, W. E., Tesmer, V. M., Liao, M. L., and Shay, J. W. (1999) Normal human telomeres are not late replicating. *Exp. Cell Res.* **251,** 492–499

17. Smith, C. D., Smith, D. L., DeRisi, J. L., and Blackburn, E. H. (2003) Telomeric protein distributions and remodeling through the cell cycle in Saccharomyces cerevisiae. *Mol. Biol. Cell* **14,** 556–570

18. Crabbe, L., Verdun, R. E., Haggblom, C. I., and Karlseder, J. (2004) Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. *Science* **306,** 1951–1953

19. Loayza, D., and De Lange, T. (2003) POT1 as a terminal transducer of TRF1 telomere length control. *Nature* **423,** 1013–1018

20. Adeli, K., Wettesten, M., Asp, L., Mohammadi, A., Macri, J., and Olofsson, S. O. (1997) Intracellular assembly and degradation of apolipoprotein B-100-containing lipoproteins in digitonin-permeabilized HEP G2 cells. *J. Biol. Chem.* **272,** 5031–5039

21. Nakatani, Y., and Ogryzko, V. (2003) Immunoaffinity purification of mammalian protein complexes. *Methods Enzymol.* **370,** 430–444

22. Washburn, M. P. (2008) Sample preparation and in-solution protease digestion of proteins for chromatography-based proteomic analysis. *Curr. Protoc. Protein Sci.* **Chapter 23,** Unit 23.6.1–23.6.11

23. Baldwin, M. A., Medzihradszky, K. F., Lock, C. M., Fisher, B., Settineri, T. A., and Burlingame, A. L. (2001) Matrix-assisted laser desorption/ionization coupled with quadrupole/orthogonal acceleration time-of-flight mass spectrometry for protein discovery, identification, and structural analysis. *Anal. Chem.* **73,** 1707–1720

24. Neubert, H., Bonnert, T. P., Rumpel, K., Hunt, B. T., Henle, E. S., and James, I. T. (2008) Label-free detection of differential protein expression by LC/MALDI mass spectrometry. *J. Proteome Res.* **7,** 2270–2279

25. Ye, J. Z., Hockemeyer, D., Krutchinsky, A. N., Loayza, D., Hooper, S. M., Chait, B. T., and de Lange, T. (2004) POT1-interacting protein PIP1: a telomere length regulator that recruits POT1 to the TIN2/TRF1 complex. *Genes Dev.* **18,** 1649–1654

26. O'Connor, M. S., Safari, A., Xin, H., Liu, D., and Songyang, Z. (2006) A critical role for TPP1 and TIN2 interaction in high-order telomeric complex assembly. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 11874–11879

27. Vasilescu, J., Guo, X., and Kast, J. (2004) Identification of protein-protein interactions using in vivo cross-linking and mass spectrometry. *Proteomics* **4,** 3845–3854

28. Li, B., Oestreich, S., and de Lange, T. (2000) Identification of human Rap1: implications for telomere evolution. *Cell* **101,** 471–483

29. Metz, B., Kersten, G. F., Hoogerhout, P., Brugghe, H. F., Timmermans, H. A., de Jong, A., Meiring, H., ten Hove, J., Hennink, W. E., Crommelin, D. J., and Jiskoot, W. (2004) Identification of formaldehyde-induced modifications in proteins: reactions with model peptides. *J. Biol. Chem.* **279,** 6235–6243

30. Lomant, A. J., and Fairbanks, G. (1976) Chemical probes of extended biological structures: synthesis and properties of the cleavable protein cross-linking reagent [35S]dithiobis(succinimidyl propionate). *J. Mol. Biol.* **104,** 243–261

31. Davidson, W. S., and Hilliard, G. M. (2003) The spatial organization of apolipoprotein A-I on the edge of discoidal high density lipoprotein particles: a mass spectrometry study. *J. Biol. Chem.* **278,** 27199–27207

32. Zang, X., and Komatsu, S. (2007) A proteomics approach for identifying osmotic-stress-related proteins in rice. *Phytochemistry* **68,** 426–437

33. Greenberg, R. A., Sobhian, B., Pathania, S., Cantor, S. B., Nakatani, Y., and Livingston, D. M. (2006) Multifactorial contributions to an acute DNA damage response by BRCA1/BARD1-containing complexes. *Genes Dev.* **20,** 34–46

