Analytical Ultracentrifugation Studies of Translin: Analysis of Protein−DNA Interactions Using a Single-Stranded Fluorogenic Oligonucleotide

S. Paul Lee,[‡] Elena Fuior,[‡] Marc S. Lewis,[§] and Myun K. Han*,[‡]

Department of Microbiology and Immunology, Georgetown University Medical Center, Washington, DC 20007, and Division of Bioengineering and Physical Science, Office of Research Services, National Institutes of Health, Bethesda, Maryland 20892

Received February 13, 2001; Revised Manuscript Received August 23, 2001

ABSTRACT: Translin is a recently identified nucleic acid binding protein that appears to be involved in the recognition of conserved sequences found at many chromosomal breakpoints. Previous reports indicate that, based on gel filtration analysis and electron microscopy of protein−DNA complexes, translin forms an octameric structure that binds the DNA. In this study, we further examine the possibility of self-association of translin and its interactions with DNA by analytical ultracentrifugation. Sedimentation velocity analysis of translin indicates that the predominant species sediments with a sedimentation coefficient of 8.5 S and has a frictional ratio, f/f_o, of 1.35; these data are consistent with the presence of an octamer with an ellipsoidal configuration; a small amount of a component with significantly higher mass is also present. Equilibrium sedimentation studies of translin at three different protein concentrations also indicate that the predominant species present is an octamer with a minor fraction of aggregated species. Neither monomer nor dimer was detected. Sedimentation equilibrium studies of translin with an FITC-labeled single-stranded oligonucleotide were performed to examine the interaction. A novel analysis method has been developed to analyze protein−nucleic acid interactions based on global fitting of scans of 280 and 490 nm to appropriate mathematical models. Utilizing this method, it was determined that the DNA binding species of translin is an octamer binding a single-stranded oligonucleotide with a ΔG° value of -9.49 ± 0.12 kcal/mol, corresponding to a dissociation constant, K_d, of 84 ± 17 nM. On the basis of this evidence and electron microscopy, it is envisioned that translin forms an annular structure of eight subunits, hydrodynamically an oblate ellipsoid, which binds DNA at chromosomal breakpoints.

Chromosomal translocations are often associated with the development of lymphoid malignancies (*1*), but less is known about the molecular mechanisms that cause interchromosomal breakage and rejoining at specific sites of the genome. Translin, also referred to as testis−brain RNA binding protein (TB-RBP),[1] appears to be involved in the recognition of conserved sequences found flanking many major chromosomal breakpoints, such as ATGCAG and GCCC(A/T)-(G/C)(G/C)(A/T) (*2*, *3*). Translin has also been observed to bind to conserved elements in the 3′ untranslated regions of several transported brain mRNA's and to attach these elements to microtubules in vitro (*4*). It has been suggested that translin, in the brain as well as in germ cells, is involved in mRNA storage and transportation (*4*−*7*). In the brains of mice, translin is predominantly detected in the nuclei and the dendrites of neurons (*8*). It has also been identified in the neural BC1 ribonucleoprotein (RNP) particle in dendrites, which has been proposed to regulate mRNA translation within dendrites (*5*).

Translin is a 228 amino acid residue protein with a highly evolutionary conserved sequence (*9*). Its relative molecular mass calculated from the cDNA sequence is 26 180 Da. Translin is primarily localized in the cytoplasm, but is present in the nucleus of lymphoid cells with rearranged immunoglobulin and T cell receptor loci (*3*). Interestingly, DNA damaging agents such as mitomycin C and etoposide promote the transportation of this protein from the cytosol into the nucleus, thus suggesting a role in DNA recombination/repair processes (*10*). It has also been suggested that translin may be important in recognizing chromosomal breakpoints in solid tumors such as alveolar rhabdomyosarcoma cell lines (*11*).

Previous studies indicate that translin forms oligomeric structures; the formation is presumably mediated through its leucine zipper domain, which spans residues 177−212. Gel filtration studies (*3*) and electron microscopy of protein−DNA complexes (*10*) suggest that the active oligomeric form of translin is an octamer. In contrast, glycerol gradient centrifugation studies of the mouse homologue TB-RBP suggest that the smallest nucleic acid binding species is a dimer (*12*). The leucine zipper domain of TB-RBP also appears to be essential for both oligomerization and DNA binding. Yeast two-hybrid studies of TB-RBP demonstrated that deletion of amino acids 176−228 disrupts the protein−

* To whom correspondence should be addressed at the Department of Microbiology and Immunology, Georgetown University, Med-Dent Building, Room NE323, 3900 Reservoir Rd., NW, Washington, DC 20007. Tel: (202) 687-5240. Fax: (202) 687-1800. E-mail: myunhan@yahoo.com.

[‡] Georgetown University Medical Center.

[§] National Institutes of Health.

[1] Abbreviations: TB-RBP, testis−brain RNA binding protein; DTT, dithiothreitol; EDTA, ethylenediaminotetraacetate; EST, expressed sequence tag; His-translin, hexahistidine-tagged translin; IPTG, isopropyl-β-D-thiogalactoside; NTA, nitrilotriacetate; PCR, polymerase chain reaction; PMSF, phenylmethylsulfonyl fluoride.

10.1021/bi010302t CCC: $20.00
Published on Web 10/26/2001

protein interaction. Subsequently, gel shift assays demonstrated an absence of DNA binding (*12*). Therefore, there is a discrepancy between the two reports, and the identity of the functional oligomeric species of translin remains to be resolved.

In the present work, we have examined the possibility of reversible self-association of translin and its protein−DNA interactions by analytical ultracentrifugation. Sedimentation velocity experiments indicate that the oligomeric form of translin is an octamer. Equilibrium sedimentation experiments also corroborated this finding. A novel data analysis method has been developed to examine the protein−DNA interactions of translin. Utilizing fluorescein-labeled single-stranded DNA, it was determined that the octameric form of translin binds DNA without the presence of other species (i.e., monomeric or dimeric forms) present. These studies further contribute to a model in which translin binds to single-stranded DNA in an annular structure to protect single-stranded DNA at chromosomal breakpoints.

