# DNA Binding and Cleavage by Human Parvovirus B19 NS1 Nuclease Domain

**Funding Source:** Research reported in this publication was supported by the National Science Foundation under the Grant No. MCB1410355, the Office of the Director, National Institutes of Health of the National Institutes of Health under award number S10OD013237, and the National Institute of General Medical Sciences of the National Institutes of Health under award numbers T32GM008659 (to JLS) and linked award numbers RL5GM118969, TL4GM118971, and UL1GM118970 (to AQ). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS, NIH, or NSF.

Jonathan L. Sanchez[1], Zachary Romero[1,2], Angelica Quinones[1,2,3], Kristiane R. Torgeson[1] and Nancy C. Horton[1*]

## Department of Chemistry and Biochemistry, University of Arizona, Tucson, Arizona 85721

**Running title**: DNA binding and cleavage by human parvovirus B19 NS1 nuclease domain

[1]Department of Chemistry and Biochemistry, University of Arizona, Tucson, AZ 85721
[2]Undergraduate Research Opportunities Consortium-minorities Health Disparity Program (UROC-MHD), University of Arizona Graduate College, University of Arizona, Tucson, AZ 85721
[3]BUILDing SCHOLARS program, University of Texas at El Paso, El Paso, TX 79968

***To whom correspondence should be addressed:**
Dr. Nancy C. Horton
Department of Chemistry and Biochemistry
Telephone: (520) 626-3828.
E-mail: nhorton@u.arizona.edu.

**Abbreviations**

AAV, Adeno Associated Virus

AUC, analytical ultracentrifugation

bp, base pair

BPB, bromophenol blue

DTT, dithiothreitol

EDTA, ethylenediaminetetraacetic acid

HEPES, 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid

HEPES-NaOH, HEPES titrated to a desired pH with NaOH

kDa, kilodalton

KO, “knock-out”, where a subset of the DNA sequence is substituted

MVM, Minute Virus of Mice

Mw, molecular weight or molecular mass in Daltons

NS1-nuc, human parvovirus B19 NS1 nuclease domain (residues 2-176)

NSBE, NS1 DNA binding element

nt, nucleotide or nucleotides

OAc, acetate

PAGE, polyacrylamide gel electrophoresis

Rep, a NS1 homologue from AAV, specifically Rep68 or Rep78

SDS, sodium dodecyl sulfate

Tris, Tris(hydroxymethyl)aminomethane

Tris-HCl, Tris titrated to a desired pH with HCl

TSS, transcription start site

v-bar, partial specific volume in ml/g

XCFF, xylene cyanol FF

**ABSTRACT:**

Infection with human parvovirus B19 (B19V) has been associated with a myriad of illnesses, including erythema infectiosum (Fifth disease), hydrops fetalis, arthropathy, hepatitis, cardiomyopathy and also possibly the triggering of any number of different autoimmune diseases. B19V NS1 is a multi-domain protein that plays a critical role in viral replication, with predicted nuclease, helicase, and gene transactivation activities. Herein we investigate the biochemical activities of the nuclease domain (residues 2-176) of B19V NS1 (NS1-nuc) in sequence-specific DNA binding of the viral origin of replication sequences, as well as those of promoter sequences including the viral p6 and the human p21, TNFα, and IL-6 promoters previously identified in NS1-dependent transcriptional transactivation. NS1-nuc was found to bind with high cooperativity and with multiple (5-7) copies to the NS1 binding elements (NSBE) found in the viral origin of replication and the overlapping viral p6 promoter DNA sequence. NS1-nuc was also found to bind cooperatively with at least 3 copies to the GC-rich Sp1 binding sites of the human p21 gene promoter. Only weak or nonspecific binding of NS1-nuc was found to the segments of the TNFα and IL-6 promoters. Cleavage of DNA by NS1-nuc occurred at the expected viral sequence (the terminal resolution site, trs), but only in single stranded DNA, and NS1-nuc was found to covalently attach to the 5' end of the DNA at the cleavage site. Off-target cleavage by NS1-nuc was also identified.

**Key words:** protein-DNA complex, nuclease, DNA binding, parvovirus, viral protein

Human Parvovirus B19 (B19V) is a ubiquitous virus infecting the majority of the human population[1, 2]. B19V, a *Parvoviridae* family member of the genus *Erythrovirus,* so named for its tropism for erythroid precursor cells[3], has been associated with a myriad of different illnesses. It was first discovered as the cause of aplastic crisis in patients with chronic hemolytic anemia[4], then as the causative agent of erythema infectiosum (Fifth disease) in 1983[5] which results in mild fever and a distinctive rash in children, and fever often with hepatitis and arthralgia in adults. B19V infection is also associated with pure red-cell aplasia from persistent infection in immune-compromised patients and hydrops fetalis in pregnant women[1]. In addition, B19V infection has also been associated with several other serious conditions such as inflammatory cardiomyopathy and the induction of autoimmune or autoimmune-like disease (short or long term)[6-9].

B19V is a single stranded nonenveloped DNA virus of 5594 nucleotides, with an internal coding region flanked by palindromic sequences capable of forming terminal hairpin structures (NCBI entry NC_000883.1). The viral genome encodes five known, and a possible sixth, protein products: VP1 and VP2 that compose the viral capsid, NS1, the main replicative protein, and two or three smaller proteins of unclear function (7.5 kDa, 11 kDa, and an open reading frame for the putative X protein)[10-12]. In addition to erythroid precursor cells, which are permissive for viral replication, B19V can infect a number of different cell types, utilizing the globoside or blood group P-antigen as a receptor[13] along with α5β1 integrin[14]. After entry, the viral DNA is shuttled to the nucleus with the host-encoded DNA repair protein Ku80[15], where replication and transcription occur concurrently following initial conversion of the viral DNA to double stranded[16, 17]. Expression is thought to occur utilizing a single viral promoter, with mRNA for the different gene products produced from different splice variants[11, 18, 19]. Binding sites for several cellular transcription factors are found in the viral promoter[20, 21], which can be transactivated by host factors as well as by the viral NS1 protein[20-22].

Viral replication utilizes cellular proteins, proposed to be coordinated by the viral NS1 protein[23-25], and is thought to make use of the terminal hairpins[23, 25] (Fig. 1). Extension of the 3' end allows for replication of the majority of the viral genome, and NS1 is thought to cleave the opposite strand (at the trs, Fig. 1)

producing a new 3' end that can be used to prime synthesis of the remaining viral DNA[24]. B19V NS1 is predicted to contain both nuclease and helicase activities, likely required in B19V replication, and also contains a C-terminal domain that may be involved in the protein's gene transcriptional transactivation activity. In addition to the transactivation of its own viral promoter, p6, NS1 has also been implicated in the transactivation of several host promoters[26-30], some of which induce the S phase of the cell cycle (in permissive cells) for production of host replicative proteins, followed by G2/M stalling, and finally apoptosis for viral release[16, 29]. In addition, the genes for the pro-inflammatory cytokines TNFα and IL-6 have also been shown to be directly transactivated by B19V NS1 and have been related to models of autoimmune disease induction[27, 30].

Despite the prevalence of B19V and its multiple associated illnesses, relatively few biochemical and no structural studies have been reported for B19V NS1. Herein we investigate the DNA binding and cleavage activity of the nuclease domain of human parvovirus B19 NS1 (NS1-nuc). We show that NS1-nuc binds specifically and cooperatively to GC-rich sequences in the human p21 promoter, as well as to those known as NSBE (for NS1 binding elements) located in the viral origin of replication (Ori), which also overlap with the viral p6 promoter. Stoichiometric and sedimentation measurements indicate that three copies of NS1-nuc bind to the GC-rich regions of the p21 promoter, and five to seven to the Ori region containing all 4 NSBE sequences. Only weak or nonspecific binding of NS1-nuc to the sequences derived from the human TNFα and IL-6 promoters was found. We also show specific cleavage at the terminal resolution site (trs) within the Ori (Fig. 1), resulting in covalent attachment of NS1-nuc to the 5' end of the DNA. We find that NS1-nuc cleaves the trs only when single stranded, and with limited sequence specificity. The limited specificity results in off-target DNA cleavage as well, which may pertain to one proposed mechanism of autoimmune disease induction by B19V involving epitope spreading[31].

## Materials and Methods

*Protein preparation*

A synthetic codon optimized gene for NS1 (coding sequence based on Genbank number ABN45789.1) was purchased (Epoch, Inc.) and used to amplify NS1 residues 2-176 (hereafter NS1-nuc), followed by ligation into the pMAL-c4E vector (New England Biolabs, Inc.) using the SacI and EcoRI restriction sites to create the expression vector for MBP-NS1-nuc. This expression construct was further modified to include: 1) a cleavage site for the protease from tobacco etch virus (TEV), inserted between the enterokinase cleavage site and the N terminus of NS1-nuc, 2) the addition of six histidine residues inserted just after the first residue of the fusion protein, and 3) three glycine codons inserted between the TEV protease cut site (TEVcs) and the N terminus of NS1-nuc to create the expression vector for hMBP-TEVcs-GGG-NS1-nuc. The glycine residues were found to be necessary to obtain full protease cleavage of the chimeric protein, but result in four additional amino acid residues (SGGG) on the N terminus of NS1-nuc following cleavage by TEV protease. TEV protease was prepared as described[32].

Purification of NS1-nuc free of the purification tag was prepared using the hMBP-TEVcs-GGG-NS1-nuc construct and Talon resin (Clonetech, Inc.) chromatography, followed by overnight cleavage (at 4ºC) with TEV protease, and finally DEAE and Heparin FPLC (GE, Inc.) to separate NS1-nuc from TEV protease and hMBP. Figure S1 shows the purity of the protein preparation using coomassie stained SDS-PAGE.

*DNA preparation*

The oligonucleotides were prepared synthetically by a commercial source and purified using C18 reverse phase HPLC or polyacrylamide gel electrophoresis and extraction. The concentration was measured spectrophotometrically, with an extinction coefficient calculated from standard values for the nucleotides[33], and where appropriate including that for fluorescein. Equimolar quantities of complementary DNA were annealed by heating to 90°C for 10 minutes at a concentration of 1 mM, followed by slow cooling to room temperature. The different DNA substrates used in binding, sedimentation and cleavage assays are shown in

Figure 2 and Figure S2.

*DNA binding affinity and cooperativity measurements*

The gel shift assay[34] was used to measure the affinity and cooperativity of NS1-nuc binding to DNA. The DNA was 5' end labeled with $^{32}P$ using T4 polynucleotide kinase (Thermo Fisher Scientific Inc.) and $\gamma$-$^{32}P$-ATP (Perkin Elmer, Inc.), then purified using gel filtration (Micro Bio-30 Spin columns, Bio-Rad Laboratories, Inc.). Assays used a constant concentration of 1 nM $^{32}P$-labeled DNA in 20 µl binding buffer: 100 mM Tris (pH 8), 150 mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, and 10% glycerol, and independent incubations were performed with varied concentrations of NS1-nuc (0.1 nM to 15 µM). These concentrations were carefully chosen to give a well-defined binding curve. Native polyacrylamide gel electrophoresis (10% 29:1 acrylamide:bisacrylamide, 89 mM Tris base, 89 mM boric acid, 2 mM EDTA) was used to separate the bound and unbound DNA. Care was taken to prevent heating of the gel, by electrophoresing at 4°C and at low voltage (190 V). Gels were loaded while undergoing electrophoresis at 300 V, and the voltage returned to 190 V five minutes after the loading of the last sample. Gels were then electrophoresed an additional 2-3 hours at $4^0$C. Autoradiography of gels was performed without drying with a phosphor image plate exposed at $4^0$C for 12-17 hours. Densitometry of phosphor image plates was performed with a PharosXF imager (Bio-rad, Inc.), and integration using Image Lab software (Bio-rad, Inc.). The data were fit using the Hill equation[35] and the software Kaleidagraph (Synergy Software):

$$A=A_{min} + (A_{max}-A_{min})[(P)^n/((K_{1/2})^n + (P)^n)]$$

where A is the fraction of DNA that is shifted (calculated from the integrated densities of shifted and unshifted bands) at a given NS1-nuc monomer concentration, $A_{max}$ is the fitted fraction of shifted DNA upon saturation of all binding sites (at very high NS1-nuc concentrations), $A_{min}$ is the fitted fraction of shifted DNA in incubations without added NS1-nuc, P is the concentration of free NS1-nuc (estimated here using the total concentration of NS1-nuc), $K_{1/2}$ is related to the apparent equilibrium dissociation constant $K_d$ ($K_d = (K_{1/2})^n$, and is the concentration of NS1-nuc giving half maximal binding), and n is the Hill

coefficient and a measure of cooperativity in the binding of NS1-nuc to DNA. All measurements were performed in triplicate, and reported as the average ± the standard deviation. Examples of gel images and data fitting are given in Figures S3-S4.

In some cases, as indicated, the equilibrium dissociation constant for NS1-nuc and DNA was also measured using fluorescence polarization[36]. Fluorescein 5' end labeled DNA (labeled on top strand only, Fig. 2) at 1 nM in 2 ml of binding buffer (100 mM Tris-HCl (pH 8.0@RT), 50-150 (as noted) mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, with 10% glycerol) at 4°C was titrated with increasing amounts of NS1-nuc (from 1 nM – 2 μM), and the anisotropy of the emitted fluorescence monitored. Excitation occurred at 494 nm in a PC1 (ISS) fluorimeter with temperature control. The emitted intensities were measured using a 50.8 mm diameter 570 nm cutoff filter with a 580-2750 nm transmittance range (ThermoOriel Inc., catalog no. 59510) and 1 mm slit widths. The anisotropy of the emitted light as a function of added NS1-nuc was fit to the Hill equation as described above for the gel shift assay, with the exception that A is the anisotropy at a given protein concentration, $A_{max}$ is the predicted anisotropy of fully bound DNA, $A_{min}$ is the anisotropy with no NS1-nuc binding.

*Stoichiometric measurements of DNA binding using the gel shift method*

Measurement of the stoichiometry of NS1-nuc binding to various DNAs was performed using the gel shift assay as described above but with 10 μM $^{32}$P-labeled DNA, and NS1-nuc concentrations up to 240 μM. Gels were imaged as described above and the stoichiometry determined from the break or end point in NS1-nuc concentration after which no further binding occurred: the % bound vs. NS1-nuc concentration was fit to two straight lines, one before and one after saturation of binding of the DNA, and their intersection used to determine the concentration of NS1-nuc that fully saturated the 10 μM DNA in the assay. Examples of gel images and data fitting are given in Figures S5-S6.

*DNA Cleavage Assays*

Single turnover kinetic measurements of DNA cleavage were performed using 5′ end $^{32}$P-labeled oligonucleotides substrates (1 nM, unless otherwise noted), under conditions of excess NS1-nuc (1 μM, unless otherwise noted). All reactions were performed at 37°C in 50 mM HEPES-NaOH (pH 7.0@RT), 150 mM NaCl and 10 mM $CoCl_2$ (unless otherwise noted). Aliquots of 5 μl were withdrawn at specific time intervals after mixing NS1-nuc and $^{32}$P-labeled DNA (100 μl total reaction volume), quenched by the addition of 5 μl of quench (80% formamide, 50 mM EDTA, 1 mg/ml XCFF dye, and 1 mg/ml BPB dye), and electrophoresed on denaturing polyacrylamide (20% acrylamide:bisacrylamide (19:1 ratio), 4 M urea, 89 mM Tris base, 89 mM boric acid, 2 mM EDTA) gels. Autoradiography of gels was performed without drying using a phosphor image plate exposed at 4°C for 12-17 hours. Densitometry of phosphor image plates was performed with a PharosXF imager (Bio-rad, Inc.), and integration using Image Lab software (Bio-rad, Inc.). The percent of cleaved DNA product formed as a function of time was determined by integrating the density of both cleaved and uncleaved DNA bands, and normalizing to the total amount cleaved. The fraction of cleaved DNA was then fit to a single exponential function to determine the single turnover rate constant of DNA cleavage using Kaleidagraph (Synergy Software):

$$\text{Percentage of product} = C_1 + C_2 * (1 - e^{-kt})$$

where $C_1$ is a constant fitting the baseline, $C_2$ is the total percent of DNA predicted to be cleaved, k is the observed rate constant ($k_{obs}$), and t is the length of incubation in minutes. Measurements were performed at least three independent times, and presented as the average ± standard deviation.

*5' Attachment of NS1-nuc to DNA*

A 3'fluorescein labeled version of the Ori1-67-top (Fig. 2B, top strand) was purchased (Sigma-Genosys, Inc.) with PAGE purification and used in DNA cleavage assays to determine if NS1-nuc covalently attaches to the 5' end upon DNA cleavage. A reaction utilizing 2.6 μM Ori1-67-top-3'fluorescein DNA, 27 μM NS1-nuc, 10 mM bis-tris propane (pH 9.5), 150 mM NaCl, 8.5

mM $MnCl_2$, in 414 µl total volume was incubated overnight at 37ºC. The reaction was quenched with 100 µl of 5x SDS-PAGE loading buffer (0.4 M Tris-HCl pH 6.8, 0.5 M DTT, 10% SDS, 50% glycerol) followed by heating to 95ºC for 5 minutes, cooling on ice, then electrophoresis on a 12% acrylamide SDS-PAGE gel. Visualization of the fluorescein utilized a Typhoon scanner (GE, Inc.) with the Y520 filter setting (band pass filter, 500-540 nm, centered on 520 nm) and the blue (488 nm) excitation laser. SDS-PAGE gels were also stained with silver stain (staining protein and nucleic acids) and coomassie stain (staining protein only).

*Analytical Ultracentrifugation*

Sedimentation velocity experiments were performed using a Beckman Coulter XL-I instrument with monochromator and interference scanning optics, automated scanning capability, and a Ti-50 rotor. For NS1-nuc without DNA: approximately 400 µl of sample containing 29 µM NS1-nuc was loaded in one of the sectors of the two-sector sedimentation velocity cells. The other sector was loaded with 425 µl of buffer (10 mM Tris-HCl (pH 8.0@RT), 150 mM NaCl, 1 mM EDTA, and 1 mM DTT). The sample was allowed to equilibrate at 4°C for at least 1 hour in the mounted rotor. The sample was centrifuged at 40,000 rpm (115,000xg) and absorbance scans were taken continuously at 280 nm for 10 hours. Sedimentation velocity measurements performed with Flo-NSBE_DNA alone, Flo-p21 dsDNA alone, and mixtures of NS1-nuc with these DNAs were performed similarly but using concentrations of 17 µM (Flo-NSBE_DNA alone), 14 µM (Flo-p21 dsDNA alone), 107 µM and 15 µM (NS1-nuc with Flo-NSBE_DNA, respectively) and 130 µM and 16 µM (NS1-nuc with p21 dsDNA, respectively). Detection utilized the absorption of fluorescein at 495 nm. Data obtained from the scans were fit to a sedimentation coefficient distribution, c(s), using SEDFIT[37, 38]. The viscosity and density of sample buffers were estimated using SEDNTERP[39]. Fitted values of $f/f_0$ were converted to diffusion coefficients and used in the Svedberg equation, along with estimated values from SEDNTERP for partial specific volume (v-bar) to compute a molecular mass (Mw). Calculated partial specific volumes of 0.72 ml/g and 0.55 ml/g were used to calculate Mw from data of NS1-nuc alone

or DNA alone, respectively. However, in the case of protein/DNA mixtures, and since only a single value for v-bar can be used with the c(s) model in SEDFIT, a v-bar of 0.72 ml/g was used and the stoichiometry determined using the fitted Mws for the complexes, NS1-nuc alone, and DNA alone determined via sedimentation using the same value for v-bar (i.e. 0.72 ml/g). An additional analysis was used for the NS1-nuc/DNA data, where the fitting of different frictional ratios (f/f0) was allowed for different s-values (the bimodal fits, Table 2). Table S1 shows the final fitted parameters from SEDFIT. To determine the error limits on the fitted Mw of the NS1-nuc/DNA complexes, the following analysis was performed. First, the sedimentation data were fit using SEDFIT to give the lowest RMSD. Then, the critical RMSD was determined (at 1 sigma level) using the statistical analysis component of SEDFIT, which is the RMSD below which different fits to the data can be considered equivalent, within one sigma or standard deviation (i.e. at a 68% confidence level). Then different values of the frictional ratio, f/f0, were systematically altered and held fixed during SEDFIT fitting of the data. The resulting calculated apparent Mw and RMSD for only those fits (and corresponding f/f0) below the critical RMSD are given in Tables S2-S3. Since one peak bears an apparent Mw equivalent to that of dsDNA, this value was used in the estimate of NS1-nuc/DNA stoichiometry, along with the apparent Mw of NS1-nuc also determined using AUC sedimentation velocity.

