Quantitative Analysis of the Binding of Simian Virus 40 Large T Antigen to DNA

Amélie Fradet-Turcotte, Caroline Vincent, Simon Joubert, Peter A. Bullock, and Jacques Archambault

Laboratory of Molecular Virology, Institut de Recherches Cliniques de Montréal, 110 Pine Avenue West, Montreal, Quebec, Canada H2W 1R7, and Department of Biochemistry, University of Montreal, Montreal, Quebec, Canada, and Department of Biochemistry, Tufts University School of Medicine, Boston, Massachusetts 02111

Received 22 February 2007/Accepted 18 June 2007

SV40 large T antigen (T-ag) is a multifunctional protein that successively binds to 5′-GAGGC-3′ sequences in the viral origin of replication, melts the origin, unwinds DNA ahead of the replication fork, and interacts with host DNA replication factors to promote replication of the simian virus 40 genome. The transition of T-ag from a sequence-specific binding protein to a nonspecific helicase involves its assembly into a double hexamer whose formation is likely dictated by the propensity of T-ag to oligomerize and its relative affinities for the origin as well as for nonspecific double- and single-stranded DNA. In this study, we used a sensitive assay based on fluorescence anisotropy to measure the affinities of wild-type and mutant forms of the T-ag origin-binding domain (OBD), and of a larger fragment containing the N-terminal domain (N260), for different DNA substrates. We report that the N-terminal domain does not contribute to binding affinity but reduces the propensity of the OBD to self-associate. We found that the OBD binds with different affinities to its four sites in the origin and determined a consensus binding site by systematic mutagenesis of the 5′-GAGGC-3′ sequence and of the residue downstream of it, which also contributes to affinity. Interestingly, the OBD also binds to single-stranded DNA with an ~10-fold higher affinity than to nonspecific duplex DNA and in a mutually exclusive manner. Finally, we provide evidence that the sequence specificity of full-length T-ag is lower than that of the OBD. These results provide a quantitative basis onto which to anchor our understanding of the interaction of T-ag with the origin and its assembly into a double hexamer.

Eukaryotic DNA replication is a complex process that is initiated by the coordinated action of several factors, including the origin recognition complex, cdt1, cdc6, and the minichromosome maintenance complex, the likely cellular replicative helicase (21, 31, 40). Our understanding of the molecular events involved in the initiation of DNA replication has greatly benefited from studies performed with simpler model systems, like budding yeast or small-DNA tumor viruses, for which the origins of replication are of well-defined nucleotide sequences (47, 57). Small-DNA viruses also offer the advantage of encoding a single initiator protein that performs multiple functions during replication. A well-studied example is large T antigen (T-ag), the initiator protein from simian virus 40 (SV40). This multifunctional protein can successively recognize the viral origin of replication, melt it, interact with the host DNA replication factors polymerase α-primase, replication protein A, and topoisomerase I, and unwind DNA ahead of the replication fork (5). Whereas the smallest form of T-ag that has ATPase and minimal unwinding activity is the hexamer (1, 56). T-ag, a member of helicase superfamilies I (18), is a modular protein that contains three distinct domains implicated in viral DNA replication (Fig. 1B). The C-terminal half of T-ag (amino acids [aa] 261 to 708) contains the helicase domain and thus can assemble into hexamers and bind single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA) nonspecifically (20, 65). The central part of the protein contains the origin-binding domain (OBD) (aa 131 to 260) which, as its name indicates, recognizes specific sequences in the origin (34, 52, 59, 67). Finally, the N-terminal region, which contains a J-domain (aa 1 to 82; reference 12), a nuclear localization signal (aa 126 to 132; references 23 and 24), and phosphorylation sites for cellular kinases, including one, Thr124, for cyclin/cdk (32, 33, 35, 38, 50), also contribute to affinity. Interestingly, the OBD also binds to single-stranded DNA with an ~10-fold higher affinity than to nonspecific duplex DNA and in a mutually exclusive manner. Finally, we provide evidence that the sequence specificity of full-length T-ag is lower than that of the OBD. These results provide a quantitative basis onto which to anchor our understanding of the interaction of T-ag with the origin and its assembly into a double hexamer.

* Corresponding author. Mailing address: Laboratory of Molecular Virology, Institut de Recherches Cliniques de Montréal (IRCM), 110 Pine Avenue West, Montreal, Quebec, Canada H2W 1R7. Phone: (514) 987-5739. Fax: (514) 987-5741. E-mail: jacques.archambault@ircm.qc.ca.

† Published ahead of print on 27 June 2007.
VOL. 81, 2007

AFFINITY OF SV40 T ANTIGEN FOR DNA

preferentially on P1 and P3 (22, 56) and is dependent on the integrity of the N-terminal domain (62, 64), the OBD (53, 64, 67) and C-terminal \( \beta \)-hairpin motif (residues 507 to 517) (16, 28, 46, 51). This motif is presumably needed to melt the region downstream of each T-ag binding site (TBS) in order to generate regions of ssDNA which promote the assembly of additional T-ag molecules into a complete double hexamer (28).

The molecular details of how two specific binding sites, P1 and P3, nucleate the assembly of a double hexamer and specifically of how T-ag transits from a monomeric sequence-specific DNA binding protein to a nonspecific oligomeric helicase are still poorly understood.

Biochemical and genetic studies have provided evidence that specific binding of the OBD to its target pentanucleotides in the origin involves two structural elements, A1 and B2, that form part of the DNA binding surface (30, 54). Our previous studies further indicated that the OBD binds to the four pentanucleotides in site II independently, without any evidence of cooperativity (59). These findings have been corroborated by nuclear magnetic resonance (NMR) and recent X-ray studies of the OBD bound to DNA (4, 6, 37), which indicated that the OBD recognizes specific bases in the major groove of DNA and that OBD molecules bound on P1 and P3, or on all four pentanucleotides in site II, do not interact with each other. In contrast, the crystal structure of the OBD without DNA showed that the protein can assemble into a left-handed spiral that is wide enough to accommodate duplex DNA in a negatively charged central channel (36). These findings prompted Meinke et al. (36) to speculate that this spiral structure reflects how the OBDs are organized within a double hexamer.