EXPERIMENTAL PROCEDURES

Plasmid Construction. The clone encoding human translin cDNA [IMAGE Consortium (LLNL) clone ID 668197] was obtained from Research Genetics (Huntsville, AL). This clone was identified by searching an EST database against the sequences reported for human cDNA, encoding translin by Aoki et al. (*3*). The coding sequence of translin was amplified by PCR using the primers TR1, 5′-GAGGGATCCTCTGTGAGCGAGATC TTCGTG-3′, and TR2, 5′-GAGGTCGACCTATTTTTCAACACAAGCTGC-3′. These primers were designed to contain single *Bam*HI and *Sal*I sites (underlined), respectively. The initial ATG encoding Met was omitted in order to prevent internal initiation when translin is expressed as a recombinant fusion protein. The PCR reaction was performed in the presence of *Taq* polymerase as follows: 30 cycles of 40 s at 94 °C, 40 s at 46 °C, and 50 s at 72 °C, followed by a final extension of 72 °C for 7 min. The purified PCR product was directionally cloned into the *Bam*HI/*Sal*I site of the pQE9 vector (Qiagen), resulting in pQE9-translin. The correct sequence of the protein was confirmed by automated sequencing on an ABI prism instrument (Vincent T. Lombardi Cancer Center Core Facility, Washington, DC).

Protein Expression and Purification. The pQE9 vector allows for the production of recombinant fusion proteins with a hexahistidine tag at the N-terminus. pQE-translin was transformed into BL21(DE3)[pREP4] *E. coli*. The histidine-tagged protein was expressed and purified according to the following procedure. A freshly saturated 5 mL overnight culture was used to innoculate 500 mL of LB medium in the presence of 100 μg/mL ampicillin and 25 μg/mL kanamycin and grown at 37 °C to an absorbance of 0.8 at 600 nm. The culture was then induced with 0.4 mM IPTG, and growth was continued for 3 h at 37 °C. The cells were harvested at 4 °C and resuspended in buffer A (50 mM sodium phosphate, pH 8.0, 1 M NaCl) supplemented with 10 mM imidazole and 5% glycerol. All subsequent purification steps were performed at 4 °C. Lysozyme (0.2 mg/mL) was added to the resuspended cells, and the suspension was incubated on ice for 1 h. Subsequently, 10 mM β-mercaptoethanol and 1 mM PMSF were added to the suspension, and the sample was sonicated on ice with a Fisher dismembrator model 300 in alternating cycles (20 s on, 2 min off) until the lysate was no longer viscous. The lysate was centrifuged at 10000*g* for 30 min to remove the cellular debris. One milliliter of 50% Ni−NTA slurry (QIAGEN) for every 4 mL of lysate was added to the supernatant solution, and it was incubated for 1 h on a rotary shaker on ice. This lysate−resin mixture was then loaded into an empty column and the resin washed with a 20× resin volume of buffer A supplemented first with 20 mM imidazole, then with 50 mM imidazole, and finally with 80 mM imidazole. It was then eluted with 250 mM imidazole−buffer A. The eluted samples were then dialyzed against 50 mM sodium phosphate, pH 7.4, 100 mM NaCl, 1 mM DTT, 1 mM EDTA, and 5% sucrose. Protein concentrations were assayed by the Bradford method, using BSA as a standard (*13*), and by absorbance, using an extinction coefficient of 16 750 M^{-1} cm^{-1}, calculated from the amino acid sequence based on the values of Perkins (*14*).

Sedimentation Velocity Experiments. Sedimentation velocity experiments were performed in a Beckman Optima XL-I analytical ultracentrifuge using a 4-hole rotor, 12 mm carbon-filled epoxy double-sector centerpieces, and sapphire windows, at a rotor speed of 55 000 rpm and at a temperature of 19 °C. The buffer was 50 mM sodium phosphate, pH 7.4, 100 mM NaCl, 1 mM EDTA, 0.2 mM DTT. Radial scans were acquired at 230 nm at 60 s intervals. Analysis of the data was performed utilizing the ultracentrifugal program SEDFIT, as described (*15*).

Sedimentation Equilibrium Studies of Translin. Sedimentation equilibrium studies were performed in a Beckman Optima XL-A analytical ultracentrifuge using an 8-hole rotor, 12 mm carbon-filled epoxy double-sector centerpieces, and quartz windows. Samples were dialyzed against the appropriate buffer overnight, and centrifuged at 14000*g* to remove large aggregates prior to loading. Sample volumes were 0.15 mL, giving a column length of approximately 4 mm.

Initial studies were performed in 50 mM sodium phosphate, pH 7.4, 100 mM NaCl, 1 mM EDTA, 0.2 mM DTT at 20 °C. Three initial loading concentrations of 3, 6, and 12 μM (corresponding to absorbances of 0.05, 0.1, and 0.2 at 280 nm) were used. The rotor speed was 10 000 rpm. The radial scans at 280 nm were recorded after 72 h.

In another experiment, intended to emphasize lower degrees of oligomerization, three lower protein concentrations of 0.43, 0.85, and 1.7 μM (corresponding to absorbances of 0.05, 0.1, and 0.2 at 230 nm) were used. The rotor speed was 10 000 rpm, the temperature was 20 °C, and scans were taken at 230 nm. The buffer was the same as before except that DTT was omitted in order to avoid excessive absorbance at 230 nm.

The sedimentation equilibrium data were edited and analyzed using MLAB software (Civilized Software, Inc., Silver Spring, MD). Absorbance data corresponding to the different loading concentrations were simultaneously fit with appropriate mathematical models using nonlinear least-squares curve-fitting.