**Results**

*Purification of recombinant B19V NS1-nuc*

The nuclease domain (residues 2-176) of B19V NS1 (NS1-nuc) was successfully purified using a synthetic codon optimized (for *E. coli*) gene in *E. coli* and an *E. coli* maltose binding protein (MBP) fusion expression system. Following initial capture of the MBP fusion protein, the NS1 nuclease domain was cleaved free from the MBP domain using recombinant TEV protease and the cleavage site engineered between the MBP and NS1 nuclease domains. In earlier versions of the fusion protein, cleavage was successful in only 50% of the protein preparation, hence several (three) glycine residues were inserted between the cleavage site and the amino terminus of the nuclease sequence. This modification resulted in

complete cleavage of the fusion protein (overnight incubation at 4ºC), and the MBP was subsequently purified from NS1-nuc using ion exchange chromatography. An overloaded lane in the SDS-PAGE gel shown in Figure S1 demonstrates the purity of NS1-nuc using coomassie stain. Purified NS1-nuc was then dialyzed, concentrated, and stored with 50% glycerol in small aliquots at -80ºC until used in assays. Use of buffer at high pH (9.5) greatly reduced solubility problems encountered with the protein, enabling relatively high concentrations (e.g. 10 mg/ml) to be achieved.

*NS1-nuc is monomeric in the absence of DNA*

In order to determine the oligomeric state of NS1-nuc, analytical ultracentrifugation sedimentation velocity was used. Figure 3A shows the resulting c(s) distribution and calculated apparent Mw (22.5 kDa), which is very close to the Mw of 20.1 kDa calculated from the protein sequence (Table 2, Fig. S8, see also Table S1 for fixed and fitted parameters). Hence NS1-nuc is monomeric in the absence of DNA.

*NS1-nuc binding to viral origin of replication DNA*

The affinity of NS1-nuc for DNA containing sequences from the B19V origin of replication (Ori) was measured using the gel shift, and in some cases the fluorescence polarization assay (Materials and Methods). The resulting binding isotherms show sigmoidicity characteristic of cooperative binding (Fig. 4), and therefore the Hill equation was used to fit the date and determine the $K_{1/2}$ and Hill coefficient n (Materials and Methods). The $K_{1/2}$ is a measure of affinity, while the Hill coefficient n is a measure of cooperativity. Table 1 gives the results of triplicate measurements for binding of NS1-nuc to the synthetic Ori containing DNA constructs Ori2-75, NSBE_DNA (and its "knock-outs" or KO)(see Fig. 2 and Fig. S2 for sequences). Ori2-75 is a 75 bp double stranded DNA that contains all four of the NSBE (for NS1 binding elements) sequences, as well as the trs (for terminal resolution site) where NS1 is implicated binding and cleaving, respectively (Fig. 1-2). NSBE_DNA is a shortened form of Ori2-75 containing only the NSBE sequences (Fig. 2). In the case of NSBE_DNA and the KO versions (where each NSBE is replaced with a random sequence, Fig. S2), the stoichiometry of binding was also determined which gives the number of NS1-nuc

bound per DNA. From the data of Table 1, it is apparent that NS1-nuc binds relatively weakly ($K_{1/2}$~1 μM), but highly cooperatively (Hill coefficient n~2-4) to the Ori DNA. In addition, little affinity is lost upon the decrease in DNA sequence from Ori2-75 to NSBE_DNA, which eliminates the trs and surrounding sequences but retains four of the NSBE sequences (Fig. 2B-C). Binding to the NSBE_DNA and KO DNA constructs was measured with both the gel shift and the fluorescence assay (GS and FPA, respectively, Table 1), and show very good agreement, with the exception of a lower $K_{1/2}$ for the NSBE_DNA using FPA (Table 1). In general, the $K_{1/2}$ for the four KO versions of NSBE_DNA tend to be higher (weaker affinity), but retain high cooperativity (Hill coefficient n values >1)(Table 1). The stoichiometry of NS1-nuc binding to NSBE_DNA and its KO sequences was also determined using the gel shift assay, and found to be 7 copies of NS1-nuc bound in the case of NSBE_DNA, and reduced by 1-2 in the case of the different KO sequences (Table 1). This suggests two copies of NS1-nuc bind per NSBE, with the exception of NSBE4, which binds only 1.

*Size of NS1-nuc/viral origin of replication DNA complexes*

Analytical ultracentrifugation (AUC) sedimentation velocity was used to investigate the Mw of the complex formed by NS1-nuc and DNA containing the viral origin of replication sequence. NSBE_DNA dsDNA was purchased with 5' fluorescein (top strand labeled only, Fig. 2A, Flo-NSBE_DNA) to allow detection during ultracentrifugation of only species containing the fluorescein labeled DNA. Figures S9-S12 show the results with Flo-NSBE_DNA dsDNA alone and NS1-nuc mixed with this DNA. In the case of the Flo-NSBE_DNA dsDNA alone, species with apparent Mw of 29 kDa and 13 kDa were determined when the value of 0.55 ml/g for the partial specific volume (v-bar) predicted for this DNA is used in the data analysis (Table 2, Table S1, Fig. S9). The predicted Mw of the duplex form is 27 kDa, which agrees well with the major peak, and the peak with half the size is likely the single stranded DNA (the fluorescein labeled strand). When mixed with NS1-nuc, two peaks were found in the calculated c(s) distribution: 166 kDa and 44 kDa (in order of peak size). When the data were fit using a bimodal model, allowing for different frictional ratios to be fitted for species with different s-values (to accommodate the

possibility of different shapes of the DNA and the NS1-nuc/DNA complexes), apparent Mw of 159 kDa, 92 kDa, and 38 kDa were found (Fig. S12, Table 2, Table S1).

To determine the stoichiometry of NS1-nuc and DNA in the complexes observed in the AUC data, the following procedure was used. First, since analysis of the data with the program SEDFIT required the use of only a single value for the partial specific volume of the macromolecular species (i.e. v-bar), and protein and DNA have very different values of v-bar, a v-bar corresponding to that of NS1-nuc (0.72 ml/g) was used. This required the estimation of the apparent Mw of the DNA when this v-bar is assumed, resulting in a value of 48 kDa (Table 2, Fig. S10). It was also noticed that the apparent Mw of one of the peaks of the c(s) distribution in the NS1-nuc/Flo-NSBE_DNA mixtures was very similar to this value (38-44 kDa, Table 2, also 37-57 kDa in the analysis of Table S2). Hence, stoichiometries were calculated using the apparent Mw of the larger species in the NS1-nuc/DNA mixtures less the apparent Mw of the DNA, and divided by the apparent Mw of NS1-nuc. Using the apparent Mw of 166 kDa for the Mw of the NS1-nuc/DNA complex (Table 2), and a value of 48 kDa for the DNA (Table 2), a stoichiometry of 5.2 NS1-nuc to 1 Flo-NSBE_DNA is found. If the apparent Mw of 44 is used for the DNA instead (as found in the NS1-nuc/Flo-NSBE_DNA c(s) distribution, Table 2), a ratio of 5.4 to 1 is found. Data from the bimodal fit results in stoichiometric ratios of 5.0 to 1 and 5.4 to 1 (Table 2).

To estimate the error limits of the stoichiometric ratios so derived, an analysis investigating the dependence of the apparent Mw (determined by SEDFIT[37]) on different values of the frictional ratio, f/f0, was performed (Table S2). The frictional ratio, f/f0, is a measure of the effect of molecular size and shape on sedimentation, and values significantly greater than 1 indicate deviations from perfect spherical shape and/or from hydration. Sedimentation velocity experiments measure the s-value of sedimenting species, and to estimate a Mw from that s-value, an estimate of f/f0 must first be determined. SEDFIT estimates f/f0 from diffusion dependent processes found in the experimental data[37], thus allowing the estimation of an apparent Mw. SEDFIT also calculates the RMSD for each fitting analysis, as well as a critical RMSD below which any fit can be considered equivalent, within a set error limit. The analyses of Table S2 show results from different fits of the experimental data when f/f0 was held constant at different values, and which

showed RMSD below the critical RMSD (at 1 sigma). The calculated apparent Mw of the largest component of the c(s) distribution ranges from 146 to 209 kDa (Table S2). The peak identified as that of the free dsDNA (Mw3, Table S2) also varies, from 37 to 57 kDa, in apparent Mw. An intermediate sized complex is identified in some of the fits and ranges from 88 to 93 kDa. The stoichiometric ratios shown in Table S2 were calculated using the apparent Mw of the free dsDNA (from the same fit, i.e. Mw3), and a value of 22.5 kDa for the apparent Mw of NS1-nuc (determined via AUC, Table 2). These ratios range from 4.8 to 6.7 copies of NS1-nuc per single copy of Flo-NSBE_DNA. Hence the stoichiometry determined by AUC sedimentation velocity is 5-7 copies of NS1-nuc per Flo-NSBE_DNA for the major NS1-nuc/DNA complex. It should be noted that this estimate assumes only a single copy of the Flo-NSBE_DNA per complex. If two copies of the DNA were present in the complex, 3 copies of NS1-nuc would be predicted to occur in the complex.

*NS1-nuc binding to p21 promoter DNA*

The binding of NS1-nuc to a sequence derived from the human p21 promoter was also investigated for affinity and cooperativity (Table 3). This 53 bp segment of the p21 promoter which was the minimum sequence implicated in gene transactivation by NS1[26] (Fig. 2D), and also contains several Sp1 binding sites, was found to bind to NS1-nuc with a $K_{1/2}$ of 5.7±0.8 μM and with high cooperativity (Hill coefficient 2.3±0.3, Table 3).

*Size of the complex formed from NS1-nuc binding to p21 promoter DNA*

In the case of the fluorescein labeled p21 dsDNA (Flo-p21 dsDNA) alone, the AUC sedimentation velocity data analysis (using the calculated v-bar of 0.55 ml/g) indicate species with apparent Mw of 34 kDa, and 14 kDa (Fig. S13, Table 2, Table S1). The predicted Mw is 32.6 kDa, in good agreement with the estimated Mw of the major peak in the c(s) distribution (Table 2). Analysis of the sedimentation velocity data of a mixture of NS1-nuc and Flo-p21 dsDNA resulted in the determination of species with apparent Mw of 115 kDa, 50 kDa, and 266 kDa (in order of peak size, Table 2, Fig. 3C, Fig. S15). Using the apparent

Mw for the DNA when v-bar of 0.72 ml/g is used (57 kDa, Table 2) and 22.5 KDa for NS1-nuc (Table 2), stoichiometries of 2.6:1 and 9.3:1 NS1-nuc/Flo-p21 dsDNA are found for the two largest species. Using the Mw computed for the peak closest 57 kDa in the NS1-nuc/Flo-p21 dsDNA analysis (50 kDa, Table 2) as that of the free DNA, gives a stoichiometry of 2.9:1 and 9.6:1 for the two largest species. A similar analysis as that described above for the NS1-nuc and Flo-NSBE_DNA AUC data was performed to determine the 1 sigma error boundary limits of the stoichiometric ratios (Table S3), resulting in the ranges 2.8-3.0:1 and 9.6-9.7:1. We interpret these values to indicate stoichiometries of 3:1 and 10:1 NS1-nuc per Flo-p21 dsDNA. The complex with 3:1 is by far the predominant complex, with the 10:1 complex being present at only an estimated 2% of the total DNA bearing species. Alternative stoichiometries of the major complex (the 3:1) are not possible, since the size of the major complex is too small to harbor more than one copy of the dsDNA with any copies of NS1-nuc. In addition, binding to single stranded DNA is not anticipated since such binding was not seen in the gel shift data. (The $K_{1/2}$ of NS1-nuc to either single strand in p21 dsDNA is estimated to be much greater than 50 μM from gel shift data using single stranded p21 DNA, *data not shown*, which is 25 fold weaker than with double stranded p21 DNA. Hence binding would favor the double stranded form, and also very little single stranded DNA was observed in the sedimentation velocity experiment of Flop21 dsDNA alone (Fig. S13).)

*NS1-nuc binding to TNFα promoter, IL-6 promoter, and random DNA sequences*

Finally, NS1-nuc binding to sequences derived from the promoters of the human TNFα and IL-6 genes was investigated (Table 3). Unlike the B19V Ori/p6 and p21 sequences described above, binding of NS1-nuc to the TNFα and IL-6 sequences was much weaker ($K_{1/2}$ of 4±3 and 28±18 μM, respectively, Table 3) and with lower cooperativity (Hill coefficients n between 1 and 2, Table 3). To determine if this binding behavior is non-sequence specific, binding was also measured to a 62 bp duplex DNA containing random sequences of variable %GC (50 and 72.6%, Fig. S2). In both cases the $K_{1/2}$ was found to be ~30 μM, similar to that found for the IL-6 promoter element sequence (Table 3). However, this is 10-fold weaker

than the binding affinity observed to the TNFα sequence. The cooperativity of binding to the random sequences differed, being low for the 50% GC DNA (Hill coefficient of 1.6±0.4, Table 3), but higher for the 72.6% (Hill coefficient of 6±2, Table 3).

*DNA cleavage*

NS1-nuc was used in assays to test for DNA cleavage (Fig. 5-6). The assays made use of 5' $^{32}$P-labeled DNA containing segments of the viral origin of replication sequence (Ori1-67-top, Fig. 2B) where NS1 is presumed to cleave during viral replication (Fig. 1). This 67 bp portion of the B19V genome has been found to be sufficient for NS1 mediated DNA replication[40]. The trs, for terminal resolution site is located 5' to 3-4 copies of a repeat sequence designated as the NS1 Binding Element (NSBE)[22](Fig. 2A-B). The first two NSBE contain exact copies of the sequence GCCGCCGG (two leftmost boxes, NSBE1 and NSBE2, Fig. 2A-C), while the sequence that follows contains GC-rich sequences that differ considerably, but contain smaller (i.e. 4 bp) portions of the repeat sequence (shown as two additional boxed regions, NSBE3 and NSBE4, Fig. 2A-B). To test if NS1-nuc is capable of site specific cleavage at the trs sequence, a synthetic oligonucleotide (Ori1-67-top, Fig. 2B) was 5' end $^{32}$P-labeled and incubated at 1 nM concentration with a high concentration of NS1-nuc (1 μM) in varied buffer conditions. Analysis of overnight incubations at 37°C in buffer containing 50 mM HEPES-NaOH (pH 7.5), 150 mM NaCl, and 10 mM divalent cation ($Ca^{2+}$, $Mg^{2+}$, $Zn^{2+}$, $Co^{2+}$, $Ni^{2+}$, or $Mn^{2+}$) shows the presence of cleavage products with $Mg^{2+}$ (red box, Fig. 5), $Co^{2+}$, $Ni^{2+}$ and $Mn^{2+}$ (Fig. 5). Cleavage with $Mn^{2+}$, $Co^{2+}$, and $Ni^{2+}$ was greater than that with $Mg^{2+}$, but also contained more than one cleavage product (Fig. 5). Buffer conditions were varied and tested in NS1-nuc cleavage assays with the optimum cleavage found to occur at pH 7.0-7.5 with either 10 mM $Mn^{2+}$ or $Co^{2+}$, and insensitive to NaCl concentrations between 0-200 mM (Fig. S7). The cleavage sites with $Mn^{2+}$ as the divalent cation were mapped using size markers containing the Ori sequence (Fig. 5), and the main cleavage site determined to be 18 nt from the 5' end, corresponding the trs site (Fig. 2B), and the minor cleavage site was found to be 30 nt from the 5' end.

In order to determine the minimal recognition sequence for trs cleavage by NS1-nuc, DNA oligonucleotides truncated at either the 5' or 3' end of Ori1-67-top were tested in DNA cleavage assays (Fig. 6). NS1-nuc was found to be capable of cleaving the trs sequence without the NSBE repeats, and even when only 6 or 9 nucleotides remain 3' or 5' of the trs site, respectively (Fig. 6B-D). Shorter oligonucleotides were not tested; instead substitutions around the trs site were investigated to map the importance of the nucleotides surround the trs site in recognition and cleavage by NS1-nuc (see below).

Timed measurements were also performed using conditions with NS1-nuc concentration in excess of the DNA concentration (1 μM NS1-nuc, 1 nM $^{32}$P-labeled DNA), and the production of cleaved DNA (Fig. 7A) fit to a first order rate constant, $k_{obs}$ (Fig. 7). Measurements were performed in triplicate (Table 4) and reveal a relatively slow rate constant compared to the rate constants for DNA cleavage by other nucleases (0.006 $min^{-1}$, for comparison, the type II restriction endonuclease EcoRV cleaves DNA with a single turnover DNA cleavage rate constant[41] of 70 $min^{-1}$), which was found to be slowed with the shorter DNAs Ori-20-top and NUC1 (Table 4, NUC1 contains the Ori1-67-top sequence but only 9 nt 5' and 6 nt 3' of the trs site). Cleavage of NUC1 is approximately 10 fold slower than that of the others tested (Table 4). Interestingly, the cleavage rate constant of the off-target cleavage (at 30 nt from the 5' end of Ori1-67-top, Fig. 2B) is very similar to that at the trs (Table 4).

To test the importance of each nucleotide (nt) around the trs in cleavage by NS1-nuc, a series of 29 nt oligonucleotides containing the Ori-top sequence (Fig. 2A-B) 18 nt 5' and 11 nt 3' of the trs were tested in cleavage assays. Each contained a single nt substitution in one of the six nt on either side of the trs (substituted base in red, Fig. 8, Ori sequence in black), and its cleavage by NS1-nuc compared to that of the unsubstituted Ori sequence. The nt -7 to -9 and +7 to +11 were also substituted collectively and tested for cleavage by NS1-nuc. Assays were performed with 1 μM NS1-nuc and 1 nM $^{32}$P-labeled DNA in 50 mM HEPES-NaOH (pH 7.5), 150 mM NaCl, and 10 mM $MnCl_2$ for 29 hours at 37°C. Analysis of the cleavage rate (when cleavage occurred) gave no systematic pattern with the substitutions (*data not shown*), but the fraction of DNA cleaved by NS1-nuc was greatly affected by some substitutions. Therefore, the fraction of

DNA cleaved, relative to that for the unsubstituted DNA, after a 29 hour incubation was calculated, normalized to that cleaved in the case of the unsubstituted 29mer, and plotted in Figure 8. Substitutions just adjacent to the trs (1 nt 5' and 2 nt 3', i.e. -1, 1, and 2 in Fig. 8) resulted in the most drastic reductions in cleavage. In addition, substitutions in the 5 nt at the 3' end of the 29mer also affected DNA cleavage by NS1-nuc (7, Fig. 8). Therefore, specificity directly at the trs appears to be limited to the sequence A|CC, with the | indicating the cleavage site. The other apparent specificity determinant occurred ≥7 nt away from the trs, and just adjacent to the NSBE sequences (not present in the 29 nt DNA substrate used in the assays). In fact, inclusion of the NSBE sequences (i.e. cleavage of Ori1-67-top), while clearly not required for cleavage by NS1-nuc, did increase the total amount of DNA cleavage to a value of approximately 7 using the same scale as that in Fig. 8 *(data not shown)*.

Figure 9 demonstrates the covalent attachment of NS1-nuc onto the 5'end of DNA after cleavage of the Ori sequence. Approximately 27 μM NS1-nuc was incubated with 2.6 μM Ori1-67-top DNA containing fluorescein at the 3'end, overnight at 37ºC in buffer containing 8.5 mM $Mn^{2+}$. The reaction was quenched with SDS-PAGE loading buffer and samples electrophoresed using SDS-PAGE. Figure 9C shows the result of scanning the gel for fluorescein fluorescence. The fluorescein labeled DNA is visible in the right-most two lanes which correspond to reactions with and without NS1-nuc, respectively. Only the lane with NS1-nuc contains a slower migrating band which we identify as NS1-nuc covalently attached to the DNA (boxed region, Fig. 9C). Figures 9A-B correspond to coomassie and silver staining, respectively, of similarly electrophoresed gels. Coomassie stains protein only, while silver stain stains both protein and nucleic acids. In Figure 9A, the free NS1-nuc is clearly visible only in lanes that contain NS1-nuc, and the lane with the sample incubated with the fluorescein labeled DNA shows the additional band (boxed region, Fig. 9A) running at the same position as that seen in the fluorescein scan (boxed region, Fig. 9C). This region is also identifiable in the silver stained gel (boxed region, Fig. 9B).

