Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 4026–4030, June 1987
Biochemistry

Consensus topography in the ATP binding site of the simian virus 40 and polyomavirus large tumor antigens

(Rossmann mononucleotide binding fold/β-sheet/ATPase)

M. K. BRADLEY†, T. F. SMITH‡, R. H. LATHROP§, D. M. LIVINGSTON†, AND T. A. WEBSTER‡

†Division of Neoplastic Disease Mechanisms, Dana–Farber Cancer Institute, and Departments of Pathology and Medicine, Harvard Medical School, Boston, MA 02115; ‡Molecular Biology Computer Research Resource, Dana–Farber Cancer Institute, Harvard School of Public Health, Boston, MA 02115; and §Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139

Communicated by George J. Todaro, February 24, 1987 (received for review November 17, 1986)

ABSTRACT **The location and sequence composition of a consensus element of the nucleotide binding site in both simian virus 40 (SV40) and polyomavirus (PyV) large tumor antigens (T antigens) can be predicted with the assistance of a computer-based pattern-matching system, ARIADNE. The latter was used to optimally align elements of T antigen primary sequence and predicted secondary structure with a "descriptor" for a mononucleotide binding fold. Additional consensus elements of the nucleotide binding site in these two proteins were derived from comparisons of T antigen primary and predicted secondary structures with x-ray structures of the nucleotide binding sites in four otherwise unrelated proteins. Each of these elements was predicted to be encompassed within a 110-residue segment that is highly conserved between the two T antigens residues 418–528 in SV40 T antigen and residues 565–675 in PyV). Results of biochemical and immunologic experiments on the nucleotide binding behavior of these proteins were found to be consistent with these predictions. Taken together, the latter have resulted in a topological model of the ATP binding site in these two oncogene products.**

Simian virus 40 (SV40) and polyomavirus (PyV) large tumor antigens (T antigens) are homologous early viral proteins essential to the initiation of viral DNA replication in permissive cells. They also participate in the regulation of early and late viral RNA synthesis, the induction of cellular DNA and RNA synthesis, and the neoplastic transformation of appropriate cell types. Each can bind specifically to its own replication origin *in vitro*, and both display ATPase activity (1). In keeping with the latter finding, both T antigens formed a specific covalent bond with 2′,3′-dialdehyde ATP (2), and SV40 T antigen formed a 5′-AMP–T adduct after incubation with ATP (3). Although the exact relationship of these nucleotide coupling phenomena to the ATPase reaction is unclear, they nonetheless indicate that the structure of both proteins includes an ATP binding site. Until now, the topography of this structure was unknown, although others have detected segments of T antigen sequence that appear to constitute part of it (4). Therefore, we have attempted to decipher the structure of this region in more detail with the assistance of a computer-based pattern-matching system (ARIADNE; ref. 5). The results of our analyses strongly predict the existence of a nucleotide binding domain with distinct topographical features in a homologous segment of the T antigens and thereby give insight into important aspects of the probable three-dimensional structure of a major domain in two large complicated proteins prior to crystallographic analysis.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Secondary Structure. Secondary structure predictions were carried out using a computer program PRSTRC (6) incorporating the method of Chou and Fasman (7). The following values were used for these predictions: α-helix former = 1.12 (threshold, 1.08; cutoff, 1.00), β-strand former = 1.23 (cutoff, 0.97), minimum β-turn value = 0.50.

ARIADNE. The structure of the descriptor and the data set created specific requirements for an algorithm to carry out a computer search. Therefore, an expert-system ARIADNE was created for this problem (5). It uses an object-oriented representation of the protein and implements in LISP using Flavors, allowing direct representation of primary sequence implied structures such as pattern elements. ARIADNE finds optimal alignments of the descriptor, allowing for gaps and insertions. An A* search was used to order multiple paths in the network of multiple secondary structure predictions.

Assays for the Inhibition of SV40 T-Antigen ATPase by T-Specific Monoclonal Antibodies. Assays were performed as described (8) using 50 ng of purified T antigen (9, 10), preincubated with antibody at 4°C for 10 min. Reaction mixtures (25 μl) containing 0.5 μM ATP (4 μCi of [γ-^{32}P]ATP; 1 Ci = 37 GBq) were incubated at 37°C for 30 min. Antibodies (mouse) were purified by immunoaffinity chromatography using goat anti-mouse IgG-Sepharose. IgG concentrations were measured by enzyme-linked immunoadsorbent assay.