The sedimentation equilibrium data of a homogeneous macromolecular sample can be fit to a model using the equation:

$$C_T(r) = C_b \exp(AM_1\delta r^2) + \epsilon \quad (1)$$

where $C_T(r)$ represents the total absorbance at radius r and C_b is the monomer absorbance at r_b, the radius at the cell bottom. A is $(1 - \bar{v}\rho)\omega^2/2RT$, where $\bar{v}$ is the compositional partial specific volume of the protein (0.7350 mL/g for translin), ρ is the density of the solvent (1.01407 g/mL), ω is the angular velocity of the rotor (in radians s^{-1}), R is the universal gas constant, T is the absolute temperature, M_1 is the compositional molecular mass of the monomer (M_1 = 26 870 Da for His-translin), δr^2 is $(r^2 - r_b^2)$, and ϵ is the baseline correction. The molecular mass and the compositional partial specific volume of translin were calculated from the amino acid sequence according to Perkins (*14*). The density of the solvent was measured in an Anton Paar DNA5000 Density Meter.

Analytical Ultracentrifugation of DNA−Protein Mixtures. To examine the interactions of translin and DNA, a fluorescein-labeled oligonucleotide, F-BCR1: 5′-TTGCAGTGAGCCGAGATAGTGCCA-F-3′ (from Genosys), was prepared. This single-stranded oligonucleotide has been previously used to demonstrate the interactions with translin (*3*). Sedimentation equilibrium studies were performed in the XL-A analytical ultracentrifuge, as described above. PBS containing 0.5 mM EDTA was used as buffer. The protein and the DNA were dialyzed overnight against this buffer. Initially, an experiment was performed at 20 °C with a rotor speed of 30 000 rpm to allow the DNA to form a gradient optimal for the determination of its reduced mass, i.e., $M(1 - \bar{v}\rho)$. For the following experiments, the samples analyzed were the following: translin alone at two concentrations, 8 and 12 μM; DNA alone at two concentration, 1.5 and 0.9 μM; and three mixtures of translin and DNA with translin octamer to DNA molar ratios of 2.3:1, 1.1:1, and 1:1.5. In these experiments at 20 °C, a rotor speed of 10 000 rpm was used for the analyses of translin alone and of the DNA−translin mixtures. In all experiments, absorbance and intensity data were scanned at 280 and 490 nm, at a radial step size of 0.002 cm, and an average of 16 scans. The molecular mass of the oligonucleotide was 8012 Da, as calculated from its compositional formula.

Data Analysis of DNA−Protein Interaction. At 490 nm, only the fluorescein label is detectable from FITC-BCR1 and, thus, only the free DNA and DNA−translin complexes could be detected. Thus:

$$A_{r,490} = A_{b,N} \exp(BF_N M_N \delta r^2) + A_{b,PN} \exp[B(F_P M_P + F_N M_N)\delta r^2] + \epsilon \quad (2)$$

where $A_{b,N}$ and $A_{b,PN}$ are the absorbance of free DNA and the absorbance of DNA−protein complex, respectively, at r_b, B is $\omega^2/2RT$, δr^2 is $(r^2 - r_b^2)$, F_N is $(1 - \bar{v}_N\rho)$, $\bar{v}_N$ is the partial specific volume of free DNA, F_P is $(1 - \bar{v}_P\rho)$, $\bar{v}_P$ is the partial specific volume of protein, and M_N and M_P are the molecular mass of DNA and protein, respectively. In every case, $F_N M_N$ had the value of $M(1 - \bar{v}\rho)$ determined in the DNA experiments. Also at 490 nm:

$$A_{b,N} = C_{b,N}E_N \text{ and } A_{b,PN} = C_{b,PN}E_{PN} = C_{b,P}E_N \quad (3)$$

where C_i and E_i are the molar concentrations and extinction coefficients of the species, respectively. Thus:

$$A_{r,490} = C_{b,N}E_{N,490} \exp(BF_N M_N \delta r^2) + C_{b,PN}E_{N,490} \exp\{B(F_P M_P + F_N M_N)\delta r^2\} + \epsilon \quad (4)$$

At 280 nm, both protein and DNA (with some F-label visible) are observable. Thus:

$$A_{r,280} = A_{b,N} \exp(BF_N M_N \delta r^2) + A_{b,P} \exp(BF_P M_P \delta r^2) + A_{b,PN} \exp\{B(F_N M_N + F_P M_P)\delta r^2\} + \epsilon = C_{b,N}E_{N,280} \exp(BF_N M_N \delta r^2) + C_{b,P}E_{P,280} \exp\{B(F_P M_P \delta r^2\} + C_{b,PN}(E_{P,280} + E_{N,280}) \exp\{B(F_N M_N + F_P M_P)\delta r^2\} + \epsilon \quad (5)$$

By definition, $K_{PN} = C_{PN}/C_P C_N$. Thus, $C_{PN} = K_{PN}C_P C_N$. Substituting this in both eqs 4 and 5 for C_{PN} gives

$$A_{r,490} = C_{b,N}E_{N,490} \exp(BF_N M_N \delta r^2) + K_{PN}C_{b,P}C_{b,N}E_{N,490} \exp\{B(F_P M_P + F_N M_N)\delta r^2\} + \epsilon \quad (6)$$

$$A_{r,280} = C_{b,N}E_{N,280} \exp(BF_N M_N \delta r^2) + C_{b,P}E_{P,280} \exp(BF_P M_P \delta r^2) + K_{PN}C_{b,P}C_{b,N}(E_{P,280} + E_{N,280}) \exp\{B(F_N M_N + F_P M_P)\delta r^2\} + \epsilon \quad (7)$$

To prevent obtaining physically nonmeaningful negative values for K_{PN} and the concentrations, it is useful to work with ln K_{PN}, ln E's, and ln C's as fitting parameters; thus, the equations become

$$A_{r,490} = \exp(\ln C_{b,N} + \ln E_{N,490} + BF_N M_N \delta r^2) + \exp\{\ln K_{PN} + \ln C_{b,P} + \ln C_{b,N} + \ln E_{N,490} + [B(F_P M_P + F_N M_N)\delta r^2]\} + \epsilon_{490} \quad (8)$$

$$A_{r,280} = \exp(\ln C_{b,N} + \ln E_{N,280} + BF_N M_N \delta r^2) + \exp(\ln C_{b,P} + \ln E_{P,280} + BF_P M_P \delta r^2) + \exp\{\ln K_{PN} + \ln C_{b,P} + \ln C_{b,N} + \ln (E_{P,280} + E_{N,280}) + [B(F_N M_N + F_P M_P)\delta r^2]\} + \epsilon_{280} \quad (9)$$

The molar extinction coefficient of the DNA−translin complex is assumed to be the sum of the extinction coefficients of the individual components in this study.