## Discussion

*Cooperative binding of NS1-nuc at the viral origin of replication/viral p6 promoter DNA sequences*

Several reports exist in the literature regarding B19V NS1 binding to DNA, either using direct binding analysis or via detection of NS1 mediated promoter transactivation. Among the purported DNA binding sites of B19V NS1 is the sequence found in the viral origin of replication (Ori), which also overlaps with the single B19V promoter (the p6 promoter)[20, 22]. NS1 has been shown to interact specifically with a series of 8 nt GC-rich regions within this region termed NSBE for NS1 binding elements (Fig. 2A-B)[22, 42]. Since the Ori sequences overlap with the p6 sequences, it is possible that these same NSBE are bound by NS1 for both viral replication functions as well as p6 promoter transactivational functions. By analogy with the parvoviruses AAV and MVM, and their Rep/NS1 proteins that share homology with B19V NS1, binding of NS1 at the NSBE sequences is predicted to function in replication by producing DNA cleavage at the trs site just adjacent to the NSBE (Fig. 6A)[43-50]. In p6 promoter transactivation, NS1 may bind to the NSBE boxes and transactivate directly, or alternatively, indirectly through cooperation with other transcription factors[20, 22] as has also been suggested for homologues of B19V NS1[51]. Switching between the two functions, viral replication and promoter transactivation, may involve post-translational modifications of parvoviral NS1 proteins[52-62]. We found that the isolated B19V NS1-nuc, prepared recombinantly from *E. coli* and therefore without any post-translational modifications, was capable of binding to double stranded DNA containing the NSBE sequences, and in the absence of other proteins (Fig. 4A-C, Table 1). NS1-nuc binds relatively weakly ($K_{1/2}$ ~ 1 μM, Table 1), but with high cooperativity (Hill coefficient ~2-4, Table 1) to this DNA. (We use the notation $K_{1/2}$, rather than $K_D$, to denote the midpoint of the transition of the binding isotherm (Fig. 4B-C) since binding is cooperative and with a stoichiometry greater than 1, as described below.) Weak, micromolar binding of an isolated parvoviral NS1 nuclease domain to DNA sequences derived from the viral origin of replication was also observed for the MVM NS1[48]. It should be noted that although binding to the NSBE sequences is relatively weak, it is still 30 fold tighter than binding of NS1-nuc to nonspecific DNA (Table 3, binding to dsDNA of random sequence is ~30 μM $K_{1/2}$). Also, the $K_{1/2}$ of NS1-nuc to the

NSBE containing DNA determined using the fluorescence anisotropy method is lower, indicating tighter binding, than that measured using the gel shift assay (Table 1). This difference is likely due to the differences in two techniques[63]; the gel shift assay involves separation of the bound DNA from unbound in order to measure their relative concentrations, while the fluorescence anisotropy assay does not. Separation of the bound and unbound DNA disrupts their equilibrium, enabling further dissociation of the DNA bound forms, hence potentially leading to an estimate for affinity that is shifted towards a higher $K_{1/2}$ (underestimating the strength of binding). This effect is expected to be greater when binding is weaker, as in the current case. In addition, the presence of the fluorophore itself may influence the binding of a protein to the labeled nucleic acid, typically strengthening it through nonspecific binding of the fluorophore to the protein[63]. Finally, the high cooperativity in DNA binding by NS1-nuc suggests protein-protein interactions on the DNA and/or alterations in the DNA structure that facilitate further binding of NS1-nuc. The $K_{1/2}$ values indicate weak affinity of individual monomeric NS1-nuc for the DNA, however in the context of full length NS1, which is likely oligomeric like its parvoviral homologues[64-70], binding should be much stronger owing to cooperative binding from the multiple NS1-nuc domains found in an oligomeric NS1.

*Five to seven copies of NS1-nuc bind to four NSBE sequences in the viral origin of replication/p6 promoter*

Since the Hill coefficient is greater than 1 (Table 1), multiple cooperative binding of NS1-nuc to the DNA is implicated. Therefore, we also measured the stoichiometry of NS1-nuc binding to the NSBE sequences. Using a gel shift method, a total of 7 copies of NS1-nuc were found to bind per DNA containing all four NSBE sequences (NSBE_DNA, Table 1). In addition, AUC sedimentation velocity was used to measure the sizes of complexes formed with NS1-nuc and fluorescein labeled NSBE_DNA (Table 2, Fig. 3B, Figs. S11-S12, Tables S1-S2). Several peaks were identified in the c(s) distribution (Fig. 3B) with the major being consistent with approximately 5-7 copies of NS1-nuc bound per NSBE_DNA, in agreement with the stoichiometry determined by gel shift, which indicated 7. A second species was also identified, which is consistent with the binding of 2 copies of NS1-nuc bound to this DNA (Table 2).

Therefore the cooperativity observed in binding (i.e. Hill coefficient n greater than 1), stoichiometry,

and AUC data all indicate multiple copies of NS1-nuc binding per NSBE containing DNA. Since NS1-nuc is monomeric in the absence of DNA (Fig. 3), a Hill coefficient above 2 suggests that at least two copies of NS1-nuc bind cooperatively. In some cases, the Hill coefficient is as high as 4, indicating cooperative binding of at least 4 copies of NS1-nuc (Table 1). The finding from the stoichiometry and AUC measurements that five to seven copies of NS1-nuc bind this DNA indicates that the cooperativity while high is not complete. Perfect or complete cooperativity would mean all copies of NS1-nuc bind simultaneously, and that no intermediate binding states exist with less. Instead, the incomplete cooperativity with the Hill coefficient less than five to seven suggests that although binding of NS1-nuc to the DNA favors the binding of additional NS1-nuc, complexes with less than five to seven copies bound to the DNA are formed at subsaturating concentrations of NS1-nuc.

To investigate NS1-nuc binding to the individual NSBE sequences, four "knock-out" (KO) versions of the DNA (Fig. S2B-E), where each NSBE was replaced with random AT sequences, were tested for NS1-nuc binding. These showed somewhat decreased affinity ($K_{1/2}$, Table 1) for NS1-nuc, with variable effects on binding cooperativity (Hill coefficients, Table 1). The largest effect occurred with the KO2 substitution, where the sequences of the second NSBE were substituted (Fig. S2C), and that resulted in a 65% increase in the $K_{1/2}$ indicating weakened binding affinity (Table 1). However, overall, the effects of the KO substitutions on binding affinity and cooperativity are difficult to interpret due to the multiple, cooperative nature of NS1-nuc binding to the DNA. In contrast, the effects of the KO substitutions on the *total number* of NS1-nuc copies bound to the DNA is striking, with reductions in 1-2 copies with each NSBE box substitution (Table 1).

*Model of NSBE binding by NS1-nuc*

Combining the binding affinity, cooperativity, and stoichiometry data, along with the crystal structure of the AAV Rep nuclease domain bound to Ori DNA[71], results in the model for NS1-nuc binding to NSBE shown in Figure 10. In the AAV Rep nuclease domain/DNA x-ray crystal structure, five nuclease domains are bound to the AAV Ori DNA containing the 5 direct RBE repeats[71]. Since each RBE repeat contains 4

bp, yet each NSBE repeat contains 8 bp, it is possible that each NSBE half-site is analogous to each RBE repeat. This would predict the binding of two NS1-nuc domains per NSBE. Indeed our NSBE "knock-out" data (NSBE KO1-KO4 dsDNA, Table 1) do in fact indicate that two NS1-nuc bind per NSBE, with the exception of the fourth NSBE, which appears to bind only one. The fourth NSBE sequence contains one GC-rich half-site and one AT-rich half site (Fig. 10). Each of the other NSBE half-sites are GC-rich, therefore, the model of Figure 10 shows the singular NS1-nuc bound to NSBE4 at the GC-rich half site. This positioning also places it closer to the other DNA bound NS1-nuc, such that favorable protein-protein interactions could potentially occur. In the AAV Rep nuclease domain/DNA structure, the nuclease domains interact sequence specifically with the RBE repeats, mostly with one of the two DNA strands, and therefore spiral around the DNA[71]. The NS1-nuc shown schematically in Figure 10 are drawn to emphasize the relative spacing seen in the AAV Rep nuclease domain/DNA structure, where the closest approaches in three dimensions between proteins are not to those bound at adjacent DNA sequences, but to those 2 binding sites away (blue to blue and red to red, Fig. 10). In addition, since every other NS1-nuc binding site is 10 bp apart, the every other NS1-nuc would be one turn apart and on the same side of the DNA (if the DNA maintains B form conformation). Although the AAV Rep nuclease domains are not in close contact with each other in the crystal structure, the amino acid sequence of B19V NS1-nuc contains several inserts relative to that of AAV Rep, and in locations of the structure that could potentially create protein-protein contacts. Favorable protein-protein interactions between NS1-nuc bound to the NSBE sequences would explain the observed cooperativity in DNA binding (Hill coefficient > 1, Table 1). Finally, rules for sequence specificity at the NSBE half sites are not obvious, but if similar to those for the AAV Rep nuclease domain, could involve a combination of NS1-nuc/DNA contacts, not all of which need be satisfied by any one copy of NS1-nuc, and could derive from more than one copy of the DNA sequence repeat to any one NS1-nuc[71].

Recently, another study was reported that investigated the binding of B19V NS1-nuc to sequences at the viral origin of replication[42]. Similar to the current study, NS1-nuc (residues 4-180) was expressed and purified from a recombinant system in *E. coli*. As in our study, it was found to bind to the region containing

the NSBE sequences, however a major conclusion from that study was that NS1-nuc bound to NSBE1 and NSBE2, but not to NSBE3 or NSBE4. The authors suggest that host factors may bind to NSBE3-4, since NSBE1-3 have been found to be critical for viral replication, and NSBE4 was required only for maximal viral replication[40]. The study also found that the bp at positions 1, 4, 6-8 of NSBE1 or NSBE2 were the most important for sequence specific binding by NS1-nuc, and the authors suggest a model where NSBE1 and NSBE2 each bind one copy of NS1. We also found that NSBE1-2 were important for NS1-nuc binding, however propose based on our stoichiometry data that each NSBE in the first three boxes (i.e. NSBE1-3) binds to two copies of NS1-nuc, and NSBE4 binds a seventh. Also, unlike this previous study, we report the $K_{1/2}$, a measure of affinity, and the Hill coefficient cooperativity factor n for binding the NSBE1-4 and the various NSBE KO by NS1-nuc.

AAV Rep68 (i.e. full length containing the helicase/ATPase domain) has been shown to bind to its Ori with 5 total copies[67], but also to form rings containing 6, 7, or 8 copies[64, 66, 70, 72]. These studies, along with structural studies, have inspired a model for Ori recognition, binding, and manipulation by parvoviral replication proteins wherein binding to the nearby repeats (i.e. NSBE in the case of B19V Ori, RBE in the case of AAV Rep) induces ring formation of the Rep protein around the double stranded DNA. The ring structure then uses its helicase/ATPase activity to cause strand separation at the nearby trs sequence, thereby making the trs competent for cleavage by a nuclease domain of the oligomeric Rep[71] (since AAV Rep nuclease domain requires the trs to be single stranded in order to be cleaved, as we have also observed here for B19V NS1-nuc). Our observation of 5-7 copies of NS1-nuc binding to the NSBE sequences is consistent with this model. Alternatively, a model where binding to DNA in a spiral fashion that induces melting at the trs via supercoiling tension and/or DNA stretching, as has been proposed for the *E. coli* origin of replication by DnaA[73], is an attractive alternative and is also consistent with the current data.

*DNA cleavage by NS1-nuc*

B19V NS1-nuc belongs to the HUH (for histidine-hydrophobic-histidine) superfamily of single stranded DNA nucleases, which are divalent cation dependent, utilize two histidine residues in the active site, as well

as a tyrosine residue that supplies the nucleophile in the phosphodiesterase reaction resulting in a 5'phosphotyrosine complex and a free 3'OH[74]. B19V NS1 is proposed to function in cleaving replicating viral DNA (Fig. 1), necessary to complete the viral genome replication. The DNA cleavage activity of B19V NS1-nuc was tested using DNA containing the expected cleavage site, namely the B19V terminal resolution site or trs (Fig. 1, Fig. 2A-B, Fig. 6), and DNA length and cleavage buffer composition were varied and tested for their effects on DNA cleavage activity of B19V NS1-nuc (Fig. 5-6, Fig. S7). The likely biologically relevant cofactor for the nuclease activity is $Mg^{2+}$, however $Mg^{2+}$ dependent nucleases can often utilize other metal ions and therefore several were tested (Fig. 5). Weak DNA cleavage activity was found with $Mg^{2+}$ (red box in Fig. 5), but more robust activity was found with $Mn^{2+}$, $Co^{2+}$, and $Ni^{2+}$ (Fig. 5). Substitution of $Mn^{2+}$ in place of $Mg^{2+}$ is known to "rescue" the $Mg^{2+}$ dependent nuclease activity of mutant enzymes, doing so by increasing both the cleavage rate as well as the affinity of the enzyme for the DNA substrate[75, 76]. In the case of NS1-nuc, truncation of the nuclease domain from the remainder of the full length protein may compromise its binding affinity for the DNA substrate, which could be partially compensated by using ions such as $Mn^{2+}$, $Co^{2+}$, and $Ni^{2+}$. The metal ion dependence reported for other HUH superfamily endonucleases[74] show DNA cleavage with $Mg^{2+}$, $Mn^{2+}$, and $Ni^{2+}$, but also $Ca^{2+}$, $Zn^{2+}$, and $Cu^{2+}$. However, work with other parvoviral NS1 enzymes[45, 77] has shown that while the full length protein can utilize $Mg^{2+}$, the isolated AAV Rep nuclease domain[77] strongly preferred $Mn^{2+}$, and little activity[78] is found with $Zn^{2+}$ or $Ca^{2+}$, as seen in our assays.

Cleavage assays performed under different pH and salt concentrations indicate that the nuclease activity of NS1-nuc is optimal at pH 7.0-7.5 with 10 mM $Co^{2+}$ or $Mn^{2+}$, and that NaCl concentration between 0-200 mM had little effect on the activity (Fig. S7). Cleavage was also tested with double stranded (*data not shown)* and single stranded versions of the viral Ori sequences (Fig. 2A-B). NS1-nuc only cleaved the DNA when single stranded, consistent with earlier reports with related nucleases of increased nuclease activity on a non-duplexed substrate[71, 74, 79, 80]. The precise cleavage site was mapped to the expected location (i.e. the trs site, Fig. 2A-B) by comparison of the electrophoretic mobility of the product of cleavage with size markers containing the same sequence of the expected product DNA if cleaved at the trs site (18 nucleotides

from the 5' end of Ori1-67-top, Fig. 2B), or one nucleotide shorter or longer (17 and 19 nucleotides, respectively, Fig. 5). The main cleavage product runs at the same position as the 18 nt size marker (Fig. 5), showing clearly that NS1-nuc targets specifically the trs site in the DNA; however at least one minor cleavage product occurs further from the 5' end, discussed below.

To identify the minimal DNA sequences required for recognition and cleavage by NS1-nuc, truncations of the trs containing DNA were tested in cleavage assays (Fig. 6). First, Figure 6A shows the 67 nucleotide single stranded DNA, Ori1-67-top (sequence shown in Fig. 2B), and Figure 6D shows that this DNA is cleaved specifically by NS1-nuc ($2^{nd}$ lane from left, Fig. 6D). Next, a DNA construct lacking the NSBE boxes and containing only 6 nt 3' of the trs (Fig. 6B) was found to be cleaved by NS1-nuc ($4^{th}$ lane from left, Fig. 6D). This indicates that the NSBE sequences are not necessary for recognition and cleavage of the trs site by NS1-nuc in single stranded DNA. A construct with the DNA 5' of the cleavage site reduced to 9 nt was also cleaved by NS1-nuc (Fig. 6C, and right-most lane, Fig. 6D). Together, this data indicate that NS1-nuc can recognize the trs site using all or a subset of the 17 nt (9 nt 5' of the trs and 6 nt 3') around the trs. Further truncations of the DNA were not tested; instead, to investigate which nucleotides were most important for recognition by NS1-nuc, substitutions in a 29 nucleotide DNA containing 18 nt 5' and 11 nt 3' of the trs were next tested for cleavage by NS1-nuc (see below).

*Rate of DNA cleavage by NS1-nuc*

The rate of DNA cleavage by NS1-nuc was also measured using timed cleavage assays (Fig. 7). Fig. 7A shows an autoradiogram of cleavage reaction products formed during a timed reaction, and Fig. 7B shows the data plotted as a function of time. The data fit well to a single exponential function giving an observed rate constant of ~0.006 $min^{-1}$ (Table 4). As can be seen in Table 4, similar rate constants were found for the different trs containing DNA substrates that differ in length, with the exception of the shortest construct, NUC1 DNA (Table 4). In addition, cleavage of an off-target (i.e. not the trs) site in the longer construct, Ori1-67-top, occurred with similar kinetics (Table 4). Although the data fit well to a single exponential function, we hesitate to ascribe this rate constant to the rate of the "chemical step" of DNA cleavage by

NS1-nuc, since the binding affinity of the nuclease domain to these DNA constructs is very weak ($K_{1/2} > 6$ μM measured without divalent cation, *data not shown*, which is similar to the 10 μM concentration of NS1 in the assays), and saturation of binding to the DNA in the assay may be incomplete. It is also interesting that the trs does not appear to be cleaved in the "off-target" cleaved DNA; this off-target cleavage product reaches a maximum value but then does not then decrease to produce the product cleaved at the trs (red, Fig. 7B). This may be due either to loss of affinity (sequence-specific or non-sequence-specific) of NS1-nuc for the DNA missing the nucleotides 3' of the off-target cleavage site (and shortening the number of nt 3' of the trs from 49 to 12), or alternatively to loss of enzyme activity over time, or perhaps both.

*Sequence specificity of DNA cleavage by NS1-nuc*

Since NS1-nuc cleaved relatively specifically at the trs sequence, substitutions around the cleavage site were tested for their effects on this cleavage activity (Fig. 8). The substitutions were made in the context of a 29 nt oligonucleotide containing 18 nt 5' of the trs and 11 nt 3' (see Fig. 2A-B for DNA sequence around the trs). The reaction conditions included 50 mM HEPES-NaOH (pH 7.5), 150 mM NaCl, 10 mM $MnCl_2$, 1 nM $^{32}$P-labeled DNA, and 1 μM NS1-nuc. $Mn^{2+}$ was used in the assays due to the relatively robust cleavage activity it imparts to NS1-nuc (Fig. 5), however it is important to remember than the substitution of $Mn^{2+}$ for $Mg^{2+}$ may also lead to loss of some DNA sequence specificity[81]. Substitutions were made either individually (the 6 nt either side of the trs) or in groups (nts -7, -8, and -9 together and nts +7 to +11 together, Fig. 8), and were made to alter the base identity to its base pairing complement, hence G was changed to C, and A to T, etc. The cleavage reaction was carried out for 29 hours at 37°C, and the fraction of total DNA cleaved was determined for each substituted DNA and normalized to that for the unsubstituted DNA (dotted line, Fig. 8). Most of the substituted oligonucleotides were cleaved to a similar or greater extent than the unsubstituted DNA, with the exception of those substituted directly around the cleavage site, as well as those at the 3' end (Fig. 8). When either the C just 3' of the trs (black sequence, position 1, Fig. 8), or the next nucleotide, also a C (position 2, Fig. 8) were changed to G, cleavage by NS1-nuc fell to nearly

undetectable levels. Substitution of the base just 5' of the cleavage site also resulted in a clear reduction in cleavage by NS1-nuc.

The data show that the DNA cleavage sequence specificity of NS1-nuc for the trs is limited to very few nucleotides around the cleavage site (Fig. 8), at least under the buffer conditions tested, which included $Mn^{2+}$. Substitution of the 5 nt at the 3' end of the 29mer, which are just before the NSBE sequences (absent in the 29mer), also greatly diminishes cleavage by NS1-nuc. This cannot merely be due to changes in non-sequence-specific affinity, since the length of the oligonucleotide does not change in the substituted DNA, merely its sequence. Therefore, although not necessary for cleavage, these sequences have a positive effect on DNA cleavage by NS1-nuc. Similarly, the inclusion of the NSBE sequences, though clearly not necessary for trs recognition and cleavage, greatly increases the percentage of DNA cleavage using this assay (to roughly a value of 7 in the scale of Fig. 8 showing % cleavage/% cleavage of unsubstituted Ori-29-top, *data not shown*). These effects may be a result of increased affinity of NS1-nuc to the DNA through sequence specific (nt +7 to +11 and/or NSBE) or nonsequence-specific interactions (the DNA with the NSBE is longer).

These results differ from those reported for AAV Rep[82], where effects on cleavage by substitutions of 7 nucleotides around the trs (5 nt 5' and 2 nt 3' of the trs) were found. However, the studies with Rep were done with the full length enzyme, rather than the isolated nuclease domain, and also used a double stranded Ori construct containing the full RBE sequences (analogous to the NSBE sequences of B19V). Additionally, the studies with AAV Rep were carried out with $Mg^{2+}$ as the divalent cation cofactor, rather than $Mn^{2+}$ as in the current B19V NS1-nuc studies, and as previously mentioned, $Mn^{2+}$ has been reported to affect sequence specificity in some $Mg^{2+}$ dependent endonucleases[81].

