AN ABSTRACT OF THE THESIS OF

Elizabeth A. Winzeler for the degree of Master of Science in Biochemistry and Biophysics presented on July 20, 1990.

Title: The Effects of pH on the Torsional Flexibility of DNA Bound to a Nucleosome Core Particle

Abstract Approved: _____

Enoch W. Small

The effects of pH on the torsional flexibility of DNA bound to a nucleosome core particle were investigated by studying the time-resolved fluorescence anisotropy decays of ethidium bromide intercalated into the DNA of the core particle. As the torsional flexibility of DNA is affected by the presence of an intercalating dye, the decays were studied at different ethidium bromide to core particle binding ratios. The anisotropy decays were collected using the method of time-resolved single-photon counting and were fit to a model developed by J. M. Schurr (Schurr, 1984) using a non-linear least squares fitting algorithm developed by the author for this purpose. It was shown that below a binding ratio of 0.1 there was no demonstrable change in the anisotropy as a function of binding ratio. Our results show, that the apparent torsional flexibility of DNA of to a nucleosome core particle is dependent on the number of base pairs of the DNA between points of attachment to the histone core. If this number is as high as 30 base pairs, then the torsional flexibility of DNA on a nucleosome core particle is as high or higher than DNA free in solution. Also, for reasonable values of N, the friction felt by the DNA on a core particle is much higher than that felt by free DNA. This indicates that the DNA on a core particle is highly constrained in its motions. The hydrogen ion concentration was shown to have a substantial effect on the fluorescent anisotropy decays, particularly in the early regions of the decay. These analyses indicated that the observed change could be attributed to either a loosening of the contacts between the DNA and the histone core, or a relaxing of the torsional flexibility of the DNA.

The Effects of pH on the Torsional Flexibility
of DNA Bound to a Nucleosome Core Particle

by

Elizabeth A. Winzeler

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Completed July 20, 1990

Commencement June 1991

APPROVED:

Redacted for Privacy
Professor of Biochemistry and Biophysics in charge of major

Redacted for Privacy
Professor of Biochemistry and Biophysics in charge of co-field

Redacted for Privacy
Dean of Graduate School

Date thesis is presented _____ July 20, 1990

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THE EFFECTS OF PH ON THE TORSIONAL FLEXIBILITY OF DNA BOUND TO A NUCLEOSOME CORE PARTICLE

INTRODUCTION

While the advances in understanding the architecture of the cell at the molecular level in the past two decades have been astounding, there still remains a great deal of speculation about how this architecture is constructed and maintained by the subcellular machinery. For example, it is well known that the nucleosome core particle is a fundamental structural component in the body of proteins and DNA which make up chromatin. It is known that the core particle consists of 146 base pairs of DNA wrapped in a negative superhelix around an octameric core of histone proteins (see Figure 1). The core particles are connected to one another by a strand of linker DNA of variable length (about 60 base pairs for chicken erythrocytes). While it is understood that this long strand of DNA, periodically twisted around globular proteins in a linear array, is somehow coiled into the ordered structure which makes up the eukaryotic chromosome, it is not understood how genes on the highly ordered structure are differentially expressed or how DNA synthesis is initiated. The problem facing the molecular biologist now is to determine the possible mechanisms which would allow for events in the life of the cell to occur.

There are a number of approaches the molecular biologist may take to elucidating the mechanism involved in such phenomena, in large part dependent on the training he or she has received. If the molecular biologist identifies himself or herself principally as a biochemist then it is natural to look for enzymes which would mediate such processes. The biophysicist, however, being a physicist at heart and therefore having a fundamental conviction that truth is beauty, may feel uncomfortable in looking for the truth in such a complex and ugly entity as an enzyme and will therefore look for simpler and more physical explanations of phenomena. While it has been the biochemist who has been most often rewarded, as very sophisticated enzyme complexes

capable of the elaborate activities have been discovered, the biophysicist has never-the-less, perhaps blinded by idealism, continued to search for elegant physical mechanisms by which subcellular events could be initiated or directed.

One of the phenomena for which a complete understanding is still lacking and which is an area of active interest to both biophysicist and biochemist is eukaryotic transcription. Discussion of this subject includes the questions of how polymerases recognize the DNA to be transcribed and what the relationship is of the histone protein component of the chromatin to the transcription process. There has been some evidence in recent years that in eukaryotes, as in prokaryotes, transcription is to some degree modulated by variation in the torsional stress of the DNA (van Holde, 1988). This scenario suggests the role of an enzyme, perhaps a eukaryotic gyrase, and many biochemists are searching for such an enzyme. Biophysicists however are investigating simpler mechanisms. Is it possible that changes in the physical environment of the core particle could induce the particle to undergo a conformational transition which would affect the torsional flexibility of the DNA? It has been established that the nucleosome core particle unfolds at both high and low ionic strengths. Could this unfolding play a possible part in the process of transcription or translation? While the significance of such transitions with respect to the process of transcription is unclear, the very existence of such transitions does warrant investigation.

There have been a number of studies of transitions in the nucleosome core particle. Researchers have investigated the effects of temperature and denaturing agents on the conformation of the core particle (Olins *et al.* 1977; McMurray *et al.*,1985) The low salt transition has been most extensively studied (Libertini and Small, 1987b; Libertini and Small, 1982; Brown *et al.* 1990a; Gordon *et al.* 1978; Libertini and Small, 1980). To a lesser extent the affects of changes in the hydrogen ion concentration on the conformation of the core particle have also been investigated (Zama *et al.* 1978b; Libertini and Small, 1984a; Libertini and Small, 1982; Muller *et al.* 1885; Gordon *et al.* 1979; Kawashima *et al.* 1982). Among some of the biophysical techniques used to study the low salt transitions are sedimentation (Zama *et al.,* 1978a; Gordon

et al., 1979), flow birefringence (Harrington, 1981), electric dichroism (Wu *et al.* 1979), intrinsic tyrosine fluorescence (Libertini and Small, 1982) and the time-dependent fluorescence anisotropy of ethidium bromide intercalated into the DNA of the core particle (Brown *et al.,* 1990a). The latter technique has been useful in documenting changes in the gross morphology of the core particle, which are easily discerned in the changes in the rotational diffusion coefficient. This is because the rates of Brownian rotation depend strongly on the size and shape of the particles. While such a method might seem to be a fitting candidate for investigating other transitions, especially those involving the DNA, there is some difficulty in determining how the data produced from such an experiment should be analyzed. The anisotropy function, if the internal motions of the probe are neglected, is generally fit to a sum of exponentials (Belford *et al.* 1972; Chuang *et al.* 1972; Ehrenberg *et al.* 1972), the lifetimes of which are related to the dimensions of the molecule undergoing rotational diffusion (Small *et al.* 1988b). Such an approach has been successfully applied by several investigators to study the low salt transition (Brown *et al.* 1990a; Small *et al.* 1990), ethidium binding (Genest *et al.* 1982) and DNA flexibility (Ashikawa *et al.* 1983). However, this is a simplistic approach for studying more subtle changes in the core particle because it does not take into account the movement of the DNA independent of the rotational diffusion of the core particle. For example, if the DNA is loosely bound to the core particle, twisting motions of the DNA may contribute to depolarization of the fluorophore resulting in a shorter apparent lifetime. Secondly, a sum of exponentials provides little information about the physical properties of the DNA. The activity of the DNA is of special interest because physical changes in the conformation of the DNA on the nucleosome core particle might also provide a mechanism by which active genes are recognized by the polymerase or other DNA binding proteins. It is easy to see that if one uses ethidium bromide, a dye which intercalates into the DNA, then the depolarization of the fluorescence signal could yield much information about the local environment of the DNA. The speed at which depolarization occurs might indicate whether the DNA on a core particle has as much torsional freedom as DNA free in solution.

Working from an expression developed by Barkley and Zimm (Barkley *et al.* 1979), J. M. Schurr and co-workers have developed a theoretical expression for the anisotropy function of DNA bound to a core particle. By fitting the actual anisotropy function to this expression, one can determine values in the torsional flexibility of DNA. In order to describe the flexing motions of the DNA, Schurr *et al.* proposed a model of the nucleosome core particle where a filament consisting of 146 rigid rods connected by torsional springs is constrained to girdle the axis of a sphere and is rigidly clamped at both ends (Schurr, 1984). The model yields the following expression for the fluorescence anisotropy decay of such a particle as a function of time:

$$r(t) = r_0 \exp(-6D_{sph}t) \left\{ \left[\frac{3}{2}\cos^2\varepsilon_0 - \frac{1}{2} \right]^2 + 3\cos^2\varepsilon_0 \sin^2\varepsilon_0 \, C_1(t) + \sin\varepsilon_0 C_2(t) \right\} \quad (1)$$

Here D_{sph} is the rotational diffusion coefficient of a sphere and ε_0 is 70.5°, the angle between the transition dipole and the helix axis (Schurr and Schurr, 1985). The first term in the expression predicts the depolarization due to rotational diffusion. The second half of the expression is dependent on the twisting correlation functions, $C_n(t)$, which are given by:

$$C_n(t) = \frac{1}{N} \sum_{m=1}^{N} \exp\left[-n^2 \sum_{l=1}^{N} d^2_l \, Q^2_{ml}[1 - \exp(-t/\tau_l)] \right], \quad (2)$$

where

$$Q_{ml} = [2/(N + 1)]^{1/2} \sin[ml\pi/(N + 1)], \quad (3)$$

$d^2_l = (k_B T/\gamma)\tau_l$, $\tau_l = \gamma/(4\alpha\sin^2[l\pi/2(N + 1)])$, N is the number of rigid rods, k_B is Boltzman's constant, and T is the absolute temperature (Schurr 1984). The twisting correlation functions are dependent on α, a constant describing the properties of the torsion spring constant between the rigid rods, and γ a frictional factor which is dependent on solvent viscosity. These functions predict the depolarization due to torsional twisting motions of the DNA. The

depolarization which would result from dye wobble, or overdamped flexing and bending motions are considered negligible. The anisotropy decay depends linearly on r_o, the anisotropy at zero time.