RESULTS

Analytical Ultracentrifugation Studies of Translin. Previous reports indicate that translin binds DNA or RNA in the form of a multimer, either as an octamer for the human protein (*3*, *10*) or as a dimer in the case of the mouse homologue, TB-RBP (*12*). Therefore, it is still unclear which form of the protein exists in solution and which is the functional form that binds DNA. To absolutely understand the molecular basis of interaction between translin and DNA, it is important to discern the forms of the protein that exist in solution and which of the existing forms are responsible for nucleic acid binding. Analytical ultracentrifugation is a reliable method to characterize the stoichiometry and dynamics of macromolecular assemblies. Both sedimentation velocity and equilibrium sedimentation experiments are rigorously based on first principles of irreversible and reversible thermodynamics, respectively, thereby eliminating the need

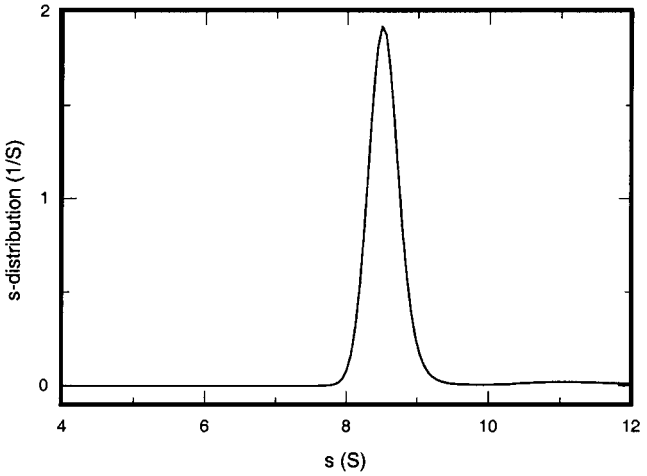

FIGURE 1: Velocity sedimentation analysis of translin. Sedimentation velocity experiments were performed in a Beckman Optima XL-A analytical ultracentrifuge using a 4-hole rotor, at a rotor speed of 55 000 rpm, and a temperature of 19 °C. The buffer was 50 mM sodium phosphate, pH 7.4, 100 mM NaCl, 1 mM EDTA, 0.2 mM DTT. Radial scans were acquired at 230 nm at 60 s intervals. Data analysis was performed using Lamm equation modeling to give mass distributions as a function of sedimentation coefficient. In addition to the main peak at 8.5 S, there is a broad distribution of a very small amount of heavier, more rapidly sedimenting aggregates.

for calibration, as is the case for other methods such as gel-filtration chromatography. Our studies focus on determining the possibility of reversible self-associating species of translin and the form of translin responsible for binding DNA by means of analytical ultracentrifugation.

The coding sequence of translin was PCR-amplified from a clone encoding human translin cDNA. The amplified PCR product was directionally cloned into the *Bam*HI/*Sal*I site of the pQE9 vector (QIAGEN) which allows for the formation of hexahistidine-tagged fusion proteins. Hexahistidine-tagged translin was affinity-purified on a nickel column under native conditions. Initial sedimentation velocity experiments were performed to determine the existing self-associating species of translin in solution. The results indicate that the protein exists predominantly as an octamer with minor fractions of higher aggregates (Figure 1). It should be emphasized that neither monomer nor dimer was detected under the experimental conditions. Further analysis of the data revealed that the sedimentation coefficient of the octameric species was 8.5 S, and the frictional ratio, f/f_0, was 1.35, indicating, in combination with the electron microscopy data, that the translin octamer behaves hydrodynamically as an oblate ellipsoid with an axial ratio of 7.5:1 (*16*).

To determine if translin undergoes reversible self-association, sedimentation equilibrium experiments were performed with three initial loading concentrations. These data were simultaneously analyzed using appropriate fitting functions to determine the best-fitting global mathematical model. Initially, the data were fit to a model using the equation: $c(r) = c_b \exp[nM_1A(r^2 - r_b^2)]$ (see Experimental Procedures). This analysis allows the determination of n, the number of subunits. The recovered n was 8.4, indicating that the major oligomeric species present were octamers with minor species of larger oligomers. These results were consistent with the data obtained from the velocity sedimentation studies. Further

