THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 37, Issue of September 10, pp. 38952–38959, 2004
Printed in U.S.A.

Insights into the Oligomeric States, Conformational Changes, and Helicase Activities of SV40 Large Tumor Antigen*

Received for publication, June 3, 2004
Published, JBC Papers in Press, July 9, 2004, DOI 10.1074/jbc.M406160200

Dahai Gai‡§, Dawei Li‡¶, Carla V. Finkielstein‡, Robert D. Ott‖, Poonam Taneja‖, Ellen Fanning‖, and Xiaojiang S. Chen‡§‡‡**

From the ‡Biochemistry and Molecular Genetics, University of Colorado Health Sciences Center, School of Medicine, Denver, Colorado 80262 and ‖Biological Sciences, Vanderbilt University, Nashville, Tennessee 37235

The large T (LT) antigen encoded by SV40 virus is a multi-domain, multi-functional protein that can not only transform cells but can also function as an efficient molecular machine to unwind duplex DNA for DNA replication. Here we report our findings on the oligomeric forms, domain interactions, and ATPase and helicase activities of various LT constructs. For the LT constructs that hexamerize, only two oligomeric forms, hexameric and monomeric, were detected in the absence of ATP/ADP. However, the presence of ATP/ADP stabilizes LT in the hexameric form. The LT constructs lacking the N- and C-terminal domains, but still retaining hexamerization ability, have ATPase as well as helicase activities at a level comparable to the full-length LT, suggesting the importance of hexamerization for these activities. The domain structures and the possible interactions between different LT fragments were probed with limited protease (trypsin) digestion. Such protease digestion generated a distinct pattern in the presence and absence of ATP/ADP and Mg^{2+}. The most C-terminal fragment (residues 628–708, containing the host-range domain), which was thought to be completely unstructured, was somewhat trypsin-resistant despite the presence of multiple Arg and Lys, possibly due to a rather structured C terminus. Furthermore, the N- and C-terminal fragments cleaved by trypsin were associated with other parts of the molecule, suggesting the inter-domain interactions for the fragments at both ends.

SV40 large T (LT)[1] antigen is a multi-functional protein that transforms cells and is important for viral DNA replication (reviewed in Refs. 1 and 2). It contains 708 amino acids that fold into several functional domains (Fig. 1) (reviewed in Refs. 1–3). The N-terminal domain containing the first 80 residues is a DnaJ homology domain that interacts with Hsc70 for remodeling the pRB-E2F complex to promote cell growth (4–7). Residues 131–250 fold into the origin binding domain (OBD) for recognizing the viral origin of replication (8–10). Following that is the helicase domain of LT that has recently been mapped to residues 251–627 (11). The functional helicase domain actually contains three separable structural domains: a zinc domain from residues 266–345; an AAA^+ module from residues 415–548 that participate in ATP binding and hydrolysis; and a globular domain containing residues 346–414 and 549–627 (11). The helicase domain that is located on the surface also interacts with other key cellular proteins (such as p53, p300/CREB-binding protein, and pol α/primase) and binds single-stranded DNA without sequence specificity, a function essential for the helicase activity (2, 12–16). The structure and functions of the C-terminal fragment (residues 628–708) are not well defined, but it is thought to be non-structured and part of it (residues 682–708) play a role in host-range determination and in viral assembly (1, 2, 17, 18). Despite the growing knowledge regarding the domain structure and functions of LT, it is not clear how different domains of LT are organized relative to each other or whether these separate domains interact with each other within one LT molecule.

LT has been reported to exist in multiple oligomeric forms including monomeric, dimeric, trimeric, tetrameric, and pentameric in the absence of $ATP+Mg^{2+}$ (19). In the presence of $ATP+Mg^{2+}$, the smaller oligomeric forms assemble into hexamers (19) or even double hexamers if the origin DNA is also present (10, 20, 21). All of the oligomeric states previously reported for LT were obtained using native gel shift assay with protein samples treated by cross-linking reagent glutaraldehyde. However, using the gel-filtration chromatography in buffers without any cross-linking reagents, a truncated LT fragment containing residues 251–627 (LT251–627) was shown to form hexamers even in the absence of $ATP+Mg^{2+}$ (11). Furthermore, this LT251–627 construct exhibited only one other oligomeric form, possibly a monomer, under the tested condition, which is close to physiological (11). It is not clear whether the behavior of only two discreet oligomeric forms in the absence of cross-linking reagents is unique to this particular LT construct or if it is the common property for other longer LT constructs including the full-length LT.

To further understand the biochemical behaviors of LT, the oligomeric states of LT polypeptides of various constructs purified from *Escherichia coli* as well as insect cells have been characterized under native buffer conditions without cross-linking reagents. In addition, the domain structures and inter-domain interactions of LT have been studied using a combina-

* This work was supported in part by a fellowship (to D. G.) from Colorado Cancer League, a postdoctoral fellowship (to C. V. F.) from the American Heart Association (0225388Z), and a Scholar Research grant (to X. S. C.) from the American Cancer Society (GMC-106782). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Present address: Molecular and Computational Biology, University of Southern California, Los Angeles, CA 90089.

¶ Present address: National Key Laboratory, China Agriculture University, Beijing 100094, China.

** Present address: Arcturus Engineering, Inc., 400 Logue Ave., Mountain View, CA 94043.

‡‡ To whom correspondence should be addressed: Molecular and Computational Biology, University of Southern California, Los Angeles, CA 90089-1340. Tel.: 213-740-5487; Fax: 213-740-8631; E-mail: Xiaojiac@usc.edu.

[1] The abbreviations used are: LT, large T; OBD, origin-binding domain; FPLC, fast-flow performance liquid chromatography system; GST, glutathione *S*-transferase; pk, peak; sPk, shoulder peak; MM, molecular mass; CREB, cAMP-response element-binding protein.

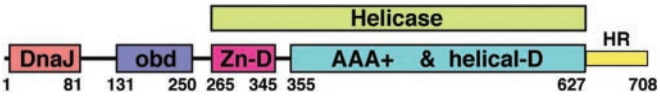


FIG. 1. **Cartoon representation of the known LT domain structures.** The amino acid numbers are indicated at the *bottom*. The functional domains are represented by *open boxes* and are labeled accordingly. The linkers between domains are represented by *thin lines*. The C-terminal domain from residue 628 to 708, which contains the host-range fragment (residues 682–708), is labeled HR (in *yellow*) for convenience. The HR domain is thought to be unstructured.