Within the restrictions of the model, analysis of the anisotropy decay can yield parameters which describe of the torsional flexibility of DNA. Such a number is rather useless without a basis for comparison. We have found its principle utility is in quantifying the extent of change as a function of some physical parameter. To be more specific, we have used it to determine whether there exists any conformational transition in the DNA on a core particle as a function of pH which can be revealed in terms of the parameters α, γ, or D_{sph}.

Why study pH? Intracellular pH is by no means constant throughout the cell cycle or within various organelles of the cell. A rapid rise in pH has been detected intracellularly using ^{31}P NMR following activation of eggs from *Xenopus laevis* (Webb and Nucitelli, 1982), and just prior to maximal DNA synthesis in yeast (Gillies, 1982). There are numerous other examples (for a review see Nucitelli and Heiple, 1982). More recently, Ober and Pardee (1986) have shown that mitogens or growth factors activate the cellular Na/H^+ antiporter system and that such activation results in a significant change in the intracellular pH. Accompanying this rise in pH, is the commencement of mitotic events, including DNA synthesis. This same researchers suggest that the tumorigenic capacities of certain cell lines may result from an inability to regulate pH. While one could argue strongly that the pH is affecting the optimal activity of an enzyme associated with the synthesis machinery, one should still consider the possibility that the change in pH causes a conformational change in the DNA, allowing polymerases to recognize the origin of replication. Recently, it has been shown that the solubility of chromatin is dependent on pH, suggesting that a rise in pH might lead to a decondensation of the chromatin prior to mitosis. Additionally, rises in pH have been correlated to increases in protein synthesis (Grandin and Charbonneau, 1989). Though, it now seems most likely that in the case of protein synthesis, this is due to pH-dependent phosphorylation of the ribosomal subunit (Chambard and Pouyssegur, 1986). It may also be true that changes

in pH may occur only in the cytoplasm and not in the nucleus where they might have some bearing on the conformation of core particles, such phenomena, none-the-less, are illustrative of the fact that fluctuation in intracellular pH occurs and may promote an activity such as protein synthesis. While it may seem that the decision to study the effects of pH on a core particle was preconceived resulting from the aforementioned considerations, in actuality the decision resulted from the chance discovery that the time resolved anisotropy decay of ethidium bound to the DNA on a nucleosome core particle was highly dependent on pH.

In this study the Schurr model will be used to investigate the changes in the anisotropy decay of nucleosome core particles in terms of the torsional rigidity of DNA on a core particle. Ethidium bromide, which intercalates between base pairs will be used as a probe. The dye is very convenient to use, and its interaction of this dye with DNA has been studied in detail. There is some question as to whether the ethidium binding affects the torsional rigidity of the DNA. Ethidium is known to cause dissociation of the core particle at high binding ratios (McMurray *et al.* 1986; McMurray *et al.* 1990b). For this reason the effect of binding ratio on the anisotropy decay will also be studied. We demonstrate that if a low enough dye to core particle binding ratio is used, the results will be independent of the binding ratio. A great deal of effort has been expended in investigating new methods of data analysis and collection. A new instrument which has not previously been described was also used. For this reason it is pertinent to discuss these aspects of the experiment in some detail.
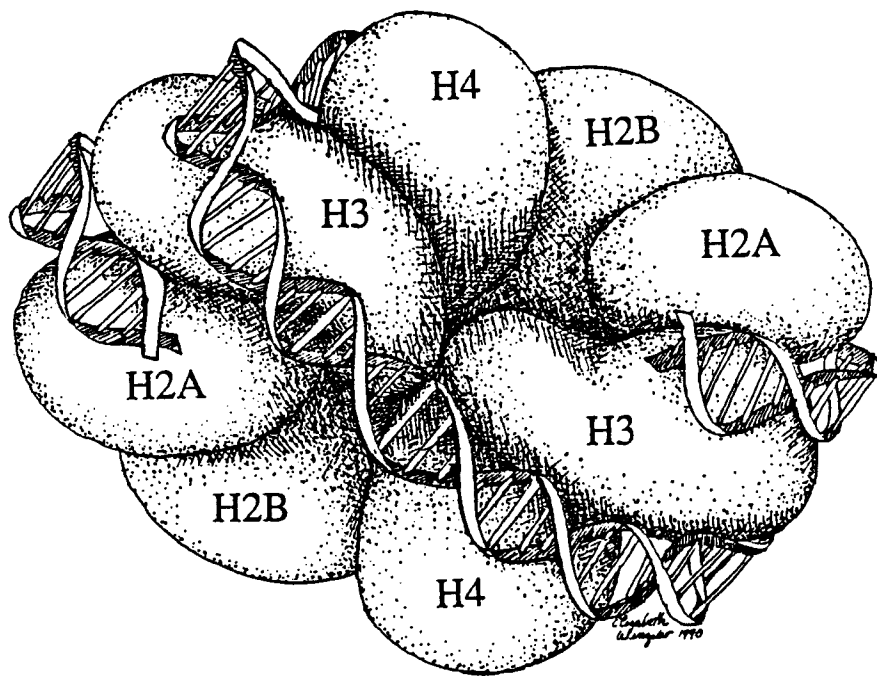
Figure 1. Rendition of a Nucleosome Core Particle Showing Relative Positions of DNA and Octameric Core of Histones.

MATERIALS AND METHODS

1. The Instrument

The experiments presented here have all been conducted using a technique known as time-resolved polarized fluorescence spectroscopy. In conducting a time-resolved fluorescence experiment, the object is to determine the number of photons emitted as fluorescence per unit time following the excitation pulse. The method used in the laboratory has been termed "time-correlated single photon counting." This method was first developed by Schuyler and Isenberg (Schuyler *et al.* 1971). More recent descriptions can be found elsewhere (Davie *et al.* 1981; O'Connor *et al.* 1984). The measurement commences when a pulse of light is emitted by the dye laser. The pulse boosts dye molecules in the sample under study to a higher energy level. Subsequentially, the dye molecule may emit a photon as fluorescence. There is a small chance that a photon emitted as fluorescence will reach the photomultiplier tube whereupon a *start* signal is generated by the photon arriving at the photomultiplier tube. The *stop* is generated by a laser pulse. The time elapsed between the start and the stop signal is converted to a voltage by a time to amplitude converter (TAC). Different voltages correspond to different memory locations in a pulse height analyzer (PHA). A value of 1 is added to some memory location every time a photon arrives, and thus a histogram showing the relative number of photons emitted as a fluorescence as a function of time is generated.

The instrumentation has three basic components: the laser which provides the source of the excitation pulse; the fluorometer which holds the sample, lenses and polarizers which allow for manipulation of either the excitation or emission intensity of polarization; and the electronics which detect the fluorescence and process the signal. Though the fluorometer has been constructed recently and has not been previously described, the laser and the electronics have been described in greater detail than will be provided here (Small and Anderson, 1988a; Small *et al.* 1984).

The Laser Source

The laser source consists of a cavity-dumped Spectra-Physics dye laser whose output is mode locked at about 82 MHz and cavity dumped at 800 kHz. The dye laser is synchronously pumped by a frequency-doubled Spectra-Physics Nd:YAG laser (Small 1989b). A special dye, rhodamine 575, is used in the dye laser to extend the response to the blue (Libertini and Small, 1987a). All experiments were performed with an excitation wavelength of 556 nm, normally not obtainable with this laser system.

The Fluorometer Optics

A diagram of the optical system is shown in Figure 2. The output of the dye laser is directed off of two front surface mirrors and through a fused silica beam splitter which diverts a fraction of the beam into a photodiode. The signal from this photodiode is used to provide the stop pulse in experiments examining time ranges less than 384 nanoseconds, however, it was more usual to use a stop pulse originating directly from the cavity dumper electronics. The intensity of the remainder of the beam is selectively diminished by the positioning of one or two neutral density filters, of varying intensity, in its pathway and by the use of a crystalline quartz half wave plate in a rotating mount. The half wave plate provides two functions. First, since the laser output is vertically polarized, the half wave plate rotates the plane of polarization generating a horizontal intensity component needed for making sensitivity corrections (described below). Also, since the half wave plate is followed by a linear polarizer, rotating the plane of polarization by rotating the plate varies the throughput of the system providing a fine adjustment of incident intensity. It is necessary to attenuate the incident light so as to minimize the possibility of a double photon event occurring (Schuyler *et al.* 1971*)*. If the intensity of the incident light and therefore the resulting fluorescence is too great then there is some probability that two or more photons will reach the photomultiplier tube within a given pulse cycle. Only the first photon will be registered when this happens, and if it happens often enough, the decay will appear to have a shorter lifetime than it actually does.

Following, the optional filters, the beam is directed off a 45 degree fused silica prism and into the half wave plate. The light then enters a box designed to allow entry of light emanating from the laser.