Analyses of CNBr Cleavage Products of ^{32}P-AMP-Labeled SV40 T Antigen. T antigen (10 μg), purified by immunoprecipitation, was adenylylated by incubation with [α-^{32}P]ATP (3). The labeled adduct was eluted from NaDodSO$_4$/polyacrylamide gels and precipitated with acetone in the presence of bovine serum albumin (50 μg), resuspended in 100 μl of a saturated solution of CNBr in 60% formic acid, and incubated for 18 hr at 25°C. Products were vacuum dried and lyophilized from water three times and analyzed by electrophoresis in a 20% NaDodSO$_4$/polyacrylamide gel with known protein standards (2).

RESULTS

Derivation of the Descriptor for the Protein Structure Pattern-Matching System. In the past, attempts to locate nucleotide binding sites in various proteins by direct comparison of primary sequences alone were not always successful. For instance, aminoacyl-tRNA synthetases all have the same generic function without enough primary sequence identity among them to permit the immediate recognition of an ATP nucleotide binding site. The x-ray structures of two of these enzymes have been solved, and the segment that

Abbreviations: SV40, simian virus 40; PyV, polyomavirus; T antigen, large tumor antigen; MNBF, mononucleotide binding fold; EF-Tu, elongation factor Tu.

binds both ATP and the relevant aminoacyl adenylate was identified in each (11, 12). This domain contains certain elements of structure originally recognized by Rossmann *et al.* (13) in a series of dinucleotide binding proteins. These workers (13) proposed that a particular combination of primary and secondary structures constitutes an important part of a consensus nucleotide binding domain termed a "mononucleotide binding fold" (MNBF; refs. 13 and 14). Its existence in the tRNA synthetases, as well as in the dinucleotide binding proteins, suggested that a computer-assisted search for those combinations of primary and predicted secondary structural characteristics, which represent a MNBF, might lead to the detection of previously unrecognized nucleotide binding domains. Operating on this hypothesis, Lathrop *et al.* (5) have recently developed a protein pattern-matching system (ARIADNE) designed to locate a Rossmann-type MNBF in proteins of disparate function whose x-ray structures have not yet been solved.

The system uses pattern descriptors composed of combinations of common primary sequence and predicted secondary structural characteristics known to constitute a MNBF in members of a particular protein class. A different descriptor was developed for each class of nucleotide binding protein studied. ARIADNE was used to provide the optimal alignment of the descriptor to the known data for a protein of interest. This system was successfully applied to a search for a consensus MNBF in a set of 12 aminoacyl tRNA synthetases (15).

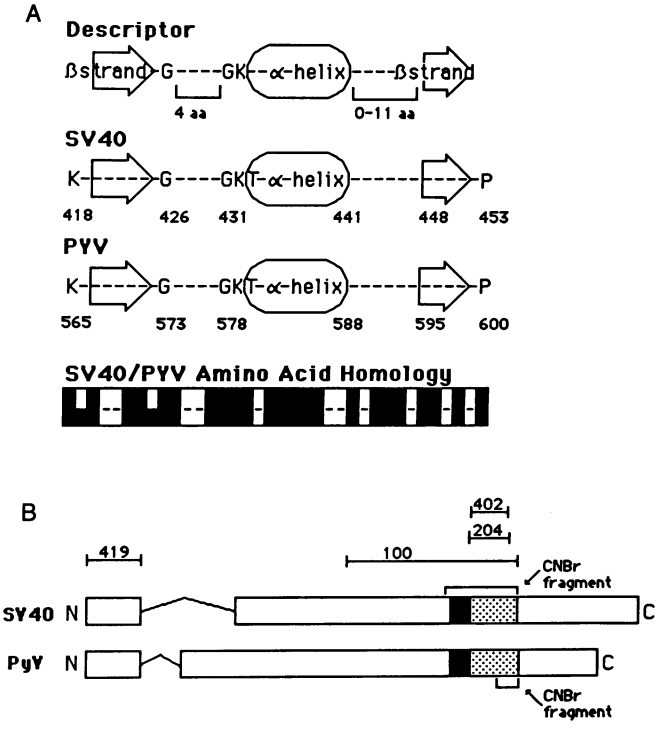
Search for MNBF in SV40 and PyV T Antigen Using ARIADNE (5). The necessary secondary structure predictions for each T antigen were derived with computer assistance (PRSTRC; ref. 6). Specific criteria for declaring a match between a given descriptor and a relevant protein have been reported elsewhere (5, 15). The descriptor used in searching for a MNBF, specific for ATP/ADP or GTP/GDP binding proteins, was derived from multiple sources: (*i*) from the work of Rossmann *et al.* (13, 14); (*ii*) from analyses of the known three-dimensional structure of the nucleotide binding sites in rabbit muscle adenylate kinase and *Escherichia coli* elongation factor Tu (EF-Tu) (16–18); and (*iii*) from biochemical and genetic analyses of the nucleotide binding properties of nematode myosin (19) and *E. coli* RecA protein (20), respectively. The following combination of primary sequence and predicted secondary structure constituted the descriptor used here:

NH_2-(β-strand)-G X X X X-(G K α-helix)-
(0–11 amino acids)-(β-strand)-COOH.