*Covalent attachment of NS1-nuc to DNA*

Previous reports using nuclease enzymes homologous to NS1 show covalent attachment to the 5' end of the DNA at the site of cleavage via an active site tyrosine residue[43, 45, 56, 83-85]. To test for this activity in recombinant B19V NS1-nuc, products of the reaction with 3'fluorophore labeled Ori1-67-top DNA (Fig.

2B) were analyzed using SDS-PAGE with staining for protein (coomassie stain, Fig. 9A), protein and DNA (silver stain, Fig. 9B), or DNA only (fluorescence, Fig. 9C). In Figure 9, the boxed region identifies the inferred position of the NS1-nuc-DNA covalent adduct in these gels. Since the sample was treated with SDS and reducing agent (DTT), and was boiled prior to electrophoresis, no non-covalent complexes should be present. The boxed species migrates more slowly than the protein alone (Fig. 9A-B), as well as the DNA only (Fig. 9C) and contains protein (Fig. 9A) and DNA containing a 3' fluorophore (Fig. 9C), as expected for a covalent complex between NS1-nuc and DNA.

*Off-target DNA cleavage by NS1-nuc*

Cleavage of sequences other than the viral origin of replication is not implicated in any known biological function of NS1, however, cleavage at non-trs sites in the host genome has been implicated in pathogenic activities of B19V and NS1[31, 86]. Such cleavages are damaging to the host genomic DNA, both in terms of the single stranded break, as well as the resulting covalent attachment of NS1 to the DNA. This damage can result in apoptosis and the production of apoptotic bodies that could stimulate the immune system to become autoreactive via the production of anti-DNA antibodies[31, 86]. Such antibodies may also be a gateway to immune reaction via epitope spreading against additional nuclear targets, leading to further loss of self-tolerance[87]. This is one of several mechanisms implicated for the induction of autoimmune disease by B19V[31]. In our DNA substrate, the off-target cleavage was mapped to 30 nucleotides from the 5'end of the Ori1-67-top substrate (Fig. 5), which also corresponds to cleavage just after the first bp of the first NSBE sequence (Fig. 2B). The cleavage sequence at this position is G|CC (| indicating cleavage site), while the sequence at the trs is A|CC. Cleavage at the G|CC sequence is consistent with the results of cleavage of DNA substituted at nt around the trs site (Fig. 8), where the two C nt just 3' of the trs (i.e. the |CC nucleotides) were most important to recognition, followed by the position just 5' of the cleavage site (A| in the trs sequence). While the identity of that base in the off-target cleavage is a G, rather than an A, its identity as a purine may confer some recognition by NS1-nuc (the substitution tested in Fig. 8 at this position, -1, was a T, rather than a G). This off-target site would presumably be masked by NS1 binding at

the first NSBE in double stranded DNA (see above), and if the NS1 nuclease is similar to the homologous AAV Rep nuclease domain, the cleavage active site would be located on a face of the enzyme distinct from the DNA sequence-specific binding face[71]. Interestingly, three other GCC sequences (and one TCC) sequence also occur in Ori1-67-top, but are not cleaved by NS1-nuc in our assays. This suggests other recognition elements exist around the trs and off-target cleavage site, or that these non-cleaved sites are somehow masked, perhaps also by alternative modes of NS1-nuc binding.

*Promoter DNA binding by B19V NS1-nuc*

B19V NS1 has been implicated in promoter transactivation of the B19V p6 promoter[20], as well as those of several host genes (p21[26], TNFα[30], IL-6[27]). The transactivation of TNFα and IL-6 genes by NS1 is another possible mechanism by which B19V could modulate the host immune system and contribute to autoimmune disease[88], while production of p21 cyclin-dependent kinase through promoter transactivation results in cell cycle arrest at G1[26]. In each case, cooperation of NS1 with a host transcription factor has been implicated: Sp1 and possibly others for the p6 promoter[20, 22], Sp1 for the host p21 promoter[26], AP-1 for the TNFα promoter[30], and NF-kB for the IL-6 promoter[27]. Direct interactions with NS1 have been shown only in the case of Sp1[22, 26] and possibly AP-1[30]. In addition, in the cases of the host promoters, the exact DNA sequences necessary and sufficient for transactivation by NS1 have been mapped to those same transcription factor binding sites[26, 27, 30]. In the p6 promoter, the critical sequences were identified as those containing the NSBE, which also coincide with Sp1 binding sites[20]. Therefore the model for NS1 mediated gene transactivation involves direct and/or indirect interaction of NS1 with the transcription factor binding sites of these promoters, as well as the possible interaction of NS1 with the host transcription factors Sp1, AP-1, and NF-κB. To investigate direct interactions of NS1 with the promoter sequences, the implicated DNA sequences were used in binding assays with NS1-nuc. Specific binding (i.e. stronger than that to random DNA sequences) was found between NS1-nuc and DNA sequences of the p6, p21, and TNFα promoters (see Fig. 2 for sequences, Fig. 4 for binding isotherms, and Tables 1 and 3 for binding constants).

Because the implicated NS1 binding sites of the p6 promoter overlap with those of the viral origin of replication[20, 22], the binding study described above with DNA shown in Figure 2A-C relates also to binding at the p6 promoter. As discussed above, binding was found to be relatively weak ($K_{1/2}$ ~ 1 μM, Table 1), but highly cooperative (Hill coefficients of 2-4, Table 1), with five-seven copies of NS1-nuc bound to the p6 promoter derived DNA.

The binding affinity of NS1-nuc to the DNA containing p21 promoter derived sequences was found to have a $K_{1/2}$ in the micromolar range (5.7±0.8 μM, Table 3), and to also be cooperative (Hill coefficients of 2.3±0.2, Table 3). Similarly, the $K_{1/2}$ and Hill coefficient for NS1-nuc binding to the AP-1 site in the TNFα promoter were found to be 4±3 μM and 2±1, respectively (Table 3). In contrast, binding of NS1-nuc to the NF-κB site of the IL-6 promoter DNA was found to be noncooperative and much weaker, and in fact equivalent to that with a random DNA sequence ($K_{1/2}$ of 28±18 μM, Table 3). The weak binding affinity to these three promoter element DNAs precluded stoichiometry measurements using the gel shift method. The observed weak binding suggests that interaction of NS1 with these DNA sequences may require the assistance of other proteins (host or viral), other parts of the NS1 protein( including those implicated in NS1 oligomerization), post-translational modifications[52-58], modification of the DNA[89], or other segments of the promoter sequences. Alternatively, NS1 may induce expression of these genes through some pathway that does not involve NS1 binding at their promoters.

Further investigation of the stoichiometry of NS1-nuc binding to the p21 promoter element DNA was performed using AUC sedimentation velocity, and resulted in the finding of 3 copies of NS1-nuc per fluorescein labeled p21 dsDNA (Fig. 3C, Table 2, Table S1, Table S3, Fig. S15-S16). An additional complex, although much lower in concentration, contained approximately 10 copies of NS1-nuc (Table 2, Table S1, Table S3). The observation of multiple copies of NS1-nuc binding to the p21 promoter element DNA is also consistent with the observed cooperativity in the DNA binding isotherms (Hill coefficient>1, Table 3).

The finding that NS1-nuc protein is capable of direct DNA binding to sequences derived from the p6, and

to a lesser extent p21 and TNFα promoters, indicates that NS1 does not necessarily require host transcription factors for promoter binding. It is even conceivable that NS1 could transactivate these promoters itself, as the C-terminal domain is implicated in transactivational activity in homologous proteins (athough the C-terminal domain sequences of these homologous NS1 proteins do not share significant sequence identity)[90, 91]. The current data does not, however, rule out the possibility that NS1 cooperates with transcription factors at these promoters at some level, either in recruitment to the promoter or promoter transactivation. Further studies will be needed to investigate these potential interactions and activities.

The role of cooperative binding between NS1-nuc at the p6 and p21 promoter sequences is also interesting. This cooperativity could serve to enhance binding affinity, bring NS1 molecules together on the DNA, and/or have some role in altering the architecture of the DNA that could influence the binding or structure of other factors at the promoter, and thereby influence replication or transcriptional events. It is also noteworthy that studies using an NS1 protein with a mutation that results in loss of its ATPase activity show that this activity is necessary for the gene transactivation activity of NS1 in some cases[26], though not all[92]. Similar results have been found with NS1 homologues[49, 91, 93]. Our studies utilize only the nuclease domain of NS1, which is missing the portion of the protein that contains the ATPase activity, however, the involvement of the ATPase activity of NS1 in transactivation does suggest some role in DNA and/or chromatin structural manipulation (or alternatively in a conformational change that may unmask a transactivational region of NS1). Such mechanisms have been suggested for related parvoviral replication proteins[49, 91, 93]. Future studies will be necessary to fully understand the mechanism of gene transactivation by B19V NS1.

**Supporting Information**

Coomassie stained SDS-PAGE of purified NS1-nuc (Figure S1), DNA sequences of additional oligonucleotides used in assays (Figure S2), images of gel shift experiments to measure binding of NS1-nuc to Ori2-75 dsDNA (Figure S3), plots of gel shift data fit to the Hill equation to determine $K_{1/2}$ and the Hill coefficient for binding of NS1-nuc to Ori2-75 dsDNA (Figure S4), Images of gel shift experiments used to

determine the stoichiometry of binding of NS1-nuc to NSBE_DNA dsDNA (Figure S5), plots of gel shift data with fits to determine stoichiometry of binding of NS1-nuc to NSBE_DNA (Figure S6), DNA cleavage experiments to demonstrate pH and salt concentration dependence of cleavage by NS1-nuc (Figure S7), screenshots of data and fits from the software SEDFIT analysis of sedimentation velocity experiments for NS1-nuc (Figure S8), Flo-NSBE_DNA dsDNA (Figure S9-S10), NS1-nuc with Flo-NSBE_DNA dsDNA (Figure S11-S12), Flo-p21 dsDNA (Figure S13-S14), and NS1-nuc with Flo-p21 dsDNA (Figures S15-S16), fitted parameters from SEDFIT analysis of AUC sedimentation velocity data (Table S1), dependence of the stoichiometric ratio on frictional ratio f/f0 for NS1-nuc/Flo-NSBE_DNA AUC sedimentation velocity data (Table S2), and dependence of the stoichiometric ratio on frictional ratio f/f0 for NS1-nuc/Flo-p21 dsDNA sedimentation velocity data (Figure S3).

## References


[1] Young, N. S., and Brown, K. E. (2004) Parvovirus B19, *N Engl J Med 350*, 586-597.

[2] Cossart, Y. E., Field, A. M., Cant, B., and Widdows, D. (1975) Parvovirus-like particles in human sera, *Lancet 1*, 72-73.

[3] Siegl, G., Bates, R. C., Berns, K. I., Carter, B. J., Kelly, D. C., Kurstak, E., and Tattersall, P. (1985) Characteristics and taxonomy of Parvoviridae, *Intervirology 23*, 61-73.

[4] Pattison, J. R., Jones, S. E., Hodgson, J., Davis, L. R., White, J. M., Stroud, C. E., and Murtaza, L. (1981) Parvovirus infections and hypoplastic crisis in sickle-cell anaemia, *Lancet 1*, 664-665.

[5] Anderson, M. J., Jones, S. E., Fisher-Hoch, S. P., Lewis, E., Hall, S. M., Bartlett, C. L., Cohen, B. J., Mortimer, P. P., and Pereira, M. S. (1983) Human parvovirus, the cause of erythema infectiosum (fifth disease)?, *Lancet 1*, 1378.

[6] Tsay, G. J., and Zouali, M. (2006) Unscrambling the role of human parvovirus B19 signaling in systemic autoimmunity, *Biochem Pharmacol 72*, 1453-1459.

[7] Colmegna, I., and Alberts-Grill, N. (2009) Parvovirus B19: its role in chronic arthritis, *Rheum Dis Clin North Am 35*, 95-110.

[8] Franssila, R., and Hedman, K. (2006) Infection and musculoskeletal conditions: Viral causes of arthritis, *Best Pract Res Clin Rheumatol 20*, 1139-1157.

[9] Kerr, J. R. (2015) The role of parvovirus B19 in the pathogenesis of autoimmunity and autoimmune disease, *J Clin Pathol 69*, 279-291.

[10] Cotmore, S. F., McKie, V. C., Anderson, L. J., Astell, C. R., and Tattersall, P. (1986) Identification of the major structural and nonstructural proteins encoded by human parvovirus B19 and mapping of their genes by procaryotic expression of isolated genomic fragments, *J Virol 60*, 548-557.

[11] Ozawa, K., Ayub, J., Hao, Y. S., Kurtzman, G., Shimada, T., and Young, N. (1987) Novel transcription map for the B19 (human) pathogenic parvovirus, *J Virol 61*, 2395-2406.

[12] St Amand, J., Beard, C., Humphries, K., and Astell, C. R. (1991) Analysis of splice junctions and in vitro and in vivo translation potential of the small, abundant B19 parvovirus RNAs, *Virology 183*, 133-142.

[13] Weigel-Kelley, K. A., Yoder, M. C., and Srivastava, A. (2001) Recombinant human parvovirus B19 vectors: erythrocyte P antigen is necessary but not sufficient for successful transduction of human hematopoietic cells,

*J Virol 75*, 4110-4116.

[14] Weigel-Kelley, K. A., Yoder, M. C., and Srivastava, A. (2003) Alpha5beta1 integrin as a cellular coreceptor for human parvovirus B19: requirement of functional activation of beta1 integrin for viral entry, *Blood 102*, 3927-3933.

[15] Munakata, Y., Saito-Ito, T., Kumura-Ishii, K., Huang, J., Kodera, T., Ishii, T., Hirabayashi, Y., Koyanagi, Y., and Sasaki, T. (2005) Ku80 autoantigen as a cellular coreceptor for human parvovirus B19 infection, *Blood 106*, 3449-3456.

[16] Bonvicini, F., Filippone, C., Manaresi, E., Zerbini, M., Musiani, M., and Gallinella, G. (2008) Functional analysis and quantitative determination of the expression profile of human parvovirus B19, *Virology 381*, 168-177.

[17] Bonvicini, F., Filippone, C., Delbarba, S., Manaresi, E., Zerbini, M., Musiani, M., and Gallinella, G. (2006) Parvovirus B19 genome as a single, two-state replicative and transcriptional unit, *Virology 347*, 447-454.

[18] Anderson, L. J., and Young, N. S. (1997) *Human parvovirus B19*, Karger, Basel ; New York.

[19] Yoto, Y., Qiu, J., and Pintel, D. J. (2006) Identification and characterization of two internal cleavage and polyadenylation sites of parvovirus B19 RNA, *J Virol 80*, 1604-1609.

[20] Gareus, R., Gigler, A., Hemauer, A., Leruez-Ville, M., Morinet, F., Wolf, H., and Modrow, S. (1998) Characterization of cis-acting and NS1 protein-responsive elements in the p6 promoter of parvovirus B19, *J Virol 72*, 609-616.

[21] Vassias, I., Hazan, U., Michel, Y., Sawa, C., Handa, H., Gouya, L., and Morinet, F. (1998) Regulation of human B19 parvovirus promoter expression by hGABP (E4TF1) transcription factor, *J Biol Chem 273*, 8287-8293.

[22] Raab, U., Beckenlehner, K., Lowin, T., Niller, H. H., Doyle, S., and Modrow, S. (2002) NS1 protein of parvovirus B19 interacts directly with DNA sequences of the p6 promoter and with the cellular transcription factors Sp1/Sp3, *Virology 293*, 86-93.

[23] Zhi, N., Mills, I. P., Lu, J., Wong, S., Filippone, C., and Brown, K. E. (2006) Molecular and functional analyses of a human parvovirus B19 infectious clone demonstrates essential roles for NS1, VP1, and the 11-kilodalton protein in virus replication and infectivity, *J Virol 80*, 5941-5950.

[24] Ozawa, K., Kurtzman, G., and Young, N. (1986) Replication of the B19 parvovirus in human bone marrow cell cultures, *Science 233*, 883-886.

[25] Berns, K. I. (1990) Parvovirus replication, *Microbiol Rev 54*, 316-329.

[26] Nakashima, A., Morita, E., Saito, S., and Sugamura, K. (2004) Human Parvovirus B19 nonstructural protein transactivates the p21/WAF1 through Sp1, *Virology 329*, 493-504.

[27] Moffatt, S., Tanaka, N., Tada, K., Nose, M., Nakamura, M., Muraoka, O., Hirano, T., and Sugamura, K. (1996) A cytotoxic nonstructural protein, NS1, of human parvovirus B19 induces activation of interleukin-6 gene expression, *J Virol 70*, 8485-8491.

[28] Duechting, A., Tschope, C., Kaiser, H., Lamkemeyer, T., Tanaka, N., Aberle, S., Lang, F., Torresi, J., Kandolf, R., and Bock, C. T. (2008) Human parvovirus B19 NS1 protein modulates inflammatory signaling by activation of STAT3/PIAS3 in human endothelial cells, *J Virol 82*, 7942-7952.

[29] Wan, Z., Zhi, N., Wong, S., Keyvanfar, K., Liu, D., Raghavachari, N., Munson, P. J., Su, S., Malide, D., Kajigaya, S., and Young, N. S. (2010) Human parvovirus B19 causes cell cycle arrest of human erythroid progenitors via deregulation of the E2F family of transcription factors, *J Clin Invest 120*, 3530-3544.

[30] Fu, Y., Ishii, K. K., Munakata, Y., Saitoh, T., Kaku, M., and Sasaki, T. (2002) Regulation of tumor necrosis factor alpha promoter by human parvovirus B19 NS1 through activation of AP-1 and AP-2, *J Virol 76*, 5395-5403.

[31] Poole, B. D., Kivovich, V., Gilbert, L., and Naides, S. J. (2011) Parvovirus B19 nonstructural protein-induced damage of cellular DNA and resultant apoptosis, *Int J Med Sci 8*, 88-96.

[32] Tropea, J. E., Cherry, S., and Waugh, D. S. (2009) Expression and purification of soluble His(6)-tagged TEV protease, *Methods Mol Biol 498*, 297-307.

[33] Fasman, G. D. (1975) *CRC Handbook of Biochemistry and Molecular Biology*, 3rd ed., CRC, Cleveland, OH.

[34] Carey, J. (1991) Gel retardation, *Methods Enzymol 208*, 103-117.

[35] Weiss, J. N. (1997) The Hill equation revisited: uses and misuses, *FASEB J 11*, 835-841.

[36] Park, C. K., Stiteler, A. P., Shah, S., Ghare, M. I., Bitinaite, J., and Horton, N. C. (2010) Activation of DNA cleavage by oligomerization of DNA-bound SgrAI, *Biochemistry 49*, 8818-8830.

[37] Lebowitz, J., Lewis, M. S., and Schuck, P. (2002) Modern analytical ultracentrifugation in protein science: a tutorial review, *Protein Sci 11*, 2067-2079.

[38] Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling, *Biophys J 78*, 1606-1619.

[39] Hayes, D. B., Laue, Y. and Philo, J. (1995-2009) SEDNTERP: Sedimentation Interpretation Program, Version 1.09, University of New Hampshire.

[40] Guan, W., Wong, S., Zhi, N., and Qiu, J. (2009) The Genome of Human Parvovirus B19 Can Replicate in Nonpermissive Cells with the Help of Adenovirus Genes and Produces Infectious Virus, *Journal of Virology 83*, 9541-9553.

[41] Martin, A. M., Sam, M. D., Reich, N. O., and Perona, J. J. (1999) Structural and energetic origins of indirect readout in site-specific DNA cleavage by a restriction endonuclease, *Nature Structural Biology 6*, 269-277.

[42] Tewary, S. K., Zhao, H., Deng, X., Qiu, J., and Tang, L. (2014) The human parvovirus B19 non-structural protein 1 N-terminal domain specifically binds to the origin of replication in the viral DNA, *Virology 449*, 297-303.

[43] Im, D. S., and Muzyczka, N. (1990) The AAV origin binding protein Rep68 is an ATP-dependent site-specific endonuclease with DNA helicase activity, *Cell 61*, 447-457.

[44] Willwand, K., Baldauf, A. Q., Deleu, L., Mumtsidu, E., Costello, E., Beard, P., and Rommelaere, J. (1997) The minute virus of mice (MVM) nonstructural protein NS1 induces nicking of MVM DNA at a unique site of the right-end telomere in both hairpin and duplex conformations in vitro, *J Gen Virol 78 ( Pt 10)*, 2647-2655.