The fluorometer optical unit used in these experiments was recently constructed and has not been previously described in the literature. The main object in the conception of the design for this fluorometer was to minimize broadening of the impulse response function. The impulse response function is the true response of the instrument to an infinitely fast excitation pulse. For the excitation pulse to be infinitely fast all photons would need to leave and arrive at exactly the same time. Because the speed they travel is fixed, discrepancies in their arrival time result when some take a slightly longer route. Some possible sources for such deviations in path length include reflections off the sides of the cuvette and imperfections in focusing lenses. If a photon from the laser were to be reflected off the far side of the cuvette due to the (15%) difference in the refractive index between quartz and water and back into the solution before encountering a fluorophore, its pathlength would be 4 mm longer than that of a photon which was not reflected. The excitation event would be broadened by 53 picoseconds in this case. Fluorescence can take a peripatetic path to the PMT as well. Fluorescence is emitted from the sample chamber as a cone of light. This light is then focused onto the cathode of the photomultiplier tube with a convex lens. Light reaching the outer edges of the lens will have traveled a greater distance than light arriving at the center. The lens is thicker in the middle in order to compensate for this as it will take longer for light to pass through the thicker part of the lens. Here, spherical and chromatic aberrations will result in some photons suffering a delay and must be minimized. Such effects may to some degree be avoided by the use of high quality lens and cuvettes and the strategic placement of baffles to absorb deviant photons. As a result of such efforts the profile of the excitation pulse as a function of time for this instrument is a clean sharp signal with a half width of about 80 ps (see Figure 4).

The first optical device within the sample chamber is a single Glan Taylor prism polarizer. Apertures were placed before and after the polarizer to reduce

stray light. The sample holder, designed to hold a cuvette with a 4mm pathlength, consists of an anodized aluminum, water cooled container with an entry and exit aperture, through which the laser beam will pass, and an aperture perpendicular to the direction of the laser beam through which fluorescence can be observed. After passing through the cuvette, the laser beam enters a light trap where it is absorbed. A double Glan Taylor prism polarizer is placed directly against the face of the cuvette holder that is at right angles to the laser beam. A filter holder is placed immediately behind the polarizer. When collecting a profile of the excitation pulse a narrow band pass filter occupied this position, its primary function being to limit the trace of unwanted fluorescent light emitted by the scatter sample (Libertini and Small, 1983). A combination of an interference filter and a glass cutoff filter was used in the collection of fluorescent decays. These filters were chosen so their bandwidth and peak transmittance would ensure that the photons recorded originate from fluorescence and not from the excitation pulse or an impurity. The light emitted from the sample is passed through an iris diaphragm aperture, whose function is to control the intensity of the image and as well as to decrease the amount of improperly polarized emissions reaching the photomultiplier tube. A lens focuses the image of the laser beam in the sample onto the photocathode of the double microchannel plate photomultiplier tube (Small, 1989b). Additionally a system of baffles has been placed within the box to prevent stray scattered light from reaching the photomultiplier tube before or after the main pulse.

The primary use of the fluorometer was in collecting fluorescence anisotropy decays. The anisotropy function will be described in greater detail in the description of data analysis procedures. For these measurements two decays are collected, one with the emission polarizer in the vertical position and one with emission polarizer in the horizontal position. The excitation polarizer is held at vertical for both these measurements. The time-dependent anisotropy function r(t) is basically a measure of the difference between the parallel and perpendicular components of the fluorescence decay divided by the total intensity as a function of time. See DATA ANALYSIS for more details. For these measurements it is critical that no light of the wrong polarization

reach the photomultiplier tube. If the emission polarizer failed to screen all but the desired orientation of light, then the resulting fluorescence would have less initial polarization. This could be misinterpreted as the result of fast motions of the fluorophore and would be most apparent at t=0. For a molecule whose absorption and emission dipoles are parallel one can compute that the theoretical maximum value of the anisotropy at time = 0 would be 0.4. While this maximum is seldom achieved for ethidium, presumably because of a non-colinearity in the absorption and emission dipoles, never-the-less, we viewed maximizing the initial anisotropy as a substantial test of the new instrument ability to allow only the correct polarization to reach the PMT. Thus, several experiments were conducted to ensure that the instrument was capable of collecting anisotropy decays whose peak value would approach 0.4.

The Fluorometer Electronics

The electronics allowing us to make these measurements have been exhaustively described, (Hutchings *et al.* 1990; Small *et al.* 1988a), but owing to changes in certain components, a rough outline of the current configuration will be presented. Figure 3 diagrams the electronics used for these measurements. Central to the process of data collection is the Hamamatsu R1564U triple microchannel plate photomultiplier tube (PMT). A thorough description of this device and its applications to fluorescence studies is given in Small (1989b). Simply stated, it works as follows: When a photon of sufficient energy is absorbed by the photocathode in an evacuated tube, the energy of the photon may be imparted to an electron which has a probability of being expelled from the surface of the photocathode, a phenomena known as the photoelectric effect. The electron is accelerated toward the anode which is positively charged. Its arrival at the anode results in a small pulse of current which will, after amplification, eventually be interpreted as the start event.

The signal from the PMT subsequently is directed into a Phillips Scientific 6955 amplifier which splits the pulse into two components, a simple charge which is ultimately used for energy windowing, which will be described in greater detail later, and a 50 Ω timing pulse which will provide the start signal

in the generation of the histogram. The timing pulse is fed into a Phillips Scientific Model 6915 constant fraction timing discriminator. Since the incoming pulses vary considerably in amplitude, a standard 5 volt timing pulse is generated at the time the incoming pulse reaches 20% of its final amplitude. This reduces time jitter. The TAC , an Ortec model 467, generates a signal whose voltage increases proportionally to the amount of time elapsed between the arrival of the start signal from the photomultiplier and the stop signal.

The stop signal may be supplied to the TAC directly by the cavity dumper or it may be supplied by a signal originating with the incidence of a fraction of the next excitation pulse from the laser upon a Spectra Physics 403B high speed photodiode. The output of the photodiode is amplified with a Phillips Scientific fast pulse amplifier Model 6954 and then fed into an Ortec model 934 Constant Fraction Discriminator (CFD). The signal from the CFD may be delayed for varying amounts of time using a Science Accessories Corporation Model 033, dual nanosecond delay box, or by its passage through variable amount of 50 Ω cable on its journey to the TAC. The delay is necessary because the photodiode event used to stop the TAC is initiated before the arrival of the fluorescent photon at the photomultiplier. Changing the delay moves the position of the histogram in the PHA. Using this method, the data are collected in reverse and are later inverted in the computer.

In our experiments, the intensity is controlled such that the probability of the photomultiplier responding to a photon is only about 0.03. per pulse cycle (Hutchings et al. 1990). However, despite this low number, there still exists a significant probability (0.0009) that two photons will reach the photomultiplier in a pulse cycle. When two photons are detected in a pulse cycle, the first one will be used to provide the start signal while the second will be ignored. The consequence of this happening regularly will be an apparently shorter decay lifetime. Energy windowing is a technique used to diminish the possibility of such errors occurring. Using this method, all photons arriving at the photomultiplier tube during a time span are converted to voltage then and integrated over that time range. If the integral is outside the window, that is, greater than some limit (corresponding to a double photon event), or less than

some limit, (corresponding to noise), a negative logic pulse is sent to the pulse height analyzer instructing it to ignore a signal arriving simultaneously from the TAC. In order to accomplish this, the charge emitted from the Phillips 6955 amplifier enters an Ortec 113 preamplifier where the charge is converted to a long duration voltage pulse. An Ortec 485 amplifier receives the pulse from the preamplifier and converts it to a bipolar pulse whose height is proportional to the amount of signal arriving at the photomultiplier for a given period of time. Finally an Ortec 420 single channel analyzer determines whether the bipolar pulse arriving from the amplifier falls within the limits of the energy window and sends an NIM logic pulse to the pulse height analyzer.

The Tracor Northern 1750 Pulse Height Analyzer (PHA) receives signals simultaneously from the Ortec 420 SCA, and the TAC and decides whether to process a signal from the TAC depending on the information received from the Ortec 420. If the signal is good, an increment is added to the memory location corresponding to the voltage of the incoming signal.