As mentioned above, double hexamer assembly also relies on the integrity of the N-terminal domain. This is exemplified by the finding that a truncated T-ag lacking the first 82 aa oligomerizes into aberrant multimeric complexes either in solution or on DNA (62). Furthermore, double hexamer assembly is regulated by phosphorylation of the N-terminal domain, specifically of Thr124 by a cyclin/cdk (15). Indeed, a mutant T-ag in which Thr124 was changed to alanine was found to assemble into double hexamers with impaired cooperativity, and these complexes dissociated more rapidly from DNA (64). In addition, the T124A T-ag was shown to require longer sequences in the EP or AT region for its assembly on single assembly units (2, 46). Interestingly, specific amino acid substitutions in the OBD, known as class 4 mutations, have been identified that also interfere with double hexamer assembly but without affecting phosphorylation of the protein (64). These findings suggested that the N-terminal domain and the OBD play coordinated functions in double hexamer assembly. This prompted Kim et al. (26) to investigate if peptides containing Thr124 could bind DNA. Strikingly, peptides spanning aa 118 to 130 were found to bind nonspecifically to DNA and inhibit double hexamer assembly in vitro, but only when nonphosphorylated (26). These results provided the first evidence that the N-terminal domain may influence the affinity of the protein for DNA.

The transition of T-ag from a monomeric sequence-specific DNA binding protein to a nonspecific dodecameric helicase is likely dictated by a number of parameters, in particular its affinity for specific pentanucleotides in site II and for nonspecific ds- and ssDNA and its propensity to oligomerize. We believe that a quantitative analysis of these interactions will provide a useful framework to understand how a double hexameric helicase is assembled from individual monomeric components and how this process is regulated by phosphorylation. It was previously reported that the interaction of the OBD with its target site can be studied using a quantitative and sensitive DNA binding assay based on fluorescence anisotropy (59). Here we have extended these studies to measure the interaction of wild-type and mutant forms of the OBD (aa 131 to 260), and of a longer T-ag fragment also containing the N-terminal domain (aa 1 to 260), with different DNA substrates. We found that the N-terminal domain does not contribute directly to DNA binding but dampens the propensity of the OBD to form large oligomers in solution and on DNA. We also report the determination of a specific T-ag consensus binding sequence and that the OBD has a substantial affinity for ssDNA, which
is 10-fold higher than its affinity for nonspecific dsDNA. These findings are discussed in light of the recent crystal structures of the OBD and models of double hexamer assembly.

**MATERIALS AND METHODS**

**Expression plasmids.** Plasmids to express the SV40 T-ag N130, OBD, and N260 protein fragments fused to a six-histidine tag and glutathione S-transferase (GST) were constructed by inserting BamHI-EcoRI-digested PCR products encoding these fragments between the BamHI and EcoRI sites of plasmid AB-401. AB-401 is a modified version of pGEX-4T-1 (GE Healthcare, formerly Amersham Biosciences) in which a sequence encoding a His tag has been inserted downstream of the GST-coding region. Primers used for amplification were designed so as to encode EcoRI and BamHI restriction sites. The following pairs of primers were used: SV40 T-ag OBD (aa 1 to 260), 5′-GGCTGAGACTCTCTTGAGTTAAAATGGTCAAGAGAG-3′ and SV40 T-ag N260 (aa 1 to 260), 5′-GGCTGAGACTCTCTTGAGTTAAAATGGTCAAGAGAG-3′ and 5′-GGGGAAATTCACTATTCTGGATTAAAATCATGGCTCC-3′; SV40 T-ag N260 (aa 1 to 260), 5′-GGCTGAGACTCTCTTGAGTTAAAATGGTCAAGAGAG-3′ and SV40 T-ag N130 (aa 1 to 130), 5′-GGCTGAGACTCTCTTGAGTTAAAATGGTCAAGAGAG-3′ and 5′-GGGGAAATTCACTATTCTGGATTAAAATCATGGCTCC-3′; and SV40 T-ag OBD (aa 131 to 260), 5′-GGCTGAGACTCTCTTGAGTTAAAATGGTCAAGAGAG-3′ and 5′-GGGGAAATTCACTATTCTGGATTAAAATCATGGCTCC-3′; and SV40 T-ag OBD (aa 131 to 260), 5′-GGCTGAGACTCTCTTGAGTTAAAATGGTCAAGAGAG-3′ and 5′-GGGGAAATTCACTATTCTGGATTAAAATCATGGCTCC-3′.

**Protein expression and purification.** Expression vectors for GST-N130, OBD, and GST-N260 were transformed into Escherichia coli BL21 (Novagen). Bacterial cultures were grown in Luria-Bertani medium supplemented with 2% glucose, 100 μg/ml of ampicillin, and 100 μg/ml of kanamycin. Bacterial cultures were grown in Luria-Bertani medium supplemented with 2% glucose, 100 μg/ml of ampicillin, and 100 μg/ml of kanamycin.

**Dynamic light scattering (DLS).** Dynamic light scattering (DLS) was performed using a DynaPro instrument (Wyatt Technology Corporation). Purified proteins (N130, OBD, and N260) were analyzed in 20 mM HEPES-NaOH (pH 7.4), 50 mM NaCl, 1 mM DTT, and 0.01% NP-40 at concentrations of 10 μM. GST and buffer alone were used as a positive and as a negative control, respectively (data not shown). Scattering data were analyzed using the Dynamics software (Protein Solutions). The hydrodynamic radius values reported in this study were obtained using the regularization algorithm in Dynamics, which performs a multiple exponential fit of the raw data and provides a size distribution analysis for samples containing two or more populations. In this study, results were an average of 20 readings of each sample.

**Cross-linking assays.** Increasing concentrations (1.25, 2.5, 5, and 10 μM) of purified N130, OBD, and N260 were incubated alone or together (N130 and N260 or OBD) in 20 mM HEPES-NaOH (pH 7.4), 50 mM NaCl, 1 mM DTT, 0.01% NP-40, and 0 to 2.5% glycerol in a total volume of 80 μl for 1 h at room temperature. Samples were then cross-linked with 0.1% glutaraldehyde (Sigma) for 10 min. Subsequently, the protein samples were mixed with 20 μl of 5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) loading buffer and boiled for 5 min. Cross-linked proteins were resolved on 10% (N260), 12% (OBD), or 15% (N130) SDS-PAGE gels and detected by silver staining or Western blotting with an anti-His tag antibody (8036, H3; Santa Cruz) as indicated. In these experiments, GST was used as a positive control (data not shown).

**Fluorescence anisotropy DNA binding assay.** Binding assays were performed by use of 15 mM fluorescein-labeled probe and the indicated concentrations of protein in 150 μl reaction mixtures with the following buffer: 20 mM HEPES-NaOH (pH 7.4), 50 mM NaCl, 0.01% NP-40, 1 mM DTT, and 0 to 2.5% glycerol. Fluorescence readings were taken after an hour of incubation at room temperature in OptiPlate-96 high-binding-affinity black 96-well plates (Perkin Elmer). The samples were then assayed on a VICTOR plate reader (Perkin Elmer) equipped with a 485-nm/535-nm filter set.