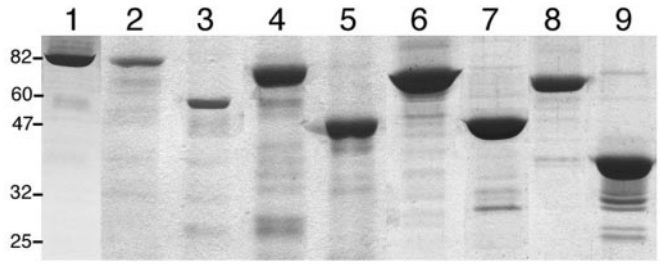


FIG. 2. **SDS-PAGE analysis of the proteins of various LT constructs expressed and purified from insect cells or *E. coli*.** *Lane 1* is the full-length LT (LT1–708) purified from insect cells. *Lanes 2–9* are LT proteins from different deletion constructs expressed and purified from *E. coli* cells. *Lanes 2–3*, *4–5*, *6–7*, and *8–9* are the GST-LT fusions or the corresponding LT proteins cleaved from the GST fusion from the constructs LT131–627, LT251–627, LT262–627, and LT303–627, respectively.

tion of limited proteolysis and gel-filtration column chromatography on the fast-flow performance liquid chromatography system (FPLC, Amersham Biosciences). Furthermore, the helicase and ATPase activities of the LT proteins have been assayed to further map the helicase domain and to establish the correlation between oligomerization and enzymatic (ATPase and helicase) activities. Here we show that LT proteins of various truncations, even though purified from *E. coli*, have similar biochemical properties and ATPase/helicase activities to the full-length LT obtained from insect cells.

EXPERIMENTAL PROCEDURES

LT Expression and Purification—Various truncated LT constructs for *E. coli* expression were constructed with a glutathione *S*-transferase (GST) fused at the N terminus of LT using pGEX-2T vector (Amersham Biosciences). In this GST-LT fusion protein, a thrombin cleavage site is placed at the boundary between GST and LT. The expression of the GST-LT fusion proteins in *E. coli* was controlled under Ptac promoter. To induce the expression of the fusion proteins, isopropyl-1-thio-β-D-galactopyranoside was added to the *E. coli* cells at a density of A_{600} = 0.4 at room temperature (25 °C). To purify the proteins, *E. coli* cells expressing the GST-LT fusion proteins were lysed by sonication in a buffer L (containing 50 mM Tris (pH 8.0), 250 mM NaCl, and 1 mM dithiothreitol). After separating the soluble supernatant of the cell lysates from the pellet by centrifugation (25,000 × *g*), the fusion proteins in the supernatant fraction were purified by passing through the glutathione affinity column (Amersham Biosciences) followed by washing the column with 10× bed volume of buffer L. To separate LT from the GST fusion, thrombin was used to cleave the GST-LT fusion. The thrombin-cleaved proteins were concentrated for further purification with gel-filtration chromatography on a Superdex-200 column with buffer S (containing 25 mM Tris-Cl (pH 8.0), 250 mM NaCl, and 1 mM dithiothreitol). The expression and purification of the full-length LT (LT1–708) from insect cells was performed as previously described (22).

Limited Trypsin Digestion and Characterization of the Digested Products—Limited trypsin digestion of LT was performed at room temperature, and the digestion was terminated at the designed time interval by adding 1 mM phenylmethylsulfonyl fluoride and an equal volume of 2× SDS sample buffer followed by immediate boiling. The trypsin to LT ratio (wt/wt) used was 1:400. If ATP or ATP+Mg^{2+} were included in the digestion, LT was first incubated with 1 mM ATP or 1 mM ATP plus 3 mM $MgCl_2$ for 30 min at room temperature before adding trypsin.

For N-terminal sequencing, the tryptic LT samples were prepared in two ways. One was to separate all of the obvious cleavage fragments shown on SDS-PAGE and then blot them to polyvinylidene difluoride membrane (Millipore) for the N-terminal sequencing of each individual fragment. This method of preparing the sample gives unambiguous sequencing results for each isolated fragment but may only obtain the N-terminal sequences of those that show up well in the SDS-PAGE gel. To avoid missing the N-terminal sequences of small cleavage fragments that do not stain well or run well on the gel, the bulk of digested LT fragments were also directly subjected to N-terminal sequencing.

For gel-filtration chromatography, prepacked Superdex-200 column (16/60) from Amersham Biosciences was used on the ΔKTA FPLC system. The column was equilibrated with 1.5× bed volume of buffer S (containing 25 mM Tris-Cl (pH 8.0), 250 mM NaCl, and 1 mM dithiothreitol) before the protein samples were injected onto the column. A flow rate of 1 ml/min (a pressure <42 p.s.i.) was used for the chromatography run.

Helicase Assay—For the helicase assay, the double-stranded DNA substrate is made from the annealed product between a single-stranded DNA oligonucleotide primer and the M13mp19 single-stranded DNA, which is adapted from previously reported M13 substrates (23–25). The oligonucleotide primer contains 60 nucleotides with the sequence of 5′-TTT TTT TTT TTT TTT TTT TTT TTT TTC GCG CGG GGA GAG GCG GTT TGC GTA TTG GGC GCC-3′. Its 5′-end was end-labeled with [γ-^{32}P]ATP by the polynucleotide kinase. The labeled primer was then annealed with M13mp19 single-stranded DNA, and the resulting double-stranded DNA substrate was purified with a Sephacryl 300HR column as described previously (23). Approximately 10 fmol of substrate DNA was incubated with ~2 pmol of LT proteins (monomer) of various constructs in a helicase buffer H containing 20 mM Tris (pH 7.5), 10 mM $MgCl_2$, 4 mM ATP, 0.1 μg/μl bovine serum albumin, and 1 mM dithiothreitol for 45 min at 37 °C. The reaction was terminated by introducing SDS to a final concentration of 0.25% and EDTA to 50 mM. The samples then were analyzed with 12% native polyacrylamide gel in Tris borate-EDTA buffer. The gel was dried, and positions of the labeled DNA were detected using autoradiography.

ATPase Assay—For the ATPase assay, 4 pmol of LT proteins (monomer) from various constructs were incubated with 1 μCi of [γ-^{32}P]ATP (Amersham Biosciences) in a buffer containing 20 mM Tris (pH 7.5) and 10 mM $MgCl_2$ for 30 min at room temperature. 20% of the reaction mixtures were spotted and dried on a thin layer chromatography (TLC) plate (Selecto Scientific). The TLC plate was placed in a developing buffer containing 2 M acetic acid and 0.5 M LiCl for 2 h. The plate then was air-dried, and the positions of the radioisotope were detected with autoradiography.

RESULTS

Protein Expression and Purification—LT proteins with different truncations have been purified from either insect cells or from *E. coli*. The full-length LT protein (LT1–708) was expressed and purified from insect cells as described previously (22). The shorter LT polypeptides were expressed and purified from *E. coli* as GST fusions at their N termini. The GST fusion proteins were cleaved by thrombin to separate GST and LT moiety. After a final purification step using Superdex-200 gel-filtration chromatography on FPLC, all of the proteins have a purity of >95% (Fig. 2). The yields of LT polypeptides from *E. coli* varied depending on the constructs. The general trend is that the longer the constructs, the lower the yield. Approximately 0.2 mg of LT131–627/liter of cell culture was obtained; however, as much as 10 mg of LT303–627 could be purified from 1 liter of cell culture.