[45] Nuesch, J. P., Cotmore, S. F., and Tattersall, P. (1995) Sequence motifs in the replicator protein of parvovirus MVM essential for nicking and covalent attachment to the viral origin: identification of the linking tyrosine, *Virology 209*, 122-135.

[46] Cotmore, S. F., Christensen, J., Nuesch, J. P., and Tattersall, P. (1995) The NS1 polypeptide of the murine parvovirus minute virus of mice binds to DNA sequences containing the motif [ACCA]2-3, *J Virol 69*, 1652-1660.

[47] Cotmore, S. F., Gottlieb, R. L., and Tattersall, P. (2007) Replication initiator protein NS1 of the parvovirus minute virus of mice binds to modular divergent sites distributed throughout duplex viral DNA, *J Virol 81*, 13015-13027.

[48] Mouw, M., and Pintel, D. J. (1998) Amino acids 16-275 of minute virus of mice NS1 include a domain that specifically binds (ACCA)2-3-containing DNA, *Virology 251*, 123-131.

[49] Christensen, J., Cotmore, S. F., and Tattersall, P. (1995) Minute virus of mice transcriptional activator protein NS1 binds directly to the transactivation region of the viral P38 promoter in a strictly ATP-dependent manner, *J Virol 69*, 5422-5430.

[50] Cotmore, S. F., and Tattersall, P. (1989) A genome-linked copy of the NS-1 polypeptide is located on the outside of infectious parvovirus particles, *J Virol 63*, 3902-3911.

[51] Krady, J. K., and Ward, D. C. (1995) Transcriptional activation by the parvoviral nonstructural protein NS-1 is mediated via a direct interaction with Sp1, *Mol Cell Biol 15*, 524-533.

[52] Nuesch, J. P., Corbau, R., Tattersall, P., and Rommelaere, J. (1998) Biochemical activities of minute virus of mice nonstructural protein NS1 are modulated In vitro by the phosphorylation state of the polypeptide, *J Virol 72*, 8002-8012.

[53] Nuesch, J. P., Dettwiler, S., Corbau, R., and Rommelaere, J. (1998) Replicative functions of minute virus of mice NS1 protein are regulated in vitro by phosphorylation through protein kinase C, *J Virol 72*, 9966-9977.

[54] Narasimhan, D., Collaco, R., Kalman-Maltese, V., and Trempe, J. P. (2002) Hyper-phosphorylation of the adeno-associated virus Rep78 protein inhibits terminal repeat binding and helicase activity, *Biochim Biophys Acta 1576*, 298-305.

[55] Chiorini, J. A., Weitzman, M. D., Owens, R. A., Urcelay, E., Safer, B., and Kotin, R. M. (1994) Biologically active Rep proteins of adeno-associated virus type 2 produced as fusion proteins in Escherichia coli, *J Virol 68*, 797-804.

[56] Han, Y., Wang, Q., Qiu, Y., Wu, W., He, H., Zhang, J., Hu, Y., and Zhou, X. (2013) Periplaneta fuliginosa densovirus nonstructural protein NS1 contains an endonuclease activity that is regulated by its phosphorylation, *Virology 437*, 1-11.

[57] Dettwiler, S., Rommelaere, J., and Nuesch, J. P. (1999) DNA unwinding functions of minute virus of mice NS1 protein are modulated specifically by the lambda isoform of protein kinase C, *J Virol 73*, 7410-7420.

[58] Nuesch, J. P., Lachmann, S., Corbau, R., and Rommelaere, J. (2003) Regulation of minute virus of mice NS1 replicative functions by atypical PKClambda in vivo, *J Virol 77*, 433-442.

[59] Farris, K. D., Fasina, O., Sukhu, L., Li, L., and Pintel, D. J. (2010) Adeno-associated virus small rep proteins are modified with at least two types of polyubiquitination, *J Virol 84*, 1206-1211.

[60] Sukhu, L., and Pintel, D. (2011) The large Rep protein of adeno-associated virus type 2 is polyubiquitinated, *J Gen Virol 92*, 2792-2796.

[61] Collaco, R., Prasad, K. M., and Trempe, J. P. (1997) Phosphorylation of the adeno-associated virus replication proteins, *Virology 232*, 332-336.

[62] Weger, S., Hammer, E., and Heilbronn, R. (2004) SUMO-1 modification regulates the protein stability of the large regulatory protein Rep78 of adeno associated virus type 2 (AAV-2), *Virology 330*, 284-294.

[63] Anderson, B. J., Larkin, C., Guja, K., and Schildbach, J. F. (2008) Using fluorophore-labeled oligonucleotides to measure affinities of protein-DNA interactions, *Methods Enzymol 450*, 253-272.

[64] Maggin, J. E., James, J. A., Chappie, J. S., Dyda, F., and Hickman, A. B. (2011) The amino acid linker between the endonuclease and helicase domains of adeno-associated virus type 5 Rep plays a critical role in DNA-dependent oligomerization, *J Virol 86*, 3337-3346.

[65] Zarate-Perez, F., Bardelli, M., Burgner, J. W., 2nd, Villamil-Jarauta, M., Das, K., Kekilli, D., Mansilla-Soto, J., Linden, R. M., and Escalante, C. R. (2012) The interdomain linker of AAV-2 Rep68 is an integral part of its oligomerization domain: role of a conserved SF3 helicase residue in oligomerization, *PLoS Pathog 8*, e1002764.

[66] Zarate-Perez, F., Mansilla-Soto, J., Bardelli, M., Burgner, J. W., 2nd, Villamil-Jarauta, M., Kekilli, D., Samso, M., Linden, R. M., and Escalante, C. R. (2013) Oligomeric properties of adeno-associated virus Rep68 reflect its multifunctionality, *J Virol 87*, 1232-1241.

[67] Mansilla-Soto, J., Yoon-Robarts, M., Rice, W. J., Arya, S., Escalante, C. R., and Linden, R. M. (2009) DNA structure modulates the oligomerization properties of the AAV initiator protein Rep68, *PLoS Pathog 5*, e1000513.

[68] Pujol, A., Deleu, L., Nuesch, J. P., Cziepluch, C., Jauniaux, J. C., and Rommelaere, J. (1997) Inhibition of parvovirus minute virus of mice replication by a peptide involved in the oligomerization of nonstructural protein NS1, *J Virol 71*, 7393-7403.

[69] Yang, B., Zhang, J., Cai, D., Li, D., Chen, W., Jiang, H., and Hu, Y. (2006) Biochemical characterization of Periplaneta fuliginosa densovirus non-structural protein NS1, *Biochem Biophys Res Commun 342*, 1188-1196.

[70] Smith, R. H., Spano, A. J., and Kotin, R. M. (1997) The Rep78 gene product of adeno-associated virus (AAV) self-associates to form a hexameric complex in the presence of AAV ori sequences, *J Virol 71*, 4461-4471.

[71] Hickman, A. B., Ronning, D. R., Perez, Z. N., Kotin, R. M., and Dyda, F. (2004) The nuclease domain of adeno-associated virus rep coordinates replication initiation using two distinct DNA recognition interfaces, *Molecular cell 13*, 403-414.

[72] Musayev, F. N., Zarate-Perez, F., Bishop, C., Burgner, J. W., 2nd, and Escalante, C. R. (2015) Structural Insights into the Assembly of the Adeno-associated Virus Type 2 Rep68 Protein on the Integration Site AAVS1, *J Biol*

*Chem 290*, 27487-27499.

[73] Duderstadt, K. E., Chuang, K., and Berger, J. M. (2011) DNA stretching by bacterial initiators promotes replication origin opening, *Nature 478*, 209-213.

[74] Chandler, M., de la Cruz, F., Dyda, F., Hickman, A. B., Moncalian, G., and Ton-Hoang, B. (2013) Breaking and joining single-stranded DNA: the HUH endonuclease superfamily, *Nat Rev Microbiol 11*, 525-538.

[75] Sam, M. D., and Perona, J. J. (1999) Mn2+-dependent catalysis by restriction enzymes: pre-steady-state analysis of EcoRV endonuclease reveals burst kinetics and the origins of reduced ativity, *J. Am. Chem. Soc. 121*, 1444-1447.

[76] Little, E. J., Babic, A. C., and Horton, N. C. (2008) Early interrogation and recognition of DNA sequence by indirect readout, *Structure 16*, 1828-1837.

[77] Yoon, M., Smith, D. H., Ward, P., Medrano, F. J., Aggarwal, A. K., and Linden, R. M. (2001) Amino-terminal domain exchange redirects origin-specific interactions of adeno-associated virus rep78 in vitro, *J Virol 75*, 3230-3239.

[78] Yoon-Robarts, M., and Linden, R. M. (2003) Identification of active site residues of the adeno-associated virus type 2 Rep endonuclease, *J Biol Chem 278*, 4912-4918.

[79] Brister, J. R., and Muzyczka, N. (2000) Mechanism of Rep-mediated adeno-associated virus origin nicking, *J Virol 74*, 7762-7771.

[80] Hickman, A. B., Ronning, D. R., Kotin, R. M., and Dyda, F. (2002) Structural unity among viral origin binding proteins: crystal structure of the nuclease domain of adeno-associated virus Rep, *Molecular cell 10*, 327-337.

[81] Hsu, M., and Berg, P. (1978) Altering the specificity of restriction endonuclease: effect of replacing Mg2+ with Mn2+, *Biochemistry 17*, 131-138.

[82] Brister, J. R., and Muzyczka, N. (1999) Rep-mediated nicking of the adeno-associated virus origin requires two biochemical activities, DNA helicase activity and transesterification, *J Virol 73*, 9325-9336.

[83] Snyder, R. O., Im, D. S., and Muzyczka, N. (1990) Evidence for covalent attachment of the adeno-associated virus (AAV) rep protein to the ends of the AAV genome, *J Virol 64*, 6204-6213.

[84] Cotmore, S. F., and Tattersall, P. (1988) The NS-1 polypeptide of minute virus of mice is covalently attached to the 5' termini of duplex replicative-form DNA and progeny single strands, *J Virol 62*, 851-860.

[85] Baldauf, A. Q., Willwand, K., Mumtsidu, E., Nuesch, J. P., and Rommelaere, J. (1997) Specific initiation of

replication at the right-end telomere of the closed species of minute virus of mice replicative-form DNA, *J Virol 71*, 971-980.

[86] Poole, B. D., Zhou, J., Grote, A., Schiffenbauer, A., and Naides, S. J. (2006) Apoptosis of liver-derived cells induced by parvovirus B19 nonstructural protein, *J Virol 80*, 4114-4121.

[87] Craft, J., and Fatenejad, S. (1997) Self antigens and epitope spreading in systemic autoimmunity, *Arthritis Rheum 40*, 1374-1382.

[88] Lehmann, H. W., von Landenberg, P., and Modrow, S. (2003) Parvovirus B19 infection and autoimmune disease, *Autoimmun Rev 2*, 218-223.

[89] Bonvicini, F., Manaresi, E., Di Furio, F., De Falco, L., and Gallinella, G. (2012) Parvovirus b19 DNA CpG dinucleotide methylation and epigenetic regulation of viral expression, *PLoS One 7*, e33316.

[90] de Souza, R. F., Iyer, L. M., and Aravind, L. (2010) Diversity and evolution of chromatin proteins encoded by DNA viruses, *Biochim Biophys Acta 1799*, 302-318.

[91] Legendre, D., and Rommelaere, J. (1994) Targeting of promoters for trans activation by a carboxy-terminal domain of the NS-1 protein of the parvovirus minute virus of mice, *J Virol 68*, 7974-7985.

[92] Moffatt, S., Yaegashi, N., Tada, K., Tanaka, N., and Sugamura, K. (1998) Human parvovirus B19 nonstructural (NS1) protein induces apoptosis in erythroid lineage cells, *J Virol 72*, 3018-3028.

[93] Li, X., and Rhode, S. L., 3rd. (1990) Mutation of lysine 405 to serine in the parvovirus H-1 NS1 abolishes its functions for viral DNA replication, late promoter trans activation, and cytotoxicity, *J Virol 64*, 4654-4660.

## Tables

**Table 1.** Fitted binding constants and stoichiometries for NS1-nuc with DNA sequences derived from the B19V origin of replication/p6 promoter.

| DNA | $K_{1/2}$ (μM) | Hill Coefficient n | Stoichiometry | Method[a] |
|---|---|---|---|---|
| Ori2-75 dsDNA | 0.937±0.015 | 4.4±0.8 | ND[b] | GS |
| NSBE_DNA dsDNA | 0.930±0.05 | 2.60±0.13 | 6.94±0.08 | GS |
| Flo-NSBE_DNA dsDNA | 0.61±0.14 | 3.1±0.9 | ND[b] | FPA |
| NSBE KO1 dsDNA | 1.09±0.14 | 3.4±0.5 | 5.27±0.08 | GS |
| NSBE KO2 dsDNA | 1.58±0.18 | 2.7±0.3 | 5.6±0.4 | GS |
| NSBE KO3 dsDNA | 1.39±0.02 | 3.7±0.6 | 5.3±0.1 | GS |
| NSBE KO4 dsDNA | 1.27±0.12 | 3.0±0.19 | 6.0±0.4 | GS |

[a]Method used to measure binding: GS, gel shift assay, FPA, fluorescence polarization assay.
[b]ND, not determined.

**Table 2.** AUC sedimentation velocity results with NS1-nuc and fluorescein labeled DNA

| Sample | V-bar (partial specific volume, ml/g) | Apparent Mw (peak size) | Interpretation |
|---|---|---|---|
| NS1-nuc | 0.72 | 22.5 kDa (95%) | Monomeric (20.1 kDa theoretical) |
| Flo-NSBE_DNA | 0.55 | 29 kDa (89%)<br>13 kDa (10%) | dsDNA (27 kDa theoretical)<br>ssDNA (14 kDa theoretical) |
| | 0.72 | 48 kDa (90.5%)<br>21 kDa (8.8%) | N/A[a] |
| NS1-nuc/ Flo-NSBE_DNA | 0.72 | 166 kDa (65%)<br>44 kDa (23%) | 5.4:1 NS1-nuc:dsDNA<br>dsDNA |
| NS1-nuc/Flo-NSBE_DNA Bimodal fit | 0.72 | 159 kDa (67%)<br>92 kDa (7%)<br>38 kDa (23%) | 5.3:1 NS1-nuc:dsDNA<br>2.4:1 NS1-nuc:dsDNA<br>dsDNA |
| p21 dsDNA | 0.55 | 34 kDa (82%)<br>14 kDa (6%) | dsDNA (33 kDa theoretical)<br>ssDNA (16 kDa theoretical) |
| | 0.72 | 57 kDa (81%)<br>23 kDa (6%) | N/A[a] |
| NS1-nuc/ Flo-p21 dsDNA | 0.72 | 115 kDa (83%)<br>50 kDa (13%)<br>266 kDa (2%) | 2.9:1 NS1-nuc:dsDNA<br>dsDNA<br>9.6:1 NS1-nuc:dsDNA |
| NS1-nuc/ Flo-p21 dsDNA bimodal fit | 0.72 | 110 kDa (76%)<br>51 kDa (12%)<br>266 kDa (2%) | 2.6:1 NS1-nuc:dsDNA<br>dsDNA<br>9.6:1 NS1-nuc:dsDNA |

[a]Not Applicable.

**Table 3.** Fitted binding constants for NS1-nuc binding to DNA sequences derived from the p21, TNFα and IL-6 promoters, as well as random sequence DNA, using the gel shift method.

| DNA | $K_{1/2}$ (μM) | Hill Coefficient n |
|---|---|---|
| p21 dsDNA | 5.7±0.8 | 2.3±0.2 |
| TNFα AP-1 43 bp dsDNA | 4±3 | 2±1 |
| IL-6 NFkB 43 bp dsDNA | 28±18 | 1.32±0.14 |
| Random 62 bp with 72.6% GC content | 30±4 | 6±2 |
| Random 62 bp with 50% GC content | 33±6 | 1.6±0.4 |

**Table 4.** Rate constants for DNA cleavage by NS1-nuc.

| DNA[a] | $k_{obs}$ ($min^{-1}$)[b] | Divalent Cation[c] | Cleavage site |
|---|---|---|---|
| Ori1-67-top | $2.3\pm0.4\times10^{-3}$ | $Co^{2+}$ | trs site |
| Ori1-67-top | $5.5\pm1.2\times10^{-3}$ | $Mn^{2+}$ | trs site |
| Ori1-67-top | $5.9\pm0.8\times10^{-3}$ | $Mn^{2+}$ | Off-target site |
| Ori1-67-top | $1.48\pm0.17\times10^{-3}$ | $Ni^{2+}$ | trs site |
| Ori-24-top | $3.2\pm1.5\times10^{-3}$ | $Co^{2+}$ | trs site |
| Ori-24-top | $1.32\pm0.17\times10^{-3}$ | $Ni^{2+}$ | trs site |
| Ori-20-top | $1.23\pm0.4\times10^{-3}$ | $Co^{2+}$ | trs site |
| Ori-20-top | $4.1\pm0.\times10^{-4}$ | $Mn^{2+}$ | trs site |
| NUC1[d] | $5.4\pm2.7\times10^{-4}$ | $Mn^{2+}$ | trs site |
| NUC1[d] | $2.26\pm0.5\times10^{-5}$ | $Ni^{2+}$ | trs site |

[a]See Figure 2, Figure 6, and Figure S2 for DNA sequences.
[b]$k_{obs}$ given as the average of three independent trials ± the standard deviation.
[c]Reaction conditions: 1 nM DNA, 1 μM NS1-nuc (unless otherwise listed), 50 mM HEPES-NaOH (pH 7 ($Co^{2+}$) or 7.5 ($Mn^{2+}$, $Ni^{2+}$)), 150 mM NaCl, 10 mM Divalent cation (as listed), 37ºC overnight.
[d]NUC1 contains the Ori sequence (see Ori1-67-top, Fig. 2B) with only 9 nt 5' and 6 nt 3' of the trs.

**Figure legends**

**Figure 1.** Hairpin primed replication mechanism of B19V. Dotted lines represent the newly synthesized DNA, trs, terminal resolution site (predicted NS1 cleavage site).

**Figure 2.** DNA sequences used in binding and cleavage assays. Boxed regions (NSBE1-4) are implicated in NS1 or transcription factor binding. ITR, inverted terminal repeat or sequences that form the fold-over hairpin structures at either end of the B19V genome.

**Figure 3.** Sedimentation velocity analytical ultracentrifugation of NS1-nuc. The c(s) distribution calculated for the NS1-nuc sedimentation velocity data using SEDFIT, with the apparent Mw determined from the fit given above the c(s) peak. The theoretical Mw determined from the protein sequence is 20.1 kDa. **B**. As in A, but NS1-nuc/NSBE_DNA data. **C.** As in A, but the NS1-nuc/p21 dsDNA data.

**Figure 4**. DNA binding by NS1-nuc. **A.** Titration of 1 nM $^{32}$P-labeled NSBE_DNA dsDNA and varied NS1-nuc protein (concentration increasing left to right) using the gel shift assay (100 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, and 10% glycerol, 4ºC). **B.** Fit (line) to integrated densities (filled circles) from A and the Hill equation giving $K_{1/2}$=930 nM and n=2.8. **C.** Binding data (filled circles) and fit using Hill equation (line) for fluorescein labeled NSBE_DNA (100 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, and 10% glycerol, 4ºC) using fluorescence polarization anisotropy. Fit gives $K_{1/2}$=550 nM and n=3.6.

**Figure 5**. Cleavage of viral origin of replication DNA. Cleavage reactions of NS1-nuc and $^{32}$P-labeled Ori1-67-top, with 10 mM metal salts as indicated and size markers from 17-19 and 28-30 nt. The size in nt of the main cleavage products is given in red adjacent to dotted lines indicating the migration pattern of the relevant size markers. The red box indicates a band corresponding to the 18 nt cleavage product expected from cleavage at the trs with $MgCl_2$. All reactions performed in buffer (50 mM HEPES-NaOH (pH 7.5), 150 mM NaCl) at 37°C for 18 hours.

**Figure 6.** DNA cleavage and binding assays with isolated NS1-nuc. **A-C**. Constructs based on the viral origin of replication (Ori1-top-67, Fig. 2B) with different degrees of truncation. trs=cleavage site. Grey boxes indicate putative NS1 binding sites (NSBE1-4), arrow indicates cut site (trs), and numbers indicate

nucleotides between cut site and ends. **D.** Denaturing PAGE of 5'end labeled DNA substrates from **A**, **B**, and **C** before and after reaction with NS1-nuc (overnight at 37°, 60 mM HEPES-NaOH (pH 7.5), 10 mM $CoCl_2$, 1 nM DNA, 1 μM NS1-nuc). Numbers indicate migration positions of different lengths of DNA. Asterisk indicates off-target (i.e. not at trs) cleavage.