**Fluorescence DNA probes and competitor oligonucleotides.** Fluorescin-labeled oligonucleotides were purchased from Invitrogen or Genset (Paris, France), with the fluoresophore attached at the 5′ end by a six-carbon linker. Duplex DNA probes were prepared by annealing each fluorescin-labeled oligonucleotide to a complementary oligonucleotide in a ratio of 1:1.5 by heating at 95°C for 5 min in DNA binding assay buffer (20 mM HEPES-NaOH [pH 7.4], 50 mM NaCl, 0.01% NP-40, and 1 mM DTT) and then slowly cooling to room temperature. Competitor oligonucleotides were purchased from Invitrogen and annealed as described above except in a ratio of 1:1. 

**K<sub>D</sub> determination.** Equilibrium dissociation constant (K<sub>D</sub>) values were obtained from direct binding isotherms, with each data point performed in triplicate, and fitted by nonlinear least-squares regression with the GraphPad Prism program to the standard Michaelis-Menten equation, D + P ⇌ DP (where D is DNA, P is the T-ag protein fragment, and DP is DNA-T-ag protein complex) with the expression of the Hill equation, where n is the number of binding sites, A<sub>D</sub> is the actual amount of DNA bound (A<sub>B</sub> is the measured amount of DNA probe bound and A<sub>T</sub> is the total amount of DNA), and K<sub>D</sub> is the dissociation constant. The Hill equation is a modification of the Michaelis-Menten equation that accounts for the binding of more than one molecule of protein to a single DNA strand.

**RESULTS**

Purification and characterization of N130, OBD, and N260 by DLS. As a first step towards determining the contribution of the N-terminal domain of T-ag to the assembly of the protein...
As these proteins were produced in bacteria at the viral origin, we expressed and purified from bacteria the N-terminal domain (N130; aa 1 to 260), the OBD (aa 131 to 353), or a protein fragment containing both the OBD and the N-terminal domain (N130, OBD, and N260 proteins at a concentration of 10 μM each curve represents an average of 20 measurements. (B) Molecular weight (MW) and types of complexes estimated from each hydrodynamic radius.

### Table 1

<table>
<thead>
<tr>
<th>Protein</th>
<th>Peak</th>
<th>Average MW</th>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>N130 (■)</td>
<td>1</td>
<td>30</td>
<td>Monomer</td>
</tr>
<tr>
<td>1a</td>
<td>5600</td>
<td>HMWC</td>
<td></td>
</tr>
<tr>
<td>N260 (△)</td>
<td>2</td>
<td>50</td>
<td>Monomer</td>
</tr>
<tr>
<td>2a</td>
<td>8000</td>
<td>HMWC</td>
<td></td>
</tr>
<tr>
<td>OBD (●)</td>
<td>3</td>
<td>20</td>
<td>Monomer</td>
</tr>
<tr>
<td>3a</td>
<td>5400</td>
<td>HMWC</td>
<td></td>
</tr>
</tbody>
</table>

FIG. 2. DLS analysis of N130, OBD, and N260. (A) The graph represents the scattering intensity (fraction of total scattering) as a function of the hydrodynamic radii obtained for the N130, OBD, and N260 proteins at a concentration of 10 μM. Each curve represents an average of 20 measurements. (B) Molecular weight (MW) and types of complexes estimated from each hydrodynamic radius.

at the viral origin, we expressed and purified from bacteria the N-terminal domain (N130; aa 1 to 130), the OBD (aa 131 to 260), or a protein fragment containing both the OBD and the N-terminal domain, N260 (aa 1 to 260). These proteins were produced as fusions with GST and a His tag, and the GST N-terminal domain, N260 (aa 1 to 260), or a protein fragment containing both the OBD and the N-terminal domain (N130, OBD, and N260 ranging from 1.25 to 10 μM were incubated in the presence or absence of 0.1% glutaraldehyde and analyzed by SDS-PAGE. GST, a known dimer, was used as a positive control (data not shown). As shown in Fig. 3A, N130 was primarily monomeric even at high concentrations. In contrast, the OBD was cross-linked primarily into dimers, trimers, and tetramers at low concentrations and into HMWC at higher ones (Fig. 3B). Similar results were obtained with an untagged OBD protein (data not shown). The analysis with N260 protein revealed that it could also assemble into dimers, trimers, and tetramers at high concentrations (Fig. 3C, lane 5) but never formed complexes as large as those generated with the OBD. These results were in agreement with those obtained by DLS. Collectively, they suggested that the N-terminal domain dampens the propensity of the OBD to form large complexes, perhaps, one may speculate, by obstructing an oligomerization interface on the OBD.

Interestingly, the cross-linking of N260 generated a species that migrated faster than the intact monomer (not cross-linked) and which was formed even at low protein concentrations (Fig. 3C; lanes 2 to 5 form the left of the gel). This faster-migrating species is possibly the result of intramolecular cross-linking events that would generate a more tightly packed protein, a phenomenon which has been reported for other proteins (14). Since no faster-migrating species were detected either with the OBD or with N130, it is possible that the one generated with N260 arises from intramolecular cross-linking between the OBD and the N-terminal domain. To investigate a possible interaction between the N-terminal domain and a region of the OBD, we asked whether N130 could be cross-linked to N260 or the OBD. This was found to be the case for N260, as determined by the appearance of new protein bands not detected when either protein was cross-linked alone. The presence of a novel band corresponding in size to a heterodimer of N130 and N260 was particularly striking (Fig. 3D, lane 5). The same protocol could not be used to determine if N130 can be cross-linked to the OBD, since both proteins are of comparable molecular weights. To circumvent this problem, we used an untagged OBD and followed its cross-linking to increasing concentrations of His-tagged N130 by Western blotting with an anti-His tag antibody. These experiments revealed the formation of hetero-oligomers, in particular dimers and trimers, between the OBD and N130 (Fig. 3E right, lanes 4 and 5). Interestingly, the formation of these hetero-oligomers was accompanied by a concomitant decrease in the assembly of the OBD into HMWC, suggesting that N130 reduces in trans oligomerization of the OBD (Fig. 3E left, lanes 3 to 5). These results provided further evidence that the N-terminal domain can associate with the OBD.