The PHA will retain a limited amount of data in memory. The data in 1024 channel memory blocks are periodically transferred to an Apple II computer which is in direct communication with an IBM model 80 microcomputer. The model 80 functions to convert the data to the permanent storage format as well as to serve as a permanent storage receptacle for the data. In addition, it contains a number of utility programs which display, process, and analyze the data. Additionally, a Tektronix PEP 303 is used for data processing.
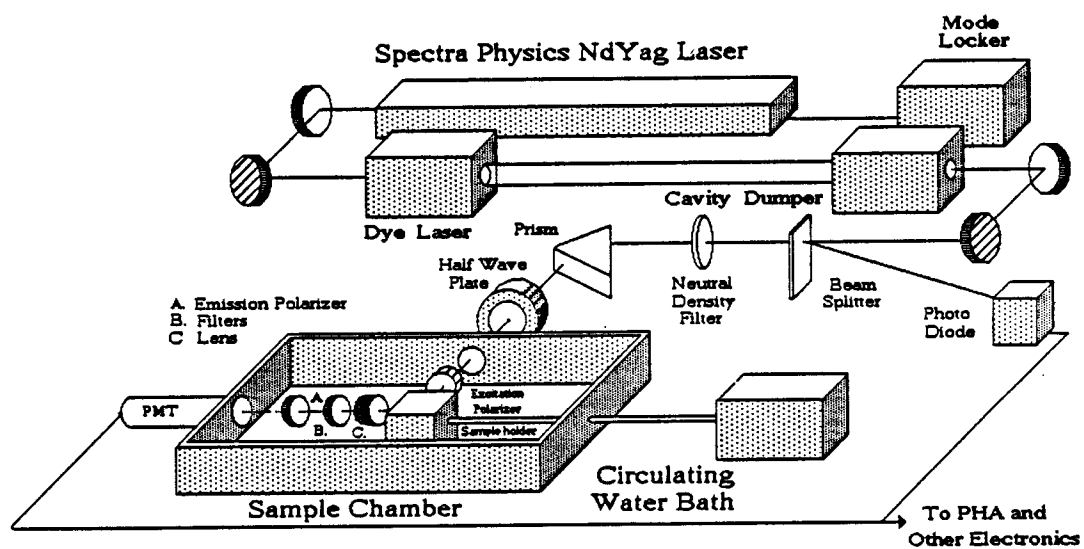
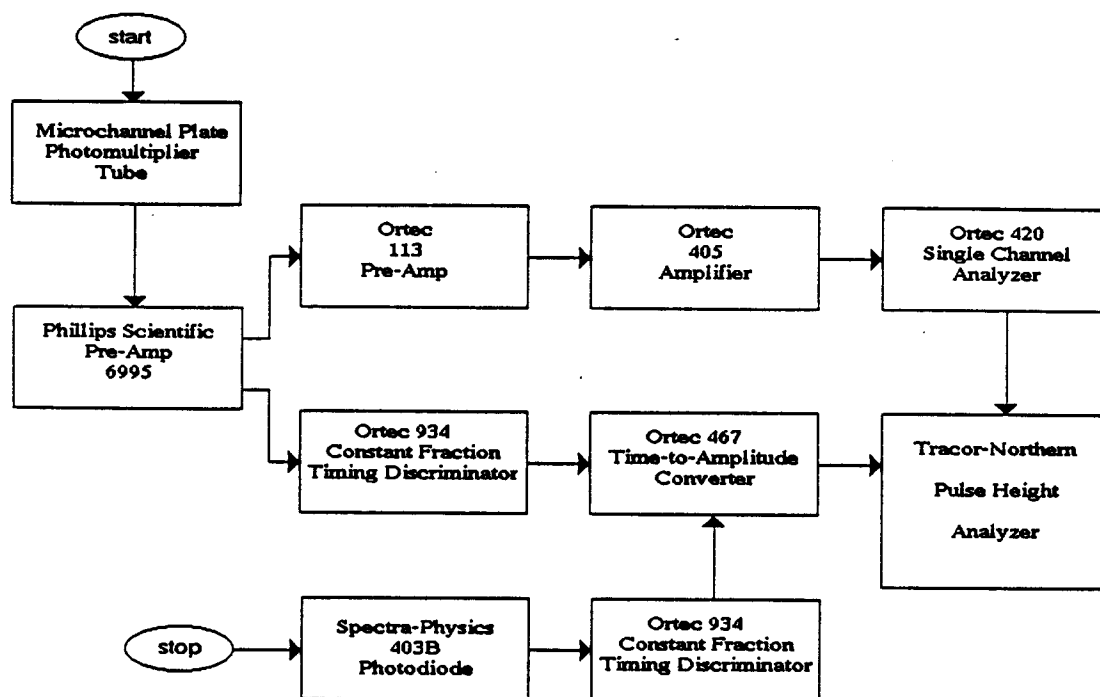Figure 2. <u>Diagram of Instrumentation Used for the Measurements Described in These Experiments.</u>

Figure 3. Diagram of Electrical Circuitry Used in our Laboratory.

2. Data Analysis

In the course of these experiments a great number of histograms depicting the fluorescence decays of ethidium bound nucleosome core particles were collected under many different conditions. The effect of changing the conditions in the sample were often immediately visible on the screen of the PHA. Showing such histograms to a fluorescence spectroscopist might elicit excitement but would probably elicit little more than confusion in the researcher whose primary interest is chromatin and not fluorescence. The challenge, therefore, is to describe these changes in terms of physically meaningful parameters which are useful to the biologist as well as the fluorescence spectroscopist. For this purpose a number of software routines were developed and tested to analyze the data. While a few of the routines existed previously the majority were written expressly for the tasks at hand and apply only to the special case of the fluorescence decay of ethidium bound to a nucleosome core particle. Most of the innumerable programs which have been written and used at least once in the course of these experiments fall into three general categories based on their function: Basically there are routines which 1) apply a numerical correction to the convolution artifact, 2) fit the data, or some derivative of it, and 3) present the data or the results of an analysis. The latter category is not relevant to the science of these experiments and shall not be discussed further. The other two, however, do deserve some elaboration.

A Numerical Correction of the Convolution Artifact:

One can see from a plot of the reference function, (see Figure 4), that while ideally one would find a single very narrow peak, in actuality one also sees a number of smaller secondary peaks. While the separation between principle peaks is long enough that the fluorescence resulting from one excitation pulse will decay to zero long before another excitation pulse is generated, the time between secondary peaks is not, resulting in continuous repetitive low level stimulation of the fluorophores. The numerical correction allows us to ameliorate the effects of the secondary peaks. An integral of each of the

secondary scatter peaks in the reference function, E(t), is calculated. The amount of fluorescence which would result from a peak of given size is calculated based on the amount resulting from the primary peak. This amount is subtracted from the perpendicular and parallel components of the decay. After numerical correction of the data, it is possible to fit the data without using the cumbersome convolution routines which will be described shortly in the context of data fitting.

Fitting the Data

a. The General Case.

For our work we have investigated fitting an expression (equation 1) developed by J. M. Schurr (Schurr 1984) for the theoretical anisotropy of a nucleosome core particle to the anisotropy data. This expression is dependent on a number of parameters, including a torsional coefficient, α; a frictional coefficient, γ; the rotational diffusion coefficient, D_{sph}; r_o, the anisotropy at time zero; and the number of base pairs, N, between binding sites, all of which may be treated as variable to achieve a good fit.

In general, measured anisotropy data may be computed from:

$$r_m(t) = \frac{f_{||}(t) - f_{\perp}(t)}{f_{||}(t) + 2 f_{\perp}(t)} = \frac{d(t)}{s(t)}. \tag{4}$$

Here f(t) is the true fluorescence which we call the impulse response function. The impulse response function is what would be measured if one had an infinitely narrow excitation pulse and an instrument which could respond infinitely fast to the fluorescence. We will assume that the impulse response is related to the observed fluorescence response by the convolution:

$$F(t) = E(t) * f(t), \tag{5}$$

where E(t) is the reference function (Small *et al.* 1989a). This is necessary because as the excitation pulse is not infinitely short, the solution will begin to fluorescence before the excitation pulse has been completed. If the data have been numerically corrected then it is assumed that $F(t) = f'(t)$ which can be used interchangeably with $f(t)$ to some extent. The prime in $f'(t)$ indicates that this is not the true instrument response, but only an approximation. One can then solve directly for the values of the parameters in equation 1 which will produce the best agreement between the theoretical expression and the anisotropy data computed with (4).

The fitting procedure, a Marquardt non-linear least squares analysis (See Marquardt (1963)), used in these experiments was adapted from Bevington (1969). Discussion will be confined to the use of the method in fitting the anisotropy directly, but the same technique can easily be adapted to fitting any function. The steps involved are as follows:

1. Using some initial values of the parameters, a_j, where a_j may be α, γ, r_0, or D_{sph} * , one calculates

$$r(t) = r_0 \exp(-6D_{sph}t) \left\{ \left[\frac{3}{2} \cos^2\varepsilon_0 - \frac{1}{2} \right]^2 + 3\cos^2\varepsilon_0 \sin^2\varepsilon_0 \ C_1(t) + \sin\varepsilon_0 C_2(t) \right\} \quad (6)$$

as well as a measure of the goodness of the fit of $r_c(t)$ to the experimental data, $r_m(t)$:

$$\chi^2 = \sum_{i=1}^{npts} \left\{ \frac{1}{\sigma^2} [\ r_m(t_i) - r_c(t_i)]^2 \right\}. \quad (7)$$

Here

---

* N is treated as variable; however it is impossible to apply the method described here for fitting N. N can only take integer values, and therefore no adequate form of the derivative is available for the computations described here. To determine N one must fit the other parameters at consecutive values of N. The analysis which produces the lowest value of χ^2 reveals the approximate value of N.

$$\sigma(t) = \frac{(f^{''}_{||}(t) + f^{''}_{\perp}(t))(f^{''}_{||}(t) + 2f^{''}_{\perp}(t))^2 + (f^{''}_{||}(t) - f^{''}_{\perp}(t))^2(f^{''}_{||}(t) + 2f^{''}_{\perp}(t))}{(f^{''}_{||}(t) + 2f^{''}_{\perp}(t))^4}, \quad (8)$$

and *npts* is the difference between the number of data points used in the fitting routine and the number of degrees of freedom. The number of degrees of freedom is equal to the number of parameters being fit in a particular analysis. By performing a first order Taylor's expansion of (1) with respect to each of the n parameters to be fit, one can approximate the $r_m(t)$, (4), collected experimentally such that:

$$r_m(t) \cong r_c(t) + \sum_{j=1}^{n} \frac{d(r_c(t))}{da_j} \delta a_j. \quad (9)$$

2. One then determines the optimal values of δa_j at which χ^2 changes minimally with respect to the parameter increment δa_j. The value of delta can be determined by taking the derivative of χ^2 with respect to δa_j, and setting it equal to zero:

$$\chi^2 = \sum_{i=1}^{npts} \left\{ \frac{1}{\sigma^2} \left[r_m(t_i) - r_c(t_i) - \sum_{j=1}^{n} \frac{d(r_c(t_i))}{da_j} \delta a_j \right]^2 \right\}, \quad (10)$$

$$\frac{d(\chi^2)}{d(\delta_k)} = -2 \sum_{i=1}^{npts} \left[\frac{1}{\sigma^2} \left[r_m(t_i) - r_c(t_i) - \sum_{j=1}^{n} \frac{d(r_c(t_i))}{da_j} \delta a_j \right] \frac{d(r_c(t_i))}{da_k} \right] = 0. \quad (11)$$

These expressions yield the following set of linear equations whose solution are the optimal values of the parameter increments.

$$
\begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \cdot \\ \cdot \\ \varphi_n \end{bmatrix} = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \cdot \\ \cdot \\ \delta_n \end{bmatrix} \begin{bmatrix} \beta_{11} \beta_{21} \cdot \cdot \beta_{n1} \\ \beta_{12} \beta_{22} \cdot \cdot \beta_{n2} \\ \cdot \quad \cdot \cdot \cdot \cdot \cdot \\ \cdot \quad \cdot \cdot \cdot \cdot \cdot \\ \beta_{1n} \beta_{2n} \cdot \cdot \beta_{nn} \end{bmatrix}, \tag{12}
$$

Where

$$
\beta_{j,k} = \sum_{i=1}^{npts} \frac{d(r_c(t_i))}{da_j} \frac{d(r_c(t_i))}{da_k} \frac{1}{\sigma^2} \quad \text{and} \tag{13}
$$

$$
\varphi_j = \sum_{i=1}^{npts} \frac{1}{\sigma^2}(r_m(t_i) - r_c(t_i)) \frac{d(r_c(t))}{da_j} . \tag{14}
$$

3. After solving for δa_j, these increments are added to the initial values of the parameters a_j. Step 1 is performed again using the new values of a_j.