Oligomeric States of LT in the Presence and Absence of ATP/ADP and Mg^{2+} Ion—It has been reported that full-length LT exists in multiple oligomeric states ranging from monomeric through hexameric as assayed by a cross-linking experiment in solutions containing glutaraldehyde (19). However, the deletion construct of LT containing residues 251–627 exhibited only two oligomeric states, hexameric and monomeric, as assayed by gel-filtration chromatography in a buffer solution without any cross-linking reagent (11). To understand whether the two-oligomeric phenomenon is unique to LT251–627 or is a general feature for other LT constructs, LT polypeptides of various constructs, including the full-length LT, were subjected to gel-filtration chro-

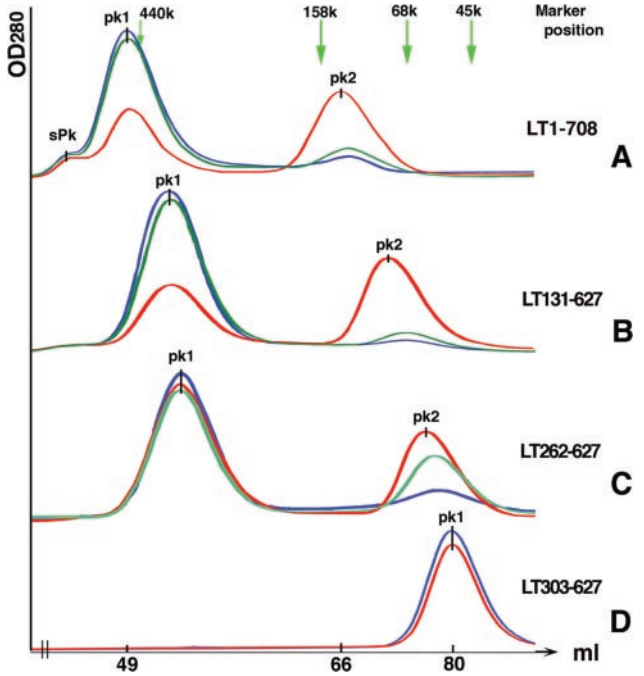


FIG. 3. **Oligomerization assay of proteins of various LT constructs in the presence and absence of ATP or ADP using Superdex-200 gel-filtration column chromatography on FPLC.** The protein peaks (*y* axis) are detected by UV absorption (A_{280}) over elution volume (*x* axis, in ml). The positions of elution peaks of the protein size markers are indicated by *green arrows* at the *top*. The LT constructs are listed on *the right side* of their corresponding elution profiles. Color usage in the profile: *red line*, elution profile without ATP/ADP; *blue line*, with ATP; *green line*, with ADP. In the absence of nucleotides, only two predominant peaks corresponding to two oligomeric states are observed for all of the constructs with the exception of LT303–627.

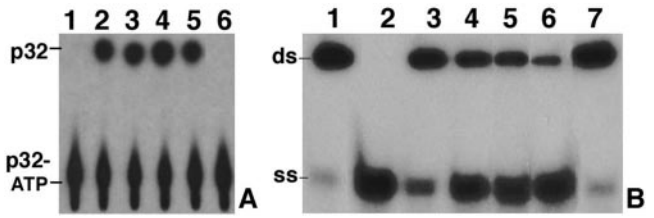


FIG. 4. **Helicase and ATPase activities of various LT constructs.** *A*, ATPase assay on TLC plate. All of the lanes contain 1 μCi of [γ-^{32}P]ATP. *Lane 1*, no protein (*lanes 2–6* contains ~4 pmol (calculated as monomers) of various LT proteins); *lane 2*, full-length LT (LT1–708) from insect cells; *lane 3*, LT131–627; *lane 4*, LT251–627; *lane 5*, LT262–627; *lane 6*, LT303–627. Only LT303–627 did not have detectable ATPase activity. *B*, helicase assay. All of the lanes contain the same amount of DNA substrate (~10 fmol). *Lanes 1* and *2*, DNA substrate alone and boiled DNA substrate alone (*lanes 3–7* contains ~2 pmol of LT polypeptides (calculated as monomers) expressed from various constructs; *lane 3*, LT1–708; *lane 4*, LT131–627; *lane 5*, LT251–627; *lane 6*, LT262–627; and *lane 7*, LT303–627. *ds*, double-stranded; *ss*, single-stranded.

matography analysis using a Superdex-200 column on the FPLC system under physiological buffer conditions.

The elution profiles of all of the tested LT constructs on the gel-filtration chromatography are shown in Fig. 3. For the full-length LT (Fig. 3*A*, *LT1–708*), two prominent elution peaks (*pk1* and *pk2*) were present in the absence of ATP+Mg^{2+} (*red line*). The relative ratio of the two-peak height (pk1:pk2) changes depending on the protein concentration; the higher the protein concentration, the larger the ratio of pk1:pk2, suggesting high protein concentration favoring pk1 formation. In addition, there is a small shoulder peak (sPk) at the void volume position (peak at ~46 ml) that contains protein aggregates (but soluble) with a large molecular mass beyond the resolution limit of this column (~700 kDa). The first elution peak (pk1) has an apparent molecular mass (MM) of 540 kDa, consistent with a hexameric form of LT with a calculated MM of 490 kDa. The second peak (pk2) has an apparent MM of 138 kDa, which is between the calculated MMs of a monomer (81.6 kDa) and a dimer (163 kDa). When ATP was added to the protein and incubated for 30 min at 25 °C before the sample was injected onto the gel-filtration column, the elution profile of LT1–708 changed dramatically (*blue line*). Now, the height of pk2 almost reduced to base line, whereas that of the hexameric pk1 increased correspondingly, suggesting that the protein in pk2 shifts into hexameric forms in the presence of ATP. ADP has a similar effect in shifting the protein in pk2 to pk1 under the test condition (*green line*).

Similar to full-length LT, two other truncated LT constructs, including LT131–627 and LT262–627, also showed two discreet peaks (pk1 and pk2) in their elution profiles in the absence of nucleotide (Fig. 3, *B* and *C*, *red lines*). Pk1 of LT131–627 has an apparent MM of 330 kDa (Fig. 3*B*), consistent with a hexameric form (calculated MM of 351 kDa). Pk2 of LT131–627 has an apparent MM of 82 kDa, which is closer to the calculated MM of a monomer (57.5 kDa) than that of a dimer (115 kDa). For the LT262–627 construct, pk1 has an apparent MM of 246 kDa (Fig. 3*C*), again consistent with a hexameric form with a calculated MM of 252 kDa. pk2 of LT262–627 has an apparent MM of 55 kDa similar to the calculated MM of a monomer (43.1 kDa). For both LT131–627 and LT262–627 constructs similar to LT1–708, the elution profiles in the presence of ATP or ADP have a reduced pk2 and increased pk1 (Fig. 3, *B* and *C*, *blue* and *green lines*), suggesting an ability of ATP or ADP to shift the proteins in the oligomeric state in peak 2 to the hexameric state in peak 1. However, for construct LT303–627, the polypeptide only eluted as a single peak regardless of the presence or absence of ATP (or ADP) treatment before chromatography (Fig. 3*D*), suggesting the lack of conformational change when ATP/ADP was added. The peak position has an apparent MM of 49 kDa similar to the calculated MM of a monomer for this fragment (38.1 kDa).