The β-strand, glycine, 4-amino acid spacer, glycine sequence (GXXXXG), followed by a functionally important α-helix (21), is present in the nucleotide binding domains of all of the proteins of the NAD/NADH, FAD/FADH, GTP/GDP, and ATP/ADP binding classes of known x-ray structure (17, 18, 22–24). Aminoacyl-tRNA synthetases (ATP/aminoacyl adenylate binding class) reveal primary sequence differences in this region (25). The above descriptor includes a lysine at the COOH terminus of the G X X X X G consensus sequence, because it has been detected in multiple, although not all, ATP/ADP and GTP/GDP binding proteins (17, 18, 24). Moreover, a number of ATP/ADP binding proteins also contained an adjoining threonine or serine. A threonine residue is present in this position in both T antigens.

To test the validity of the descriptor for detecting consensus features of a predicted MNBF in the T antigens, we simultaneously applied it to an analysis of the predicted structure of 53 unrelated proteins (15, 26). This set included 27 proteins whose crystal structures are known and was chosen to represent four major structural classes (27). A strong match was detected between the descriptor and both T antigens as well as 2 out of the 53 control proteins. These were *Clostridium* flavodoxin, a known FMN-binding protein (14, 28), and the α chain of tubulin, shown to have an ATP binding site (29). There were no false positives.

The predicted MNBF in SV40 T antigen extends from lysine-418 to proline-453 and contains two β-strands (Fig. 1*A*). The analogous MNBF for PyV T antigen extends from lysine-565 to proline-600. Notably, these 36 SV40 and PyV T antigen residues lie within a region of strong primary sequence similarity (30).



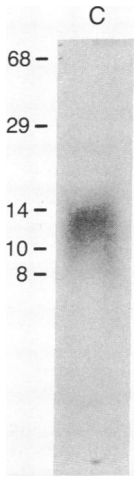


FIG. 1. (*A*) The region in SV40 and PyV T antigen that matches the descriptor for the ATP/ADP MNBF. Arrows and ellipses include the residues predicted to form β-strands and α-helices, respectively (6). Identical amino acids in the two T antigens are indicated by solid bars at the bottom. Half solid bars denote alignment of lysine and arginine. (*B*) Alignment of the two T antigens according to their primary sequence similarities. Open bars represent coding sequences of the T antigen genes with residues predicted to be in the MNBFs shown as solid bars and sequences also predicted to contribute to the formation of the nucleotide binding domain shown as stippled bars. 2′,3′-Dialdehyde [^{32}P]ATP-labeled CNBr fragments are shown in brackets and residues presently known to include epitopes for certain SV40 T-antigen-specific monoclonal antibodies are indicated. (*C*) Autoradiogram of a CNBr digest of [α-^{32}P]AMP-labeled SV40 T antigen. Labeling, cleavage, and analysis were performed as described. Molecular weight markers are shown as $M_r \times 10^{-3}$.

Search for Additional Consensus Features of the T-Antigen Nucleotide Binding Site. There is reason to believe that a β-sheet must have a minimum of four or five β-strands to form a stable structure (27). The MNBF descriptor includes only two. From analyses of the topography of well-characterized proteins from different nucleotide binding groups, five or six β-strands (which includes the two β-strands of the MNBF) are present in the nucleotide binding domains of several. Among them are *E. coli* tyrosyl-tRNA synthetase (12), porcine H_4 lactate dehydrogenase (23), the *E. coli* translation factor EF-Tu (18), and rabbit muscle adenylate kinase (17). Given these examples, we searched for and identified, in the two T antigens, a 110-residue segment that contains five predicted β-strands and the 36 amino acids of the predicted MNBF. There is 64% similarity between these two primary sequences, and the lengths and locations of the five β-strands in one are nearly identical to those in the other. For SV40 T antigen, this sequence extends from lysine-418 to methionine-528 and for PyV T antigen, it extends from lysine-565 to methionine-675 (Fig. 1*B*). Proximal to the most NH_2-terminal and beyond the most COOH-terminal predicted β-strand of this proposed β-sheet, the spacing of predicted β-strands is not in register between the two T antigens. Therefore, in evaluating components of a consensus T-antigen ATP binding site, only a 110-residue region was considered.