If χ^2 decreases, steps 2 and 3 are performed again. When successive applications of steps 2 and 3 fail to produce any decrease in χ^2 the search is terminated assuming that the minimum in the χ^2 surface has been found. Using clean data and analytical expressions for the derivatives, χ^2 will approach a minimum in 3 or 4 iterations.

b. Global Analysis

Occasionally one will work with a set of data which contains insufficient information to allow accurate determination of a particular parameter. Such a situation is encountered when analyzing very short time ranges, an area rich in information about the torsional and frictional coefficients but deficient with respect to information about the rotational diffusion coefficient. Under such circumstances a technique termed global analysis (Beechem *et al.*,1986; Eisenfeld *et al.* 1979; Beechem and Gratton, 1988; Small *et al.*, 1989a) is used.

This method can be used when several of the datasets to be fit theoretically have one or more parameters in common. For instance, one would expect that the number of base pairs between points of attachment would be constant regardless of the study conditions. One also might predict that 2 or 3 fold changes in the binding ratio of ethidium to core particle would have little effect on the rotational diffusion coefficient. One could therefore link the information on the diffusion coefficients from a number of files to obtain, if not a better fit, a truer value of this parameter. Thus, as an alternative to analyzing a number of datasets sharing one or more characteristics independently and averaging the recovered shared characteristics, one can analyze the data globally.

Global analysis is easily accomplished. The number of parameters returned by this program is variable, depending on the number the decays analyzed as well as the similarity of the decays. Each decay to be analyzed has four parameters associated with it. Each parameter has three descriptive variables associated with it: The first is the type of parameter, *i.e.* whether it is α, γ, r_0 or D_{sph}. The second is whether it is to be considered constant or variable. Values of Constants are passed to the program at the start and values of variables are returned at the end. The third is the linkgroup to which a parameter belongs. This is nothing more than the address in the computer's memory of the value of the parameter. If two datasets share a characteristic such as the rotational diffusion coefficient then the value of D_{sph} as well as the address of this value is common to the two. If the characteristic is unique to the decay, then the address is unique as well. Once these assignments have been made it is possible to construct a matrix equation as described above. Parameters whose values are constant are not included in the matrix. If there are L possible combinations of linkgroup and type for parameters whose value is variable then there will be L parameter increments δ_j; L unknowns, a_j; and L rows in each matrix. The jth and kth element in the B matrix and jth element in the Φ (equations 13 and 14) matrix are respectively given by

$$\beta_{j,k} = \sum_{m=1}^{n} \sum_{i=1}^{npts} \frac{d(r_c(t))}{da_j} \frac{d(r_c(t))}{da_k} \frac{1}{\sigma^2} , \qquad (15)$$

$$\varphi_j = \sum_{m=1}^{n} \sum_{i=1}^{npts} \frac{1}{\sigma^2}(r_m(t_i) - r_c(t_i)) \frac{d(r_c(t))}{da_j} , \qquad (16)$$

where M is the number of datasets analyzed in the program. It should be apparent that for each combination of linkgroup and pair, the derivative for a dataset will be zero unless one of the parameters associated with that dataset is variable and described by that combination. The rest of the fitting procedure is identical to that described above.

Such descriptions may seem unduly complicated. There are easier ways of thinking of the problem. For an independent analysis of a given dataset, a matrix equation, (12), is produced as described above. In order to work globally one simply sums all of the matrix equations from all the datasets included in the study to create a global matrix equation analogous to the one given in (12). More elegant and theoretical discussions of this technique can be found elsewhere (Beechem et al.,1986; Eisenfeld et al. 1979; Beechem and Gratton, 1988; Small et al., 1989a). Though time consuming, this method of analysis produces excellent results.

c. Deconvolution

As we are interested in the very fast motions of the DNA we have occasionally collected data over very short time ranges, usually about 10 nanoseconds. In this case a significant part of the fluorescence decay overlaps the measured excitation pulse, narrow as it may be. Applying a numerical correction to the data is not sufficient in this case and because if it is not done F(t) it is no longer considered equivalent to f(t). It is no longer possible to compute the anistoropy data directly from $F_{\parallel}(t)$ and $F_{\perp}(t)$. In order to more fully

understand the decay at its initial stages, it is necessary to take a convoluted approach to determining the best values of the parameters α, γ, D_{sph} or r_0. The method is essentially the same as the general case but with a few modifications.

From definition (5) it follows that

$$F_{||}(t) - S\,F_{\perp}(t) = E(t)*f_{||}(t) - E(t)*f_{\perp}(t), \tag{17}$$

which is equivalent to

$$F_{||}(t) - SF_{\perp}(t) = E(t)*(f_{||}(t) - f_{\perp}(t)), \tag{18}$$

and which may alternately be written as

$$D(t) = E(t)*d(t). \tag{19}$$

This may be arranged, using the expression $d(t) = r(t)s(t)$ to yield,

$$D(t) = E(t)*(r(t)s(t)). \tag{20}$$

If an expression for $s(t)$ was known it would now be possible to solve for the α, γ, *etc.* using the method outlined above after substituting $D(t)$ for $r(t)$. Generally, a sum of exponentials is sufficient for representing $s(t)$ such that

$$D(t) = E(t)*\left\{ r(t) \sum_{i=1}^{n} a_i e^{-t/\tau} \right\}. \tag{21}$$

First appropriate values of a_i and τ_i must be determined. Just as $D(t) = E(t) * d(t)$, it can easily be shown that

$$S(t) = E(t)*s(t) = F_{||}(t) - SF_{\perp}(t) = \int_0^t E(t-u) \sum_{i=1}^{n} a_i e^{-u/\tau}\, du, \tag{22}$$

which can also be solved using the Marquardt method for a_i and τ_i. This method has the added advantage of being somewhat insensitive to choice of time zero, but unfortunately, is limited by the time required to make the analyses, and the memory capacity of the compiler used.

3. Data Collection: Theory and Methods

The anisotropy, $r(t)$, is an indication of the difference between the amount of horizontally and vertically polarized light, $d(t)$, relative to the total amount of light, $s(t)$, as a function of time. This function reveals the average angular displacement of a fluorophore which occurs between absorption and subsequent emission of a photon. It is therefore useful for determining physical properties affecting the movement of the fluorophore on the time scale of the fluorescence event as well as the hydrodynamic properties of the fluorophore. The measured anisotropy is given by:

$$r(t) = \frac{f_{||}(t) - f_{\perp}(t)}{f_{||}(t) + 2 f_{\perp}(t)}. \tag{23}$$

Here $f_{\perp}(t)$ is the number of perpendicularly polarized photons emitted at time t after excitation and $f_{||}(t)$ is the number of photons with a parallel polarization. $f_{\perp}(t)$ is related to the observed fluorescence $F_{\perp}(t)$ by the convolution:

$$F_{\perp}(t) = \frac{f_{\perp}(t) * E(t)}{S}, \tag{24}$$

where $E(t)$ is the excitation function and S is a correction factor known as the sensitivity correction. *Accurate determination of the sensitivity correction is one of the most critical aspects of data collection* and a large amount of effort was devoted to determining the most correct approach to measuring it during the course of these experiments. For these reasons it is relevant to discuss the sensitivity correction which must be used in calculating the anisotropy function as well as the different methods by which it can be measured.

If a perfect instrument existed, one could generate an anisotropy function by collecting a fluorescent decay with the emission polarizer in the vertical position, and then without making any other changes to the instrument except for the rotation of the emission polarizer to the horizontal position, collect a

second fluorescent decay and finally make the necessary computations. The object is to know the difference between the parallel, $F_{||}(t)$, and perpendicular, $F_{\perp}(t)$, components of the fluorescence as a function of time. Science is never as easy as this however. The first problem one encounters is that the amount of the vertical component of the decay in photons per second is likely to be twice as much as the horizontal component. While this is an acceptable condition at low counting rates where the PHA is operating below capacity, it is not at higher rates. The ability of the PHA to process every signal received from the TAC is limited such that the relationship between the number of counts received and the number of counts added to the histogram is not linear at count rates above about 5 kHz. It is advantageous to work at much higher count rates than this, typically at about 20 kHz so as to achieve high quality data in a minimal amount of time. It is preferable to minimize the amount of time spent in collection to reduce the possibility of instrumental variation over the course of the measurement. Secondly, it is unknown as to whether the instrument has the same sensitivity when the polarizers are placed in two different positions. For our new instrument, using an isotropic solution such as ethidium in acetone, the ratio of $F_{||}$ to $F_{\perp}$ will be about 1.03, instead of 1.00. This means that the instrument is 3% more sensitive while the emission polarizer is in the vertical orientation than in the horizontal.