**Figure 7.** Timed cleavage assay using Ori1-top-67 DNA. **A.** Image of denaturing gel showing increase in cleaved DNA with increasing time. **B.** Integrated band intensities (points) and fits (lines) to data derived from **A** with cleavage at the trs in blue and blue filled squares, and the off-target cleavage in red and red filled circles. Fits utilized the single exponential function described in Materials and Methods.

**Figure 8**. Plot of percent cleavage of DNA sequences containing single site substitutions (red) around the trs site within the Ori-29-top ssDNA. Cleavage occurs between -1 and 1. The wild type sequence is shown in black. Data are normalized to cleavage of unsubsituted Ori-29-top (a 29 nt single stranded oligonucleotide with 18 nt 5' and 11 nt 3' of the trs, see Fig. 2A for sequence around trs). Reactions performed with 1 μM NS1-nuc, 1 nM DNA, 50 mM HEPES-NaOH (pH 7.5), 150 mM NaCl, and 10 mM $MnCl_2$ at 37°C for 29 hours.

**Figure 9.** Covalent attachment of NS1-nuc to 3'fluorescein labeled Ori1-top-67 following cleavage analyzed on 12% SDS-PAGE. **A**. Coomassie stain, **B**. Silver stain, **C.** Fluorescein scan.

**Figure 10.** Model for NS1-nuc binding to the NSBE sequences in the B19V Ori DNA.

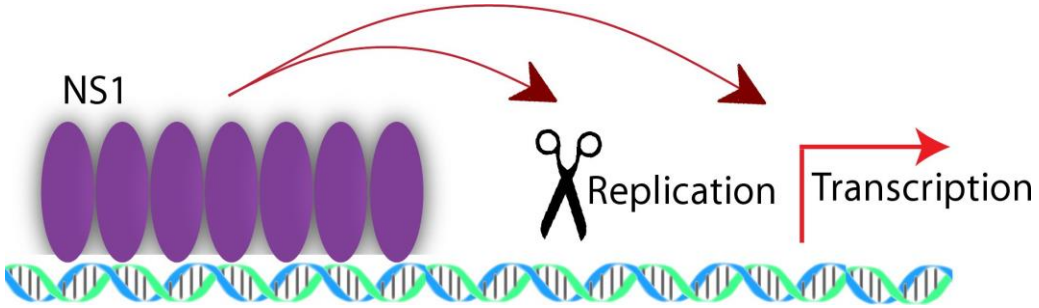
**For Table of Contents Use Only**

**DNA Binding and Cleavage by Human Parvovirus B19**

**NS1 Nuclease Domain**

**Jonathan L. Sanchez, Zachary Romero, Angelica Quinones, Kristiane R. Torgeson and**

**Nancy C. Horton**



**Supporting Information for**

# DNA Binding and Cleavage by Human Parvovirus B19 NS1 Nuclease Domain

**Jonathan Sanchez, Zachary Romero, Angelica Quinones, Kristiane R. Torgeson and Nancy C. Horton**

Figure S1 contains an image of a coomassie stained SDS-PAGE gel showing the purity of the purified NS1-nuc protein. Figure S2 shows the sequences of DNA used in the binding and stoichiometry assays. Figures S3-S4 show gel images and data fitting, respectively, from gel shift experiments used to measure binding affinity and cooperativity. Figures S5-S6 show gel images and data fitting, respectively, for stoichiometry of binding measurements made using the gel shift assay. Figure S7 shows an image of an autoradiogram of urea-PAGE gels used to separate and visualize the cleavage products of reactions of NS1-nuc with DNA containing the trs (terminal resolution site, the cleavage site within the viral origin of replication). These reactions were performed in various buffer conditions to test their effect on DNA cleavage by NS1-nuc. Figures S8-S16 are screen shots from data analysis of sedimentation velocity AUC studies using the SEDFIT software. Table S1 provides the final fitted parameters from analysis of AUC sedimentation velocity data using SEDFIT. Tables S2-S3 provide fitted and fixed parameters from SEDFIT analysis of the AUC sedimentation velocity data for the NS1-nuc/DNA mixtures to determine the range or error of the computed stoichiometric ratios of NS1-nuc/DNA. In this analysis, the critical RMSD value (at 1 sigma) was determined using the best fit of the data, then the f/f0 is fixed to different values and the fitted RMSD and Mw determined using SEDFIT. Only fits with RMSD below the critical RMSD (indicating equivalent fitted parameters within 1 sigma error) are shown.

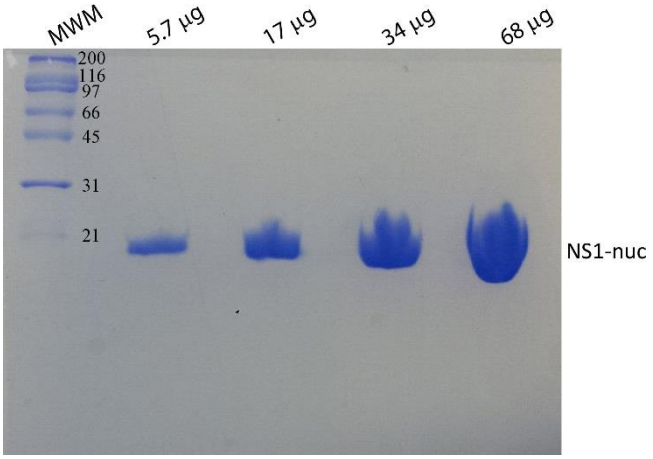



**Figure S1.** Coomassie stained gel of purified recombinant NS1-nuc (size marker, MWM, in leftmost lane with sizes in kDa as marked, and quantities of NS1-nuc given above for each lane). Migration of NS1-nuc indicated to the right.

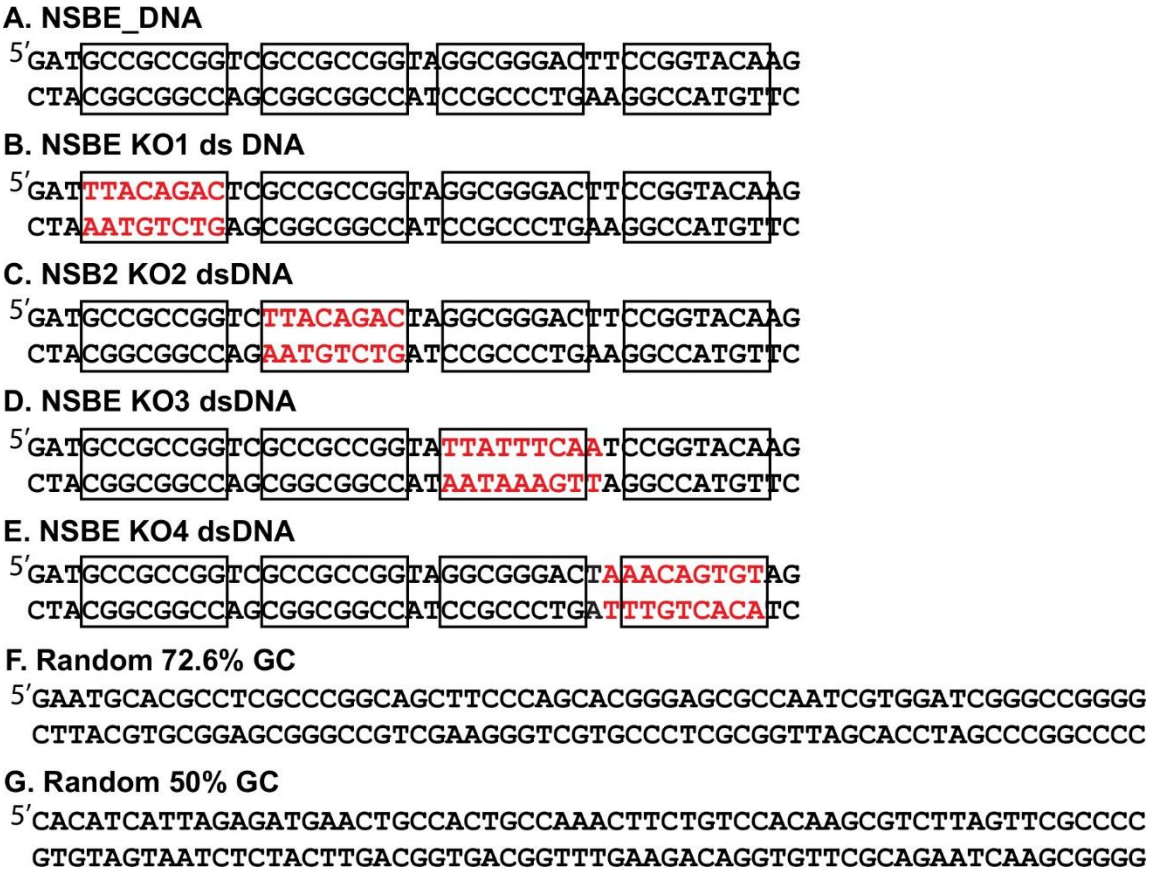



**Figure S2.** Additional DNAs for binding analyses.

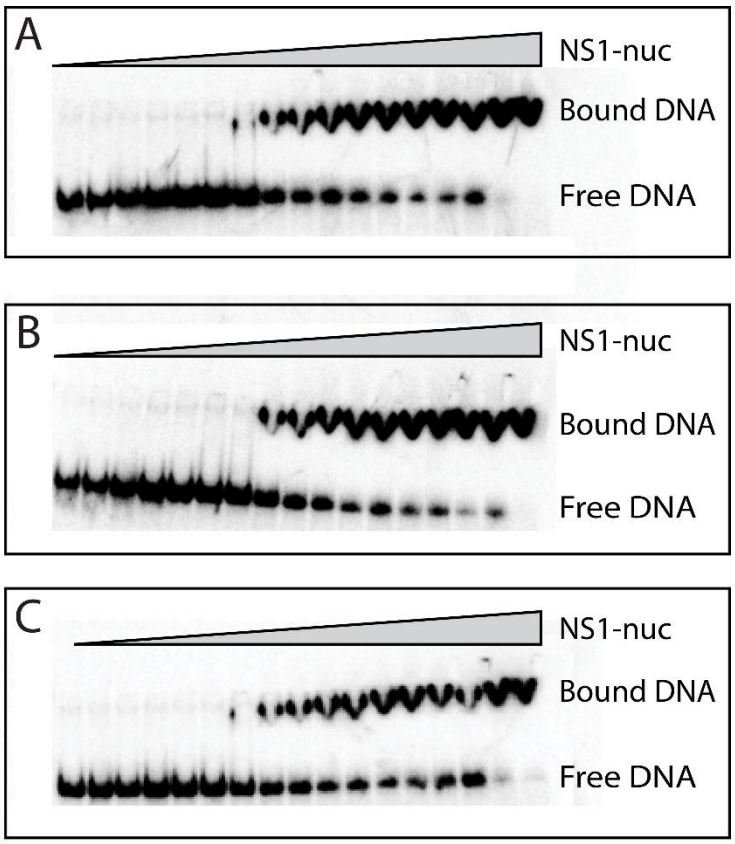



**Figure S3. Gel shift images of NS1-nuc and $^{32}$P labeled Ori2-75 dsDNA.** Each lane contains 1 nM $^{32}$P labeled Ori2-75 ds DNA and varied concentrations of NS1-nuc (increasing in lanes from left to right). A-C each represents a unique trial.

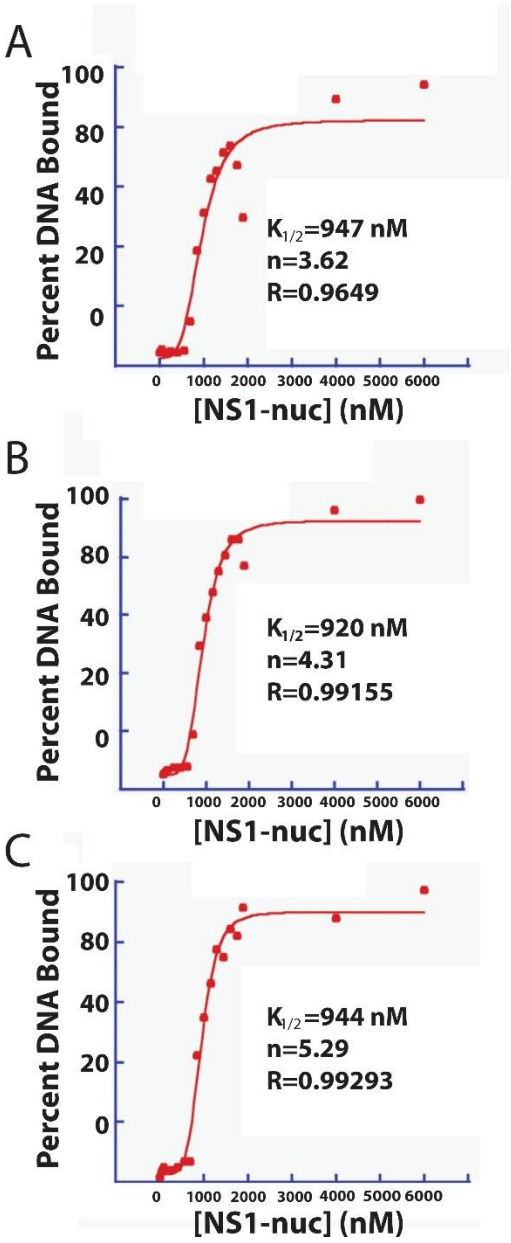



**Figure S4. Fits to data from gel shift images of NS1-nuc and $^{32}$P labeled Ori2-75 dsDNA.** A-C each represents a unique trial.

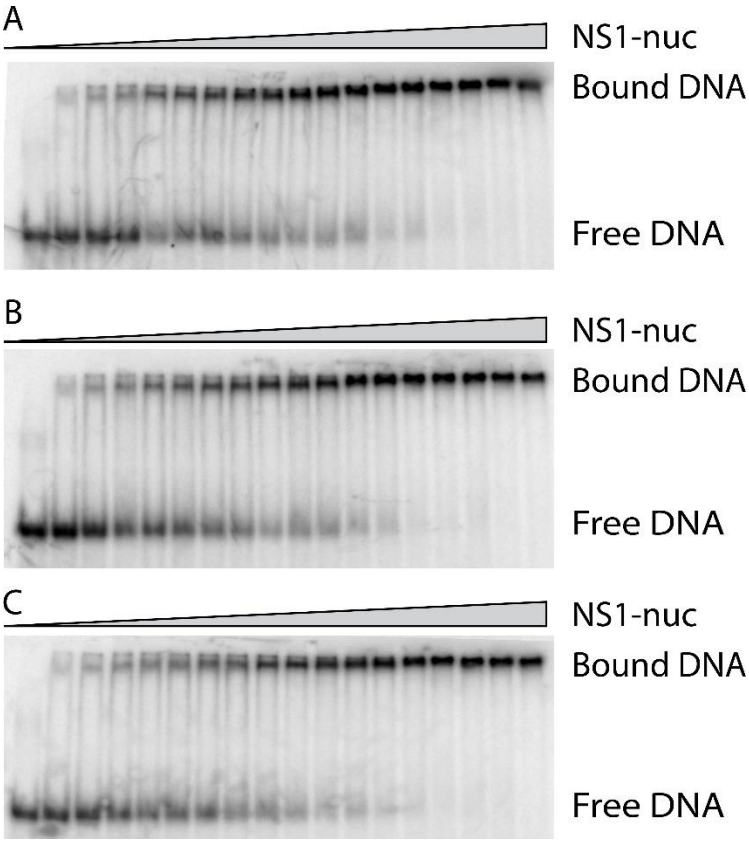



**Figure S5. Stoichiometry gel shifts with 32P labeled NSBE_DNA dsDNA.** Each lane contains 10 μM $^{32}$P labeled NSBE_DNA dsDNA and varied concentrations of NS1-nuc, increasing left to right. A-C represent independent trials.

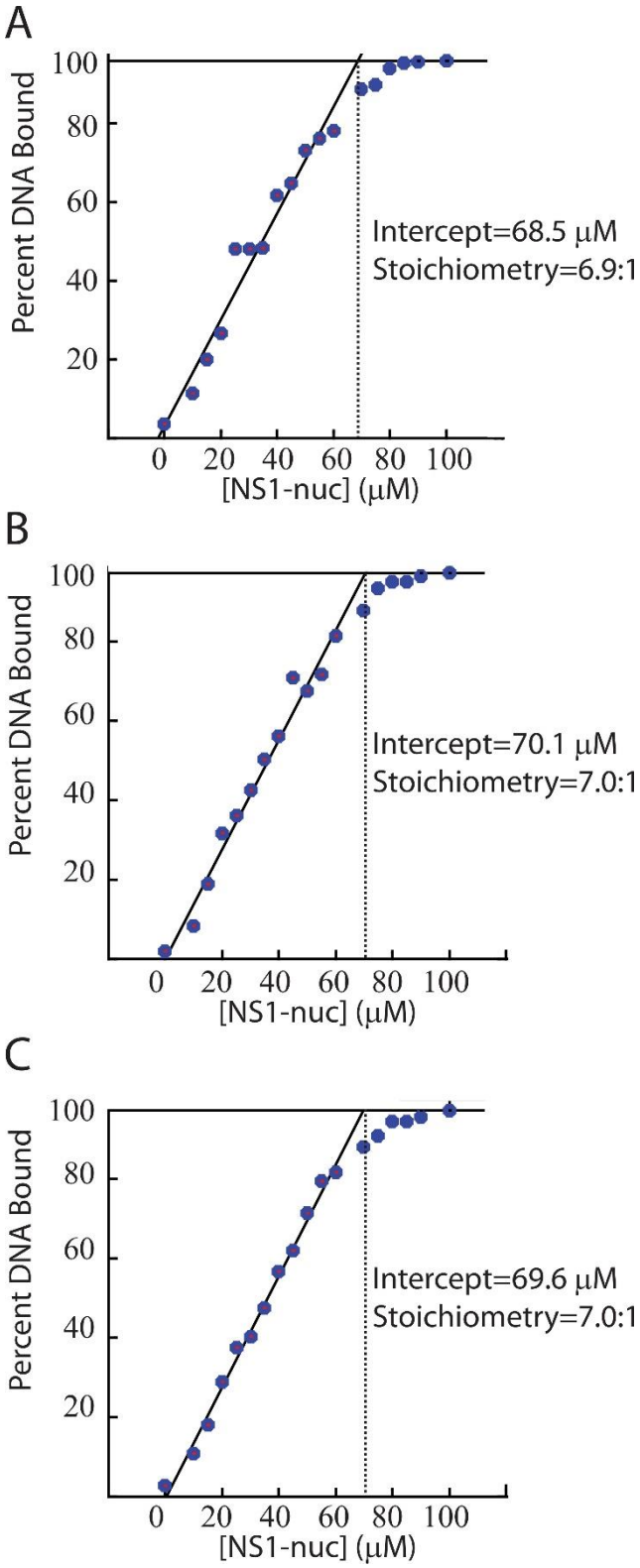



**Figure S6. Stoichiometry of NS1-nuc binding to NSBE_DNA dsDNA data fits.** DNA was 10 μM. A-C each represent unique trials.

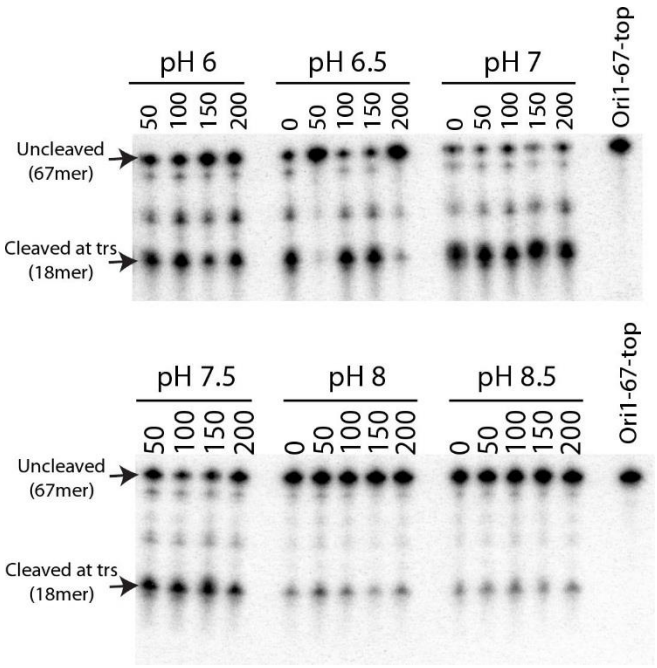



**Figure S7. Optimization of DNA cleavage by NS1-nuc.** Reactions of Ori1-67-top with 10 mM $CoCl_2$ and varied pH (50 mM HEPES-NaOH) and NaCl concentrations (mM) (1 μM NS1-nuc, 1 nM $^{32}$P-labeled Ori1-67-top, overnight at 37ºC).