### DNA binding properties of purified N130, OBD, and N260.

The DNA binding properties of these purified proteins were then characterized using an assay based on FP (17, 29, 59). In comparison to other common methods to monitor DNA binding, such as electrophoretic mobility shift assays, FP assays offer the advantages of being performed in solution and at equilibrium and hence are ideally suited for affinity measurements (17, 29). FP DNA binding assays were used to determine the affinity of the three proteins for a duplex oligonucleotide probe containing one pentanucleotide (TBS; 5'-GAGGC-3'). As a control for specificity, a similar probe in which the TBS

---

**Fig. 1C**: As these proteins were produced in bacteria at the viral origin, we expressed and purified from bacteria the N-terminal domain (N130; aa 1 to 260), the OBD (aa 131 to 353), or a protein fragment containing both the OBD and the N-terminal domain (N130, OBD, and N260 proteins at a concentration of 10 μM). Each curve represents an average of 20 measurements. (B) Molecular weight (MW) and types of complexes estimated from each hydrodynamic radius.
was inactivated by five mutations (5'-AGAAT-3') was used (Fig. 4A). Throughout this study, we will when necessary refer to the TBS probe as “specific DNA” and to the mutant probe as “nonspecific DNA.” Titration of the purified OBD and N260 proteins resulted in a dose-dependent increase in polarization (expressed as an anisotropy gain) with the TBS probe (Fig. 4B), indicative of their binding to DNA. In contrast, titration of purified N130 had no effect on polarization, demonstrating that this protein does not bind significantly to DNA, even at concentrations as high as 20 μM (Fig. 4B and data not shown). From a series of binding isotherms similar to those shown in Fig. 4B, \(K_D\) values of 28 ± 4 nM and 63 ± 6 nM were obtained for the binding of N260 and OBD proteins to the TBS-containing probe (Fig. 4C), respectively. Binding of these two proteins to a control probe lacking a TBS was much weaker, confirming that both bind DNA in a sequence-specific manner. Interestingly, a significant increase in anisotropy was reproducibly observed at high concentrations of the OBD (>500 nM) (Fig. 4B), an observation which we made with several nonspecific probes of differing sequences (data not shown). This effect is likely due to nonspecific binding of the OBD to DNA and/or its ability to oligomerize at high concentrations (>500 nM), as determined above by DLS and cross-linking studies. This property of the OBD is also likely the reason why its binding to the TBS-containing probe did not reach a plateau at high protein concentrations. To rule out a possible effect of the His tag, we performed similar experiments with an untagged OBD and confirmed that it has affinities similar to those of the tagged protein for the one TBS probe (\(K_D\) values of 24 ± 9 nM) and for nonspecific DNA (data not shown). In contrast to the OBD, N260 showed little binding to the control probe, and its binding to the TBS-containing probe was clearly saturable. Collectively, these results indicate that the N260 and OBD proteins bind to a single TBS with comparable affinities, which diverge by less than threefold. However, their binding to nonspecific DNA appears to be more dissimilar. Namely, the significant gain in anisotropy observed at high concentrations of the OBD suggests that more molecules of the OBD are bound on the nonspecific DNA probe than for N260.

We then investigated the impact of the length of the probe on the binding of the N260 and OBD proteins to a single TBS. These studies showed that both proteins could not bind to a probe containing 8 bp (Table 1) but could do so on a 12-bp probe. However, maximum binding affinity was reached only with probes greater than 16 bp (Table 1). These observations are in agreement with the crystal structure reported by Bochkareva et al. (4) as well as Meinke et al. (37), which demonstrated that the OBD spans approximately 10 bp when bound to dsDNA. Moreover, the fact that the same length of DNA is needed for binding of the OBD and the N260 protein supports the idea that the N-terminal domain of the protein does not significantly interact with DNA.

Influence of pH and ionic strength on the binding of the OBD and N260 to DNA. The binding of polyomavirus T-ag to origin DNA has been shown to increase at lower pH (43). We used our FP assay to determine if acidity also increases the affinities of the SV40 T-ag OBD and N260 proteins for specific as well as nonspecific DNA. Assays were performed as before, except that the final pH of the binding reactions was 6.8 rather
This lower pH increased binding of the OBD and N260 to a single TBS-containing probe only slightly, by approximately twofold (Fig. 5A and B). In contrast, binding to a nonspecific DNA probe (lacking a TBS) was drastically increased by 25- to 40-fold. We also investigated the effect of increasing the ionic strength on the affinities of the proteins for specific DNA. As expected, increasing the NaCl concentration reduced the affinities of the OBD and N260 proteins for the TBS-containing probe (Table 1). Specifically, 100 mM NaCl reduced DNA binding affinity 10- to 20-fold, and a further increase to 150 mM almost completely abolished DNA binding of both proteins to the probe. Error bars are not visible on the graph, as they are smaller than the symbols. (C) $K_D$ values for the N130, OBD, and N260 proteins for the specific and nonspecific probe were calculated from the data by nonlinear regression and fitting to a one-binding-site equilibrium. The values and standard deviations reported were calculated from over 20 independent binding isotherms, with each data point performed in triplicate.

TABLE 1. Effect of probe length and ionic strength on the DNA binding affinities of the OBD and N260 proteins

<table>
<thead>
<tr>
<th>Probe or buffer constituent</th>
<th>Probe length or ionic strength</th>
<th>$K_D$ (nM) for:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>OBD</td>
</tr>
<tr>
<td>5'-CCGAGGCATTCGAGAATTGAGC-3'</td>
<td>8 bp</td>
<td>&gt;2,500</td>
</tr>
<tr>
<td>5'-CCGAGGCATTCGAGAATTGAGC-3'</td>
<td>12 bp</td>
<td>41 ± 3</td>
</tr>
<tr>
<td>5'-CCGAGGCATTCGAGAATTGAGC-3'</td>
<td>16 bp</td>
<td>26 ± 2</td>
</tr>
<tr>
<td>5'-CCGAGGCATTCGAGAATTGAGC-3'</td>
<td>20 bp</td>
<td>21 ± 2</td>
</tr>
<tr>
<td>NaCl</td>
<td>50 nM</td>
<td>22 ± 2</td>
</tr>
<tr>
<td></td>
<td>100 nM</td>
<td>269 ± 39</td>
</tr>
<tr>
<td></td>
<td>150 nM</td>
<td>&gt;2,500</td>
</tr>
<tr>
<td></td>
<td>200 nM</td>
<td>&gt;2,500</td>
</tr>
</tbody>
</table>

$^a$ The $K_D$ values for the OBD and N260 proteins for the indicated probes (with TBS underlined and in boldface) at 50 mM NaCl (top) and for a 24-bp probe at the indicated concentrations of salt (NaCl; bottom) were calculated from binding isotherms performed as described for Fig. 4. Standard deviations were calculated from two independent binding isotherms, with each data point performed in triplicate.
affinities of the OBD and N260 proteins for each of these four sites. Affinity measurements were made with the FP assay described above but using unmodified duplex DNAs carrying one of the four pentanucleotides as competitors. Because we wanted to determine the affinity for each binding site in the context of the origin, we used competitor versions of the 31-bp site II in which three of the four pentanucleotides were inactivated by mutations (Fig. 6). A version of site II with all TBS mutated served as a negative control. For each competitor DNA, increasing amounts were used to inhibit the binding of the OBD and N260 to a TBS-containing probe in order to determine an IC50, from which an apparent Ki was derived (see Materials and Methods) (Fig. 6). Three observations were made. First, we confirmed that the binding affinities of the OBD and N260 proteins for P3 measured in these competition experiments (Ki values) were comparable, within twofold, to those obtained in direct binding experiments to fluorescent probes (Kd values; Fig. 4), as expected if the fluorescein moiety on the probe does not impair binding. Second, we observed that the affinities of N260 for each of the four pentanucleotides were consistently two- to threefold lower than those measured for the OBD (Fig. 6), consistent with the data from Fig. 4C. Third, we reproducibly found that both N260 and the OBD bind with low affinities to P2 and to a lesser extent to P4, compared to what was seen for P1 and P3, despite the fact that all four sites have identical sequences. This last finding provides a rational explanation for why double hexamers have been found to preferentially assemble on sites 1 and 3 rather than 2 and 4 (22).