The next question was whether it was possible to predict the order of these five β-strands with respect to one another in space. As an aid in this analysis, we have used a graphic method derived from a study of the crystal structure of *E. coli* methionyl tRNA synthetase (11). In a schematic diagram of the nucleotide binding site, β-strands were represented as a planar array of arrows, each pointed toward its COOH terminus. α-Helices were represented as cylinders, and intervening sequences or loops were shown as solid lines. We have constructed similar diagrams from the crystal structures of the four nucleotide binding proteins noted earlier (Fig. 2). These diagrams were not meant to convey the details of how each residue is oriented with respect to its neighbors, but, in general, how each structural component relates to the others within the nucleotide binding domain. Note that the core elements of the MNBF are present in each (β-strand, G X X X X G, α-helix, β-strand). We have assigned each of five β-strands in the sheet a Roman numeral and defined its position in the planar array in the following order: V, IV, I, II, III (left to right). In the primary sequence, β-strands were labeled alphabetically—A–E or F (NH_2 to COOH). In all proteins considered, β-strand A occupied position I in the β-sheet. Given that observation, A was predicted to occupy a similar position in the T antigens. The spatial order of the other four β-strands in the model proteins varied (Fig. 2). The order E-D-A-B-C was detected in both aminoacyl-tRNA synthetase (25) and H_4 lactate dehydrogenase (23); B-C-A-D-E was the case for EF-Tu (18); and E-D-A-C-B was noted in the adenylate kinase binding domain (17).

Three elements of structure are presently known to be involved in the placement of the nucleotide in the site of three of the four proteins in Fig. 2. First, in keeping with observations on the role of the dipole moment of α-helices in the binding of phosphate moieties in the active sites of enzymes (21), it appears that the α-phosphate of a nucleotide approaches a glycine proximal to the NH_2 terminus of the MNBF α-helix (Fig. 3). Second, the magnesium ion (Mg^{2+}), known to be complexed with the β- or γ-phosphate of a purine nucleotide substrate, has been located in the structures of adenylate kinase and EF-Tu (17, 18). In both cases the magnesium–phosphoryl group is positioned near a negatively charged residue at the COOH-terminal end of the β-strand in position IV—e.g., aspartate-119 (D-119) at the COOH terminus of β-strand D in adenylate kinase and aspartate-80 (D-80) at the COOH terminus of β-strand C in EF-Tu (Figs. 2 and 3). Moreover, in adenylate kinase, β-strand D contains a cluster of hydrophobic residues (17, 24). Third, a negatively charged residue at the COOH terminus of the β-strand in position II is known to form a hydrogen bond with the 2′-OH of the adenine ribose in multiple NAD/NADH binding proteins (31). In porcine H_4 lactate dehydrogenase (23), the 2′-OH of the adenine ribose interacts with aspartate-53 (D-53)

