The phosphoCTD-interacting domain of Topoisomerase I

Jianhong Wu, Hemali P. Phatnani [1], Tao-Shih Hsieh, Arno L. Greenleaf *

Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA

ABSTRACT

The N-terminal domain (NTD) of *Drosophila melanogaster* (*Dm*) Topoisomerase I has been shown to bind to RNA polymerase II, but the domain of RNAPII with which it interacts is not known. Using bacterially-expressed fusion proteins carrying all or half of the NTDs of *Dm* and human (*Homo sapiens*, *Hs*) Topo I, we demonstrate that the N-terminal half of each NTD binds directly to the hyperphosphorylated C-terminal repeat domain (phosphoCTD) of the largest RNAPII subunit, Rpb1. Thus, the amino terminal segment of metazoan Topo I (1–157 for *Dm* and 1–114 for *Hs*) contains a novel phosphoCTD-interacting domain that we designate the Topo I–Rpb1 interacting (TRI) domain. The long-known in vivo association of Topo I with active genes presumably can be attributed, wholly or in part, to the TRI domain-mediated binding of Topo I to the phosphoCTD of transcribing RNAPII.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Topoisomerase I in metazoan cell nuclei associates with transcriptionally active regions of the genome [1,2], and this attribute maps to the N-terminal domain (NTD) of the protein [3], a non-catalytic segment of variable sequence ranging from about 200 to 400 amino acids in length. The NTD of *Drosophila* Topo I (amino acids 1–315 of the 972-residue protein) is capable of associating with RNA polymerase II (RNAPII) as judged by pull down experiments [3], suggesting a mechanism for targeting Topo I to active genes. Moreover, Topo I was found by proteomics approaches to bind directly to the hyperphosphorylated CTD (PCTD) of Rpb1, the largest subunit of RNAPII [4,5]. Taken together these facts suggest that the NTD can mediate the association of Topo I with elongating RNAPII via direct binding to the PCTD. Herein we confirm this suggestion for fruit fly and human Topo I, and we define the subsegment of the NTD that manifests PCTD binding.

2. Materials and methods

2.1. Plasmids and proteins

Three segments of human Topo I N-terminus, 1–199, 1–114, and 104–199, were constructed by PCR with cDNA as template, and primer pairs CCACTAGTATGAGTGGGGACCACCTCCACA/CCGTCGACCCTCTTCTTTCTTCGGCTTCTT, CACTAGTATGAGTGGGGACCAC

CTCCACA/CCGTCGACGTGGTGGACT AGAGAAGCCA, and GCACTAGTAAGGAGAAGGAAAATGGC/CCGTCGACCCTCTTCTTTCTTCGGCTTCTT were used, respectively. The underlined sequences in the oligo nucleotides correspond to SpeI and SalI recognition sites, respectively. The fragments were inserted into pET41a (Novagen) with SpeI and SalI.

Similarly, the primer pairs GCCAAACTAGTATGAGTGGGGATGTGGCTGC/CGAAGTCGACATCATTGTAGTTCATGGTGC, GCCAAACTAGTATGAGTGGGGATGTGGCTGC/CGAAGTCGACCCGACTGGATGAACTGCT, and GCCAAACTAGTCACAAGTCCTCGTCCAGC T/CGAAGTCGACAT CATTGTAGTTCATGGTGC were used for constructing 1–315, 1–157, and 158–315 of *Drosophila* Topo I, respectively.

All constructs were confirmed by DNA sequencing before they were transformed into the expression strain BL21(DE3) (Novagen). Affinity purification of protein with Ni–NTA (Qiagen) resin was carried out according to the procedure provided by the manufacturer.

2.2. Phosphorylation of GST-CTD fusion protein, and Far Western blotting

A ßGal-*Dm*CTD fusion protein [6], or a GST-yeast CTD fusion protein [7] was exhaustively phosphorylated, either with or without radioactive ATP, using purified yeast CTDK-I [8]. Far Western blotting (also "reverse" Far Western blotting) was performed as described [8].

3. Results

We constructed GST fusion proteins carrying part or all of the N-terminal Domain (NTD) from Topo I of *Drosophila melanogaster* or *Homo sapiens* (illustrated in Fig. 1A). The fusion proteins were

---

* Corresponding author. Fax: +1 919 684 8885.
  E-mail address: arno.greenleaf@duke.edu (A.L. Greenleaf).
[1] Present address: Department of Biochemistry & Molecular Biophysics, Columbia University College of Physicians and Surgeons, New York, NY, USA.