To combat the problems outlined we have adopted another strategy for collecting our anisotropy decays. First, count rate is adjusted so that the amount of data will be similar for both the perpendicular and parallel decays. $F_{||}(t)$ and $F_{\perp}(t)$ are collected at a count rate of 20 kHz for either 1500 or 1000 seconds. This gives the form of the function for each decay with respect to time, but obscures the ratio of relative amounts of each decay. Secondly, three normalization factors are calculated. The first, referred to as R, is the true ratio between $F(t)$ and $F_{\perp}(t)$ which would exist if the response of the PHA were linear with respect to count rate. This can be determined for a given sample by collecting a series of decays with the excitation polarizer vertically oriented and alternating the position of the emission polarizer. It is important to make this measurement at a low count rate and to make no changes in the instrumentation, especially to components which would affect the count rate,

except for the rotation of the polarizer. The normalization factor, referred to as R, can be computed from I_{vh}/I_{vv} where I represents the intensity calculated from the sum of the total counts per time unit, and the two subscripts denote the orientation of the excitation and emission polarizers, respectively. In the v orientation the polarizer transmits vertically polarized light; in the h orientation it transmits horizontally polarized light.

The second is a normalization factor which corrects for differences between optical sensitivity of the instrument to different orientations of the polarizers. There are two methods of determining this correction factor, sometimes called the G factor. One can collect two or more decays with the excitation polarizer in the horizontal position alternating the position of the emission polarizer between vertical and horizontal. The excitation polarizer is located at right angles to the emission polarizer and so if the excitation polarizer's orientation is horizontal, the decays collected with the emission polarizer at either orientation will be a measurement of the perpendicular component of the decay. Therefore, if any differences exist between the two decays it means that the instrument is more sensitive to the detection of fluorescence with one orientation of the emission polarizer than with the other. The G factor is then given by $I_{h,v}/I_{h,h}$. Another method is to collect two decays of a second sample such as ethidium in acetone with the excitation polarizer vertical; one each with the emission polarizer vertical and horizontal. As ethidium in acetone will depolarize very rapidly, the perpendicular and parallel components of the decay as such should be equal. The G factor can then be taken to be I(free ethidium)$_{vv}$/I(free ethidium)$_{vh}$. The G factor obtained with this method, usually 1.03, agrees well with that obtained through the rotation of the excitation polarizer but has the advantage of eliminating a possible source of error caused by the physical movement of the excitation polarizer.

The third normalization factor corrects for the fact that $F_{||}(t)$ and $F_{\perp}(t)$ should contain a different number of total counts even though we have made an attempt to collect the same number in each. The total correction factor to be used in the calculation of the anisotropy, hitherto referred to as the sensitivity correction, **S**, is a product of the three normalization factors as given by:

$$S = \frac{\text{total counts}_{||}}{\text{total counts}_{\perp}} R\,G, \tag{25}$$

where *total counts*$_{||}$ and *total counts*$_{\perp}$ are the total number of counts collected for the decay with the emission polarizer in the parallel and perpendicular orientations, respectively. The accuracy to which we can analyze a decay is limited by the precision with which we can collect the sensitivity correction. If there are fluctuations in the laser power level over the course of the measurement of R or G, the correction may be in error. Errors in determining the sensitivity correction will result in a different apparent r_o as well as lifetimes. We have investigated several methods for determining R and G. Most of the techniques used generally address the issue of laser fluctuation by collecting many measurements for time spans short enough that laser power remains constant. As the mechanism for doing this has not been automated and involves manual rotation of the polarizer, an event taking 10 seconds or more, this method is cumbersome and not terribly reliable. In the future it would probably be most advantageous to develop a system in which a record of the power level of the laser over the period in which the R or G factor are collected would be provided and could be used to correct the data.

The method for collecting R or G which allows the most rapid switching between orientations and minimal drift in laser intensity involves the use of a Hewlett Packard 5381A 80 MHz frequency counter which receives a signal from the pulse height analyzer when an event is recorded there. The counter may be set to return at an average frequency of 10 seconds. The average frequency is recorded twice before rotating the polarizer. This allows for a large number of points to be collected with minimal dead time. Its major disadvantage is that the photon counter is only able to accurately measuring the photons arriving when they do so at a very low frequency, about 1 kHz. At 3 kHz, the counter is only able to count 98% of the events in a given interval, The reason for this is that the PHA sends signals in live time while the counter operates in real time which is always shorter than live time except at very low frequencies. The effect of this is almost a 2% error in the sensitivity correction

as I_{vv} and I_{vh} are typically measured at 3000 and 1200 counts per second, respectively.

This problem is overcome by using the pulse height analyzer to count the data. The pulse height analyzer registers a nearly linear increase in events per unit time to laser intensity per unit time up to 5000 Hz. The physical distance between the pulse height analyzer and the sample chamber however increases the amount of time and work needed to collect accurate measurements. We have lately collected data for each orientation of the polarizer at 30 second intervals for 300 seconds on the pulse height analyzer.

The data from each of the intervals, regardless of their time span and origin, may be analyzed in several ways. One can sum them, which provides an average value, or, one can examine the integral of each interval in relation to the others, the object being to determine the value in the series which most closely represents the true value and to discriminate and discard any values which seem excessively high or low. This is done using a sine function M estimate, a robust estimator of location (Andrews $et.$ $al.$, 1988). Here, a function $\psi(x)$ is defined as

$$\psi(x) = \begin{cases} \sin(x/2.1) & |x| < 2.1\,\pi \\ 0 & \text{otherwise} \end{cases} . \tag{26}$$

The following equation,

$$\sum_{i=1}^{10} \psi\left(\frac{x_i - T}{s}\right) = 0 , \tag{27}$$

in which s is the median of the absolute deviation about T, is solved for T (Andrew et $al.$, 1988). Another method would be to use the median value.

The different approaches used to collect the data, as well as the different methods used to analyze them do not address the possibility of instrumental

error. It is probable that the emission polarizer allows some fraction of the wrong polarization to enter the photomultiplier tube. The effect of this would be to lower the product of the ratios I_{vv}/I_{vh} and I_{hh}/I_{hv}. In installing and aligning the polarizers our object was to maximize this product such that if one examined a sufficiently short time range it would approach the theoretical maximum of 2.5 which would correspond to an r_0 <0.4. The observed usual values of r_0 were usually about 0.35 to 0.36. Secondly, in collecting any series of measurements we were careful to avoid making any changes in the alignment. The value of lifetimes and other parameters obtained from the analysis of the series, though perhaps somewhat in error, due to the presence of instrumental error, would thus all be affected to the same degree and a comparison between different results would be valid.

Though of lesser importance than the sensitivity correction but still worthy of some discussion is the reference function, E(t). The reference function shows the relative number of photons which arrive at the fluorescent sample as a function of time. The reference function is used in either performing a numerical correction for the convolution artifact, or for deconvoluting the data as is described in the next section.

There are several reasons for collecting a reference function. One can see from a plot of the reference function (Figure 4) that while ideally one would find a single very narrow peak at time zero, that in actuality one sees a large primary peak of approximately 80 ps duration. Ideally, the laser flash would be infinitely sharp in time. The instrument response function would then be a delta function excitation. As it is not infinitely thin, some molecules will have begun to fluoresce prior to the completion of the excitation pulse. This results in a convolution artifact. Here, it is impossible to determine where time = 0 is because the instrument response function is convolved with the reference function. To determine the profile of the instrument response function, it is necessary to use the information in the reference function to perform a deconvolution as described below.

Secondly, one observes a number of smaller secondary peaks in the reference function whose separation is a function of frequency of the mode locker. The action of the mode locker causes the Nd:Yag laser to produce pulses of light with a frequency of 41 MHz. These pulses are too narrowly spaced to observe an excitation pulse of anything with a fluorescence lifetime of much more than 1 nanosecond. For this reason a cavity dumper is used. This device is constructed to allow one pulse of light to pass through but to deflect all others in some variable period. For our operations the period is set to about 1.25 ms. As a result the cavity dumper produces excitation pulses with a 0.8 MHz frequency. The secondary peaks result from the leakage of the mode-locked pulse, beside the primary pulse, out of the cavity dumper. These secondary peaks, while usually no more than 1/200 of the intensity of the primary peak, demand recognition. They cause a low level of excitation which is especially visible after the fluorescence induced by the primary peak has all but disappeared. In order to compensate for the effects of the secondary peaks, either through numerical correction or through deconvolution, as described below, it is necessary to identify their location. Their frequency can also be used to compute the number of nanoseconds per channel.

The function can be collected by replacing the fluorescent sample with a solution which scatters light or with one in which the fluorescent lifetime is short enough to be almost indistinguishable from scattered light. The advantage of using the latter is that one can choose a fluorophore which can be excited and which will fluoresce at the same wavelengths as the fluorophore in the sample and therefore minimal modification need be made to the optical system. Generally, for these measurements, a diluted erythrocin sample was used. Erythrocin is a fluorescent molecule with approximately the same emission wavelength as ethidium. It has an extremely short lifetime (about 83 ps) such that the shape of the fluorescence decay function is almost identical to that obtained from a scatter sample. Data were collected for 120 sec at 20 MHz.

4. Sample Preparation

The nucleosome core particles were kindly prepared by Dr. Louis Libertini or David Brown as previously described (Libertini and Small, 1980).

