

A serine/arginine-rich nuclear matrix cyclophilin interacts with the C-terminal domain of RNA polymerase II

Jean-Pierre Bourquin, Igor Stagljar, Pascal Meier, Peter Moosmann, John Silke, Thomas Baechi¹, Oleg Georgiev and Walter Schaffner*

Institut für Molekularbiologie, Abteilung II, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland and ¹Elektronenmikroskopisches Zentrallaboratorium, Universität Zürich, Gloriastrasse 30, CH-8028 Zürich, Switzerland

Received March 3, 1997; Revised and Accepted April 7, 1997

DDBJ/EMBL/GenBank accession no. X99717

ABSTRACT

The largest subunit of RNA polymerase II shows a striking difference in the degree of phosphorylation, depending on its functional state: initiating and elongating polymerases are unphosphorylated and highly phosphorylated respectively. Phosphorylation mostly occurs at the C-terminal domain (CTD), which consists of a repetitive heptapeptide structure. Using the yeast two-hybrid system, we have selected for mammalian proteins that interact with the phosphorylated CTD of mammalian RNA polymerase II. A prominent isolate, designated SRcyp/CASP10, specifically interacts with the CTD not only *in vivo* but also *in vitro*. It contains a serine/arginine-rich (SR) domain, similar to that found in the SR protein family of pre-mRNA splicing factors, which is required for interaction with the CTD. Most remarkably, the N-terminal region of SRcyp includes a peptidyl-prolyl *cis-trans* isomerase domain characteristic of immunophilins/cyclophilins (Cyp), a protein family implicated in protein folding, assembly and transport. SRcyp is a nuclear protein with a characteristic distribution in large irregularly shaped nuclear speckles and co-localizes perfectly with the SR domain-containing splicing factor SC35. Recent independent investigations have provided complementary data, such as an association of the phosphorylated form of RNA polymerase II with the nuclear speckles, impaired splicing in a CTD deletion background and inhibition of *in vitro* splicing by CTD peptides. Taken together, these data indicate that factors directly or indirectly involved in splicing are associated with the elongating RNA polymerases, from where they might translocate to the nascent transcripts to ensure efficient splicing, concomitant with transcription.

INTRODUCTION

The C-terminal domain (CTD) of the largest subunit of RNA polymerase II (pol II) consists of tandem repeats of a heptapeptide motif with the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser. The CTD is highly conserved among different species (1,2) and is essential for *in vivo* functions (3–6). The phosphorylation status of the CTD appears to be regulated throughout the transcription cycle. There is good *in vivo* evidence that the CTD of the elongating form of pol II is hyperphosphorylated (pol IIO) (7), while pol II that is associated with the pre-initiation complex is essentially unphosphorylated (pol IIA) (8,9). Multiple intermediate forms between pol IIA and pol IIO exist and the CTD could be involved in different pre- and post-initiation functions.

The basal transcription factors TBP and TFIIE interact with the unphosphorylated form of the CTD directly (10,11). Phosphorylation by TFIIF-associated CTD kinase might disrupt such interactions and thereby promote initiation and elongation of transcription (12). Recent evidence indicates that a transcriptionally active form of pol II exists as a holoenzyme, composed of core pol II, a large multiprotein complex termed the mediator or SRB complex (suppressors of RNA polymerase B) which interacts with the CTD and possibly some additional components, including general transcription factors and components of the SWI-SNF complex (13,14), in line with the demonstrated importance of the CTD for regulated transcription (15). An active pol II holoenzyme complex has been isolated from mammalian cells (16,17) and was reported to contain essentially the pol IIA form (16).

The CTD has been proposed as a binding site for splicing factors (18). Splicing is thought to occur, at least in part, co-transcriptionally: spliceosomes have been visualized at active sites of transcription on spread chromatin (19–21) and there is evidence for splicing of nascent mRNAs that are still associated with chromatin (22,23). Accordingly, several gene loci were shown to coincide with their specific nascent and spliced

*To whom correspondence should be addressed. Tel: +41 1 257 49 11; Fax: +41 1 363 85 02; Email: wschaffn@molbio2.unizh.ch

transcripts (24–26). Components of the transcription and splicing machineries form a fine fibrogranular reticulum that connects 20–50 larger domains, the so-called nuclear speckles (20,26). The nuclear speckles are retained in nuclear matrix preparations (27,28) and are rich in splicing factors, such as members of the serine/arginine-rich (SR) family (28–30). Most pre-mRNA synthesis occurs in defined foci which are distinct from the nuclear speckles, at sites corresponding to the previously described perichromatin fibrils (21,31,32). This finding is consistent with the idea that the speckles could be a site for storage and assembly of splicing factors (29). However, transcription and splicing of several genes were found to be in close spatial association with the nuclear speckles, indicating that these structures might also have a functional role in mRNA processing (24).

Intron-containing reporter genes under the control of a pol III promoter are not spliced, suggesting that splicing could be dependent on transcription by pol II (33). Pol II was detected in the nuclear speckles and could be co-immunoprecipitated with splicing factors (34–36). Consistent with these data, several SR proteins have been isolated that interact with the CTD (37). It has recently been demonstrated that a set of 3'-processing factors also interact with the CTD directly and that deletion of the CTD affects both RNA splicing and 3'-processing *in vivo* (38).

We screened a human cDNA library for proteins interacting with the CTD of mammalian pol II using the yeast two-hybrid system. We isolated a nuclear matrix protein, SRcyp, containing an SR domain that is distinct from the CTD binding proteins previously described (37). According to a nomenclature to be proposed by Jeffrey Corden (Johns Hopkins University, personal communication) for all published and unpublished CTD-associated SR-like proteins isolated by different laboratories so far, our SRcyp would be CASP10. SRcyp/CASP10 is found in the nuclear speckles, suggesting that it could be a component of splicing factor complexes that bind the CTD, thereby linking RNA processing to transcription.