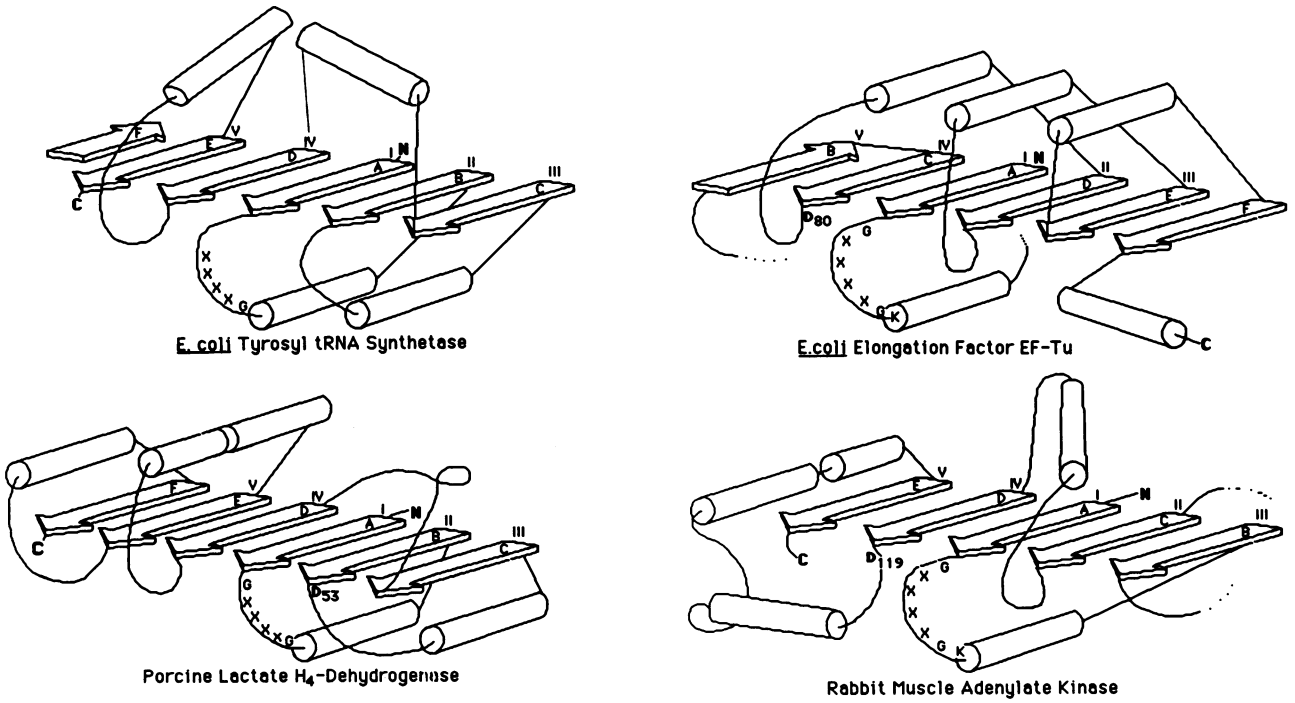


FIG. 2. Schematic diagrams of the known nucleotide binding sites of proteins from four different nucleotide binding classes. Arrows, β-strands that are pointed in the COOH-terminal direction; cylinders, α-helices. N and C indicate the amino and the carboxyl ends of the partial sequences presented. β-Strand letters (A–E) and position numbers (I–V) are described in the text.

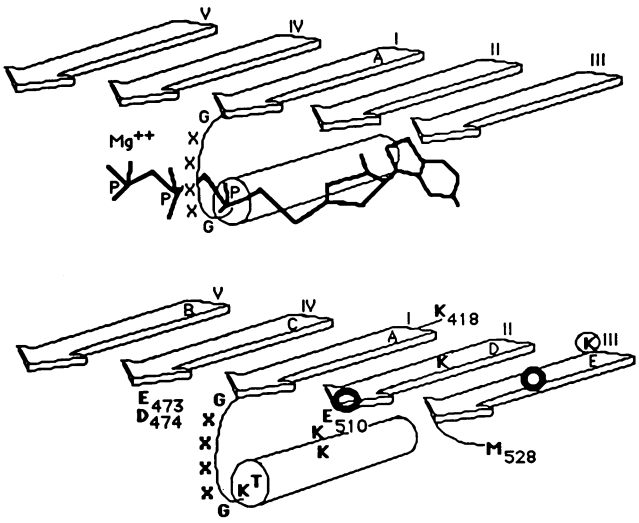


FIG. 3. (*Upper*) A simplified model of a nucleotide binding site–purine nucleoside triphosphate complex. This diagram summarizes certain common features of the complex formed between adenylate kinase and EF-Tu and their respective nucleotide ligands. The three elements of interest are the proximity of the α-PO_4 of the nucleotide to the amino end of the α-helix linked to β-strand I (A) (18), the β or γ-PO_4–Mg^{2+} complex to a negatively charged residue linked to the COOH terminus of β-strand IV, and the 2'- or 3'-OH of the ribose to the COOH terminus of β-strand II. (*Lower*) A simplified diagram of the proposed nucleotide binding site of SV40 T antigen. β-Strands are labeled as described in the text. Boldface letters denote T antigen residues at the amino and carboxyl termini and those predicted to be involved in the interaction with ATP. Four unnumbered lysines (K), one of which is believed to couple to 2',3'-dialdehyde ATP, are shown in their approximate positions in and near predicted β-strands D and E. The boldface circle at the COOH terminus of β-strand D represents one site of mutation in T antigen (dl2462), and that in the middle of β-strand E indicates another (C11-A). The mutation in C2 is indicated by a circle around the lysyl residue adjacent to the NH_2 terminus of β-strand E. The orientations of the residues extending from β-strands C, D, and E and assignment of β-strands B and E to the positions shown here were arbitrary and do not constitute formal predictions. A similar diagram can be drawn for the ATP binding site of PyV T antigen.

at the COOH terminus of β-strand B. Furthermore, in the case of adenylate kinase and EF-Tu, the ribose moiety lies close to and is oriented toward the COOH terminus of the β-strand in position II (17, 18).