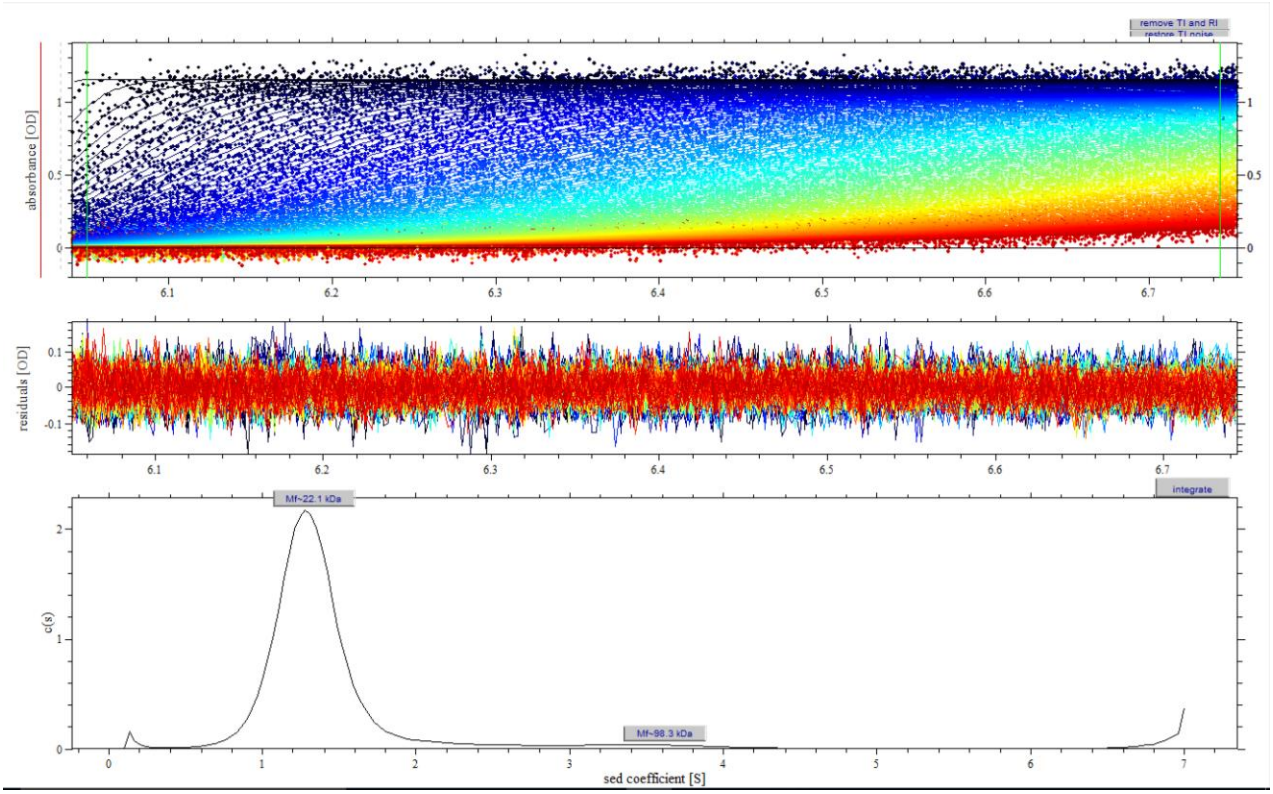



**Figure S8. AUC sedimentation velocity analysis of NS1-nuc (29 µM).** Shown is a screenshot of SEDFIT (using the predicted v-bar value of 0.72 ml/g). The top panel shows the experimental absorbance data collected in each scan (points) as well as their fits to a continuous c(s) model (lines), with earlier scans in darker blue, followed by later runs in cyan, yellow, orange, and red. The middle panel shows residuals from the fitting of the experimental data to a c(s) distribution, which is shown in the bottom panel. The apparent Mw calculated from SEDFIT are shown above each peak in the c(s) distribution in the bottom panel. See Table S1 for fixed and fitted parameters.

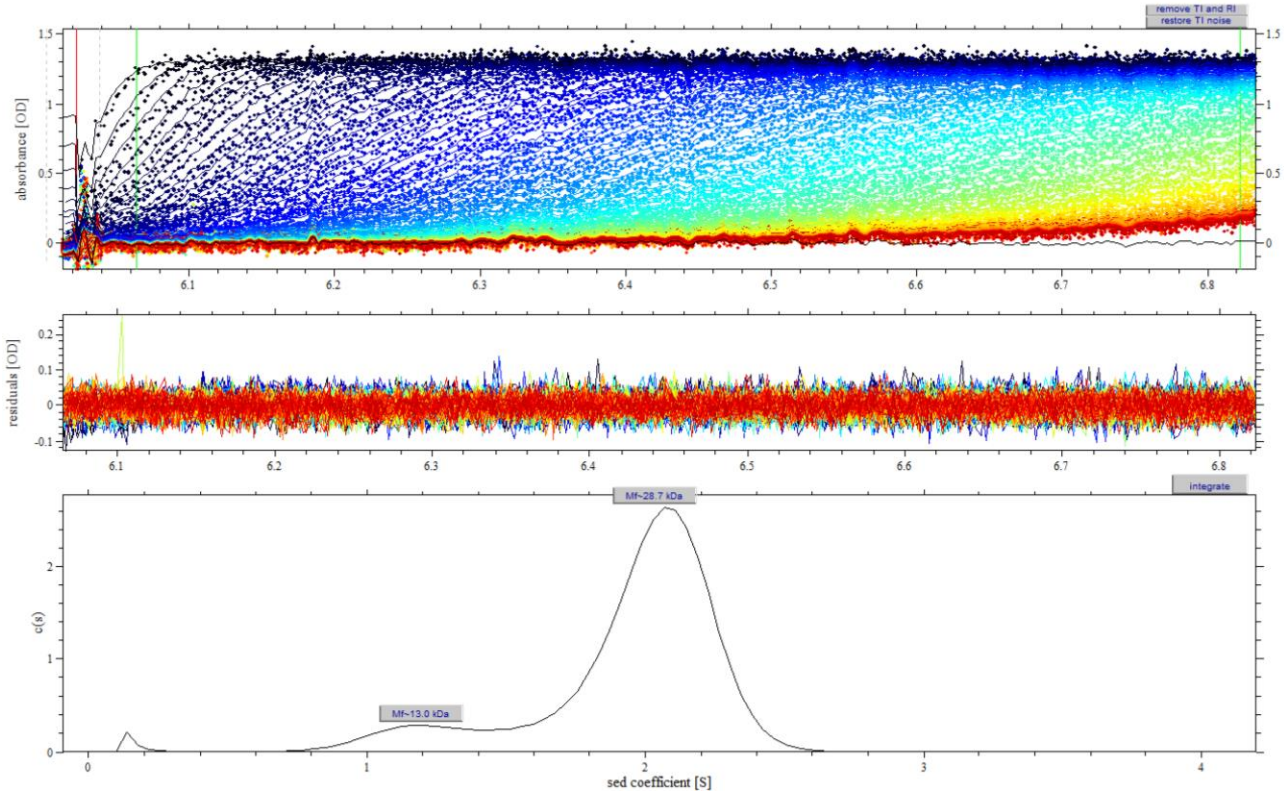



**Figure S9. AUC sedimentation velocity analysis of fluorescein labeled NSBE_DNA dsDNA (17 µM).** Shown is a screenshot of SEDFIT (using the predicted v-bar value of 0.55 ml/g). The top panel shows the experimental absorbance data collected in each scan (points) as well as their fits to a continuous c(s) model (lines), with earlier scans in darker blue, followed by later runs in cyan, yellow, orange, and red. The middle panel shows residuals from the fitting of the experimental data to a c(s) distribution, which is shown in the bottom panel. The apparent Mw calculated from SEDFIT are shown above each peak in the c(s) distribution in the bottom panel. See Table S1 for fixed and fitted parameters.

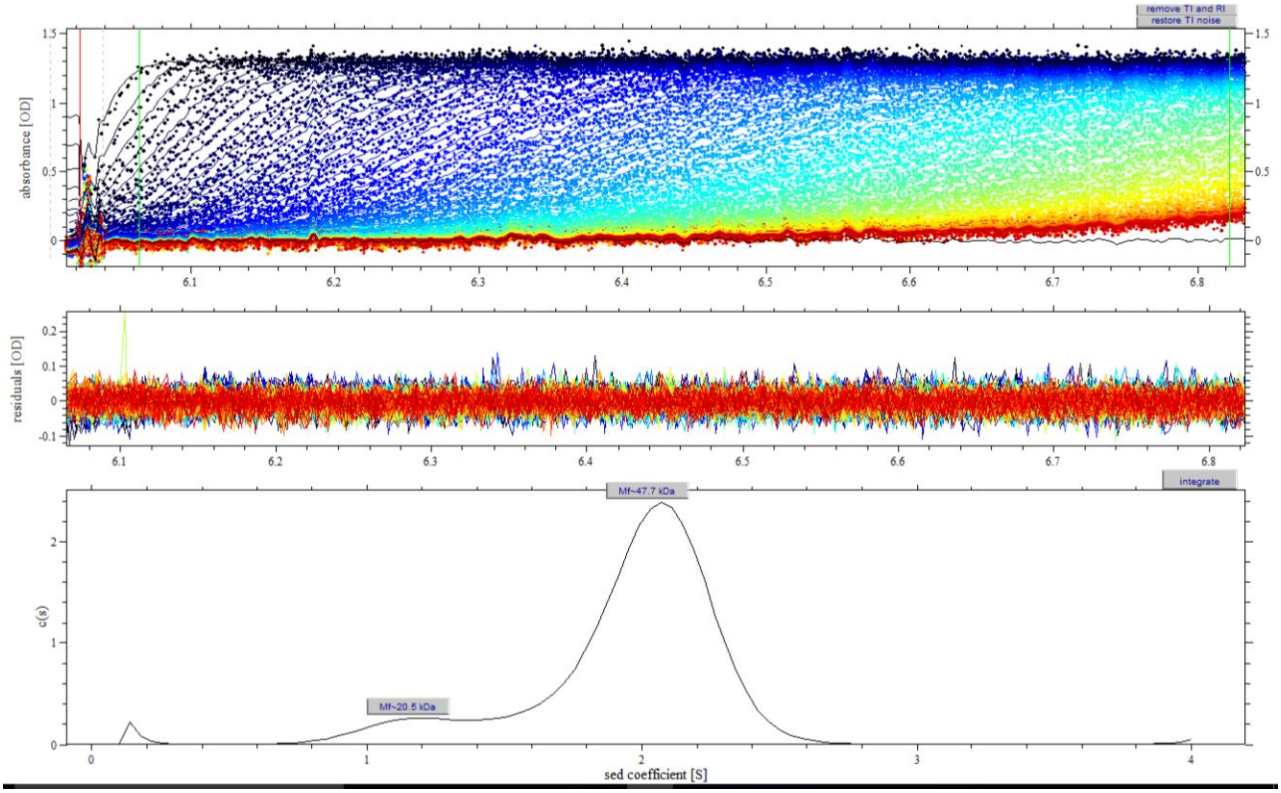



**Figure S10. AUC sedimentation velocity analysis of fluorescein labeled NSBE_DNA dsDNA (17 µM).** Shown is a screenshot of SEDFIT (using the predicted v-bar value of 0.72 ml/g). The top panel shows the experimental absorbance data collected in each scan (points) as well as their fits to a continuous c(s) model (lines), with earlier scans in darker blue, followed by later runs in cyan, yellow, orange, and red. The middle panel shows residuals from the fitting of the experimental data to a c(s) distribution, which is shown in the bottom panel. The apparent Mw calculated from SEDFIT are shown above each peak in the c(s) distribution in the bottom panel. See Table S1 for fixed and fitted parameters.

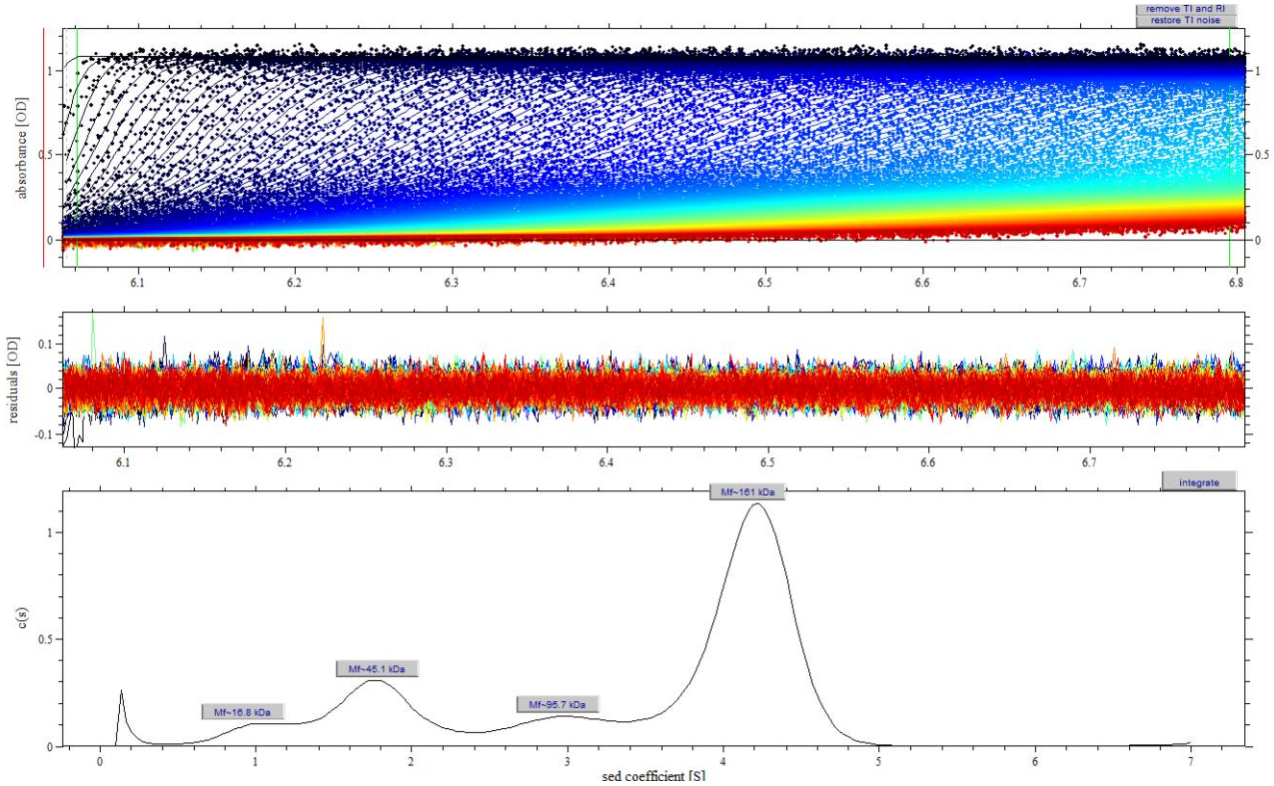



**Figure S11. AUC sedimentation velocity analysis of NS1-nuc (107 µM) mixed with fluorescein labeled NSBE_DNA dsDNA (15 µM).** Shown is a screenshot of SEDFIT (using the predicted v-bar value of 0.72 ml/g). The top panel shows the experimental absorbance data collected in each scan (points) as well as their fits to a continuous c(s) model (lines), with earlier scans in darker blue, followed by later runs in cyan, yellow, orange, and red. The middle panel shows residuals from the fitting of the experimental data to a c(s) distribution, which is shown in the bottom panel. The apparent Mw calculated from SEDFIT are shown above each peak in the c(s) distribution in the bottom panel. See Table S1 for fixed and fitted parameters.

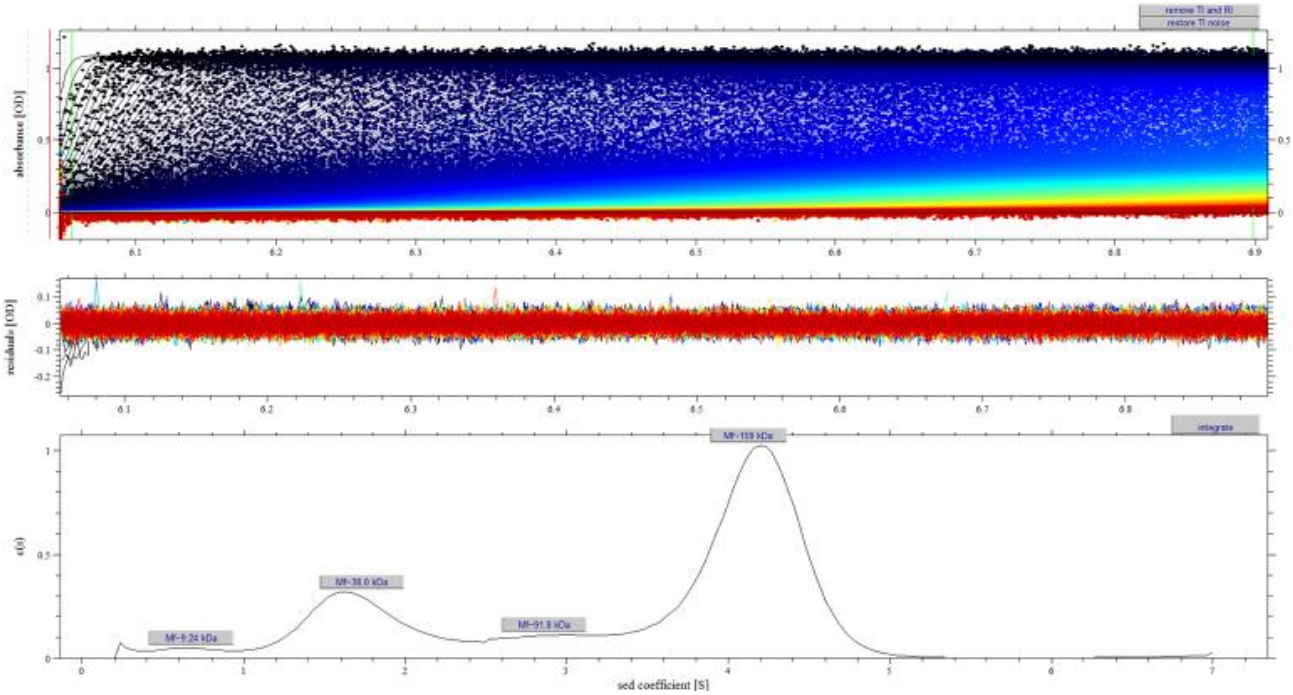



**Figure S12. AUC sedimentation velocity analysis of NS1-nuc (107 µM) mixed with fluorescein labeled NSBE_DNA dsDNA (15 µM), fit using different f/f0 for different s-values.** Shown is a screenshot of SEDFIT (using the predicted v-bar value of 0.72 ml/g). The top panel shows the experimental absorbance data collected in each scan (points) as well as their fits to a continuous c(s) model (lines), with earlier scans in darker blue, followed by later runs in cyan, yellow, orange, and red. The middle panel shows residuals from the fitting of the experimental data to a c(s) distribution, which is shown in the bottom panel. The apparent Mw calculated from SEDFIT are shown above each peak in the c(s) distribution in the bottom panel. See Table S1 for fixed and fitted parameters.