MATERIALS AND METHODS

Plasmids

The *XhoI*–*XbaI* fragment from the following GAL4 (1–93) fusion constructs, GAL4–CTD52x (39), GAL4–nuc(170–277) and GAL4–cbl2(352–469) (40), were transferred into pGG25ΔX (41). The yeast expression cassette was transferred as a *Bam*HI fragment into pRS314 (42) to generate pRS314–CTD, pRS314–nuc(170–277) and pRS314–cbl2(352–469). SRcyp mutants were generated by PCR and cloning in pGAD424 (Clontech). c24 was transferred as a *Bgl*III fragment into the *Bam*HI site of pGAD424. SRcyp was cloned in pCATCH*, a variant of pCATCH (43). SRcyp was reconstructed by insertion of the fragments *Nco*I–*Bgl*III and *Bgl*III–*Xho*I from two overlapping cDNAs into the pCATCH* *Eco*RV and *Xho*I sites. In pCATCHV the FLAG epitope is replaced by an 11mer VSV epitope from vesicular stomatitis virus G protein (44), by inserting the following annealed oligonucleotides into the *Hind*III and *Bam*HI sites of pCATCH: 5'-AGCTCGGTAC-CGAGCGGATCACCACCATGTATACAGACATAGAGATGA-AACCG-ACTTGAAAGGGAAGCTTACG-3' and 5'-GATCC-GTAAGCTTCCCTTTCCAAGTCGGTTCATCTCTATGTCT-GTATACATGGTGGTGGATCCGAGCTCGGTACCG-3'.

Yeast two-hybrid screen

The GAL–CTD fusion construct did not activate transcription in the yeast strain Y153 and was used to screen a human peripheral blood lymphocyte cDNA library tagged to the GAL4 activation domain in pACT as described (45), testing 2×10^6 transformants for histidine prototrophy and positive β -galactosidase staining on selective drop-out medium containing 25 mM 3-aminotriazole (Sigma). The insert junctions of 36 positive clones were sequenced. The interactions of positive clones with unrelated baits and of SRcyp mutants with GAL–CTD were tested under the same conditions for at least three independent colonies.

Full-length cloning

The λ ACT cDNA library (45) and a λ GT10 cDNA library from BJA-B cells were screened with two random primed fragments from SRcyp. The inserts were sequenced on both strands using specific primers and nested deletions (Pharmacia nested deletion kit).

Dephosphorylation of the yeast extracts

Samples of 10 μ g total yeast extract (46) were diluted to a volume of 100 μ l with 100 mM Tris–HCl, pH 9.6, 2 mM MgCl₂, 0.1 mM ZnCl₂ and treated with different concentrations of calf intestinal phosphatase (NEB) for 10 min at room temperature in the absence or presence of phosphatase inhibitors (5 mM sodium fluoride, 5 mM sodium vanadate, 5 mM EDTA and 5 mM EGTA). GAL–CTD was detected by immunoblotting with the anti-CTD monoclonal antibodies (mAbs) H5 and H14 and polyclonal anti-PRP20 antibody. Detection was by ECL (Amersham).

GST protein production and precipitation

A GST fusion protein to a positive clone (c24) and GST alone were produced and bound to glutathione–CL4B Sepharose (Pharmacia) according to the manufacturer's recommendations (lysis buffer, 10 mM Tris, pH 8.0, 500 mM NaCl, 10% glycerol, 0.1% Tween 20 supplemented with 1 mM PMSF). The resin was washed four times with 1 ml lysis buffer and twice with 1 ml binding buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.25% NP40, 1 mM sodium fluoride, 1 mM sodium vanadate, 1 mM PMSF). For precipitations, 20 μ l bed volumes of GST fusion protein-bound resin were mixed with 300 μ g yeast whole cell extract prepared from Y153 containing pRS314–GAL(1–93) or pRS314–CTD in a final volume of 200 μ l binding buffer and incubated overnight on a rotator at 4°C. The resin was washed five times with binding buffer and the proteins were eluted by boiling in SDS–PAGE sample buffer. Immunoblotting was as above.

Immunofluorescence assay

Transfected HeLa cells were grown on 14 mm glass coverslips. Inhibition of pol II was with α -amanitin (50 μ g/ml) for 8 h as described (47). Fixation was with 1.75% paraformaldehyde for 15 min and permeabilization with 0.5% Triton X-100 in phosphate-buffered saline for 2.5 min at room temperature. Cells were labelled with 3 μ g/ml anti-FLAG M2 (Eastman Kodak), rabbit polyclonal antiserum 49 anti-VSV 1:500 and anti-SC35 (48) 1:4, using standard techniques and mounted in Mowiol. Nuclear matrix preparations were extracted with Triton X-100, DNase I (1000 U/ml; Boehringer Mannheim) and ammonium sulphate according to the literature (47,49). Cells were viewed on

a Leica TCS Confocal Laser Scanning Microscope (Leica, Heidelberg). The digitized micrographs were merged with Imapris software (Bitplane, Zurich).

RESULTS

Two-hybrid screen with the mammalian CTD

A fusion of the CTD of mouse pol II (52 repeats) to the GAL4 DNA binding domain (amino acid residues 1–93) (39) was used to screen a human peripheral blood lymphocyte cDNA library fused to the GAL4 activation domain in yeast strain Y153 (45). The specificity of the interaction of the positive CTD interacting clones was verified by transforming the isolated clones in Y153 strains expressing either GAL–CTD (pRS314-CTD), GAL DNA binding domain [pRS314-GAL(1–93)] or a fusion with acidic regions of the cytoplasmic oncogene cbl [pRS314-cbl2(352–469)] and of nucleolin [pRS314-nuc(170–277)], which are highly acidic and do not activate transcription (40). Reporter gene activation by the positive clones was strictly dependent on the presence of the CTD and the positive clones did not interact with the GAL4 DNA binding domain alone or the GAL4 DNA binding domain fused to highly charged portions of unrelated proteins. The GAL4 activation domain alone activated the reporter genes when co-transformed with GAL–CTD (data not shown). Expression levels of the GAL4 fusion constructs in yeast strain Y153 were compared by electrophoretic mobility shift assays and found to be similar in the four strains (data not shown). Sequencing of the cDNA inserts revealed multiple overlapping clones of four different types of cDNAs (see Discussion).