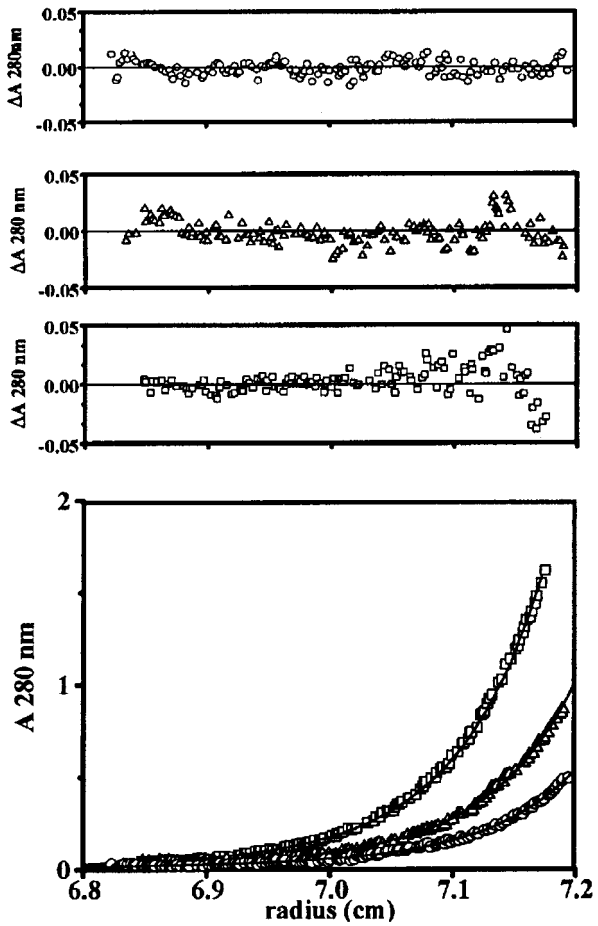

FIGURE 2: Equilibrium sedimentation analysis of translin. The sedimentation equilibrium studies were performed with 3, 6, and 12 μM translin in 50 mM sodium phosphate, pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.2 mM DTT at 20 °C. The loading concentrations of 3 (□), 6 (△), and 12 (○) μM correspond to absorbances of 0.05, 0.1, and 0.2 at 280 nm. The rotor speed was 10 000 rpm. The radial scans at 280 nm were recorded after 72 h. These data were fit with the model: $C_i(r) = C_{b1,i} \exp(AM_1\delta r^2) + C_{b2,i} \exp(AM_2\delta r^2) + \epsilon_i$, where $i = 1, 2, 3$ and indicates cell number for global fitting of the data. The distributions of the residuals for each of these fits are shown.

analyses indicate that the sedimentation data were best fit by a two-component model as judged by a minimum value for the sum of squares (SS) values and an appropriate distributions of the residuals. The optimal distributions from the two-component analysis were predominantly octamer ($M_r \sim 8M_1$) with a small fraction of large aggregates ($M_r \sim 70M_1$) (Figure 2). Excluding the large aggregated species from the analysis results in a monomer−octamer model with a quality of fit that was very significantly inferior. When monomer, octamer, and aggregated species were included, a very small negative concentration value was obtained for monomer. In this particular case, the fitting without constraining all concentration values to be positive was deliberate since it gave us an indication of the validity of the model. Since the concentration value obtained for the monomer was both very small and negative, and since negative concentration values have no physical reality, this indicated that monomeric species were not present or were not detectable under the experimental conditions and that the model was correct. If a large negative concentration had been obtained, it would have indicated that the model was incorrect. When

the data were fit with a model which had no monomer, the sum of squares did not differ significantly from that of the model including the monomer. Similar results were obtained when attempts were made to include dimer or tetramer as a third component. Analysis of scans at 230 nm also yielded the same results. Thus, the results obtained from both the velocity and equilibrium sedimentation studies indicate that translin is best described as a nonreversible oblate ellipsoidal octamer with a small amount of large aggregated species present. These data are in good agreement with previous results utilizing gel-filtration (*3*) and electron microscopy (*10*).

Studies of Translin−DNA Interactions Using FITC-Labeled Oligonucleotide. Sedimentation equilibrium is an excellent method to assess the stoichiometry of binding and the binding constant of translin to DNA. However, sedimentation equilibrium has not been widely used for the characterization of the interaction of mixtures of interacting components unless their absorption spectra are clearly different because of the greater complexity of the analyses. In the case of mixtures of nucleic acids and protein, if only one wavelength is used, it is often difficult to analyze the data due to their closely overlapping absorption spectra. For protein−DNA interactions, a very critical but laborious approach is to use a multiwavelength method. This procedure exploits the fact that the absorbance of DNA is markedly greater than that of proteins at 250 nm while the absorbance of proteins becomes higher than that of nucleic acids as 230 nm is approached (*17, 18*). An alternative and markedly easier approach, employed in this study, is the use of fluorescently labeled DNA. This approach has the advantage of specifically monitoring only the DNA by scanning at the absorbance of the fluorescent label without possible complications of hyper- or hypochromicity effects at 260 nm. In addition, the protein is invisible (no absorbance being monitored), allowing only the free DNA and the protein−DNA complex to be monitored without interference from free protein. For this study, a single-stranded fluorescein 5-isothiocyanate-labeled 24mer oligonucleotide, FITC-BCR1, was utilized with purified translin.

The initial sedimentation equilibrium studies of DNA alone, performed at 20 °C and at 30 000 rpm, were for the purpose of obtaining optimal experimental values for $M_N(1 - \bar{v}\rho)$. Figure 3 represents a global analysis of sedimentation equilibrium data of FITC-BCR1. The data were well fit as an ideal single species, described by the equation:

$$C_T(r) = c_b \exp(FBM_n\delta r^2) + \epsilon$$

where M_n is the molecular mass of FITC-BCR1. With the known value of M_n and the measured value of ρ, the partial specific volume calculated from the global fitting of the data at 260, 280, and 490 nm was 0.523 ± 0.001 mL/g. This is in agreement with the values generally reported for partial specific volumes of DNA, approximately 0.55 mL/g (*19*). In addition, from these results, the extinction coefficients of DNA at 260 and 280 nm were determined, based on the value at 490 nm. These values for the wavelengths of 260, 280, and 490 nm are 155 520, 293 920, and 72 000 M^{-1} cm^{-1}, respectively.

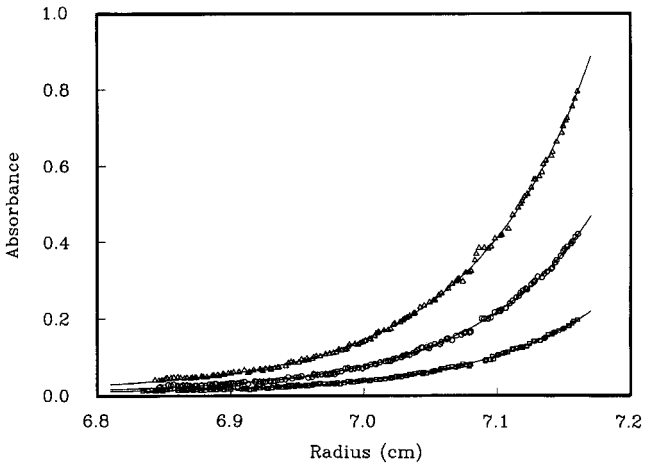

FIGURE 3: Equilibrium sedimentation analysis of FITC-labeled DNA. A solution of 1 μM FITC-BCR1 in a buffer containing 50 mM sodium phosphate, pH 7.4, 100 mM NaCl, and 1 mM EDTA is shown as absorbance as a function of radial position at equilibrium at 20 °C and at 30 000 rpm. The data shown are for the scans at 260 nm (triangles), 280 nm (circles), and 490 nm (squares). All of the data scanned at these three wavelengths were simultaneously fitted using a model for an ideal single species of DNA to obtain the value of the reduced molecular mass, $M(1 - \bar{v}\rho)$.