Even though both ATP and ADP can shift the LT proteins of various constructs from the oligomeric state in pk2 to the hexameric state in pk1, the residual pk2 height in the presence of ADP (*green lines* in Fig. 3) is always higher than that in the presence of ATP under the same nucleotide concentration (*blue lines* in Fig. 3), suggesting a less efficient hexamerization of LT in the presence of ADP. This phenomenon is consistent with the previous finding that ADP has lower affinity for LT than ATP (26). This difference between ATP and ADP becomes more evident with the shorter construct, LT262–627. Whereas the peak heights of pk2 for LT1–708 and LT131–627 were reduced to almost background level when ADP was present (Fig. 3, *A* and *B*), the residual pk2 of LT262–627 was quite large in the presence of ADP (Fig. 3*C*, *green line*), suggesting that the hexameric forms of the shorter LT construct is not very stable even in the presence of ADP. Another observation is that the residual pk2 positions in the presence of ATP or ADP in Fig. 3 are slightly later (~0.8–1.5 ml) than the pk2 positions in the absence of nucleotides for all of the constructs with the exception of LT303–627 (Fig. 3, *A–C*). Because a later elution peak means a smaller apparent MM, the presence of nucleotides (ATP/ADP) may trigger a structural change that leads to a more compact conformation, thus behaving as a smaller MM in gel filtration.

ATPase and Helicase Activities of Different Deletion Constructs—To test whether the various deletion LT constructs from *E. coli* have enzymatic activities, ATPase and helicase assays were performed. With the exception of LT303–627, all of

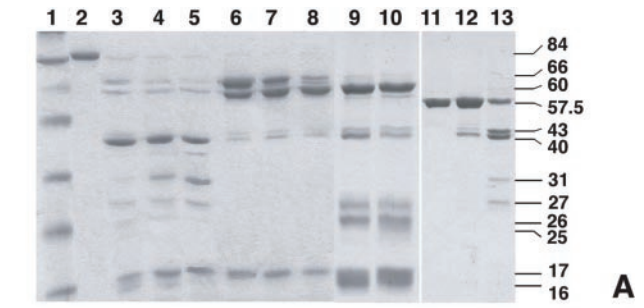
the tested LT constructs including wild type LT from insect cells showed comparable level of ATPase activity (Fig. 4*A*). Similarly, the helicase assay showed that all of the LT constructs with the exception of LT303–627 had helicase activity (Fig. 4*B*, *lanes 3–7*). Judging from the band intensity of the unwound single-stranded DNA product in Fig. 4*B*, the helicase activities of the deletion constructs are somewhat comparable with those of the full-length LT.

The Characterization of LT Domain Structure—To investigate the domain structure of the full-length LT as well as to probe the domain interactions of LT, we used a combination of methods including limited protease (trypsin) digestion in the presence and absence of ATP plus/minus Mg^{2+} ion, N-terminal sequencing of the cleavage fragments, and gel-filtration chromatography of the trypsinized LT protein. The full-length LT (LT1–708) from insect cells and LT131–627 purified from *E. coli* were first treated with trypsin in the absence or presence of ATP and Mg^{2+} for different lengths of time followed by SDS-PAGE analysis. Distinct tryptic band patterns generated in the presence or absence of ATP and Mg^{2+} ion were shown on the denaturing SDS-PAGE gel (Fig. 5*A*). In the absence of ATP and Mg^{2+}, 10 tryptic bands were detected including four major bands with apparent MMs of 40, 31, 27, and 17 kDa and five minor bands at 66-, 60-, 43-, 26-, and 25-kDa positions (Fig. 5*A*, *lanes 3–5*). However, in the presence of ATP+Mg^{2+}, there are only three major tryptic bands with apparent MMs of 66, 60, and 17 kDa plus two minor bands at the 40/43-kDa position (Fig. 5*A*, *lanes 6–8*).

Interestingly, the trypsin digestion with ATP but without Mg^{2+} yielded a somewhat different band pattern (Fig. 5*A*, *lanes 9* and *10*) from that either with ATP+Mg^{2+} (Fig. 5*A*, *lanes 6–7*) or without ATP+Mg^{2+} (Fig. 5*A*, *lanes 3–5*). This pattern with only ATP but without Mg^{2+} showed fewer tryptic bands than that derived without ATP+Mg^{2+}, but it showed more bands than that with ATP+Mg^{2+}. Similar effects of ATP and Mg^{2+} on changing trypsin digestion pattern were also observed for LT131–627 (Fig. 5*A*, *lanes 11–13*).

To learn the trypsin cleavage sites on LT that indicate the accessibility of the sites in the folded structure, the tryptic fragments obtained with/without ATP+Mg^{2+} were N-terminally sequenced and the sequencing results are shown in Fig. 5*B*. Based on the N-terminal sequencing results as well as the apparent MM of the tryptic fragments on SDS-PAGE, a trypsin cleavage map showing the cleavage positions on LT polypeptide, as well as the resulting fragments, was derived (Fig. 5*C*). In the presence of ATP+Mg^{2+}, the cleavage by trypsin occurred only at two locations: residue 131 and 651 (indicated by *arrows* in Fig. 5*C*). A possible third cleavage site at 126 is deduced by the presence of a smaller 16-kDa fragment that is co-N-terminal with the 17-kDa fagment (Fig. 5, *B* and *C*). Whereas the cleavage in the absence of ATP and Mg^{2+} was observed to occur at seven locations (*red* and *black arrows* in Fig. 5*C*), the estimated MMs of the cleavage products on SDS-PAGE are mostly consistent with the calculated MMs based on the amino acid composition of these products.

Interactions between Different LT Domains—Three of the four major structural and functional domains of LT have been solved separately using either x-ray crystallography or NMR methods. These three domains with known structures are (counting from the N terminus) as follows: the DnaJ homology domain (residues 1–117) (27); the OBD domain (residues 131–260) (9); and the helicase domain (residues 251–627) (11), leaving the C-terminal HR fragment (residues 628–708), the only major domain, with missing structural information. Despite extensive efforts, the following important issues still remain to be addressed: how these four domains are organized in an intact LT molecule; whether the four domains interact with each other; and if they do, how the four domains interact with each other intramolecularly in a monomeric LT and intermolecularly in an oligomeric LT.



No ATP		Plus ATP and Mg++	
tryptic fragment	N-term Sequences	tryptic fragment	N-term sequences
43,000	131-KVEDPKD	66,000	131-KVEDPKD
40,000	131-KVEDPKD	60,000	131-KVEDPKD
31,000	131-KVEDPKD		(137-DFPSELL) *
27,000	131-KVEDPKD	17,000	N-Blocked
26,000	513-HLNKRTQ	16,000	N-Blocked
25,000	518-TQIFPPG		
17,000	N-Blocked		
16,000	N-Blocked		

Additional N-terminal sequences obtained by sequencing the bulk trypsin digested LT without gel separation:

1. 348-RVDSLQL 2. 401-MDSVVYD 3. 484-DLPSGQG 4. 555-HCLERSE

B

* the sequence was detected at very low level.