A serine/arginine-rich CTD binding protein

We focused on a family of three overlapping cDNA clones which contain a region of homology with SR domains of RNA splicing factors (Fig. 2, clones c19, c24 and c26). The full-length cDNA, isolated from a human B-lymphocyte cDNA library, spans 2695 nt and contains an open reading frame of 754 amino acids (Fig. 1A and B). Translational stops are found in all three frames upstream of the first AUG and the length of the cDNA is consistent with a single band of ~3.0 kb in Northern blots. The SRcyp gene is expressed in all cell types we have tested to date (data not shown).

A 177 amino acid stretch in the N-terminal portion has a high identity to the cyclophilin-type peptidyl-prolyl *cis-trans* isomerases (PPIase), while the remainder of the protein is highly charged. An acidic serine-rich region flanked by stretches of basic residues is adjacent to the N-terminal PPIase and an SR domain is found in the C-terminal portion of SRcyp (Fig. 1A). The region from amino acid 540 to 639 contains 14% serine, 29% arginine, 13% SR dipeptides and 12% RS dipeptides (underlined and in bold in Fig. 1B). Shorter stretches of SR repeats are found outside the SR homology region. The SR domains of SRcyp, of the CTD binding proteins isolated by the group of J.Corden (37) and of two further serine/arginine-rich CTD binding proteins isolated in our screen all have a similar serine, arginine and SR/RS peptide content (S.Tanner, J.P.Bourquin and W.Schaffner, unpublished results). These sequence features suggest that these proteins are related to the SR protein family of splicing factors. From position 180 to 754 of SRcyp, 71.3% of the residues are Glu, Asp, Lys, Arg, Ser or Thr and positively and negatively charged residues

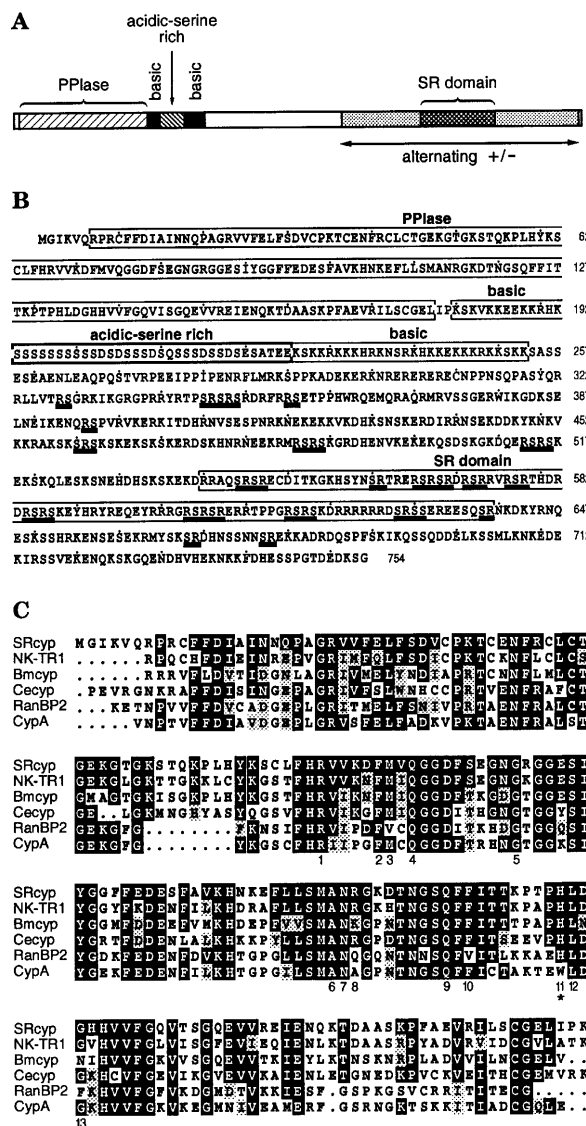


Figure 1. Sequence analysis of human SRcyp. (A) Schematic drawing of the domains of SRcyp. PPIase, peptidyl-prolyl *cis/trans* isomerase; alternating +/-, alternating charged amino acids. (B) Conceptual translation of the full-length cDNA. Amino acids are numbered on the right. (C) Comparison of the PPIase domains of SRcyp, NK-TR1, Bmcp, *C.elegans* protein CEESF41R (Cecyp), RanBP2 and human cyclophilin A (CypA). The gaps introduced to maximize the homology are represented by dots. Identical amino acid positions are shaded in black, conservative changes in grey. The conserved amino acid residues in the cyclosporin A binding pockets are indicated by numbers below the sequence alignment. Sequence alignments were performed using PILEUP (GCG, Wisconsin) and edited with BOXSHADE.

alternate regularly. Additional negative charges might be introduced by phosphorylation on serine and threonine residues. Such an alternation of positive and negative charges has been proposed to be a sequence feature of splicing factors (50).