Binding Studies

r (EtBr/CP)	EtBr (μM)	CP (μM)
.0025	.012	5.2
.005	.024	5.2
.01	.051	5.2
.025	.129	5.2
0.05	.258	5.2
.1	.258	2.2
.3	.258	.76
.6	.243	.40
1.2	.243	.20
.03	.15	5.1

Table 1. <u>Concentrations Ethidium and Core Particles</u>. Solutions to be used in binding studies were prepared according to these specifications.

The A_{260} of the core particle solution used in these measurements was 138. Nine samples of varying ethidium to base pair ratios, r, were prepared. NaCl, ethidium bromide and core particles, respectively, were added to distilled water to give the final concentrations listed above in Table 1 in 10 mM NaCl. Samples were allowed to equilibrate before collection to permit binding equilibrium to occur (Small *et al.* 1989a).

$F_{||}$ and $F_{\perp}$ were collected for 1000 seconds each at 20 kHz. Data for the R and G factor were collected by measuring the integral of counts over 10 second integrals on the photon counter. The mean of 16 measurements was determined using an Andrews sine function M estimate for each orientation of the polarizer. For the G factor, a solution of ethidium in acetone was used. All

measurements were performed at 20 °C. A solution of dilute erythrocin in water was used to collect the excitation function E(t).

A numerical correction of the convolution artifact was applied and the data were analyzed as outlined for the general case. Global analysis was used for the data originating from the samples with a binding ratio of less than 0.05 to determine best values of γ and D_{sph}. This data was analyzed beginning 2 ns after the peak of the excitation function. Data from samples with a binding ratio of greater than 0.05 was fit from the peak of the sum function. All data were fit out to 170 ns.

Torsional Flexibility

Two data sets were collected on two different days using two different time ranges: 0.3764, and 0.0928 ns/channel. Samples consisted of 10 mM NaCl, 0.0154 μM EtBr, 0.5125 μM core particles ($A_{260} = 205$) in glass distilled water. Sample were allowed to equilibrate for 1 hour before study. All measurements were performed at 20° C. Approximately 40 x10^6 counts were collected for F and $F_\perp$ in each case. A solution of dilute erythrocin in water was used to collect the reference function E(t). The G correction was determined using a solution of ethidium in acetone. Both the R and G factors were collected in alternating 30 second intervals for a total of 300 seconds for each unique arrangement of the polarizers. The mean value of each sequence was selected using an M sine estimate.

Values of α, γ, D_{sph}, and r_o, were determined which satisfied the equality:

$$F_\parallel(t) - SF_\perp(t) = \int_0^t E(t-u)r(u)f(u)du, \qquad (28)$$

where f(t) is the total fluorescence decay, assumed to take the form of a sum of exponentials

$$f(t) = \sum_{i=1}^{n} a_i e^{-t/\tau} \qquad (29)$$

and where r(u) is given by equation (1) using a the non linear least squares routine previously described,   The parameters a_i and τ_i were similarly determined using the expression

$$F_{\parallel}(t) + SF_{\perp}(t) = \int_{0}^{t} E(t-u) \sum_{i=1}^{n} a_i e^{-t/\tau} du. \qquad (30)$$

We assume that the ethidium decay is represented by a distribution of lifetimes which can be accurately approximated by a sum of 3 or 4 exponentials (Libertini and Small, 1989). In general the ability of the analysis routine to return suitable values of a_i and τ_i was the limiting factor in achieving an excellent fit, especially for the short time data. The data were analyzed under several different conditions. It is time consuming and virtually impossible to find the best fit value of N computationally because, first of all, it is an integer such that there is no satisfactory form of the derivative, and secondly, excellent values of χ^2 can be obtained for many values of N. For these reasons, the best fit values of α, γ, D_{sph}, and r_o were obtained for various values of N in separate sets of analyses. These results confirmed that justification for preferring one value of N over another cannot be obtained from the values of χ^2. However, low values of N gave slighter better values of χ^2 and were much less time consuming. For these reasons, N of 15 was used in subsequent analyses. For short time data two set analyses were performed, one in which D_{sph} was held equal to the values obtained for long time analysis for the same sample preparation and one in which D_{sph} was allowed to wander. This was done under the belief that the decay covered an insufficient time range (80 ns) to allow resolution of this parameter. Additionally, as a control, best-fit values of α, γ, D_{sph}, and r_o were obtained by fitting equation 4 directly to equation 1. Here it was necessary to apply a numerical correction for the convolution artifact.

pH Studies

The core particles were concentrated to a final OD at 260 nm of 205 in a 1 mM Tris/HCl buffer using a Centricon-30 miniconcentrator. For the 10 mM pH measurements, samples contained 5 µL core particles (OD = 205) in 1mM Tris HCL, 10.5 µL of 2.43 µM EtBr, 10 µL of 100 mM buffer and 74.5 µL glass distilled water. For the 100 mM ionic strength pH measurement samples contained 5 µL core particles (OD = 205) in 1mM Tris HCL, 6.3 µL of 2.43 µM EtBr, and 88.6 µL of 100 mM buffer. For measurements of pH<5, a sodium acetate buffer was used; for 5<pH<7 a MES buffer was used; for pH = 7, the buffer was MOPS and for pH>7 the buffer used was Tris HCL. Samples were allowed to equilibrate for at least one hour. All measurements were conducted at 20 C. pH was determined following the measurements using a Digi-sense LCD pH meter model # 5994.

40×10^6 counts were collected for each orientation of the polarizer for the solutions with 100 mM ionic strength and 25×10^6 counts were collected for the 10 mM solutions. The G factor was determined using a solution of ethidium bromide in acetone. The integrals used in the computation of the R and G factors were determined by calculating the sin function M-estimate of a series of ten integrals, collected for each orientation of the emission polarizer at 3 kHz for 30 seconds. Data was analyzed as described above and by fitting the anisotropy data directly to expression 4 as described in DATA ANALYSIS.

Figure 4. <u>Histogram of Data.</u> Figure shows $F_{||}(t)$, $F_\perp(t)$ and $E(t)$. $F_{||}(t)$ and $F_\perp(t)$ contain 40×10^6 each. $E(t)$ was collected using a dilute erythrocin solution.
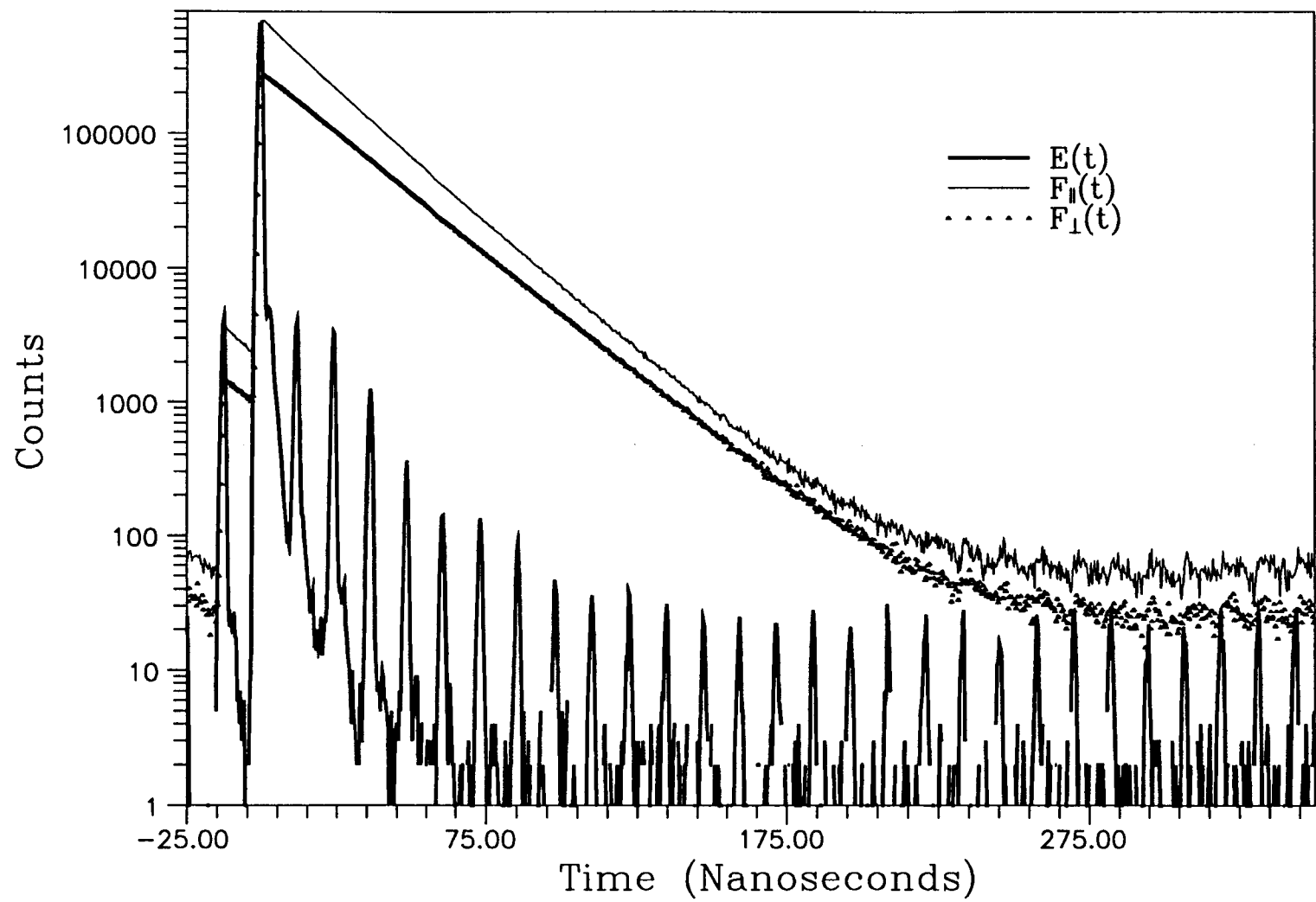
Figure 4

RESULTS

1. The Effects of Ethidium Binding on the Torsional Flexibility of DNA

In order to determine the flexibility of DNA on a nucleosome core particle as near as possible to its native conformation it is important that the conformation is disturbed as little as possible through the addition of the dye. Ethidium bromide binds to DNA by intercalation between the base pairs. The effect of binding is to lengthen, stiffen and unwind (Cantor and Schimmel, 1980) a strand of DNA. It is thus predictable that the value of the parameters recovered in our experiments would be dependent on the concentration of ethidium used and that the lower the ratio of ethidium to core particle, the more closely our results will indicate the state of the DNA on the native particle. However, at a ratio of less than 0.01 ethidium per core particle, it is typically difficult to obtain an adequate count rate. This necessitates very long collection times. Secondly, the most serious disadvantage is the failure of the ethidium fluorescence to mask emission by other processes. These include Raman scattering from water and low-level fluorescence of impurities in the solution. These emissions have an effect of distorting the early part (primarily less than 2 nanoseconds) of the anisotropy decay, a region which yields much information. It was our first objective in this study to determine a reasonable level where the disadvantages incurred by using a minimal amount of ethidium as a probe are minimized with respects to the advantages of using a low ratio.