For optimal determination of the association constant and the stoichiometry of the interaction of translin and FITC-BCR1, three different mixtures of translin and DNA were globally analyzed. In sample 1, 11.2 μM translin was mixed with 0.6 μM 3′F-BCR1 (Figure 4A); in sample 2, 10.4 μM translin was mixed with 1.2 μM 3′F-BCR1 (Figure 4B); and in sample 3, 9.6 μM translin was mixed with 1.8 μM 3′F-BCR1 (Figure 4C). Thus, the molar ratios of the octameric protein to the DNA of these three mixtures are 2.3, 1.1, and 0.7, respectively. As previously mentioned, this approach allows only the DNA to be measured at a wavelength in a spectral region where the protein does not absorb. Thus, the data scanned at 490 nm represent concentration distributions of free DNA and protein−DNA complexes without the contribution of free protein. The mathematical model can be written for the DNA alone, reflecting the presence of both free and bound DNA: $A_{r,490} = A_{b,N} \exp(BF_NM_N\delta r^2) + A_{b,PN} \exp\{B(F_PM_P + F_NM_N)\delta r^2\} + \epsilon$. On the other hand, the data collected at 280 nm reflect concentration distributions of free DNA, free protein, and their complex. Accordingly, the mathematical model can be written as: $A_{r,280} = A_{b,N} \exp(BF_NM_N\delta r^2) + A_{b,P} \exp(BF_PM_P\delta r^2) + A_{b,PN} \exp\{B(F_NM_N + F_PM_P)\delta r^2\} + \epsilon$. Utilizing these models, appropriately rewritten to include the equilibrium constant as described previously, sedimentation equilibrium data were globally analyzed, and the results of these analyses are shown in Figure 4. The data scanned at 280 and 490 nm are presented with their corresponding fits for a model where an octamer binds DNA with a 1:1 stoichiometry. The insets represent the distributions of their residuals and show good fits. From the global analysis, we have calculated the distributions of free DNA, free protein, and their complex from 280 nm data (Figure 5) and the distributions of free and complexed DNA from 490 nm data (Figure 5 insets). Both of these clearly illustrate varying levels of free DNA and that these form very shallow gradients at 10 000 rpm; thus, the gradients observed in sedimentation equilibrium profiles are due primarily to the

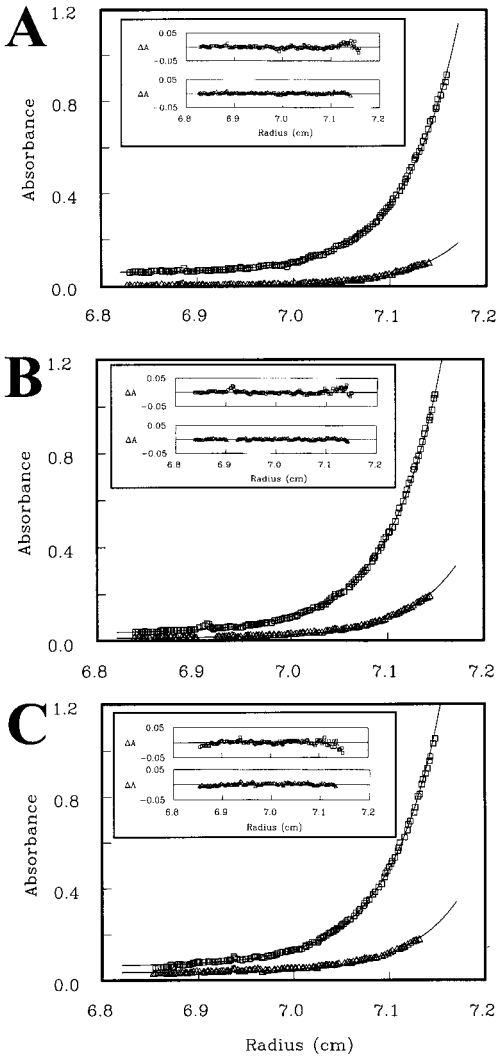

FIGURE 4: Equilibrium sedimentation analysis of the interaction of translin with FITC-labeled DNA. Equilibrium distributions at 280 nm (squares) and 490 nm (triangles) for mixtures of translin and FITC-DNA with loaded molar ratios of (A) 2.3:1 octamer to DNA, (B) 1.1:1 octamer to DNA, and (C) 1:1.5 octamer to DNA and the fitting lines for a model with a stoichiometry of 1:1 are shown. The insets show the distributions of the residuals.