SRcyp is related to a group of large cyclophilins

With 37.8% amino acid identity (71% within the cyclophilin-like domain, Fig. 1C), SRcyp appears to be most closely related to NK-TR1, which is a myeloid-specific nuclear protein containing

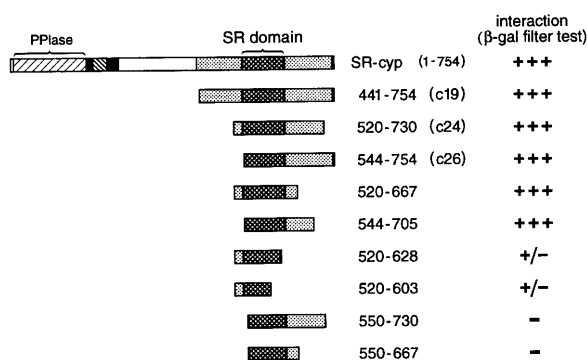


Figure 2. Mapping of the CTD interaction domain of SRcyp with SRcyp mutants in a GAL-CTD two-hybrid assay. The sequence boundaries of the clones isolated in the yeast screen (c19, c24 and c26) and the deletion mutants are shown with the corresponding amino acid positions indicated on the right. The qualitative intensity of the β -galactosidase staining, detected in a filter test for three independent colonies, is indicated: +++, an intense blue staining appearing within 1 h; +/-, a pale blue signal observed after several hours.

three SR domains. NK-TR1 has PPIase, protein folding and chaperone activities and is important for myeloid cell differentiation and cytotoxic function of natural killer cells (51,52). Similar sequence motifs are present in Bmcp from *Brugia malayi*, a filarial nematode (53), and a putative protein from *Caenorhabditis elegans*, encoded by the cDNA CEESF41R. All four cyclophilin-like proteins are highly charged outside the PPIase motif and share regions enriched for serine, arginine and lysine residues and a serine-rich stretch. Their PPIase domains differ from other cyclophilins by an eight amino acid insertion (position 52 of the SRcyp sequence) and a substitution of a histidine for the conserved tryptophan residue in the cyclosporin A binding site (indicated by * in Fig. 1C; 53). The Trp→His substitution is also found in the PPIase domain of RanBP2, a large nucleopore protein that has chaperone activity involving proline isomerization (54). The occurrence of a serine/acidic-rich stretch alternating with blocks of basic residues (180–253) resembles sequence motifs found in nuclear localization signal binding proteins (55). From these findings it is evident that SRcyp is related to a conserved family of large cyclophilin-like proteins.

The SR domain is required for the interaction of SRcyp with the CTD

SR domains have been implicated in protein–protein interactions (30,56). Since all the SRcyp fragments isolated in the two-hybrid screen contained the SR domain, we tested whether this domain was required for binding to CTD. A series of SRcyp deletion mutants was tested for CTD interaction in the two-hybrid assay (Fig. 2). An SRcyp fragment containing amino acids 520–667 is sufficient for an interaction which is comparable in strength to that observed with the full-length protein. An N-terminal deletion up to position 550 abolished the interaction with the CTD completely, while a SRcyp fragment starting at position 544 still retained the ability to interact with the CTD. We conclude that the interaction domain is contained within amino acids 544–667. This portion of SRcyp includes the SR domain (540–639), as defined by sequence comparison. We also tested if SRcyp required the full-length CTD for interaction. We found significant but weaker interaction with the proximal conserved 22 heptapep-

tide repeats of the CTD and with the 26 repeats of the yeast CTD (a kind gift of R. Young, MIT, Cambridge, MA; data not shown).

The mammalian CTD fused to the GAL4 DNA binding domain is phosphorylated in yeast

To investigate the phosphorylation state of the CTD in the context of the GAL4 DNA binding domain fusion, we took advantage of two antibodies that recognize different phosphoepitopes on the CTD (34). mAb H5 recognizes the 240 kDa Pol IIO form, while mAb H14 binds multiple intermediate phosphorylation forms between 220 and 240 kDa. Yeast whole cell extracts from Y153 strains expressing GAL-CTD were analysed by immunoblotting with mAb H5 and H14. Both mAbs detected a specific band of apparent size 70 kDa in extracts from strains expressing GAL-CTD. As described previously (34), this signal depends on the presence of phosphoepitopes on the CTD, because treatment of the extracts with calf intestinal phosphatase resulted in its dose-dependent reduction (Fig. 3). Complete loss of the signal was observed when using a longer incubation time with phosphatase and similar results were obtained using phage λ protein phosphatase (data not shown). The signal was not affected when phosphatase inhibitors were added to the reaction or in the absence of phosphatase, confirming that this effect is due to dephosphorylation. Unexpectedly, we did not observe an increase in mobility of the GAL-CTD fusion protein with dephosphorylation, although this is consistent with an earlier observation describing the effect of increasing CTD phosphorylation *in vitro* as an initial retardation followed by an increase in electrophoretic mobility on SDS–polyacrylamide gels (57). Taken together, these results imply that a significant amount of the GAL-CTD construct is phosphorylated in strain Y153 under the selective conditions used in the screen. However, we cannot exclude the possibility that a fraction of the GAL-CTD fusion is not phosphorylated. A single band of ~210 kD is recognized by both H5 and H14 and most likely represents the large subunit of yeast pol II, which is consistent with previous observations (58). The H5 signal disappears before the H14 signal with dephosphorylation, which is consistent with earlier observations with pol II using these antibodies (34).

The CTD interacts with SRcyp *in vitro*

A portion of SRcyp corresponding to a positive clone (c24) was expressed as a GST fusion protein, immobilized on glutathione–Sepharose beads and the binding of GAL-CTD from yeast extracts was tested. A significant proportion of GAL-CTD was retained by GST–c24 (Fig. 4, lane 3), but not by GST alone (lane 5), indicating that the CTD interacts with c24 directly. Monoclonal antibodies (H5 and H14) that recognize CTD phosphoepitopes on the hyperphosphorylated form of pol II (36) detected the GAL-CTD fusion used here. While the antibodies preferentially detect the phosphorylated form of pol II, we cannot exclude the possibility that a fraction of unphosphorylated CTD interacts with GST–c24 in this assay.