In SV40 T antigen (Fig. 3), β-strand C is followed by glutamate-473 (E-473) and aspartate-474 (D-474), and in PyV T antigen, β-strand C is followed by glutamate-620 and aspartate-621. In addition, β-strand D is followed by glutamate-510 (E-510) in SV40 and by glutamate-657 in PyV. Considering these observations in light of the details of Mg^{2+} and ribose binding in the above-noted proteins, it seemed reasonable to propose that β-strands C and D occupy positions IV and II in the predicted T-antigen β-sheets. However, from these considerations alone, it is not possible to predict which of these two β-strands occupies which position.

Biochemical evidence suggesting a specific orientation of β-strands C and D with respect to A can be derived from results of 2',3'-dialdehyde ATP labeling of T antigen. When PyV T antigen was incubated with 2',3'-dialdehyde [α-^{32}P]ATP, which forms a Schiff base to a lysyl residue, 2',3'-dialdehyde ATP was found to be linked to a lysine in a peptide extending from glycine-637 to methionine-675 (32). In SV40 T antigen, the nucleotide was found to be linked to a peptide extending from valine-413 to methionine-528 (32). These peptides contain part or all of the proposed 110-amino acid nucleotide binding region. Given the PyV result and the fact that the two T-antigen sequences are highly homologous in this region, it is likely that, in SV40 T antigen, 2',3'-dialdehyde ATP is coupled to a lysine present in an analogous 39-residue sequence—i.e., between glycine-490 and methionine-528. There are four lysines in this 39-residue sequence in each protein. One lies within β-strand D, two immediately adjoin its COOH terminus, and another lies at the COOH terminal end of the loop structure connecting β-strands D and E (Fig. 3). This observation suggests that the ribosyl linkage is to a specific residue present in β-strand D or the loop extending from it, leading one to predict that β-strand D occupies position II in the β-sheet and, by default, that β-strand C occupies position IV. Moreover, as is the case for the β-strand in position IV in adenylate kinase, β-strand C of the T antigens contains a cluster of hydrophobic residues. The predicted order of these β-strands would then be identical to that in EF-Tu, C-A-D. Presently, there is no experimental evidence that can be used to predict the locations of β-strands B and E in the sheet.

Experimental Evidence Supporting the Model and the Prediction that the SV40 T-Antigen Nucleotide Binding Domain Contains the 110-Residue Region Described Here. First, incubation of the protein with ATP and Mg^{2+} resulted in the formation of a specific AMP–T adduct, with the nucleotide linked to the protein by a phosphodiester bond between its α-phosphate and a specific serine (3). Exhaustive CNBr cleavage of [^{32}P]AMP-labeled T antigen yielded a 12- to 14-kDa labeled peptide in five separate experiments (e.g., see Fig. 1*C*). No smaller fragments were detected. In the SV40 T-antigen sequence, the largest predicted CNBr peptide (116 residues), extends from valine-413 to methionine-528 and is the same size as the [^{32}P]AMP-labeled fragment ($\simeq$13 kDa). The next largest peptide would contain 67 residues ($\simeq$7.5 kDa). The AMP-bound fragment, therefore, appears to be composed of the same 100 residues as the proposed SV40 T-antigen nucleotide binding region.

Immunological evidence for the existence of a specific SV40 T-antigen nucleotide binding region was derived from studies of the effect of various T-specific monoclonal antibodies on T-antigen ATPase activity (Table 1; Fig. 1*B*). In keeping with earlier reports, papovavirus antibody 204 (PAb204) (33, 34) and PAb402 (35, 36) specifically inhibited the activity of the purified protein. Comparable amounts of a third antibody, PAb100 (37), also inhibited the enzyme when tested in parallel with PAb204 and PAb402. By contrast, PAb419 failed to inhibit at IgG concentrations comparable to those of the other three antibodies. In independent assays of the relative affinity of T antigen for each of these antibodies, that for PAb419 was found to be equivalent to that of the other three (data not shown). These results are consistent with the suggestion that sequences within the predicted 110-residue nucleotide binding region are important for ATPase activity (Table 1), presumably by contributing to the integrity of its nucleotide binding site.