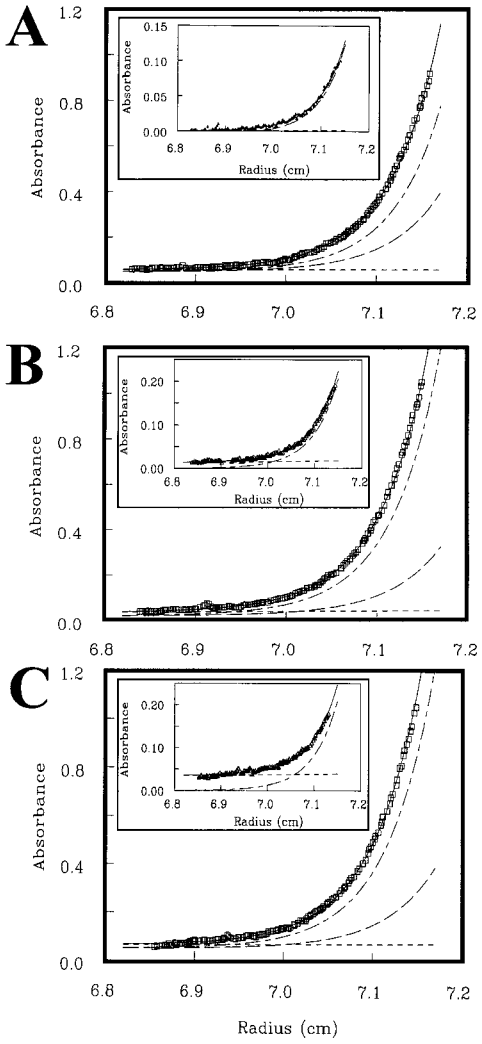

FIGURE 5: Equilibrium distributions of free and complexed translin and FITC-labeled DNA. The data and solid fitting lines for the three protein−DNA molar ratios illustrated in this figure at 280 nm are the same as those shown at this wavelength for those molar ratios in Figure 4. Here, the short dashed lines show the distribution of free DNA; the long dashed lines show the distribution of free octamer; the alternating long and short dashed lines show the distribution of the octamer−DNA complex. The solid lines fitting the data thus represent the sum of these distributions. In each case, the inset shows the same data and solid fitting line for 490 nm where only FITC-DNA is visible. Thus, the short dashed lines and the alternating long and short dashed lines show only the distributions of the free and complexed FITC-DNA, respectively.

DNA−protein complex and, secondarily, to free protein only at 280 nm; they are due primarily to DNA−protein complex at 490 nm. The quality of fits of the 1:1 stoichiometry model to all of the data presented with differing molar ratios of protein and DNA (Figures 4 and 5) clearly demonstrates that this is the optimal model. The calculated Gibbs standard free energy change ($\Delta G°$) for the protein−DNA interaction is -9.49 ± 0.12 kcal/mol, which corresponds to the dissociation constant, K_d, of 84 ± 17 nM. Since the mathematical model for fitting the data and obtaining the value of ln K from which this was calculated was a nonlinear model, the values of the parameter standard deviations, returned by the Levenberg−Marquardt algorithm, are approximations based on a linear model. To assess the validity of these values, we performed a 1000-iteration Monte Carlo simulation using the returned fitting parameters from the global analysis to generate data sets which then had normally distributed error whose mean magnitude was equal to the returned root-mean-square error added to each data set. The standard error of $\Delta G°$ from this analysis was ± 0.08 kcal/mol, a significantly smaller value than the value obtained from the global analysis. The corresponding error in the value of the dissociation constant thus is ± 12 nM. These results are relevant to considering the values of the free nucleic acid seen in Figure 5. In the worst case illustrated by Figure 5A, the global fitting gave absorbencies of free nucleic acid at the cell bottom of 0.013 ± 0.003 at 280 nm and 0.006 ± 0.001 at 490 nm. While these values are at the limits of detectability in the XL-A at these wavelengths, they are very highly constrained by the nature of the global fitting procedure, and their nominal standard errors are small enough that they must be considered as being reasonably probable from a statistical point of view. In Figure 5B,C, the levels of free DNA are well above the limits of detectability so that their values are not subject to this uncertainty, and since they are considered in the global fitting procedure, they

contribute to validation of the value obtained for the data shown in Figure 5A.

It should be noted that these results were obtained using eqs 8 and 9, which have implicit constraints so that no negative concentrations can be obtained. When we used the concentrations so obtained as initial parameter estimates for fitting using eqs 6 and 7, identical results were obtained. However, because of the large number of parameter values involved in global fitting of three data sets, unless initial parameter estimates which were close to the final values were used we frequently observed termination of the fitting by attainment of false local minima with meaningless results. Thus, efficient fitting requires the use of appropriate constraints which are implicit in the model, but which can then be verified by use of an unconstrained model.

When another model where two DNA binding sites per octameric protein was attempted, the data could not be successfully analyzed by that proposed model, increasing our confidence in our model fitting procedures and our interpretation of the results. Therfore, these data indicate that translin forms predominantly octamers in solution with the octamer form of the protein responsible for its DNA binding activity.

DISCUSSION

Translin is a 228 amino acid single-stranded DNA binding protein that binds to specific DNA sequences at recombination hot spots associated with chromosomal translocations (*2, 3, 10*). The mouse TB-RBP shows 99% amino acid identity to translin (*2*). The three amino acids that differ between human translin and mouse TB-RBP are amino acids 49 (alanine to threonine), 66 (glycine to serine), and 226 (valine to glycine). Translin contains two distinct domains: a putative transmembrane helix, which may facilitate interactions between TB-RBP and microtubule proteins (*5*); and a leucine zipper domain in the carboxy terminus (*2, 3*). Given this amino acid homology between the two proteins, it is expected that the functions of the two proteins would be similar. There is compelling evidence that suggests the leucine zipper plays an important role in protein−protein interactions of translin which may be critical for its DNA binding activity (*3, 12, 20*). Mutational analyses of truncated forms of TB-RBP containing amino acids 1−204, 1−175, 1−116, 1−90, and 1−43 are inactive in DNA or RNA binding activities (*12*). These results indicate that the C-terminus of TB-RBP is essential for DNA or RNA binding. However, mutant TB-RBP 177−228 which contains the entire leucine zipper does not bind DNA or RNA, indicating that the leucine zipper alone is not the sole determinant for DNA or RNA binding activity (*12*). Studies with the yeast two-hybrid system demonstrated that the leucine zipper is essential for dimerization (*12*). While a translin octamer is detected by electron microscopy when translin binds to single-stranded DNA (*10*), a dimer of TB-RBP is also reported based on studies with nonreducing SDS−PAGE, glycerol concentration gradient, and gel shift assays (*12*). Although there is a 99% amino acid homology between the two proteins, it is apparent that the two do not function identically. Thus, it is not clear what species of the protein exist in solution and which of these forms is the functional unit for DNA binding activity.

To address these questions, we purified the human translin under native conditions, avoiding any unfolded aggregated proteins, and performed the described analytical ultracentrifugation studies. Both sedimentation velocity and equilibrium studies indicated that translin exists as an octamer without reversible monomers or dimers being detectable under the experimental conditions. To further confirm our interpretation, we have performed gel filtration experiments to isolate the octameric protein only. The chromatograph determined that other forms of translin were not present (data not shown). In addition, monomer was not detectable in the sedimentation studies repeated with the isolated translin octamer. In light of our experimental results, we conclude either that translin predominantly exists as a nondissociating octameric protein or that the association is so strong that free monomer is not detectable under the experimental conditions at a wavelength of 280 nm and also at a wavelength of 230 nm where the sensitivity is approximately an order of magnitude greater.