Determination of a consensus T-ag binding sequence. The observation that P1 and P3 are bound with affinities differing from those for P2 and P4, despite having the same sequence, prompted us to investigate if the nucleotides immediately flanking each 5'-GAGGC-3' site could also contribute to affinity (Fig. 7A). To test this, we performed a systematic mutagenesis of the two nucleotides immediately upstream and downstream of P3, respectively. In addition, we also mutated each nucleotide of the 5'-GAGGC-3' sequence in order to determine the contribution of each to affinity and generate a consensus T-ag binding sequence, a study which has never been done systematically and quantitatively. We performed this systematic analysis on P3 in the context of a site II oligonucleotide in which only P3 had not been inactivated (Fig. 7A). The different mutated duplex DNAs were used in competition experiments, and their Ki values against the OBD and N260 were determined (Fig. 7A). This study revealed that (i) the nucleotide immediately 5' of P3 (C) and the first position of the pentanucleotide (G) can be changed to any of the three other nucleotides without a drastic effect on binding affinity, (ii) the second position (A) can be changed to a cytosine or thymine but not to a guanine, (iii) the third and fourth positions (GG) cannot be substituted by any of the other three nucleotides without a drastic loss of affinity, (iv) the fifth position (C) cannot accommodate an adenine or a thymine, and (v) the nucleotide immediately 3' of the pentanucleotide can be changed to an adenine or a thymine but a guanine at this position reduces binding. Based on this last finding, we infer that the presence of a guanine immediately 3' of P2, rather than a cytosine for all three other pentanucleotides, is likely the reason why this site is bound with the lowest affinity (Fig. 6). Collectively, these results indicate that the preferred binding site for the OBD is 5'-GAGGC-3', followed by A, C, and

<table>
<thead>
<tr>
<th>P4</th>
<th>P3</th>
<th>P2</th>
<th>P1</th>
<th>P0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S'</td>
<td>P4</td>
<td>P3</td>
<td>P2</td>
<td>P1</td>
</tr>
<tr>
<td>K_i (nM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OBD</td>
<td>47 ± 7</td>
<td>31 ± 5</td>
<td>202 ± 34</td>
<td>65 ± 8</td>
</tr>
<tr>
<td>N260</td>
<td>110 ± 14</td>
<td>31 ± 5</td>
<td>202 ± 34</td>
<td>65 ± 8</td>
</tr>
</tbody>
</table>

FIG. 6. Affinities of the OBD and N260 proteins for each individual pentanucleotide in site II. Schematic representations of site II and competitor DNAs are shown on the left. All competitors contain only one of four pentanucleotides (P1 to P4; 5'-GAGGC-3'). P0 is a competitor in which all four pentanucleotides were mutated to 5'-AGAAAT-3'. Arrows represent the direction of the 5'-GAGGC-3' sequence. Apparent Ki values measured for each competitor DNA are indicated on the right. Ki and standard deviations were calculated from the IC50 of each competitor DNA obtained in triplicate in assays performed at 250 nM of protein and 15 nM of one TBS probe.
recently by NMR that the OBD can bind to poly(dT)25 ssDNA within site II. OBD, binds independently to each of the four pentanucleotides stream of it. These results strongly suggest that N260, like the P4 was not significantly changed by the presence of P3 down-strands shown in Fig. 4A). Titration of increasing concentra-

sequence (these probes correspond to the fluorescein-labeled -GAGGC-3'H11032 probes, one containing and the other lacking a 5

affinities of the proteins for two 24-nucleotide-long ssDNA (45). We used FP to measure the affinity of the OBD and N260 for different ssDNA substrates and assess the role of the N-terminal domain on this interaction. We first measured the affinities of the proteins to a single TBS-containing probe are

values for the OBD and N260

1D, and their affinities for a single TBS-containing probe are

sequence-specific DNA binding protein to a nonspecific helicase

that specifically affect its binding to the origin or its phosphor-
ylation-dependent assembly into functional double hexamers. The availability of a sensitive FP assay that monitors binding at
equilibrium prompted us to reexamine the effect of some of these substitutions on the binding of the OBD and N260 proteins to a single TBS. We focused our attention on substitutions that do not lie directly in the DNA binding surface but rather on those in other regions of the OBD or in the N-terminal domain and for which the mechanism of action is not completely understood. The locations of these amino acid substitutions in the structure of the OBD (37) are shown in Fig.

1D, and their affinities for a single TBS-containing probe are described below and reported in Table 2.

Substitutions of series 120 and 123. Previous studies have found that dephosphorylation of Ser120 and Ser123 by protein phosphatase 2A increases the affinity of T-ag for the origin and stimulates its assembly into double hexamers (3, 27, 39, 49, 50, 60, 61). Accordingly, phosphorylation of S120 and S123 by casein kinase I was found to inhibit viral DNA replication (8). Since our FP assays are performed with recombinant proteins

T. However, significant binding can be observed on the follow-
ing consensus sequence: 5'-N(ACT)GG(CG)3' (Fig. 7B). Inter-
estingly, inspection of the sequence of site II revealed the presence of two additional TBS in addition to the four already known pentanucleotides (Fig. 7C). Determining whether these two potential TBSs are of any biological significance will neces-
sitate further experimentation.

Binding affinities of the OBD and N260 for pairs of pen-
tanucleotides. We previously reported that the OBD does not bind to a palindromic combination of two pentanucleotides, either P1 and P3 or P2 and P4, with increased affinity (59). These results led us to suggest that monomers of the OBD bind to these sites independently rather than cooperatively, a suggestion that was recently supported by the two crystal struc-
tures of the OBD bound to DNA (4, 37). To verify if the same was true for N260 and extend these studies to other combina-
tion of sites, we measured the affinities of N260 and the OBD for the three different types of combinations of pentanucleo-
tides that can be found within site II. These combinations cor-
respond to inverted repeats with a spacing of 7 bp (such as the combination of sites P3 and P1), inverted repeats with a spac-
ing of 1 bp (such as sites P3 and P2), and direct repeats with a spac-
ing of 1 bp (such as sites P4 and P3) (Fig. 8). As can be seen in Fig. 8, the presence of P1 or P2 did not enhance the affinities of the OBD or N260 for P3. Similarly, binding to P4 was not significantly changed by the presence of P3 downstream of it. These results strongly suggest that N260, like the OBD, binds independently to each of the four pentanucleotides within site II.