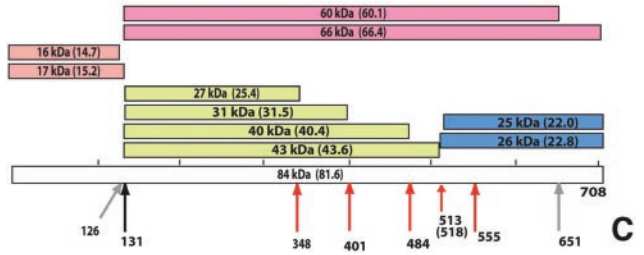


FIG. 5. **The results of limited protease digestion of LT.** *A*, SDS-PAGE analysis of the trypsin digestion of LT in the presence or absence of ATP plus or minus Mg^{2+}. *Lanes 1–10* are LT1–708 from insect cell, and *lanes 11–13* are LT131–627 from *E. coli*. *Lanes 2* and *11* did not contain trypsin. *Lanes 3*, *4*, and *5* (without ATP) or *lanes 6*, *7*, and *8* (with ATP+Mg^{2+}) were LT1–708-treated with trypsin for 20, 40, and 60 min at 25 °C. *Lanes 9* and *10* (with ATP but without Mg^{2+}) were LT1–708-treated with trypsin for 20 and 40 min at 25 °C. *Lanes 12* and *13* were LT131–627-treated with trypsin in the presence of both ATP+Mg^{2+} or ATP alone. *B*, the results of the N-terminal sequencing of the LT fragments generated by limited trypsin digestion in the presence or absence of ATP+Mg^{2+}. The various tryptic fragments are separated by SDS-PAGE and then blotted on a polyvinylidene difluoride membrane for N-terminal sequencing. At the same time, the bulk of digested LT fragments were directly subjected to N-terminal sequencing to obtain the N-terminal sequences of small fragments that do not stain well on the gel or are too small to show up on the gel. The 60-kDa fragment obtained in the presence of ATP+Mg^{2+} contains a major sequence starting at 131, but it also contains a minor sequence at 137 (within *parentheses*), suggesting that an inefficient cleavage also occurred at residue 137. *C*, a diagram showing the trypsin cleavage sites on LT polypeptide. *Open box* represents the whole LT polypeptide. *Arrows* in *red* and *black color* indicate all of the observed cleavage sites in the presence and absence of ATP. The *red arrows* indicate the cleavage sites that are accessible in the absence of ATP+Mg^{2+} but are protected in the presence of ATP+Mg^{2+}. The *arrows* in *gray* indicate the cleaved sites deduced based on the size of the trypsin fragment on the SDS-PAGE and the N-terminal sequencing result. The *number* given at each arrow position indicates the first residue number at the N terminus of the cleaved products. The fragments generated by trypsin digestion in the presence or absence of ATP+Mg^{2+} are represented by *shaded boxes*. The apparent MM of each cleavage product is shown inside the corresponding boxes with the calculated MM within *parentheses*.

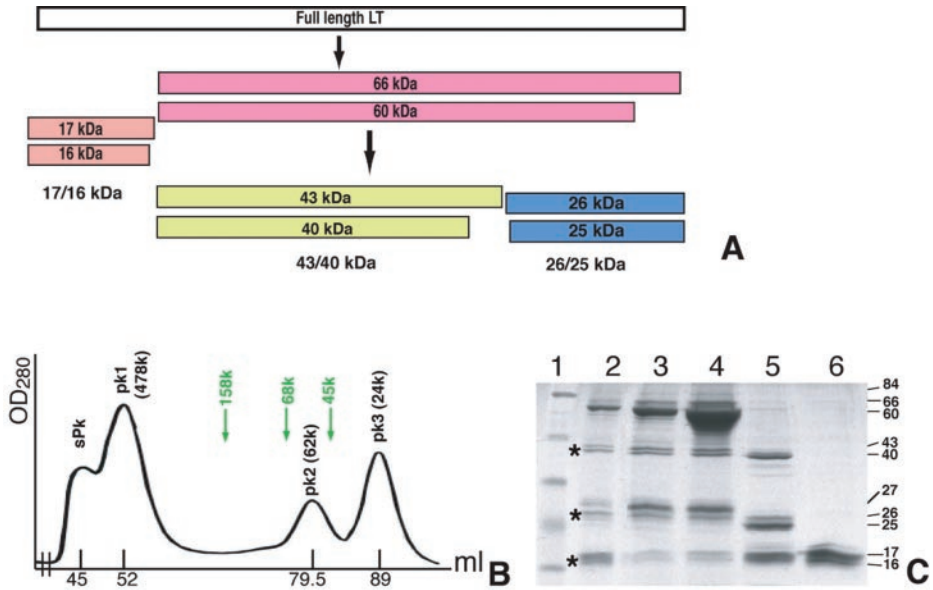


FIG. 6. **The association of cleaved fragments of LT after trypsin digestion.** *A*, diagram representing the trypsin-digested product of LT in the presence of ATP but in the absence of Mg^{2+} ion (see *lane 2* in Fig. 6*C*). The trypsin digestion was performed in the presence of 1 mM ATP (no Mg^{2+}) with a trypsin:LT ratio of 1:400 (wt/wt) for 20 min at 25 °C. *B*, the gel-filtration chromatographic profile of the full-length LT treated by trypsin in the presence of ATP but in the absence of Mg^{2+} ion in Superdex-200 column chromatography on FPLC. The gel-filtration chromatography should separate the various tryptic fragments into different peaks according to their MMs, unless these fragments still associate with each other after trypsin cleavage. The elution profile shows that there are three prominent peaks (*pk1*, *pk2*, and *pk3*) within the resolution range of the column plus a small void volume peak (or shoulder peak, *sPk*). The elution volume (ml) of the peaks is indicated on the *horizontal axis*, and the corresponding apparent MMs of the three peaks are labeled above those peaks. The positions of the molecular standards are indicated by *arrows* (in *green*). *C*, the SDS-PAGE analyses of each of the fractions shown in *B*. *Lane 1* is the marker; *lane 2* is the trypsin-digested LT before loading on to the column; *lanes 3*, *4*, *5*, and *6* are the fractions taken from sPk, pk1, pk2, and pk3 in *B*. The *stars* on the *left side* of *lane 2* indicate the position of the three groups of trypsitic fragments.

To address some aspects of the above issues, which are important for understanding the complex functions of LT, limited protease (trypsin) treatment was employed to cleave the native LT protein at the accessible sites followed by gel-filtration column chromatography to test whether the cleaved products migrate as individual fragments or as multiple fragment complex that are associated with each other. The rationale for this experiment is that, if the individual fragments elute in separate peaks according to their sizes, these fragments are not interacting with each other. Conversely, if the individual fragment elute together with other parts of the molecule in a peak with apparent MMs larger than the individual fragments, it suggests that the fragments interact with each other.