SRcyp is a nuclear matrix-associated protein and co-localizes with RNA splicing factors at nuclear speckles

Transient transfections in HeLa cells and COS cells with SRcyp tagged either with the 8mer FLAG or the 11mer VSV epitope and subsequent immunostaining revealed a nuclear localization with a

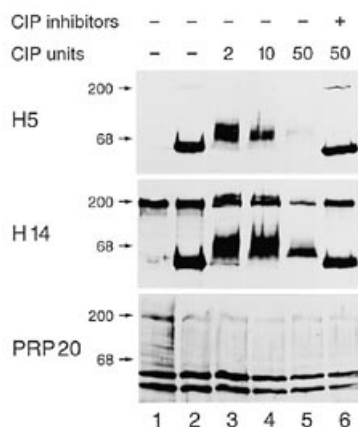


Figure 3. The mammalian CTD fused to the GAL4 DNA binding domain is phosphorylated in yeast. Yeast whole extracts from Y153 strains with or without the GAL-CTD expression vector were incubated with the indicated calf intestinal phosphatase (CIP) concentrations (in units) in the presence or absence of phosphatase inhibitors. After separation on a 7% SDS-polyacrylamide gel, the proteins were immunoblotted using the monoclonal antibodies H5 and H14 directed against different phosphoepitopes on the CTD as indicated. Absence of significant proteolytic degradation of the samples was monitored by visualizing the yeast PRP20p protein with a polyclonal antiserum. The expected 54 and 52 kDa bands have similar intensities in all lanes. The 200 kDa band with mAbs H5 and H14 corresponds to the expected size of the yeast pol II largest subunit. Lane 1, Y153 whole cell extracts; lanes 2–6, whole cell extracts of Y153 expressing Gal-CTD. Molecular mass markers in kDa are indicated on the left.

characteristic subnuclear distribution in speckles and an associated fine reticular nucleoplasmic staining (Fig. 5a and b). SRcyp was also detected in nuclear matrix preparations, which is expected for a component of the nuclear speckles (Fig. 5c and d). The nuclear distribution of SC35, a well-characterized pre-mRNA splicing factor of the SR family (48), is often used to define the nuclear speckles compartment. We used the VSV-tagged SRcyp construct in double staining experiments with anti-SC35 and an anti-VSV rabbit polyclonal antisera. Representative nuclei visualized by confocal laser scanning microscopy are depicted in Figure 5e–j. SRcyp co-localizes with SC35 at the nuclear speckles (Fig. 5e–g). Due to the faint background staining with anti-VSV, we cannot be sure that the weaker staining pattern outside the speckles entirely coincides with the SC35 distribution. Upon inhibition of RNA transcription or RNA splicing, the SC35-labelled speckles are known to undergo a redistribution from irregularly shaped interconnected speckles to enlarged round unconnected structures (47,59). We found a similar redistribution of SRcyp when pol II transcription was inhibited by α -amanitin and this pattern fully overlapped with the enlarged rounded SC35 speckles (Fig. 5h–j). Similar results were obtained with COS cells and in double labelling with mAb Y12, an anti-Sm antibody that recognizes a common epitope on different spliceosomal proteins (data not shown). These results suggest that SRcyp is involved in some aspects of RNA processing, possibly by linking a post-transcriptional mechanism to transcription by pol II.

DISCUSSION

Using the two-hybrid system, we have isolated a serine/arginine-rich protein that interacts with the CTD of pol II. SR domains are found mostly in pre-mRNA splicing factors (30). SRcyp is

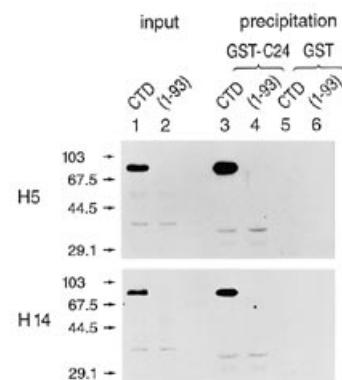


Figure 4. CTD binds to SRcyp *in vitro*. Yeast whole cell extract from Y153 expressing GAL4-CTD or GAL4(1–93) was incubated with GST or with GST fused to c24 (SRcyp amino acids 520–730) purified on glutathione-CL4B resin. After five washes proteins were eluted by boiling in sample buffer, separated on a 12% SDS-polyacrylamide gel and analysed by immunoblotting with the mAbs H5 and H14 as indicated. Aliquots of 10 μ g yeast extract were loaded as input control. Lane 1, GAL4-CTD; lane 2, GAL4(1–93). GST-c24 was mixed with GAL4-CTD (lane 3) or with GAL4(1–93) (lane 4). GST alone was mixed with GAL4-CTD (lane 5) or with GAL4(1–93) (lane 6). CTD is retained only by GST-c24 (lane 3). Molecular mass markers in kDa are indicated on the left.

associated with the nuclear matrix and co-localizes with splicing factors at the nuclear speckles, a nuclear compartment rich in splicing factors. A form of the CTD that is detected by an antibody specific for phosphoepitopes on pol II O binds to SRcyp *in vitro*. The presence of a PPIase domain in SRcyp suggests that this protein could affect protein folding and the CTD with its high proline content suggests itself as a target. These findings indicate that SRcyp could facilitate formation of macromolecular complexes linking elongating pol II to the splicing machinery.

The other proteins isolated in our two-hybrid screen contain sequence features that are mostly found in RNA processing molecules. One contains an RNA recognition motif and a region rich in glutamic acid/arginine repeats, similar to a motif described in pre-mRNA splicing factors (50), which is functionally involved in RNA splicing. The other two contain a CTD binding region that is very similar to that described for rA1 and rA9, two SR proteins isolated by virtue of their CTD interaction (37). One of our clones could be a homologue of rA9 (80% identity over the cloned region), whereas the second is a large SR protein that is clearly distinct from rA1 and rA9 outside the conserved CTD binding region (69).