More support for the proposed location of the nucleotide binding site is provided by the results of genetic analyses of SV40 T antigen function. A truncation at amino acid 399 in dl1055 (1→399; ref. 38) results in a stable nondenatured

Table 1. Inhibition of SV40 T-antigen ATPase by T-specific monoclonal antibodies

PAb	Epitope present between residues	IgG required for 50% inhibition, μg
419	1–82	>5.00
100	254–524	0.28
204	448–509	0.55
402	450–500	0.55

Assays were performed as described. PAb, papovavirus antibody.

mutant T antigen, which has lost all detectable ATPase activity. Therefore, residues COOH-terminal to proline-399 are essential for this function. Two other deletion mutants, dl2420 (1→585; ref. 39) and dl1061 (1→593; ref. 38), are also defective in ATPase activity, implying that sequences distal to methionine-528 might be important for this function. Two single-site mutations, in dl2462 (leucine-509 deleted; ref. 39), and in C11A (proline-522 → serine; ref. 40), affecting residues within β-strands D and E, respectively (Fig. 3), are also ATPase defective. By contrast, a third mutant, C2, bearing a conservative substitution (lysine-516 → arginine; ref. 40) adjacent to the NH_2 terminus of β-strand E (Fig. 3), retained ATPase activity. The knowledge that dl2462 and C11A T antigen are also defective in replication initiation activity, is consistent with the strongly suspected link between ATP hydrolysis and the replication function of T antigen (38–40).

DISCUSSION

Evidence has been presented that strongly suggests that the known nucleotide binding site of two papovaviral T antigens is, at least in part, composed of a set of consensus structural elements. The latter comprises two overlapping units—a Rossmann–MNBF and a five-strand β-sheet. Of the four proteins that served as models in the search for a consensus β-sheet within the nucleotide binding domain, the structure proposed for the T antigens most closely resembles that of the GTP/GDP binding protein EF-Tu (18). In this regard, it may be worth noting that the nucleotide binding region of another transforming protein, p21 ras (41, 42), also shares significant primary sequence and prediced secondary structural homology with EF-Tu.

A simple interpretation of the data described here would suggest that there is one ATP binding site in each T antigen. These results also imply that the ATPase active site and the ATP binding site are identical or overlap and that the *in vitro* adenylylation of T requires some of the same protein sequence as the ATPase function. They do not rule out the formal possibility that another nucleotide binding region, lacking homology with that described here, exists in either of these proteins. In addition, it is possible that portions of the relevant 110-residue segment can assume more than one conformation (22, 24, 43).

The results presented here are wholly consistent with evidence suggesting that nucleotide binding and/or hydrolysis are linked to the replication initiation function of T antigen (38–40). In this regard, it should be noted that others have previously detected significant homology between a portion of the SV40 and PyV T-antigen nucleotide binding region and a specific segment of multiple predicted papillomavirus E1 gene products (4), including the replication protein BPV-1-E1 (44).

Finally, these experiments result in a model whose structure can be tested both biochemically and genetically. In principle, such a model can help to clarify the mechanism and functional significance of nucleotide binding and hydrolysis by and coupling to T antigen.

This work was supported by National Institutes of Health Grants CA38069 and RR02275.