To examine the interactions of fragments coming from a different portion of LT (*i.e.* N and C termini), it is important to select a suitable condition for trypsin digestion so that a moderate number of fragments are generated. The digestion condition used in *lanes 9–10* of Fig. 5*A* in which LT was digested into three major fragments was chosen for this experiment. The critical feature of this digestion condition is that trypsin treatment is performed in the presence of ATP but in the absence of Mg^{2+} ion. Such trypsin digestion of LT generated the cleavage products that contained the N-terminal 16/17-kDa fragment (LT1–130) and the C-terminal 66/60-kDa fragment that are further cleaved into 43/40 fragments, a 27-kDa fragment, and the C-terminal 25/26-kDa fragment (Fig. 6, *A* and *C*, *lane 2*). When this trypsin-digested LT was subjected to the gel-filtration chromatography on Superdex-200 column, an elution profile with three prominent elution peaks (pk1, pk2, and pk3) and an sPk at the void volume was obtained (Fig. 6*B*). Pk1 has an apparent MM of 478 kDa, slightly smaller than the calculated MM of a full-length LT hexamer (~490 kDa), suggesting that LT proteins in this peak may be in a hexameric form. Interestingly, SDS-PAGE of this peak (Fig. 6*C*, *lane 4*) showed that it contains all of the trypsin-cleaved fragments that are present in the sample before being loaded on to the Superdex-200 column, suggesting that these cleaved fragments interact with each other in the protein complex. Similarly, the sPk, which contains LT in an oligomeric form(s) with an apparent MM of over 700 kDa, also contains all of the trypsin cleavage fragments (Fig. 6*C*, *lane 3*).

Further evidence that the trypsin-cleaved fragments associate with each other is that pk2 in Fig. 6*B* also contains three groups of fragments with apparent MMs of 17/16 kDa, 43/40 kDa, and 26/25 kDa, respectively (Fig. 6*C*, *lane 5*). These three fragments cover the entire region of LT, *i.e.* 16/17-kDa fragments (residues 1–130) at the very N terminus, the 43/40-kDa fragments (residues 131–484/513) in the middle, and the 26/25-kDa fragments (residues 514/518–708) at the very C terminus (Fig. 6*A*). Nonetheless, the third elution peak, pk3, contains only one part of the LT, the N-terminal 17/16-kDa fragments (Fig. 6*C*, *lane 6*), suggesting that some of the cleaved N-terminal fragments may dissociate from the C-terminal portions.

DISCUSSION

The oligomeric forms, domain interactions, and the functional activities of ATPase and helicase of different lengths of LT polypeptides prepared from insect and *E. coli* cells have been studied. Interestingly, only two oligomeric forms, hexameric and monomeric, for all of the LT constructs with the exception of LT303–627 were detected in the absence of ATP/ADP using gel-filtration chromatography. The two oligomeric forms of different LT constructs appear to equilibrate with each other in solution in a certain ratio. However, when ATP/ADP was added to the protein solution, only the hexameric form was present, suggesting that ATP/ADP promotes LT hexamerization. All of the LT constructs capable of hexamerization, including those lacking the N- and C-terminal domains, showed comparable ATPase and helicase activities, suggesting that both the N (residues 1–259) and C termini (residues 628–708) are

dispensable for these enzymatic activities. Limited protease (trypsin) treatment cleaved LT at multiple sites, generating several fragments. However, the cleavage products generated by trypsin differed in the presence and absence of ATP/ADP and Mg^{2+} ion. The presence of both ATP/ADP and Mg^{2+} ion protected several trypsin cleavage sites on LT, possibly due to the formation of the stable hexameric LT when ATP/ADP was present. The trypsin-cleaved LT fragments located at different regions of the full-length LT polypeptide appeared to remain associated with each other in solution, suggesting the presence of interdomain interactions between these fragments.

The detection of only two stable oligomeric forms of LT in the absence of nucleotides is in contrast to the previous thinking that LT exists in multiple oligomeric forms including monomers, dimers, trimers, tetramers, and pentamers in the absence or presence of nucleotides and that hexamers are formed only in the presence of nucleotides (19). The discrepancy in detecting different oligomeric forms of LT could be due to the difference in the detection methods used in these studies. All of previous studies employed native gel shift assay of the LT protein that was cross-linked by glutaraldehyde, a cross-linking reagent that can cross-link proteins that get close to each other through primary amine groups. Glutaraldehyde can also form polymers of itself in solution, which in turn can cross-link proteins over relatively long distances. However, in our studies, no cross-linking reagents were used and the oligomeric forms were assayed by gel-filtration chromatography on an FPLC system under buffers similar to physiological conditions.

Gel-filtration chromatography on FPLC can separate protein molecules through the column based on their MMs. The proteins (or complexes) with larger MMs will elute out of the column earlier in the chromatography process, and the proteins with smaller MMs will elute later. When LT proteins were analyzed by gel-filtration chromatography under native buffer conditions, only two elution peaks representing two oligomeric forms were detected for LT proteins in the absence of ATP/ADP. However, the early peak (pk1) was the only dominant species if LT was incubated with ATP/ADP before subjecting to gel-filtration chromatography. This early peak of LT with or without ATP/ADP treatment has an apparent MM consistent with a hexameric LT. Nonetheless, the late peak (pk2) has an apparent MM somewhere between the calculated MM of a monomer and that of a dimer, raising the issue of whether the protein in pk2 is a monomer or a dimer.

Even though the precise oligomeric form could not be determined for the protein in pk2, several lines of evidence listed below suggest that pk2 represents a monomeric form of LT. First, the shape of a full-length LT molecule is expected to be rather elongated based on the known structures (9, 27–29), whereas its ring-shaped hexamer is more or less globular. Proteins with an elongated shape usually behave "larger" than their actual MM in gel-filtration chromatography. This may explain why the pk2 position for full-length LT has an apparent MM larger than the calculated MM of a monomer (Fig. 3*A*). However, as a deletion construct gets shorter, the pk2 position gets closer to the MM of a monomer as in the case of LT131–627 (Fig. 3*B*) and LT262–627 (Fig. 3*C*), possibly because the shorter constructs become less elongated and thus behave more closely to the real size of the protein. For example, the pk2 of the longest construct LT1–708 has an apparent MM of 138 kDa, 69% more than the calculated MM of a monomer with 81.6 kDa. However, LT262–627, which is less elongated (11) due to the lack of the extra domains (DnaJ and OBD, residues 627–708), has a pk2 with an apparent MM of 55 kDa (Fig. 3*C*), only 27% more than the calculated MM of a monomer (43.1 kDa). Second, the shortest construct (LT303–627) that cannot hexamerize (11) has a single elution peak regardless of whether ATP/ADP is present or not (Fig. 3*D*). The LT303–627 in this peak, which is equivalent to the pk2 of other LT constructs, is considered to be monomeric (11). Here we showed that LT303–627 had neither helicase activity (Fig. 4*B*, *lane 7*) nor any ATPase activity (Fig. 4*A*, *lane 6*) despite the fact that LT303–627 contains an intact ATP binding domain (residues 355–627). The crystal structure of LT251–627 reveals that the residue(s) from a neighboring molecule is important for ATP hydrolysis (11). Thus, a dimer or higher oligomeric forms of LT303–627 should be able to hydrolyze ATP. Therefore, a lack of ATPase activity of LT303–627 is consistent with a monomeric form.