Binding of the OBD and N260 to ssDNA. It was shown recently by NMR that the OBD can bind to poly(dT)25 ssDNA (45). We used FP to measure the affinity of the OBD and N260 proteins for different ssDNA substrates and assess the role of the N-terminal domain on this interaction. We first measured the affinities of the proteins for two 24-nucleotide-long ssDNA probes, one containing and the other lacking a 5'-GAGGC-3' sequence (these probes correspond to the fluorescein-labeled strands shown in Fig. 4A). Titration of increasing concentra-
tions of proteins resulted in a dose-dependent gain in aniso-
tropy that was much larger for the OBD than for N260 and that was independent of the presence of the 5'-GAGGC-3' se-
quence. These results indicated that the OBD does bind to ssDNA in a nonspecific manner (Fig. 9A). Furthermore, the larger gain in anisotropy observed for the OBD suggests that at the highest concentrations tested, more OBD proteins are bound to the probe than N260. When N130 was incubated with the same probe, no binding was detected (data not shown), indicating that the N-terminal domain of T-ag does not bind to ssDNA.

We then characterized further the binding of the OBD and N260 to ssDNA. We found that the magnitude of the aniso-
tropy gain was a function of the probe length, as expected if the OBD and N260 bind to ssDNA in a nonspecific manner (Fig.

9B and data not shown). We also found that binding to ssDNA of increasing length was increased at lower pH, similar to what we observed for duplex DNA (Fig. 5B and data not shown). Even at this lower pH, binding to a probe of 8 or 12 nucleotides was weaker than on longer probes of 16 and 20 nucleotides. The affinities of the OBD for the 16- and 20-nucleotide-long probes at pH 6.8 were measured to be 741 and 370 nM, re-

spectively. These results with ssDNA probes were reminiscent of those obtained with nonspecific duplex DNA (compare Fig.

5 and 9C), suggesting that the OBD might use similar surfaces to interact with ss- and dsDNA. If so, we would expect that binding to both types of nucleic acids might be mutually ex-
clusive. We tested this possibility by investigating if binding of N260 and the OBD could be competed by ssDNA. As can be seen in Fig. 9D, binding to a duplex DNA probe carrying a single TBS could be efficiently inhibited by ssDNA. These results indicate that N260 and the OBD bind to ss- and dsDNA in a mutually exclusive manner. Remarkably, both proteins showed an affinity for ssDNA 10-fold greater than that for nonspecific dsDNA. These results have important implications when the mechanism by which T-ag transitions from a se-

quence-specific DNA binding protein to a nonspecific helicase is considered (see Discussion).

DNA binding affinities of mutant OBD and N260 proteins. Several amino acid substitutions in T-ag have been identified that specifically affect its binding to the origin or its phosphor-
ylation-dependent assembly into functional double hexamers.
made in *E. coli* and hence not phosphorylated, we investigated if the double amino acid substitutions of S120 and S123 for aspartic acids (S120D/S123D) as phosphomimetics would have an effect on binding affinity. Under our assay conditions, the S120D/S123D substitutions reduced only slightly the affinity of N260 for the one TBS probe (Table 2), indicating that negative charges at positions 120 and 123 are not sufficient to reduce the interaction of the OBD with its target site.

Substitutions of threonine 124 and class 4 substitutions. Phosphorylation of Thr124 by a cyclin/cdk was previously shown to be essential for viral DNA replication (15, 32, 33, 35, 38, 48) and to affect double hexamer assembly in vitro (2, 46, 64) (see the introduction). Similarly, specific amino acid substitutions in the OBD, known as class 4 mutations, have been identified that interfere with double hexamer assembly but without affecting phosphorylation of the protein (64). An interpretation of these findings is that the N-terminal domain and the OBD cooperate during the double hexamer assembly in a phosphorylation-dependent manner. We produced and purified N260 mutant proteins carrying substitutions of T124 for either an alanine (T124A) or for aspartic or glutamic acid as phosphomimetics (T124D and T124E). We also produced N260 and OBD mutant proteins carrying one of five different class 4 substitutions: H148N, K167R, Q213H, L215V, and F220Y (described in reference 67). In the FP assay, all of the T124 substitutions had little effect on the affinity of N260 and the OBD for specific DNA (Table 2). Similar results were obtained for class 4 substitutions. The two substitutions with the greatest effects were H148N and K167R, which reproducibly decreased binding but by less than twofold (Table 2). Importantly, we also observed very little effect of T124 and class 4 substitutions on the binding of the OBD or N260 to pairs of binding sites, as was done for the wild-type proteins as

![Image](http://jvi.asm.org/article-pdf/91/10/9170/3383340/9170_FRADET-TURCOTTE_ET_AL_J.VIROL)

FIG. 9. Binding affinity of wild-type OBD and N260 for ssDNA. (A) Binding isotherms obtained with ssDNA probes either containing (filled symbols) or lacking (open symbols) a 5'-GAGGC-3' sequence and with the indicated protein. (B and C) Binding isotherms generated from the binding of the OBD to ssDNA probes of the indicated length (in nucleotides [nt]) at pH 7.4 (B) or pH 6.8 (C). (D) Apparent $K_i$ values for single-stranded and double-stranded nonspecific DNA competitors of the indicated lengths against the OBD and N260 proteins. $K_i$ and standard deviations were calculated from the $IC_{50}$ of each competitor DNA obtained in triplicate with assays performed with 250 nM of protein and 15 nM of one TBS probe.