1. Griffin, B. E. (1981) in *Molecular Biology of the DNA Tumor Viruses*, ed. Tooze, J. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 61–124.
2. Clertant, P. & Cuzin, F. (1982) *J. Biol. Chem.* **257,** 6300–6305.
3. Bradley, M. K., Hudson, J., Villaneuva, M. S. & Livingston, D. M. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 6574–6578.
4. Clertant, P. & Seif, I. (1983) *Nature (London)* **311,** 276–279.
5. Lathrop, R., Webster, T. & Smith, T. F. (1987) *Commun. ACM*, in press.
6. Ralph, W., Webster, T. & Smith, T. F. (1987) *Comput. Appl. Biosci.*, in press.
7. Chou, P. Y. & Fasman, G. S. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol.* **47,** 45–148.
8. Bradley, M. K., Griffin, J. F. & Livingston, D. M. (1982) *Cell* **28,** 125–134.
9. Simanis, V. & Lane, D. P. (1985) *Virology* **144,** 88–100.
10. Dixon, R. A. F. & Nathans, D. (1985) *J. Virol.* **53,** 1001–1004.
11. Zelwer, C., Risler, J. L. & Brunie, S. (1982) *J. Mol. Biol.* **155,** 63–81.
12. Bhat, T. N., Blow, D. M. & Brick, P. (1982) *J. Mol. Biol.* **158,** 699–709.
13. Rossmann, M. G., Moras, D. & Olsen, K. W. (1974) *Nature (London)* **250,** 194–199.
14. Rossmann, M. G., Lilijas, A., Braden, C. I. & Banaszak, L. J. (1975) in *The Enzymes*, ed. Boyer, P. D. (Academic, New York), Vol. 11, pp. 61–102.
15. Webster, T., Lathrop, R. & Smith, T. (1987) *Commun. ACM*, in press.
16. Pai, E. F., Sachsenheimer, R. H. & Schultz, G. E. (1977) *J. Mol. Biol.* **144,** 37–45.
17. Fry, D. C., Kuby, S. A. & Mildvan, A. S. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 907–911.
18. la Cour, T. F. M., Nyborg, J., Thirup, S. & Clark, B. F. C. (1985) *EMBO J.* **4,** 2385–2388.
19. Szilagyi, L., Balint, M., Sreter, F. A. & Gergely, J. (1979) *Biochem. Biophys. Res. Commun.* **87,** 936–945.
20. Kawashima, H., Horii, T., Ogawa, T. & Ogawa, H. (1984) *Mol. Gen. Genet.* **193,** 288–292.
21. Hol, W. G. J., Van Duijnen, P. T. & Berensen, H. J. C. (1978) *Nature (London)* **273,** 443–446.
22. Evans, P. R. & Hudson, P. J. (1979) *Nature (London)* **279,** 500–504.
23. Grau, U. M. (1982) in *The Pyridine Nucleotide Coenzymes*, eds. Everse, J., Anderson, T. & You, K.-S. (Academic, New York), pp. 135–187.
24. Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. (1982) *EMBO J.* **1,** 945–951.
25. Blow, D. M. & Brick, P. (1985) in *Biological Macromolecules and Assemblies*, eds. Jurnak, F. A. & McPherson, A. (Wiley, New York), Vol. 2, pp. 441–469.
26. George, D. G., Barber, W. C. & Hunt, L. T. (1986) *Nucleic Acids Res.* **14,** 11–15.
27. Richardson, J. S. (1981) *Adv. Protein Chem.* **34,** 168–340.
28. Burnett, R. M., Darling, G. D., Kendall, D. S., LeQuesne, M. E., Mayhew, S. G., Smith, W. W. & Ludwig, M. L. (1974) *J. Biol. Chem.* **249,** 4383–4394.
29. Zabrecky, J. R. & Cole, R. D. (1985) *Arch. Biochem. Biophys.* **225,** 475–481.
30. Griffin, B. E. & Dilworth, S. M. (1983) *Adv. Cancer Res.* **39,** 183–268.
31. Wierenga, R. K., Terpstra, P. & Hol, W. G. J. (1986) *J. Mol. Biol.* **187,** 101–107.
32. Clertant, P., Gaudray, P., May, E. & Cuzin, F. (1984) *J. Biol. Chem.* **259,** 15196–15203.
33. Clark, R., Lane, D. P. & Tjian, R. (1981) *J. Biol. Chem.* **256,** 11854–11858.
34. Mole, S. E. & Lane, D. P. (1985) *J. Virol.* **54,** 703–710.
35. Harlow, E., Crawford, L. V., Pim, D. C. & Williamson, N. M. (1981) *J. Virol.* **39,** 861–869.
36. Schwyzer, M., Tai, Y., Studer, E. & Michel, M. R. (1983) *Eur. J. Biochem.* **137,** 303–309.
37. Gurney, E. G., Harrison, R. O. & Fenno, J. (1980) *J. Virol.* **34,** 752–763.
38. Clark, R., Peden, K., Pipas, J. M., Nathans, D. & Tjian, R. (1983) *Mol. Cell. Biol.* **3,** 220–228.
39. Cole, C. N., Tornow, J., Clark, R. & Tjian, R. (1986) *J. Virol.* **57,** 539–546.
40. Manos, M. M. & Gluzman, Y. (1985) *J. Virol.* **53,** 120–127.
41. McCormick, F., Clark, B. F. C., la Cour, T. F. M., Kjeldgaard, M., Norskov-Lauritsen, L. & Nyborg, J. (1986) *Science* **230,** 78–82.
42. Sigal, I. S., Gibbs, J. B., D'Alonzo, J. S. & Scolnick, E. M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4725–4729.
43. Anderson, C. M., Zucker, F. H. & Steitz, T. A. (1979) *Science* **204,** 375–380.
44. Lusky, M. & Botchan, M. R. (1984) *Cell* **36,** 391–401.