The stoichiometry and the natural logarithm of the binding constant for the interactions of translin and DNA were obtained from equilibrium sedimentation studies utilizing an FITC-labeled single-stranded DNA. This approach allows the sedimentation equilibrium profile of the oligonucleotide to be scanned at a wavelength of 490 nm in a spectral region where the protein does not absorb, and we were able to calculate the contributions of free and bound DNA from the sedimentation equilibrium profile without using the absorbance at 260 nm, thus avoiding possible complications due to hyper- or hypochromicity effects. Such a two-wavelength analysis places significantly greater constraints on the fitting model and thus ensures greater confidence in the stoichiometry and the accuracy of the value of the natural logarithm of the association constant. From this study, we found that translin octamer binds single-stranded DNA with a 1:1 stoichiometry and the Gibbs free energy change, ΔG°, for the interaction is -9.49 ± 0.12 kcal/mol, which correspond to a dissociation constant, K_d, of 84 ± 17 nM.

Since the results of the mouse TB-RBP studies, which indicated a dimer as the minimal DNA binding species, were not obtained by analytical ultracentrifugation or by any other highly quantitative technique which assesses association in free solution, we do not feel that we can compare our results with the results of these studies. However, our results are in agreement with previous observations reported with electron microscopy (*10*). This suggests that the leucine zipper plays an important role in the formation of the octameric protein which facilitates DNA or RNA binding by translin. In turn, our results may explain the previous mutational analysis studies that show the deletion mutants do not bind DNA. It is conceivable that deletion mutants incapable of forming the proper octameric structure may not support DNA or RNA binding activity.

Previous electron microscopy studies have demonstrated that the native form of translin is a ring-shaped structure with the DNA binding domain being created within the ring structure (*10*). Utilizing both sedimentation velocity and sedimentation equilibrium, we have confirmed that translin exists as an octamer in solution with the only functional DNA binding form of translin being an octamer. These results support the model in which translin forms an annular

structure of eight subunits free in solution capable of binding free single-stranded DNA ends at chromosomal breakpoints. The data obtainable at the present do not permit us to differentiate between the two possibilities of the annular structure surrounding the DNA or, alternatively, the annular structure being entirely external to the DNA. It is envisioned that chromosomal breakage and exposure of the free DNA ends would have to precede binding by the translin octamer. This is further supported by the fact that only one DNA binding site per translin octamer can be fit in the model. Further analysis of the translin octamer with DNA substrates more representative of those that would be encountered by translin in vivo, such as those at the chromosomal breakpoints, will need to be further examined to better understand the role of translin binding at recombination hot spots and chromosomal translocations.

ACKNOWLEDGMENT

We express our appreciation to Dr. Peter Schuck, MIR, DBEPS, ORS, NIH, for performing the Lamm equation analyses of the velocity sedimentation data.

REFERENCES

1. Rabbitts, T. H. (1994) *Nature 372*, 143−149.
2. Wu, X. Q., Gu, W., Meng, X. H., and Hecht, N. B. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 5640−5645.
3. Aoki, K., Suzuki, K., Sugano, T., Yasaka, T., Nakaharam, K., Kuge, O., Omori, A., and Kasai, M. (1995) *Nat. Genet. 10*, 167−174.
4. Han, J. R., Gu, W., and Hecht, N. B. (1995) *Biol. Reprod. 53*, 707−717.
5. Han, J. R., Yiu, G. K., and Hecht, N. B. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 9550−9554.
6. Muramatsu, T., Ohmae, A., and Anzai, K. (1998) *Biochem. Biophys. Res. Commun. 247*, 7−11.
7. Taira, E., Finkenstadt, P. M., and Baraban, J. M. (1998) *J. Neurochem. 71*, 471−477.
8. Wu, X. Q., Lefrancois, S., Morales, C. R., and Hecht, N. B. (1999) *Biochemistry 38*, 11261−11270.
9. Aoki, K., Inbazawa, J., Takahashi, T., Nakahara, K., and Kasai, M. (1997) *Genomics 43*, 237−241.
10. Kasai, M., Matsuzaki, T., Katayanagy, K., Omori, A., Maziarz, R. T., Strominger, J. L., Aoki, K., and Suzuki, K. (1997) *J. Biol. Chem. 272*, 11402−11407.
11. Chalk, J. K., Barr, F. G., and Mitchell, C. D. (1997) *Oncogene 15*, 1199−1205.
12. Wu, X. Q., Xu, L., and Hecht, N. B. (1998) *Nucleic Acids Res. 26*, 1675−1680.
13. Bradford, M. M. (1976) *Anal. Biochem. 72*, 248−254.
14. Perkins, S. J. (1986) *Eur. J. Biochem. 157*, 169−180.
15. Schuck, P. (2000) *Biophys. J. 78*, 1601−1619.
16. Svedberg, T., and Pederson, K. O. (1940) *The Ultracentrifuge*, p 41, Clarendon Press, Oxford, U.K. (Johnson Reprint Corp., NY).
17. Lewis, M. S., Shrager, R. I., and Kim, S. J. (1994) in *Modern Analytical Ultracentrifugation: Acquisition and interpretation of data for biological and synthetic polymer systems* (Schuster, T. M., and Laue, T. M., Eds.) pp 94−115, Birkhaueser, Boston.
18. Kim, S. J., Tsukiyama, T., Lewis, M. S., and Wu, C. (1994) *Protein Sci. 3*, 1040−1051.
19. Durchshlag, H. (1986) in *Thermodynamic data for biochemistry and biotechnology* (Hinz, H. J., Ed.) p 108, Springer-Verlag, New York.
20. Hohjon, H., and Singer, M. F. (1996) *EMBO J. 15*, 630−639.

BI010302T