34. America, A. H., and Cordewener, J. H. (2008) Comparative LC-MS: a landscape of peaks and valleys. *Proteomics* **8,** 731–749

35. Khurts, S., Masutomi, K., Delgermaa, L., Arai, K., Oishi, N., Mizuno, H., Hayashi, N., Hahn, W. C., and Murakami, S. (2004) Nucleolin interacts with telomerase. *J. Biol. Chem.* **279,** 51508–51515

36. Pollice, A., Zibella, M. P., Bilaud, T., Laroche, T., Pulitzer, J. F., and Gilson, E. (2000) In vitro binding of nucleolin to double-stranded telomeric DNA. *Biochem. Biophys. Res. Commun.* **268,** 909–915

37. Ishikawa, F., Matunis, M. J., Dreyfuss, G., and Cech, T. R. (1993) Nuclear proteins that bind the pre-mRNA 3′ splice site sequence r(UUAG/G) and the human telomeric DNA sequence d(TTAGGG)n. *Mol. Cell. Biol.* **13,** 4301–4310

38. Tanaka, E., Fukuda, H., Nakashima, K., Tsuchiya, N., Seimiya, H., and Nakagama, H. (2007) HnRNP A3 binds to and protects mammalian telomeric repeats in vitro. *Biochem. Biophys. Res. Commun.* **358,** 608–614

39. LaBranche, H., Dupuis, S., Ben-David, Y., Bani, M. R., Wellinger, R. J., and Chabot, B. (1998) Telomere elongation by hnRNP A1 and a derivative that interacts with telomeric repeats and telomerase. *Nat. Genet.* **19,** 199–202

40. Ford, L. P., Suh, J. M., Wright, W. E., and Shay, J. W. (2000) Heterogeneous nuclear ribonucleoproteins C1 and C2 associate with the RNA component of human telomerase. *Mol. Cell. Biol.* **20,** 9084–9091

41. Eversole, A., and Maizels, N. (2000) In vitro properties of the conserved mammalian protein hnRNP D suggest a role in telomere maintenance. *Mol. Cell. Biol.* **20,** 5425–5432

42. Seimiya, H., Sawada, H., Muramatsu, Y., Shimizu, M., Ohko, K., Yamane,

K., and Tsuruo, T. (2000) Involvement of 14-3-3 proteins in nuclear localization of telomerase. *EMBO J.* **19,** 2652–2661

43. Forsythe, H. L., Jarvis, J. L., Turner, J. W., Elmore, L. W., and Holt, S. E. (2001) Stable association of hsp90 and p23, but Not hsp70, with active human telomerase. *J. Biol. Chem.* **276,** 15571–15574

44. Holt, S. E., Aisner, D. L., Baur, J., Tesmer, V. M., Dy, M., Ouellette, M., Trager, J. B., Morin, G. B., Toft, D. O., Shay, J. W., Wright, W. E., and White, M. A. (1999) Functional requirement of p23 and Hsp90 in telomerase complexes. *Genes Dev.* **13,** 817–826

45. Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E., and Lingner, J. (2007) Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **318,** 798–801

46. Zhong, Z., Shiue, L., Kaplan, S., and de Lange, T. (1992) A mammalian factor that binds telomeric TTAGGG repeats in vitro. *Mol. Cell. Biol.* **12,** 4834–4843

47. Bianchi, A., Smith, S., Chong, L., Elias, P., and de Lange, T. (1997) TRF1 is a dimer and bends telomeric DNA. *EMBO J.* **16,** 1785–1794

48. van Overbeek, M., and de Lange, T. (2006) Apollo, an Artemis-related nuclease, interacts with TRF2 and protects human telomeres in S phase. *Curr. Biol.* **16,** 1295–1302

49. Tagwerker, C., Flick, K., Cui, M., Guerrero, C., Dou, Y., Auer, B., Baldi, P., Huang, L., and Kaiser, P. (2006) A tandem affinity tag for two-step purification under fully denaturing conditions: application in ubiquitin profiling and protein complex identification combined with in vivo cross-linking. *Mol. Cell. Proteomics* **5,** 737–748

50. Zhang, S., Hemmerich, P., and Grosse, F. (2004) Nucleolar localization of the human telomeric repeat binding factor 2 (TRF2). *J. Cell Sci.* **117,** 3935–3945

51. Déjardin, J., and Kingston, R. E. (2009) Purification of proteins associated with specific genomic loci. *Cell* **136,** 175–186