A. N-terminal Fusions
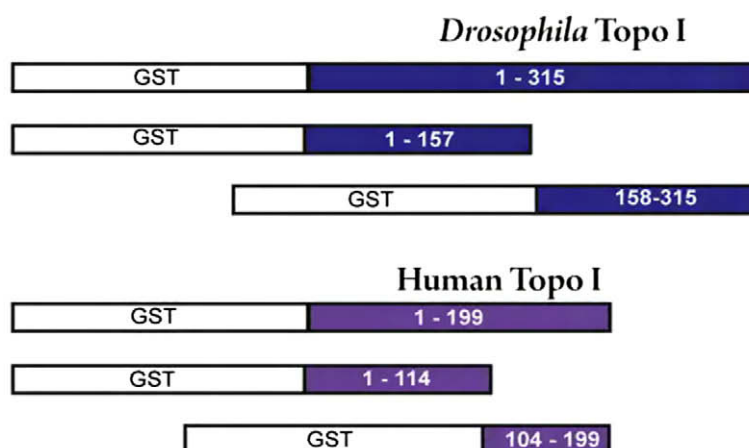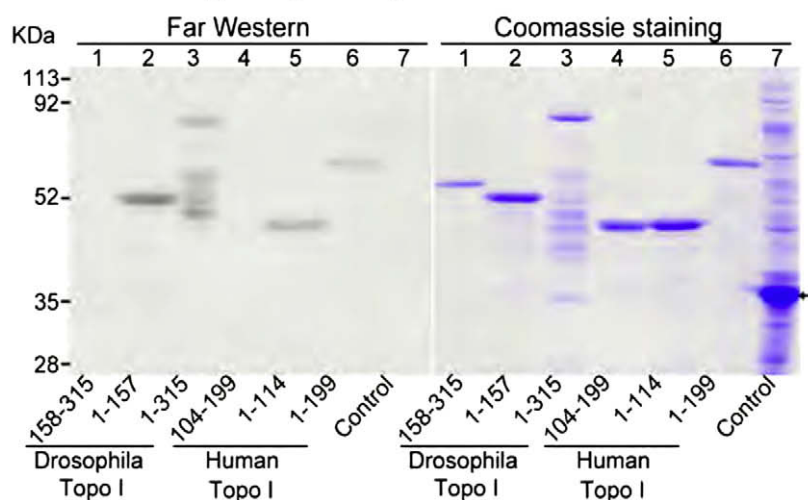

B. Binding to phosphoCTD

Fig. 1. N-terminus of Topo I physically interacts with hyperphosphorylated CTD (PCTD). (A) Diagram of GST fusion constructs for *Drosophila* and human Topo I N-termini. Open (white) portions represent GST tags. Numbers in colored portions indicate amino acid residues. (B) Purified Topo I polypeptides were subjected to SDS–PAGE, transferred to nitrocellulose membrane, and probed (Far Western) with ^{32}P-labeled ßGal-CTD fusion protein. A duplicate SDS gel was stained with Coomassie Blue. The Control lane (lane 7) was loaded with lysate of BL21(DE3) harboring pET41a, which was induced with 1 mM IPTG for 5 h in 30 °C. Arrow indicates a fusion protein encoded by pET41a (GST plus an additional 98 amino acids).

subjected to SDS–PAGE and either stained with Coomassie blue or blotted to nitrocellulose and given a chance to renature. The blot was reacted with a ^{32}P-labeled phosphoCTD probe, a ßGal-CTD fusion protein exhaustively phosphorylated by yeast CTDK-I (catalytic subunit = Ctk1), and autoradiographed. As shown in Fig. 1B (*left*, Far Western), the ^{32}P-labeled PCTD probe bound to full length NTD from both *Drosophila* and human Topo I (lanes 3 and 6, respectively). The *Drosophila* NTD had apparently suffered some proteolysis during expression and purification (lanes 3), but this afforded a demonstration of binding specificity, as only a subset of the fragments (*right*, stained gel) bound the PCTD probe (*left*). Very low background binding was achieved in the experiment, as indicated by the probe's lack of reactivity with GST or any bacterial proteins (lanes 7).