### TABLE 2. DNA binding affinities of mutant OBD and N260 proteins

<table>
<thead>
<tr>
<th>Amino acid substitution</th>
<th>$K_i$ for one TBS (nM) for:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OBD</td>
</tr>
<tr>
<td>None (wild type)</td>
<td>28 ± 4</td>
</tr>
<tr>
<td>Phosphorylation site substitutions</td>
<td></td>
</tr>
<tr>
<td>S120D/S123D</td>
<td>NA</td>
</tr>
<tr>
<td>T124A</td>
<td>NA</td>
</tr>
<tr>
<td>T124D</td>
<td>NA</td>
</tr>
<tr>
<td>T124E</td>
<td>NA</td>
</tr>
<tr>
<td>Class 4 substitutions</td>
<td></td>
</tr>
<tr>
<td>H148N</td>
<td>38 ± 10</td>
</tr>
<tr>
<td>K167R</td>
<td>30 ± 5</td>
</tr>
<tr>
<td>Q213H</td>
<td>28 ± 2</td>
</tr>
<tr>
<td>L215V</td>
<td>24 ± 3</td>
</tr>
<tr>
<td>F220Y</td>
<td>25 ± 6</td>
</tr>
<tr>
<td>Oligomerization substitu</td>
<td></td>
</tr>
<tr>
<td>tions F151A</td>
<td>23 ± 6</td>
</tr>
<tr>
<td>F151Y</td>
<td>25 ± 5</td>
</tr>
<tr>
<td>F183A</td>
<td>44 ± 10</td>
</tr>
<tr>
<td>F151A/F183A</td>
<td>28 ± 4</td>
</tr>
</tbody>
</table>

$K_i$ values and standard deviations were obtained from at least three independent binding isotherms as described in Fig. 4 and with each data point done in triplicate. For each amino acid substitution, the ratio of the $K_i$ value of N260 over that of the corresponding OBD is indicated on the right.
shown in Fig. 8 (data not shown). From these studies, we conclude that T124 and class 4 mutations do not affect binding to the origin but rather affect a later step in double hexamer assembly. It is also noteworthy that a mutant T-ag carrying the T124E substitution was shown to be replication defective by an unknown mechanism (50). Our data suggest that this defect is not related to origin binding.

**Substitutions in a putative OBD-OBD interface.** A recent crystal structure of the OBD has suggested that it can assemble into a left-handed spiral through two complementary binding surfaces (36). To determine the importance of these interaction surfaces for DNA binding, we produced mutant proteins carrying substitutions predicted to disrupt spiral formation. Specifically, phenylalanines 151 and 183 (F151, F183) of the OBD and N260 proteins were each mutated to alanine or tyrosine to maintain the aromatic character of the residue. We also characterized the effect of combining both substitutions, F151A/F183A. In the FP assay, these substitutions reduced by less than half the affinities of N260 or the OBD for the TBS probe. Throughout this study, we found that N260 binds to DNA with an affinity two- to threefold lower than that of the OBD and that nearly all of substitutions failed to alter this difference (Table 2). The only exception we observed was for F183A, which, either alone or in combination with F151A, appears to reduce this affinity difference between N260 and the OBD. Furthermore, the $K_D$ values suggest that it does so primarily by increasing the affinity of N260 for DNA, albeit very slightly. For all three substitutions tested, we also observed very little effect on the binding of the OBD or N260 to pairs of binding sites (data not shown). The fact that these substitutions in a putative OBD-OBD interface have little effect in our assays is consistent with the idea that spiral formation occurs at a step subsequent to origin binding.

**DNA binding affinities of full-length T-ag.** We purposely performed all of the studies presented above with fragments of T-ag lacking the C-terminal helicase domain in order to investigate only the DNA binding characteristics of the OBD and avoid any possible confounding effect from the nonspecific interaction of the helicase domain with DNA (20, 65). Nevertheless, we still wished to compare the binding specificity of N260 to that of the complete T-ag protein. We therefore characterized the binding of full-length T-ag, produced in insect cells by use of a baculovirus system, in our FP assay. Specifically, we measured its binding to a 24-bp probe containing a single TBS and to a control probe lacking a specific binding site. These studies were performed in the presence or absence of 5′-adenyl-β, γ-imidodiphosphate (AMP-PNP), which promotes oligomerization of T-ag. In the absence of AMP-PNP, increasing concentrations of T-ag generated a dose-dependent gain in anisotropy (Fig. 10A) that was larger than the one generated by N260. This large anisotropy gain is likely due to T-ag being a bigger protein than N260. Interestingly, T-ag also bound substantially to the nonspecific probe much more than N260. It is entirely possible that this nonspecific binding occurs through the helicase domain of the protein. Regardless, these findings indicate that the specificity of full-length T-ag, as defined by its ability to discriminate between specific and nonspecific DNA, is less than that of N260. $K_D$ values for the binding of T-ag to the specific and nonspecific probes were estimated to be $170 \pm 22$ nM and $494 \pm 38$ nM, respectively.

These values contrast with those obtained with N260 under the same conditions, which were $63 \pm 6$ nM and >2,500 nM, respectively (Fig. 4C). The addition of AMP-PNP increased the binding of T-ag to both specific and nonspecific DNA (Fig. 10B), to the point of becoming too tight to be accurately measured under these assay conditions. AMP-PNP also gave rise to a greater gain in anisotropy, as would be expected if multimers (for example, hexamers) of T-ag are binding to the probe. Regardless of the precise stoichiometry, the important point from the above-described studies is that full-length T-ag has a substantially lower affinity for nonspecific DNA, in contrast to N260. This raises the question of how full-length T-ag finds the origin in vivo in the presence of a vast excess of competitor genomic DNA. Perhaps there exist mechanisms to reduce the nonspecific binding of the protein in vivo (see Discussion).

**DISCUSSION**

In this study, we have measured the affinity of wild-type and mutant versions of the T-ag OBD and of a larger N-terminal fragment, N260, for different DNA substrates by fluorescence anisotropy as well as determined the oligomeric states of these proteins in solution by DLS and cross-linking assays. These biophysical studies have provided a quantitative basis onto
which to anchor our understanding of the interaction of T-ag with the origin and suggested a role for the N-terminal domain of the protein in regulating the oligomerization of the OBD.

Effect of the N-terminal domain on the OBD. Earlier studies suggested that the N-terminal region of T-ag modulates the interaction and/or assembly of the protein at the origin (see the introduction). Our characterization of purified N130 indicated that this T-ag fragment is primarily monomeric in solution and does not bind with any significant affinity to ds- or ssDNA. In addition, we found that N260 had affinities comparable to those of the OBD for all DNA substrates tested, which consistently differ by less than threefold. Hence, we conclude that the N-terminal domain has only a modest effect on the affinity of the OBD for specific and nonspecific ss- and dsDNA and thus that its role in double hexamer assembly must occur at a step subsequent to origin binding and/or be dependent on the C-terminal helicase domain, which is missing from our protein fragments (see below).