When the two elution peaks of LT obtained in the absence of nucleotides are isolated and subjected to another round of gel-filtration chromatography separately, the same two peaks reappear with a similar ratio, suggesting that the hexameric peak 1 and monomeric peak 2 equilibrate with each other. No intermediate oligomeric forms, such as dimer, trimer, and so forth, were detected under the tested conditions. This behavior of equilibrium between only two oligomeric forms (hexamers and monomers) suggests that the hexamerization is a highly cooperative process. However, when the hexamers are formed in the presence of ATP/ADP, they do not disassemble into monomers even after prolonged incubation (16 h) in buffers without ATP/ADP. However, the hexamers will partially disassemble into monomers after 2 days of incubation in a buffer without ATP/ADP (data not shown). These biochemical behaviors are similar for all of the LT constructs capable of hexamerizing. A functional assay demonstrated that all of the LT polypeptides capable of hexamerization have ATPase and helicase activities. The shortest construct that still hexamerizes is LT262–627. This construct showed the helicase activity that is at least comparable with and may be even higher than that of the full-length LT (Fig. 4*B*).

Previous studies using limited trypsin treatment have identified the site at between Lys^{130}-Lys^{131} in LT as the highly sensitive cleavage site (30, 31). Trypsin is a protease that cleaves at the C terminus of Arg or Lys. However, only those Arg or Lys residues located in linkers, turns and loops, and flexible regions are sensitive to limited trypsin cleavage in a folded protein. As a result, limited trypsin digestion is a frequently used biochemical method for probing the structure of a protein. In this report, we showed that limited trypsin treatment of LT produced different digestion patterns in the presence or absence of ATP+Mg^{2+} (Fig. 5*A*) with fewer digested fragments in the presence of ATP+Mg^{2+}, suggesting that nucleotide binding onto LT protected some of the trypsin sensitive sites, possibly through triggering hexamerization by the nucleotide binding. However, because LT forms hexamers even in the absence of nucleotides and Mg^{2+}, the greater trypsin cleavage in the absence of nucleotides is consistent with the gel-filtration result that suggests that the hexamers without nucleotide dissociate readily to monomers.

The time course of trypsin digestion suggests that some sites are more readily cleaved by trypsin than others, probably reflecting the relative accessibility of these sites. It appears that the early cleavage sites are at 131, 513/518, 484, and 651 followed by late cleavage at residues 384 and 401 (Fig. 5, *A* and *C*). In the presence of ATP+Mg^{2+}, the only early cleavage site is at residue 131 followed by a late cleavage site at residue 651, suggesting that the region around 131 is the most flexible part in LT. Protease treatment by three other proteases (chymotrypsin, elastase, and V8) in the presence or absence of nucleotides also had only one predominant cleavage in the region around 131 (data not shown), providing further evidence that the re-

gion between J-domain and OBD (residues 85–135) is flexible. When LT was treated by trypsin in the presence of nucleotide but in the absence of Mg^{2+}, two other cleavage sites at 484 and 513/518 in addition to 131 became sensitive to cleavage (Fig. 5, *A* and *C*), suggesting that Mg^{2+} may also play a role in stabilizing the hexamers of LT, possibly by enhancing ATP binding.

The result of the trypsin-sensitive sites on LT obtained in this studies is similar to that obtained independently by Weisshart *et al.* (32). In the two independent studies, the trypsin-digested patterns on SDS-PAGE appear to be slightly different and the apparent MM estimate for some of the digested fragments differs somewhat, possibly because of the differences in gel concentration and the molecular markers used in both studies. Nonetheless, the same N-terminal sequences were obtained for those fragments that were sequenced at the N terminus by Edman degradation in the two studies. There is only one discrepancy in the assignment of the cleavage sites. The discrepancy occurs at residue 698 where cleavage between 698 and 699 was deduced by Weisshart *et al.* (32) but not detected in this study. However, a closer examination of the cleavage pattern in our study reveals the possible presence of this cleavage site. A faint band with an apparent MM of ~64 kDa appeared after a 60-min incubation (Fig. 5, *lane 8*, between 66- and 60-kDa bands), which could be the result of further cleavage of the 66-kDa band at residue 698 (a 64-kDa product). Because the 64-kDa band is a minor product that was not visible in the 20- and 40-min incubation lanes (Fig. 5, *lanes 6* and *7*), the cleavage at residue 698 should be relatively inefficient.

Surprisingly, despite the presence of multiple Arg and Lys residues in the C-terminal HR fragment (residues 628–708), only one trypsin-sensitive site (residue 651) was present with a potential second weak cleavage site at 698, suggesting that this fragment may actually be structured contrary to the previous belief of an unstructured HR C terminus. Alternatively, because small pieces of peptide cleaved from the protein may not be detected in the SDS-PAGE system, it is possible that we may have missed some potential trypsin cleavage sites in this HR fragment. However, if that is the case, these sites should not be considered efficient cleavage sites (or not very accessible). Otherwise, the fragments containing the C terminus, such as the 66-, 60-, 26-, and 25-kDa fragments present in the gel, should not be stable enough to be observed within the time frame of our trypsin treatment. Furthermore, the sequencing results of the bulk of the trypsin-digested products did not yield sequences at the C-terminal HR fragment (Fig. 5*C*). A similar conclusion was obtained from another independent study by Weisshart *et al.* (32).

The trypsin-cleaved LT fragments in the presence of ATP but in the absence of Mg^{2+} occupy the three major regions of LT: the N-terminal 17/16-kDa fragments; the central 43/40-kDa fragments; and the C-terminal 26/25-kDa fragments as shown in a denaturing SDS-PAGE gel (Fig. 6*C*, *lane 2*). However, these major fragments did not separate from each other in gel-filtration chromatography in native buffer conditions, rather they co-eluted in peaks with MMs much larger than the size of the individual fragments (Fig. 6*B*, *pk1*, *pk2*, and *sPk*, and the *gel* in 6*C*). These results suggest that the domains from the three major regions of LT (the N-terminal 17/16-kDa, the central 40/43-kDa, and the C-terminal 26/25-kDa fragment) interact with each other. However, the detection of some of the N-terminal 17/16-kDa fragments in pk3 (Fig. 6*B*) as free form suggests that this N-terminal fragment (residues 1–130) does not interact strongly with other parts of LT so that some of it can dissociate to become free form under the assayed buffer condition. The interactions between the fragments (or domains) ranging from the N terminus to the C terminus of LT is represented in a diagram in Fig. 7, which is drawn based on the known structures of several domains of LT (9, 11, 27). The similar conclusion of the interdomain interactions within a LT molecule was also obtained by Weisshart *et al.* (32) using a battery of monoclonal antibodies to immunoprecipitate the trypsin-digested LT.