The most N-terminal half of the *Drosophila* NTD (1–157) also bound to the PCTD probe, and the amount of binding (left) relative to the amount of protein (right) was about the same as for full length NTD. Correspondingly, the other half of the *Drosophila* NTD (158–315) showed no detectable binding to the PCTD probe (lanes 1). In a similar vein, the most N-terminal fragment of human

Topo I (lanes 5, 1–114) bound as well as the full length NTD (1–199), whereas the slightly overlapping fragment, 104–199, showed little binding (lanes 4). These results demonstrate that the phosphoCTD-interacting domain (PCID) of Topoisomerase I resides in the N-terminal half of the NTD.
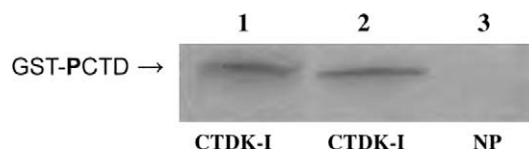

Fig. 2. "Reverse" Far Western blot probed with labeled N-terminus of human Topo I. The blot was prepared by SDS–PAGE of 1 µg of unlabeled phosphorylated or non-phosphorylated GST-CTD fusion protein per lane, followed by blotting to nitrocellulose [8]. The protein segment comprising amino acids 1–199 of human Topo I was labeled by ^{32}P with casein kinase II and used as a probe. Lanes 1 and 2: GST-CTD fusion protein phosphorylated by CTDK-I; lane 3: non-phosphorylated GST-CTD fusion protein.
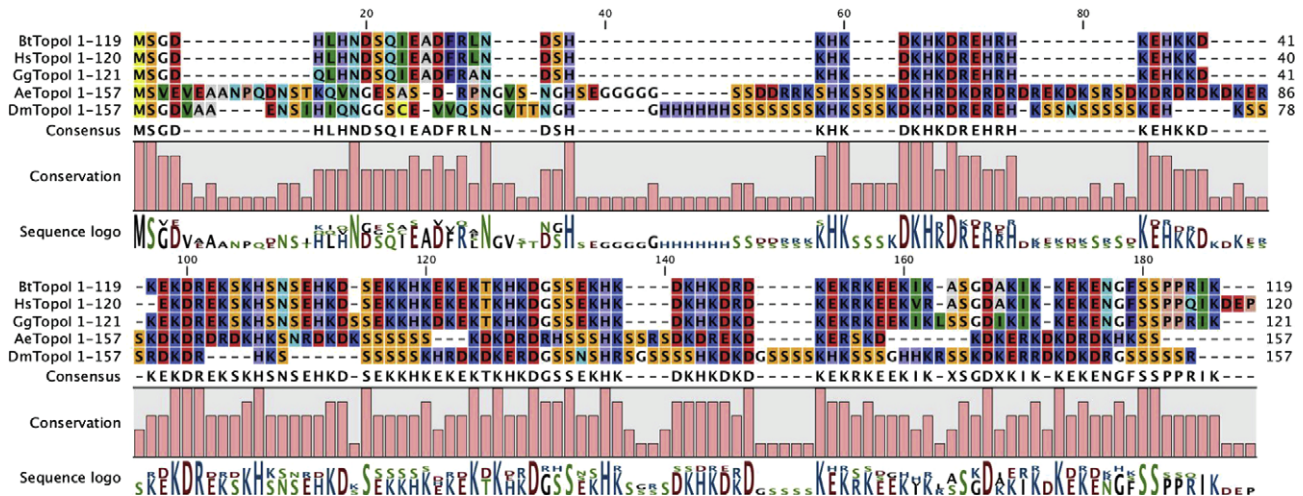
Fig. 3. N-terminal segments of Topoisomerases I aligned. The indicated segments of Topo I from *Bos taurus* (Bt), *Homo sapiens* (Hs), *Gallus gallus* (Gg), *Aedes aegypti* (Ae), and *Drosophila melanogaster* (Dm) were aligned using the Alignment program within CLC Combined Workbench, v 3.0.3 www.clcbio.com, using a Gap open cost of 6, a Gap extension cost of 0.2, and an End gap cost of Free.

The PCTD binding of the NTD from human Topo I was investigated further, using a "reverse" Far Western approach [8]. A GST-CTD fusion protein, after phosphorylation in the absence of radioactivity, was subjected to SDS–PAGE and blotted to nitrocellulose. The blot was then probed with native NTD labeled with ^{32}P. As shown in Fig. 2, the NTD bound to GST-CTD phosphorylated by yeast CTDK-I (duplicates, lanes 1 and 2), but it did not bind to the non-phosphorylated GST-CTD fusion protein (lane 3, NP).