SR proteins form an expanding family of pre-mRNA splicing factors that are thought to be involved in early splicing events such as splice site selection and early spliceosomal assembly (30). SR domains are necessary for specific interactions between different SR proteins (56). We have shown that the SR domain of SRcyp is required for the interaction with the CTD. A two-hybrid screen with Clk/Sty, a protein kinase, identified several SR proteins as specific substrates (60), among which was a clone identical to our SRcyp (61). None of the six other SR proteins described (60) that bind Clk/Sty interact with the CTD in our assay (data not shown). Thus interaction of the CTD with SRcyp cannot simply be explained by the high frequency of positive charges in the SR domain, but clearly requires a specific sequence. In addition, the other SR CTD binding proteins that we

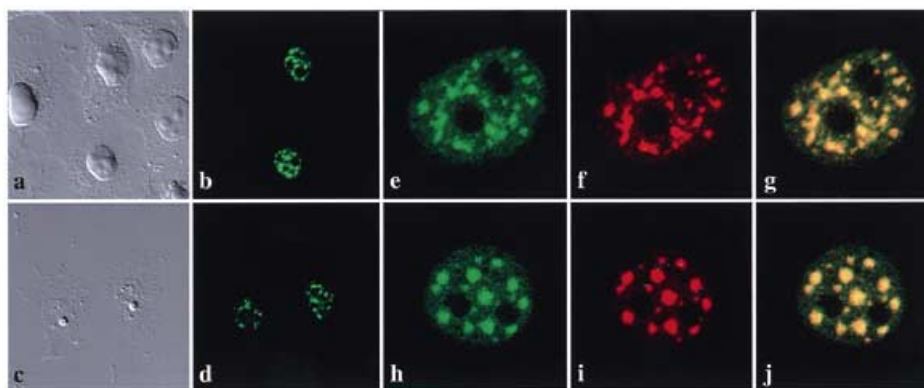


Figure 5. Nuclear localization of SRcyp and its co-localization with the splicing factor SC35. Transfected HeLa cells were stained by indirect immunofluorescence using the mAb M2 against the FLAG epitope (anti-FLAG, FLAG epitope on SRcyp) (b and d), a polyclonal rabbit antiserum against the VSV epitope (anti-VSV, VSV epitope on SRcyp) (e and h) and an anti-SC35 mAb (anti-SC35) (f and i). The superimposition of the two signals is shown in (g) and (j), where the alignment of green and red signals appears yellow. The effect of transcriptional inhibition by α -amanitin is shown in panels (h)–(j). (a), (b) and (e)–(j), normal fixation procedure, (c) and (d), nuclear matrix preparation. For comparison, differential interference contrast images are shown in (a) and (c). (a), (b), (c) and (d), 40 \times ; (e)–(j), 100 \times .

and others have isolated do not interact with the CTD through their SR domains, but rather through either of two other types of conserved protein domains (37,69). Thus, the earlier speculation that splicing factors of the SR protein family would interact with the CTD through their SR domains (18) has not proven generally valid. Instead, SR-like proteins seem to bind to the CTD either via the SR domain or via another domain, which probably allows for an intricate network of interactions.

SRcyp is a nuclear matrix component. The nuclear matrix is a fibrogranular network in the interchromatin space that consists mainly of RNA and RNA processing factors and is thought to form the structural basis for coupled transcription and RNA processing (27,62). Recent studies identified subsets of large SR-related proteins in the nuclear matrix that are possibly involved in splicing (50,63,64). Epitope-tagged SRcyp has an apparent molecular mass of ~110 kDa, which is within the size ranges of these large SR-related proteins. SRcyp shows the same subnuclear distribution as the splicing factor SC35 and the speckled pattern is affected by inhibition of pol II by α -amanitin, as described previously (47). This identifies SRcyp as a component of the nuclear speckles. The function of this nuclear compartment, which contains most of the splicing factors, is still controversial. Speckles are thought to represent a dynamic structure with potential functions in storage, assembly and recycling of RNA processing components (29). Most of the newly synthesized RNA is, however, found in defined foci outside and at the periphery of SC35 domains (21,31,32), which occur in close spatial association with the nuclear speckles (26). Consistent with the finding of Wansink *et al.* (31), we observed that certain sites of pol II transcription coincide with the weaker SRcyp staining, outside the speckles (data not shown). Recently, a rat homologue of SRcyp (93% identity) that localizes to the nuclear speckles has also been cloned. This protein co-immunoprecipitates with other SR proteins and exhibits PPIase activity *in vitro* (M.Mortillaro and R.Berezny, personal communication).

Increasing evidence supports the idea that the CTD is a binding site for RNA processing factors: several SR proteins specifically interact with the CTD (37,69). The hyperphosphorylated form of pol II can be co-immunoprecipitated with monoclonal antibodies against different splicing factors, suggesting that a complex of splicing factors is associated with the CTD *in vivo* (35,36,65).

Recently it was shown that the 3'-processing factors CstF and CPSF also specifically interact with the CTD. Accordingly, a complete truncation of the CTD results in a significant decrease in mRNA splicing and 3'-processing *in vivo* (38). Similarly, overexpression of the 52 heptapeptide repeats of the CTD interferes with RNA splicing *in vivo* (66). These results confirm the previous observation of inhibition of an *in vitro* splicing reaction with CTD peptides (37). Thus, the CTD could be crucial for the formation of larger complexes of RNA processing factors on elongating pol II and could play a central role in the coupling of RNA processing to transcription. The assembly of such complexes could be modulated by changes in the CTD phosphorylation status and/or in its conformation. Peptidyl-prolyl isomerization affects the local conformation of polypeptides (67) and might therefore provide a regulatory mechanism to modify the conformation of the CTD, which has a repetitive proline-rich amino acid sequence. Alternatively, the formation of specific protein complexes on the CTD might require some associated nuclear chaperone functions. Chaperone activity of a cyclophilin, Cyp40, has recently been demonstrated (68). Specific chaperones could be involved in maintaining the CTD in a certain folded state that could favour particular protein-protein interactions. As mentioned before, SRcyp specifically interacts with Clk/Sty (61), an SR protein kinase member of a family of protein kinases involved in regulation of RNA splicing (56,60). This finding raises the question of whether phosphorylation of SRcyp by Clk/Sty could affect the interaction of pol II with the CTD, as part of a regulatory mechanism of early splicing events. Further experiments will have to address the question of whether the CTD is a specific substrate for the PPIase activity of SRcyp and to what extent this nuclear cyclophilin influences transcription and RNA processing *in vivo*.