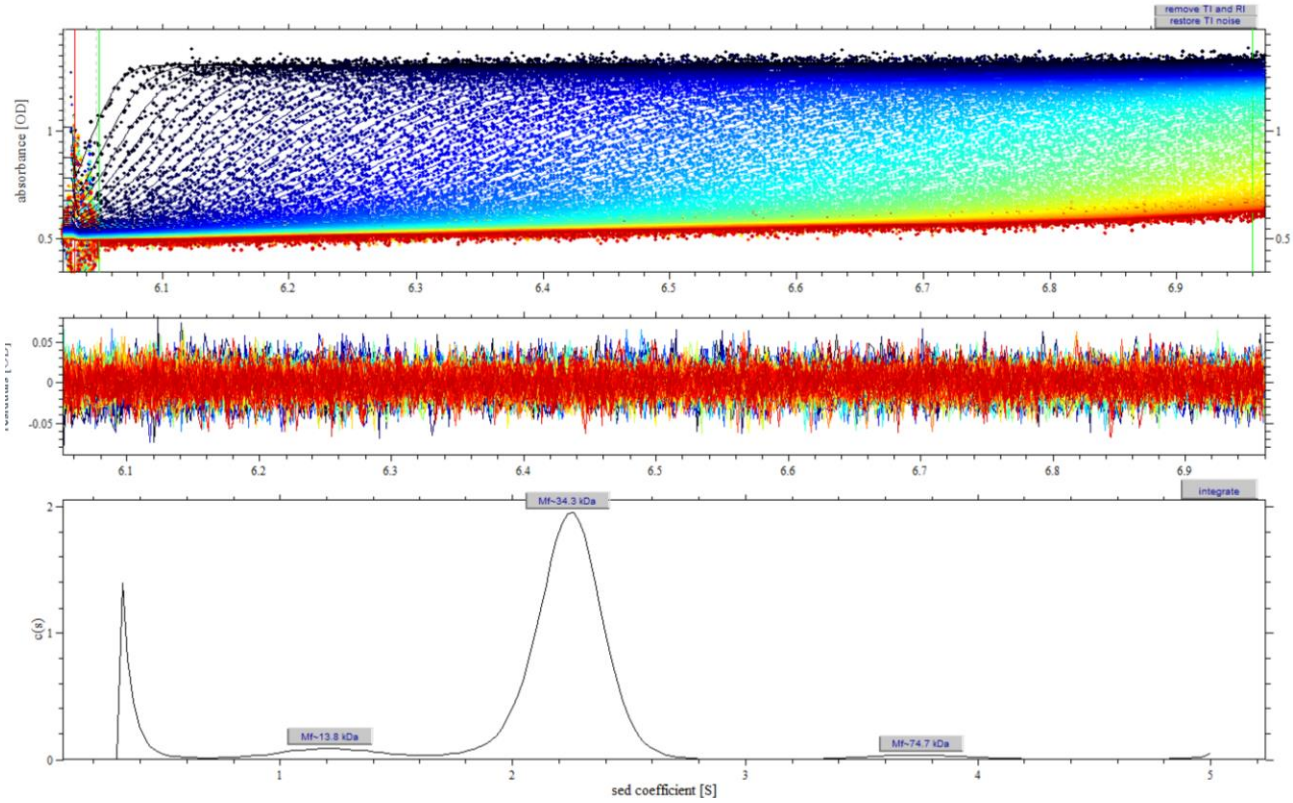



**Figure S13. AUC sedimentation velocity analysis of fluorescein labeled p21 dsDNA (14 µM).** Shown is a screenshot of SEDFIT (using the predicted v-bar value of 0.55 ml/g). The top panel shows the experimental absorbance data collected in each scan (points) as well as their fits to a continuous c(s) model (lines), with earlier scans in darker blue, followed by later runs in cyan, yellow, orange, and red. The middle panel shows residuals from the fitting of the experimental data to a c(s) distribution, which is shown in the bottom panel. The apparent Mw calculated from SEDFIT are shown above each peak in the c(s) distribution in the bottom panel. See Table S1 for fixed and fitted parameters.

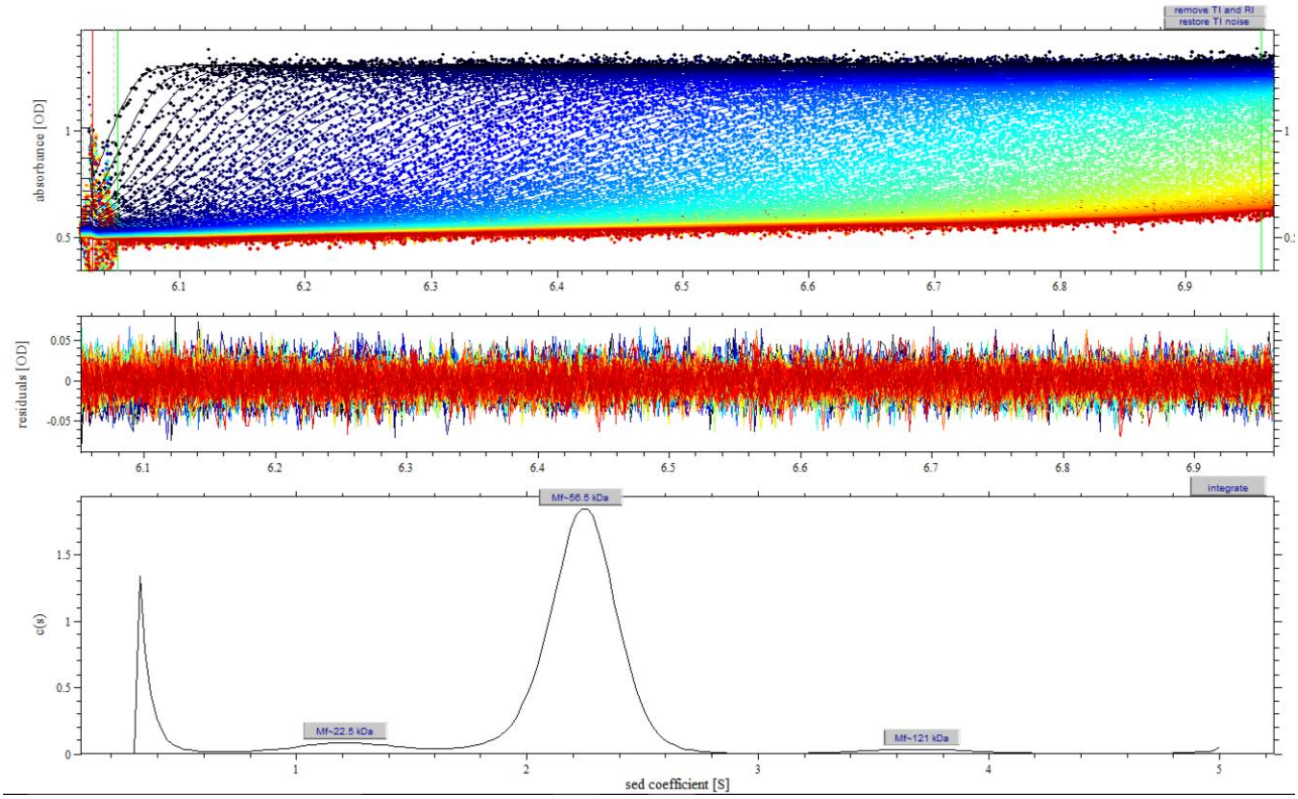



**Figure S14. AUC sedimentation velocity analysis of fluorescein labeled p21 dsDNA (14 µM).** Shown is a screenshot of SEDFIT (using the predicted v-bar value of 0.72 ml/g). The top panel shows the experimental absorbance data collected in each scan (points) as well as their fits to a continuous c(s) model (lines), with earlier scans in darker blue, followed by later runs in cyan, yellow, orange, and red. The middle panel shows residuals from the fitting of the experimental data to a c(s) distribution, which is shown in the bottom panel. The apparent Mw calculated from SEDFIT are shown above each peak in the c(s) distribution in the bottom panel. See Table S1 for fixed and fitted parameters.

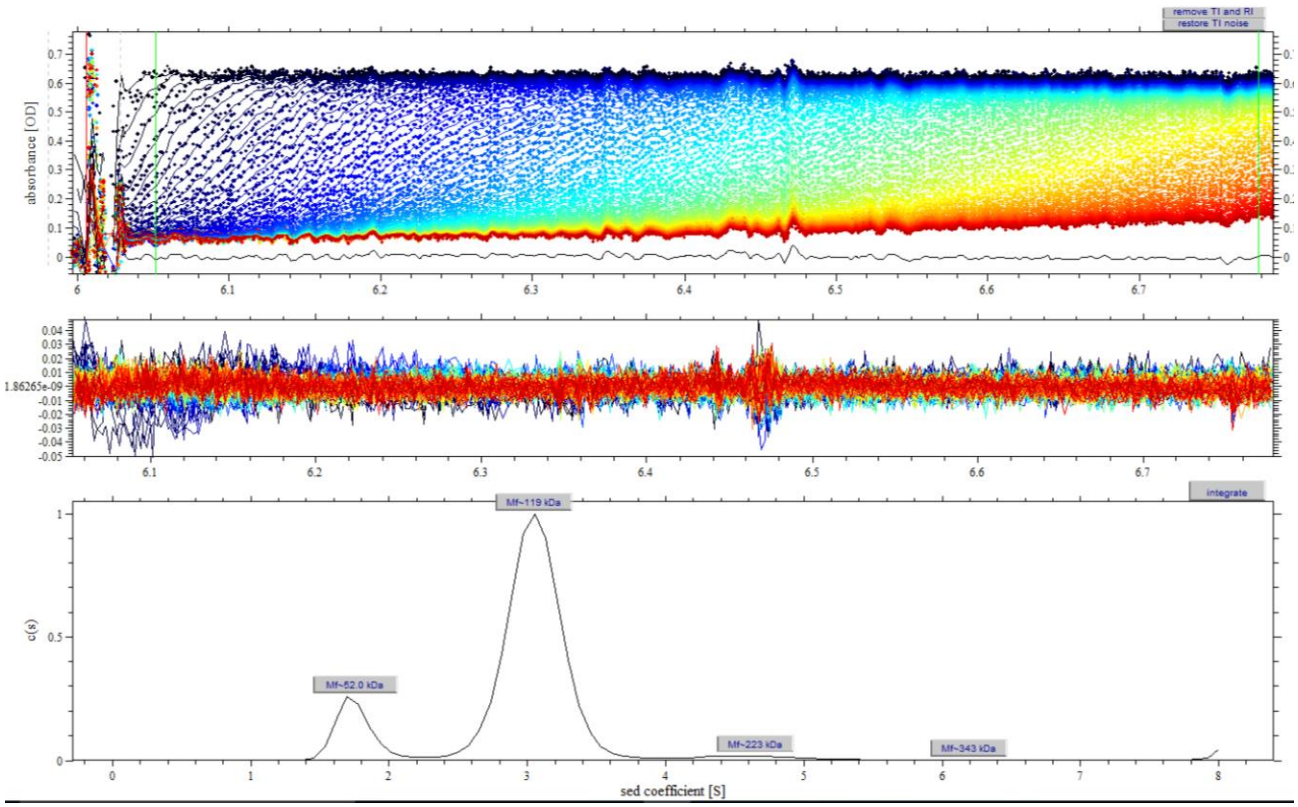



**Figure S15. AUC sedimentation velocity analysis of NS1-nuc mixed with a mixture of NS1-nuc (130 µM) and fluorescein labeled p21 dsDNA (16 µM).** Shown is a screenshot of SEDFIT (using the predicted v-bar value of 0.72 ml/g). The top panel shows the experimental absorbance data collected in each scan (points) as well as their fits to a continuous c(s) model (lines), with earlier scans in darker blue, followed by later runs in cyan, yellow, orange, and red. The middle panel shows residuals from the fitting of the experimental data to a c(s) distribution, which is shown in the bottom panel. The apparent Mw calculated from SEDFIT are shown above each peak in the c(s) distribution in the bottom panel. See Table S1 for fixed and fitted parameters.

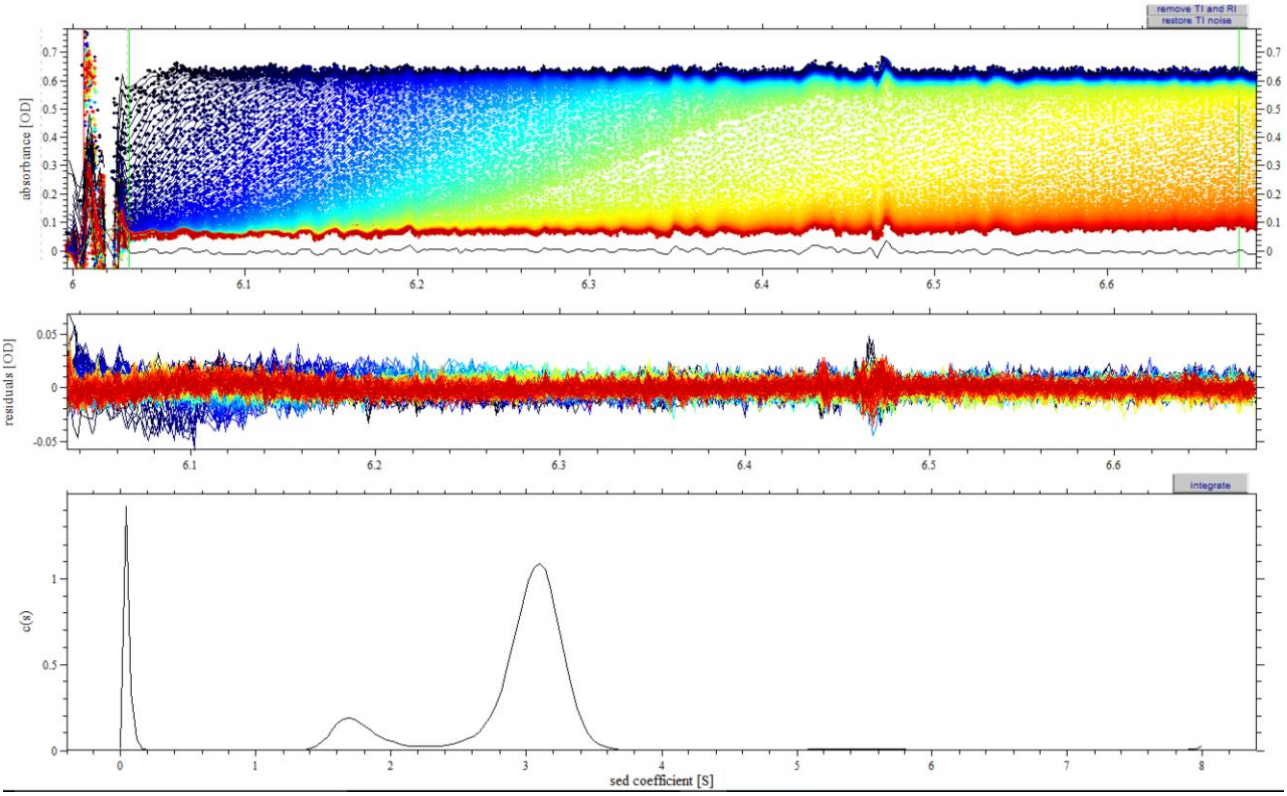



**Figure S16. AUC sedimentation velocity analysis of NS1-nuc mixed with a mixture of NS1-nuc (130 µM) and fluorescein labeled p21 dsDNA (16 µM) using different f/f0 for different s-values.** Shown is a screenshot of SEDFIT (using the predicted v-bar value of 0.72 ml/g). The top panel shows the experimental absorbance data collected in each scan (points) as well as their fits to a continuous c(s) model (lines), with earlier scans in darker blue, followed by later runs in cyan, yellow, orange, and red. The middle panel shows residuals from the fitting of the experimental data to a c(s) distribution, which is shown in the bottom panel. The apparent Mw calculated from SEDFIT are shown above each peak in the c(s) distribution in the bottom panel. See Table S1 for fixed and fitted parameters.

**Table S1**. Fitted parameters from SEDFIT analysis of AUC sedimentation velocity data.

| **Fluorescein labeled DNA** | **Protein** | **V-bar (held fixed) (ml/g)** | **Density (held fixed)** | **Viscosity (held fixed)** | **Frictional ratio f/f0 (fit)** | **Sw(20,w) of Main peak (fit) (S)** | **Mw Main peak (fit) (kDa)** | **RMSD of fit** |
|---|---|---|---|---|---|---|---|---|
| | NS1-nuc | 0.72 | 1.00 | 0.01567 | 1.53 | 2.034 | 22.5 | 0.0436 |
| NSBE_DNA | | 0.55 | 1.00 | 0.01567 | 1.94 | 3.188 | 28.8 | 0.0282 |
| NSBE_DNA | | 0.72 | 1.00 | 0.01567 | 1.56 | 3.177 | 47.7 | 0.0282 |
| NSBE_DNA | NS1-nuc | 0.72 | 1.00 | 0.01567 | 1.75 | 6.481 | 166 | 0.0229 |
| NSBE_DNA | NS1-nuc | 0.72 | 1.00 | 0.01567 | 1.56 (0.2-2.5S)<br>1.70 (0.2-7S) | 6.486 | 159 | 0.0227 |
| P21 dsDNA | | 0.55 | 1.00 | 0.01567 | 1.98 | 3.505 | 34.3 | 0.0163 |
| P21 dsDNA | | 0.72 | 1.00 | 0.01567 | 1.58 | 3.509 | 56.5 | 0.0160 |
| P21 dsDNA | NS1-nuc | 0.72 | 1.00 | 0.01567 | 1.85 | 4.813 | 115 | 0.0082 |
| P21 dsDNA | NS1-nuc | 0.72 | 1.00 | 0.01567 | 1.85 (0-2.4S)<br>1.79 (2.4-8S) | 4.808 | 110 | 0.0082 |

**Table S2.** Dependence of the stoichiometric ratio on frictional ratio f/f0 for NS1-nuc/Flo-NSBE_DNA sedimentation velocity data.

| f/f0 | RMSD[a] | Mw1 | Mw2 | Mw3 | Ratio species 1[b] | Ratio species 2[c] |
|---|---|---|---|---|---|---|
| 1.6 | 0.022906 | 146 | 87.9 | 37.3 | 4.8 | 2.2 |
| 1.7 | 0.022888 | 159 | 93.3 | 42.9 | 5.2 | 2.2 |
| 1.75 | 0.022886 | 166 | NO[d] | 43.6 | 5.4 | N/A[e] |
| 1.8 | 0.022886 | 166 | NO[d] | 45.5 | 5.4 | N/A[e] |
| 1.9 | 0.022897 | 180 | NO[d] | 49.2 | 5.8 | N/A[e] |
| 2.0 | 0.022916 | 194 | NO[d] | 52.9 | 6.3 | N/A[e] |
| 2.1 | 0.02294 | 209 | NO[d] | 57.2 | 6.7 | N/A[e] |

[a]Data in this table show only those fits with RMSD below the critical RMSD (at 1 sigma) calculated by SEDFIT for the AUC sedimentation velocity data using the same fitting parameters.

[b]Ratio is calculated as follows: (Mw1-Mw3)/22.5, where Mw1 is the apparent Mw of species 1 derived from SEDFIT analysis of the AUC sedimentation data, and Mw3 is the apparent Mw of species 3 and assumed to be that of the free DNA derived from SEDFIT using the same fitting parameters. 22.5 is used as the apparent Mw in kDa of each copy of NS1-nuc determined by AUC (Table 2). Note: using a value of 20.1 kDa for NS1-nuc gives a range of 5.4-7.6.

[c]Ratio is calculated as follows: (Mw2-Mw3)/22.5, where Mw2 is the apparent Mw of species 2 derived from SEDFIT analysis of the AUC sedimentation data, and Mw3 is the apparent Mw of species 3 and assumed to be that of the free DNA derived from SEDFIT using the same fitting parameters. 22.5 is used as the apparent Mw in kDa of each copy of NS1-nuc determined by AUC (Table 2).

[d]NO, not observed.

[e]N/A, not applicable.

**Table S3.** Dependence of the stoichiometric ratio on frictional ratio f/f0 for NS1-nuc/Flo-p21 dsDNA sedimentation velocity data

| f/f0 | RMSD[a] | Mw1 | Mw2 | Mw3 | Ratio species 1[b] | Ratio species 2[c] |
|---|---|---|---|---|---|---|
| 1.8 | 0.008160 | 111 | 266 | 48.6 | 2.8 | 9.7 |
| 1.85 | 0.008167 | 115 | 267 | 50.2 | 2.9 | 9.6 |
| 1.9 | 0.008167 | 120 | 271 | 52.3 | 3.0 | 9.7 |

[a]Data in this table show only those fits with RMSD below the critical RMSD (at 1 sigma) calculated by SEDFIT for the AUC sedimentation velocity data.

[b]Ratio is calculated as follows: (Mw1-Mw3)/22.5, where Mw1 is the apparent Mw of species 1 derived from SEDFIT analysis of the AUC sedimentation data, and Mw3 is the apparent Mw of species 3 and assumed to be that of the free DNA, derived from SEDFIT using the same fitting parameters. 22.5 is used as the apparent Mw in kDa of each copy of NS1-nuc determined by AUC (Table 2). Note: using a value of 20.1 kDa for NS1-nuc gives a range of 3.1-3.4.

[c]Ratio is calculated as follows: (Mw2-Mw3)/22.5, where Mw2 is the apparent Mw of species 2 derived from SEDFIT analysis of the AUC sedimentation data, and Mw3 is the apparent Mw of species 3 and assumed to be that of the free DNA, derived from SEDFIT using the same fitting parameters. 22.5 is used as the apparent Mw in kDa of each copy of NS1-nuc determined by AUC (Table 2).