4. Discussion

We show that a phosphoCTD-interacting domain (PCID) resides in the N-terminal segment of Topoisomerase I from *D. melanogaster* and *H. sapiens*. This PCID binds to a CTD exhaustively phosphorylated by yeast CTDK-I; we surmise therefore that it binds to CTD repeats phosphorylated in a Ser2,5P pattern [9] and potentially to repeats phosphorylated in other patterns. The presence of this PCID in the NTD of Topo I provides a molecular explanation for the ability of this domain to target an NTD-ßGal fusion protein to transcriptionally active genes in vivo [3]. More importantly, the long-known in vivo association of Topo I with active genes presumably can be attributed, wholly or in part, to the NTD–PCTD interaction. While we do not know the disposition of the Topo I catalytic domain relative to the "downstream" or "upstream" side of RNAPII, we point out that the length and flexibility of the CTD could probably accommodate either orientation (e.g., Fig. 10 in [5]).

Comparing the amino acid sequences of Topo I N-terminal segments from three vertebrates (Bt, Hs, and Gg) and two insects (Ae and Dm) reveals several homologous regions shared by these organisms (Fig. 3). Both insect NTDs contain numerous Serine runs and a few other inserts not found in the vertebrate proteins. The conserved N-terminal regions are very rich in charged residues, and a fair number of completely conserved positions contain Arg, Lys, Asp, or Glu. On the other hand, 11 of the 26 completely conserved amino acids (as defined by this alignment, and ignoring the initial Met Ser sequence) are either His, Asn or Ser. The absence of conserved aromatic and hydrophobic residues distinguishes the Topo I PCID from previously-described CID, WW, FF, and SRI domains of other proteins that associate with elongating RNAPII (see, e.g., [10]). Further experiments will be required to describe more fully the attributes of the Topo I PCID at the molecular level; in the meantime we propose to call this novel PCTD-binding entity the Topo I–Rpb1 interacting ("TRI") domain.

References

[1] D.S. Gilmour, G. Pflugfelder, J.C. Wang, J.T. Lis, Topoisomerase I interacts with transcribed regions in *Drosophila* cells, Cell 44 (1986) 401–407.
[2] D.S. Gilmour, S.C. Elgin, Association of topoisomerase I with transcriptionally active loci in *Drosophila*, NCI Monogr. 1987 (1987) 17–21.
[3] W.L. Shaiu, T.S. Hsieh, Targeting to transcriptionally active loci by the hydrophilic N-terminal domain of *Drosophila* DNA topoisomerase I, Mol. Cell. Biol. 18 (1998) 4358–4367.
[4] S.M. Carty, A.L. Greenleaf, Hyperphosphorylated C-terminal repeat domain-associating proteins in the nuclear proteome link transcription to DNA/chromatin modification and RNA processing, Mol. Cell. Proteomics 1 (2002) 598–610.
[5] H.P. Phatnani, J.C. Jones, A.L. Greenleaf, Expanding the functional repertoire of CTD phosphoCTD and RNA polymerase II: novel phosphoCTD-associating proteins in the yeast proteome, Biochemistry 43 (2004) 15702–15719.
[6] J.M. Lee, A.L. Greenleaf, A protein kinase that phosphorylates the C-terminal repeat domain of the largest subunit of RNA polymerase II, Proc. Natl. Acad. Sci. USA 86 (1989) 3624–3628.
[7] D.P. Morris, J.M. Lee, D.E. Sterner, W.J. Brickey, A.L. Greenleaf, Assaying CTD kinases in vitro and phosphorylation-modulated properties of RNA polymerase II in vivo, Methods 12 (1997) 264–275.
[8] H.P. Phatnani, A.L. Greenleaf, Identifying phosphoCTD-associating proteins, in: D.R. Schoenberg (Ed.), Methods in Molecular Biology: mRNA Processing and Metabolism, Humana Press, Totowa, NJ, 2004, pp. 17–28.
[9] J.C. Jones, H.P. Phatnani, T.A. Haystead, J.A. MacDonald, S.M. Alam, A.L. Greenleaf, C-terminal repeat domain kinase I phosphorylates Ser2 and Ser5 of RNA polymerase II C-terminal domain repeats, J. Biol. Chem. 279 (2004) 24957–24964.
[10] H.P. Phatnani, A.L. Greenleaf, Phosphorylation and functions of the RNA polymerase II CTD, Genes Dev. 20 (2006) 2922–2936.