Interestingly, preparations of N260 were more monodisperse than those of the OBD in DLS experiments, suggesting that the N-terminal domain reduces the multimerization of the OBD under our assay conditions. This was also observed in glutaraldehyde cross-linking assays, which showed that the OBD, unlike N260, could be cross-linked into HMWC. We believe that the tendency of the OBD to self-associate under our assay conditions accounts for its ability to generate larger anisotropy gains than N260 upon binding to nonspecific ds- and ssDNA in FP assays. We cannot rule out that the N-terminal domain stabilizes the OBD and prevents its aggregation in solution. However, we note that at low concentrations (<5 μM), the OBD is mostly monodisperse and monomeric, as assessed by DLS and cross-linking, respectively (data not shown), arguing against it being aggregated at the concentrations used in our assays. A recent crystal structure has suggested that the OBD can oligomerize into a left-handed spiral (36). Whether the multimerization of the OBD that we observe in solution or on DNA is related to spiral formation is unknown. If it were, it would raise the possibility that the N-terminal domain regulates spiral formation by masking an oligomerization interface on the OBD. Along these lines, we have found that glutaraldehyde cross-linking of N260 generates a faster-migrating species on SDS-PAGE gels, which we interpreted as a more compact form of the protein. This phenomenon was not observed with the OBD and N130, suggesting that it was due to intermolecular cross-linking of the N-terminal domain against the OBD. In support of this notion, we found that N130 could be cross-linked to either N260 or the OBD. These results indicate that the N-terminal domain can interact in trans with another T-ag molecule, albeit at high concentrations, as would be the case for two T-ag molecules bound next to each other on DNA. Collectively, our findings support the notion that the N-terminal domain of T-ag regulates oligomerization of the OBD, in agreement with the finding by Weisshart et al. (62) that truncation of the first 82 aa from full-length T-ag generates a protein that forms aberrant multimers in solution and on DNA.

Interaction of the OBD and N260 with their target pentanucleotide binding site. Our systematic and quantitative analysis of the importance of each nucleotide of the 5'-GAGGC-3' sequence led to the determination of the following consensus TBS: 5'-N(ACT)GG(CG)-3'. This consensus highlights the importance of the GGC triplet for high-affinity binding, in agreement with two recent crystal structures of the OBD bound to DNA showing that 70% of the protein-DNA contacts occur through these three nucleotides (4, 37). Our analysis also indicated that the nature of the first residues has little effect on affinity. This is in agreement with previous in vivo studies which showed that mutation of this residue from a G to an A reduced viral DNA replication by 5- to 10-fold, whereas mutations at positions 3, 4, and 5 completely abolished it (11). These findings are in good agreement with our in vitro data showing that changing position 1 from a G to an A reduces binding by approximately 5-fold, whereas mutations at positions 3, 4, and 5 lower affinity by 50- to 200-fold. These data also suggest that a relatively small reduction in DNA binding affinity (fivefold) is sufficient to impair DNA replication in vivo. Our binding analysis also showed that position 2 can be mutated without a drastic loss in affinity, although not to a guanine, whose amide group may sterically exclude a water molecule found in the structure of Bochkareva et al. (4) from hydrogen bonding with both the OBD and DNA. Our analysis also suggested that T-ag could interact, albeit more weakly, with two noncanonical sequences, both 5'-TCGGC-3', present in the origin (Fig. 7C). Bochkareva et al. (4) also provided evidence for a noncanonical binding site in the origin, in their case 5'-ATGGC-3'. Whether these noncanonical sequences are of any functional significance will require further experimentation. However, we note that mutation of these two sites independently to 5'-TCGGC-3' (where underlining indicates the mutation) was previously shown to reduce viral DNA replication in vivo by 40%, whereas an identical substitution in P4 (to 5'-GAGGC-3') completely abolished it (11). These results point to these noncanonical sites being of an importance lesser than that of P1 to P4. We also observed that the affinity of the OBD for its target 5'-GAGGC-3' sequence is influenced to some extent by the nucleotide located immediately downstream of it. As a result, the OBD has a higher affinity for P1 and P3 than for P2 and P4 in site II, a finding that may explain in part why double hexamers assemble preferentially on P1 and P3. Mutagenesis of the residue downstream of P3 indicated that a G, as found in P2, is deleterious to binding. Satisfyingly, the structure of Meinke et al. (37) revealed that stabilizing van der Waals interactions occur between the OBD and the nucleotide complementary to that located 3' of the 5'-GAGGC-3' sequence. Finally, we found that the OBD requires additional sequences on each side of the TBS for high-affinity binding and that its interaction with specific and nonspecific DNA is dramatically increased at lower pH and reduced at higher ionic strength. These results are consistent with both crystal structures (4, 37), which showed that the protein spans not only the 5'-GAGGC-3' sequence but also the upstream two and the downstream three nucleotides and that the interaction of the OBD with DNA is stabilized through a number of hydrogen bonds with the phosphate backbone (4, 37). Thus, our analysis has provided experimental evidence for many of the OBD-DNA interactions observed in the crystal structures.

N260 and the OBD bind independently to pairs of pentanucleotide binding sites. We previously reported that the OBD binds to palindromic pairs of TBS, either P1/P3 or P2/P4, independently, suggesting that OBD molecules bound at these
sites do not interact significantly with each other (59). This was recently confirmed by crystal structures of the OBD bound to P1/P3 or to all four TBSs (4, 37). In this study, we have obtained similar results for N260 and for other pairs of binding sites, namely, P2/P3 and P3/P4. We also obtained similar results with mutant proteins carrying amino acid substitutions known or suspected to impair protein-protein interactions important for double hexamer assembly, including changes at Ser120, Ser123, and Thr124, as well as substitutions in the OBD found in class 4 mutants or predicted to disrupt spiral assembly. Since none of them had any deleterious consequence, we surmise that they affect later steps of double hexamer assembly. These results are consistent with those of others indicating that Thr124 and class 4 substitutions, in the context of full-length T-ag, do not affect origin binding and hexamer assembly but rather specifically affect the formation of a double hexamer (2, 64).

Binding of N260 and the OBD to ssDNA. Remarkably, we found that N260 and the OBD bound to ssDNA in a sequence-independent manner with a substantial affinity of 100 to 200 nM (Fig. 9D), approximately 10-fold more tightly than to non-specific dsDNA (∼1 to 2 μM; Fig. 9D) and only 5-fold or less more weakly than to any of the four pentanucleotides in site II (Fig. 6). We also found that binding to ssDNA was mutually exclusive with binding to a specific TBS, in agreement with recent NMR data indicating that many of the same A1 and B2 residues are used for binding ss- and dsDNA (45). Thus, we imagine that the sequence-specific contacts made between the OBD and each pentanucleotide during initiation are eventually lost at the expense of binding to ssDNA regions generated upon double hexamer assembly and melting of the origin.

Sequence specificity of full-length T-ag. Interestingly, we determined that full-length T-ag produced in insect cells with a baculovirus expression system, which allows phosphorylation of the T-ag molecules assemble into double hexamers that are selective for ssDNA. A protein-DNA complexes. 

ACKNOWLEDGMENTS

We thank Daniel Simmons (University of Delaware) for the gift of class 4 mutant DNA, Albert M. Berghuis and Magdalena Korczynska (McGill University) for the use of the DLS instrument and advice on data analysis, and Steve Titolo for advice on FP data analysis. We also thank Andrew Bohm and Gretchen Meinke (Tufts University School of Medicine) for communicating results prior to publication.

REFERENCES