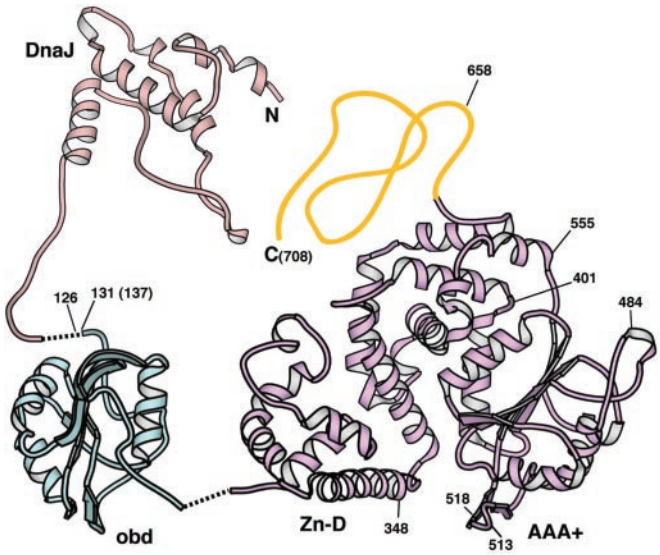


FIG. 7. **A diagram showing the assembly of the known structures of several LT domains and the location of the trypsin sensitive sites (in the absence of ATP+Mg^{2+}) in the manually assembled three-dimensional structure.** The domains are arbitrarily assembled in such a way that the possible interactions between the domains are reflected. The positions of the trypsin-sensitive sites are indicated by *thin lines* with the residue numbers at the cleavage sites. The three domains DnaJ, OBD, and helicase (including Zn-D and AAA+ domain) are taken from the published x-ray and NMR structures (9, 11, 27). The C-terminal HR domain (residues 628–708), the structure of which is unknown, is drawn in *thick yellow line*. This domain is positioned in the molecule to reflect the possibility that it may be folded toward the N terminus and interact with the N-terminal domains of LT. The polypeptides between domains are represented by *dashed lines*. The known structures are drawn to scales.

REFERENCES

1. Fanning, E., and Knippers, R. (1992) *Annu. Rev. Biochem.* **61,** 55–85
2. Simmons, D. T. (2000) *Adv. Virus Res.* **55,** 75–134
3. Ali, S. H., and DeCaprio, J. A. (2001) *Semin. Cancer Biol.* **11,** 15–23
4. DeCaprio, J. A. (1999) *Biologicals* **27,** 23–28
5. Saenz-Robles, M. T., Sullivan, C. S., and Pipas, J. M. (2001) *Oncogene* **20,** 7899–7907
6. Sullivan, C. S., Cantalupo, P., and Pipas, J. M. (2000) *Mol. Cell. Biol.* **20,** 6233–6243
7. Sullivan, C. S., and Pipas, J. M. (2002) *Microbiol. Mol. Biol. Rev.* **66,** 179–202
8. Wun-Kim, K., Upson, R., Young, W., Melendy, T., Stillman, B., and Simmons, D. T. (1993) *J. Virol.* **67,** 7608–7611
9. Luo, X., Sanford, D. G., Bullock, P. A., and Bachovchin, W. W. (1996) *Nat. Struct. Biol.* **3,** 1034–1039
10. Bullock, P. A. (1997) *Crit. Rev. Biochem. Mol. Biol.* **32,** 503–568
11. Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J., Fanning, E., Jochimiak, A., Szakonyi, G., and Chen, X. (2003) *Nature* **423,** 512–518
12. Lin, H. J., Upson, R. H., and Simmons, D. T. (1992) *J. Virol.* **66,** 5443–5452
13. Wu, C., Edgil, D., and Simmons, D. T. (1998) *J. Virol.* **72,** 10256–10259
14. Ott, R. D., Rehfuess, C., Podust, V. N., Clark, J. E., and Fanning, E. (2002) *Mol. Cell. Biol.* **22,** 5669–5678
15. Eckner, R., Ludlow, J. W., Lill, N. L., Oldread, E., Arany, Z., Modjtahedi, N., DeCaprio, J. A., Livingston, D. M., and Morgan, J. A. (1996) *Mol. Cell. Biol.* **16,** 3454–3464
16. Lill, N. L., Grossman, S. R., Ginsberg, D., DeCaprio, J., and Livingston, D. M. (1997) *Nature* **387,** 823–827
17. Pipas, J. M. (1985) *J. Virol.* **54,** 569–575
18. Spence, S. L., and Pipas, J. M. (1994) *J. Virol.* **68,** 4227–4240
19. Dean, F. B., Borowiec, J. A., Eki, T., and Hurwitz, J. (1992) *J. Biol. Chem.* **267,** 14129–14137
20. Bullock, P. A., Seo, Y. S., and Hurwitz, J. (1991) *Mol. Cell. Biol.* **11,** 2350–2361
21. Borowiec, J. A., Dean, F. B., Bullock, P. A., and Hurwitz, J. (1990) *Cell* **60,** 181–184
22. Hoss, A., Moarefi, I., Scheidtmann, K. H., Cisek, L. J., Corden, J. L., Dornre-

iter, I., Arthur, A. K., and Fanning, E. (1990) *J. Virol.* **64,** 4799–4807
23. Poplawski, A., Grabowski, B., Long, S. E., and Kelman, Z. (2001) *J. Biol. Chem.* **276,** 49371–49377
24. Chong, J. P., Hayashi, M. K., Simon, M. N., Xu, R. M., and Stillman, B. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 1530–1535
25. Shechter, D. F., Ying, C. Y., and Gautier, J. (2000) *J. Biol. Chem.* **275,** 15049–15059
26. Huang, S. G., Weisshart, K., and Fanning, E. (1998) *Biochemistry* **37,** 15336–15344
27. Kim, H. Y., Ahn, B. Y., and Cho, Y. (2001) *EMBO J.* **20,** 295–304
28. Egelman, E. H. (1996) *Structure* **4,** 759–762
29. Valle, M., Gruss, C., Halmer, L., Carazo, J. M., and Donate, L. E. (2000) *Mol. Cell. Biol.* **20,** 34–41
30. Schwyzer, M., Weil, R., Frank, G., and Zuber, H. (1980) *J. Biol. Chem.* **255,** 5627–5634
31. Wun-Kim, K., and Simmons, D. T. (1990) *J. Virol.* **64,** 2014–2020
32. Weisshart, K., Friedl, S., Taneja, P., Nasheuer, H., Schlott, B., Grosse, F., and Fanning, E. (2004) *J. Biol. Chem.* **279,** 38943–38951