ACKNOWLEDGEMENTS

We thank Steven Elledge (Houston) for providing us with the Y153 strain and the pACT cDNA library, Meinrad Busslinger (Vienna) for the λ GT10 cDNA library, Tom Maniatis (Boston) for the anti-SC35 antibody, Jean-Claude Perriard (Zürich) for the anti-VSV rabbit polyclonal antiserum, Steven Warren for the H5 and H14 antibodies, Rick Young for a yeast CTD clone and Stefan

te Heesen for the anti-PRP20 antibody. We are grateful to Martin Hoechli for training in the use of the confocal microscope, to Fritz Ochsenbein for assistance with the figures and to Olav Zilian, Matthias Gstaiger, Alcide Barberis and Richard Eckner for useful suggestions. We are indebted to Christopher Hovens for continuous discussions and help with the project. This work was supported by the Kanton Zurich and the Schweizerischer Nationalfonds and by a grant of the Schweizerische Akademie der Medizinischen Wissenschaften to J.-P.B.

REFERENCES

- Corden, J.L., Cadena, D.L., Ahearn, J.M.J. and Dahmus, M.E. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 7934–7938.
- Allison, L., Moyle, M., Shales, M. and Ingles, C. (1985) *Cell*, **42**, 599–610.
- Nonet, M., Scafe, C., Sexton, J. and Young, R. (1987) *Mol. Cell. Biol.*, **7**, 1602–1611.
- Allison, L.A., Wong, J.K.C., Fitzpatrick, V.D., Moyle, M. and Ingles, C.J. (1988) *Mol. Cell. Biol.*, **8**, 321–329.
- Bartolomei, M.S., Halden, N.F., Cullen, C.R. and Corden, J.L. (1988) *Mol. Cell. Biol.*, **8**, 330–339.
- Zehring, W.A., Lee, J.M., Weeks, J.R., Jokerst, R.S. and Greenleaf, A.L. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 3698–3702.
- O'Brien, T., Hardin, S., Greenleaf, A. and Lis, J.T. (1994) *Nature*, **370**, 75–77.
- Corden, J.L. (1993) *Curr. Opin. Genet. Dev.*, **3**, 213–218.
- Dahmus, M.E. (1995) *Biochim. Biophys. Acta*, **1261**, 171–182.
- Usheva, A., Goldring, A., Maldonado, E., Goldring, A., Lu, H., Houbavi, C., Reinberg, D. and Aloni, Y. (1992) *Cell*, **69**, 871–881.
- Maxon, M., Goodrich, J. and Tjian, R. (1994) *Genes Dev.*, **8**, 515–524.
- Svejstrup, J.Q., Vichi, P. and Egly, J.M. (1996) *Trends Biochem. Sci.*, **21**, 346–350.
- Koleske, A.J. and Young, R.A. (1995) *Trends Biochem. Sci.*, **20**, 113–116.
- Bjorklund, S. and Kim, Y.J. (1996) *Trends Biochem. Sci.*, **21**, 335–337.
- Gerber, H.P., Hagmann, M., Seipel, K., Georgiev, O., West, M.A., Litingtung, Y., Schaffner, W. and Corden, J.L. (1995) *Nature*, **374**, 660–662.
- Ossipow, V., Tassan, J.P., Nigg, E.A. and Schibler, U. (1995) *Cell*, **83**, 137–146.
- Maldonado, E., Shiekhatter, R., Sheldon, M., Cho, H., Drapkin, R., Rickert, P., Lees, E., Anderson, C.W., Linn, S. and Reinberg, D. (1996) *Nature*, **381**, 86–89.
- Greenleaf, A.L. (1993) *Trends Biochem. Sci.*, **18**, 117–119.
- Beyer, A.L. and Osheim, Y.N. (1988) *Genes Dev.*, **2**, 754–765.
- LeMaire, M. and Thummel, C. (1990) *Mol. Cell. Biol.*, **10**, 6059–6063.
- Fakan, S. (1994) *Trends Cell Biol.*, **4**, 86–90.
- Bauren, G. and Wieslander, L. (1994) *Cell*, **76**, 183–192.
- Wuarin, J. and Schibler, U. (1994) *Mol. Cell. Biol.*, **14**, 7219–7225.
- Xing, Y., Johnson, C.V., Dobner, P.R. and Lawrence, J.B. (1993) *Science*, **259**, 1326–1330.
- Zhang, G., Taneja, K.L., Singer, R.H. and Green, M.R. (1994) *Nature*, **372**, 809–812.
- Xing, Y., Johnson, C.V., Moen, P.T.J., McNeil, J.A. and Lawrence, J. (1995) *J. Cell Biol.*, **131**, 1635–1647.
- van Driel, R., Wansink, D.G., van Steensel, B., Grande, M.A., Schul, W. and de Jong, L. (1995) *Int. Rev. Cytol.*, **162A**, 151–189.
- Nickerson, J.A., Blencowe, B.J. and Penman, S. (1995) *Int. Rev. Cytol.*, **162A**, 67–124.
- Spector, D.L. (1993) *Annu. Rev. Cell Biol.*, **9**, 265–315.
- Fu, X.D. (1995) *RNA*, **1**, 663–680.
- Wansink, D.G., Schul, W., van der Kraan, I., van Steensel, B., van Driel, R. and de Jong, L. (1993) *J. Cell Biol.*, **122**, 283–293.
- Jackson, D.A., Hassan, A.B., Errington, R.J. and Cook, P.R. (1993) *EMBO J.*, **12**, 1059–1065.
- Sisodia, S.S., Sollner-Webb, B. and Cleveland, D.W. (1987) *Mol. Cell. Biol.*, **7**, 3602–3612.
- Bregman, D.B., Du, L., van der Zee, S. and Warren, S.L. (1995) *J. Cell Biol.*, **129**, 287–298.
- Mortillaro, M.J., Blencowe, B.J., Wei, X.Y., Nakayasu, H., Du, L., Warren, S.L., Sharp, P.A. and Berezney, R. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 8253–8257.
- Kim, E., Du, L., Bregman, D.B. and Warren, S.L. (1997) *J. Cell Biol.*, **136**, 19–28.
- Yuryev, A., Patturajan, M., Litingtung, Y., Joshi, R.V., Gentile, C., Gebara, M. and Corden, J.L. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 6975–6980.
- McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M. and Bentley, D.L. (1997) *Nature*, **385**, 357–361.
- Seipel, K., Georgiev, O., Gerber, H.P. and Schaffner, W. (1993) *Nucleic Acids Res.*, **21**, 5609–5615.
- Seipel, K., Georgiev, O. and Schaffner, W. (1992) *EMBO J.*, **11**, 4961–4968.
- Gill, G. and Ptashne, M. (1987) *Cell*, **51**, 121–126.
- Sikorski, R.S. and Hieter, P. (1989) *Genetics*, **122**, 19–27.
- Georgiev, O., Bourquin, J.P., Gstaiger, M., Knoepfel, L., Schaffner, W. and Hovens, C. (1996) *Gene*, **168**, 165–167.
- Soldati, T. and Perriard, J.C. (1991) *Cell*, **66**, 277–289.
- Durfee, T., Becherer, K., Chen, P.L., Yeh, S.H., Yang, Y., Kilburn, A.E., Lee, W.H. and Elledge, S.J. (1993) *Genes Dev.*, **7**, 555–569.
- Kaiser, C., Michaelis, S. and Mitchell, A. (1994) *Methods in Yeast Genetics*. Cold Spring Harbor Laboratory Press, Plainview, NY.
- Huang, S., Deerinck, T.J., Ellisman, M.H. and Spector, D.L. (1994) *J. Cell Biol.*, **126**, 877–899.
- Fu, X.D. and Maniatis, T. (1990) *Nature*, **343**, 437–441.
- Fey, E.G., Krochmalnic, G. and Penman, S. (1986) *J. Cell Biol.*, **102**, 1654–1665.
- Neugebauer, K.M., Stolk, J.A. and Roth, M.B. (1995) *J. Cell Biol.*, **129**, 899–908.
- Rinfret, A., Collins, C., Menard, R. and Anderson, S.K. (1994) *Biochemistry*, **33**, 1668–1673.
- Giardina, S.L., Coffman, J.D., Young, H.A., Potter, S.J., Frey, J.L., Ortaldo, J.R. and Anderson, S.K. (1996) *Blood*, **87**, 2269–2274.
- Page, A.P., Landry, D., Wilson, G.G. and Carlow, C. (1995) *Biochemistry*, **34**, 11545–11550.
- Ferreira, P.A., Nakayama, T.A., Pak, W.L. and Travis, G.H. (1996) *Nature*, **383**, 637–640.
- Xue, Z., Shan, X., Lapeyre, B. and Melese, T. (1993) *Eur. J. Cell Biol.*, **62**, 13–21.
- Manley, J.L. and Tacke, R. (1996) *Genes Dev.*, **10**, 1569–1579.
- Zhang, J. and Corden, J.L. (1991) *J. Biol. Chem.*, **266**, 2297–2302.
- Warren, S.L., Landolfi, A.S., Curtis, C. and Morrow, J.S. (1992) *J. Cell Sci.*, **103**, 381–388.
- O'Keefe, R.T., Mayeda, A., Sadowski, C.L., Krainer, A.R. and Spector, D.L. (1994) *J. Cell Biol.*, **124**, 249–260.
- Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J.L., Bell, J.C. and Duncan, P.I. (1996) *EMBO J.*, **15**, 265–275.
- Nestel, F.P., Colwill, K., Harper, S., Pawson, T. and Anderson, S.K. (1996) *Gene*, **180**, 151–155.
- Berezney, R., Mortillaro, M.J., Ma, H., Wei, X. and Samarabandu, J. (1995) *Int. Rev. Cytol.*, **162A**, 2–66.
- Blencowe, B., Nickerson, J., Issner, R., Penman, S. and Sharp, P. (1994) *J. Cell Biol.*, **127**, 593–607.
- Blencowe, B.J., Issner, R., Kim, J.S., McCaw, P. and Sharp, P.A. (1995) *RNA*, **1**, 852–865.
- Vincent, M., Lauriault, P., Dubois, M.F., Lavoie, S., Bensaude, O. and Chabot, B. (1996) *Nucleic Acids Res.*, **24**, 4649–4652.
- Du, L. and Warren, S.L. (1997) *J. Cell Biol.*, **136**, 5–18.
- Galat, A. and Metcalfe, S.M. (1995) *Prog. Biophys. Mol. Biol.*, **63**, 67–118.
- Freeman, B.C., Toft, D.O. and Morimoto, R.I. (1996) *Science*, **274**, 1718–1720.
- Tanner, S., Stagljar, I., Georgiev, O., Schaffner, W. and Bourquin, J.-P. (1997) *J. Biol. Chem.*, **378**, in press.