It has been reported (Erard *et al.*, 1979; McMurray and van Holde, 1990a) that the binding of ethidium bromide to the nucleosome core particles in the absence of any linker DNA, at low ionic strengths may be a highly cooperative process.   This significantly enhances the probability that a number of nucleosomes may have more than one ethidium per core particle, even at very low binding ratios. McMurray and van Holde (McMurray *et al.* 1990a) have fit the data from spectrophotmetric measurements of the concentrations of bound and unbound ethidium to the theoretical expressions of McGhee and von Hippel

(1975) and have obtained values of 4.5, 140 and 2.4 x 10^4 M^{-1}, respectively, for the number of base pairs excluded by the dye, the cooperativity parameter, ω, and the binding constant, K_a, using this conditional probability model.

Cooperative binding implies that the binding of the first ethidium disrupts the structure to facilitate the binding of the next dye residue. Thus, the ethidium is less likely to probe the state of the native core particle. Also, cooperativity implies that the binding of the second residue occurs near the first. If two ethidiums are close to one another then it is likely that radiationless energy transfer will occur between them. If such transfer occurs at a rate comparable to the excited state lifetime, then depolarization of the fluorescence will occur. The anisotropy decay will be distorted by this process and appear to have lifetime components which are too short.

Recent work in our laboratory (Libertini and Small, preliminary results) indicate that the binding studies of Erard *et al.*(1979) and McMurray and van Holde (1990) may be in error. The binding of ethidium to the core particles may *not* be cooperative. Since this is preliminary work I will take the more pessimistic view and discuss the implications of my studies in terms of the literature reports.

Using the parameters from McMurray and van Holde (1990b) and the McGhee and von Hippel (1975) expressions, one can predict the average fraction, v, of ethidium per core particle. McMurray and van Holde have also shown that when a high enough binding ratio is achieved the nucleosome dissociates, though for our measurements, we are not concerned with such high ratios. McMurray and van Holde (McMurray *et al.*1990a) have also shown, using MPE footprint analysis, that upon reaching the binding ratio where dissociation occurs, only one half of the core particles have ethidium bound. Therefore, it is more likely that the fraction of ethidium bound to a particular core particle is actually twice as high as the average number, v. Thus, at a binding ratio of 1.2 ethidium/core particle, the average fraction of ethidium per core particle is 0.7, but it is more likely that one half of the nucleosomes have no ethidium bound and the other have an average fraction of 1.4.

Alternatively, using the parameters provided in the McMurray paper and a very simple model of cooperative ethidium binding which does not consider exclusion: If

$$K_a = \frac{[BP\;EtBr]}{[BP][EtBr]},$$ (31)

and if the base pair concentration = 5.85 x10^{-5} M and the EtBr concentration = 2.43x10^{-7} M, corresponding to an EtBr/Core Particle ratio of 1.2, one can calculate that 40% of the core particles have at least one ethidium bound. Using

$$K_a{}^2\omega = \frac{[BP\;EtBr_2]}{[BP][EtBr]^2},$$ (32)

and the same initial values of base pair and ethidium concentrations, one can calculate that about 20% of the core particles will have a second ethidium bound.

These calculations indicate that there is significant probability that some of the nucleosomes will have two or more ethidium moieties bound and could therefore yield information which would not describe the characteristics of the particles with a single ethidium bound. Also, if enough ethidium is bound to a few molecules to cause dissociation, the recovered rotational diffusion coefficient would be affected, reflecting the drastic change in the geometry of some of the molecules. These predictions are substantiated by the results given in Tables 2 and 3. The data are broken up into two categories. At very low binding ratio (see Table 3) the changes in the parameters are not necessarily due to changes in the binding ratio r because substantially different concentrations of ethidium were used for each measurement. This means that there were differing amounts of scattering and impurity fluorescence present in each measurement. The amount of scatter is constant regardless of the binding ratio but is masked at r greater than 0.03. The data from prior to 2.5 ns after the excitation were ignored in this analysis in attempts to diminish the varying

effects of scatter which is often found in the first 2 nanoseconds of decay at very low binding ratio. One consequence of such an approach, because the data were not analyzed using a convolution routine, is to lower the apparent r_o. It should be emphasized that for this reason the results in Tables 3 and 2 cannot be compared. Secondly, in the analysis whose results are shown in Table 2, D_{sph} and γ were fixed. This approach allows us to measure the effects using different binding ratios in terms of changes of two parameters, α and r_o. It was impossible to do this and obtain suitable values of χ^2 in the analyses whose results are given in Table 3, presumably because of the effects of scatter and impurities in these samples. In this analysis, all five parameters were allowed to wander. The results of the analyses given in Table 2 provide a more valid basis for comparison of the effects of binding ratio than does Table 3 as the amount of ethidium was held relatively constant in the measurements used for these analyses.

Some of the more significant observations drawn from the results shown in Table 2 and 3 are that initial anisotropy r_o, and the torsional coefficient α, both decrease as binding ratio increases. At binding ratios greater than 0.3 the rotational diffusion coefficient, D_{sph} is substantially lower, apparently corresponding to a less compact structure. The torsional coefficient is in all cases lower than the value of 3.8×10^{-12} dyn cm that was reported by Thomas *et al.* (1980) for free DNA. It is possible, however, given the results in Table 3, that if one extrapolated the data to a zero binding ratio one would find values which are similar to that of free DNA. It should be noted that for all of these analyses N was held constant at 15. The validity of this approach will be discussed in more detail below. It will be shown that there is a linear relationship between α and N (see figure 7). Therefore, what seems to be an apparent increase in α as a function of binding ratio might just as well be a increase in N which would correspond to DNA coming off the core particle. As a final consideration, it should be stated that the higher values of α at the lowest binding ratios might be artifactual, resulting from the very low concentration of ethidium used. For example, at the lowest binding ratio Raman scattering and impurity fluorescence are visibly apparent in the data.

Genest *et al.*(1982), have made similar observations about the effect of binding ratio on the fundamental anisotropy, r_o. The lowest binding ratio at which they were able to collect data was 0.15 ethidium per core particle. Using measurements of the steady state anisotropy as well as measurement of the fluorescent anisotropy decays, they were able to show substantial changes in the recovered lifetimes, as well as decreases in r_o and the steady state anisotropy between the binding ratios of 0.15 and 1.5. These researchers attribute these phenomena to fast excitation energy transfer between adjacent ethidium. If such an observation is correct, it would suggest that if energy transfer occurs when there is on the average less than 0.1 ethidium per core particle then there is a higher than random probability that binding is occurring at adjacent sites and is therefore cooperative. Genest *et al.* (1981) have proposed that in the native core particle only a short segment of the DNA is accessible to the ethidium molecules and that the binding of a few ethidiums to this segment induces the accessibility of another segment such that the binding of 8 or 9 ethidium molecules is sufficient to make 93% percent of the DNA accessible to binding.

To avoid such undesirable consequences, we have sought to work with binding ratios in the range of 0.03 to 0.05. Figure 5 shows that there is very little change in the shape of the decay between binding ratios 0.01 and 0.1, but significant change at ratios higher than this. At a binding ratio of 0.03, if a high concentration of core particles is used, there is more than sufficient intensity to obtain high quality data.

r	N	D_{sph} x 10^{-5} s^{-1}	r_o	α x 10^{12} dyn cm	γ x 10^{21} dyn cm s	χ^2	EtBr (μM)
0.05	15	10.5	0.350	3.10	1.32	0.94	0.258
0.1	15	10.5	0.350	2.56	1.32	1.15	0.258
0.3	15	10.7	0.346	1.85	.935	1.02	0.277
0.6	15	7.69	0.335	.352	7.08	1.07	0.243
1.2	15	7.64	0.323	.317	5.94	0.86	0.243

Table 2. Recovered Parameters as a Function of Higher Binding Ratio. First channel was considered to occur at the peak of the sum file. All parameters were allowed to wander with the exception of N. χ^2 reflects fit from first channel to 170 ns.

r	N	D_{sph} x $10^{-5}\,s^{-1}$	r_o	α x 10^{12} dyn cm	γ x 10^{21} dyn cm s	χ^2	EtBr (μM)
0.0025	15	10.00	0.336	3.94	5.0	1.14	0.006
0.005	15	10.00	0.348	2.84	5.0	1.32	0.024
0.01	15	10.00	0.342	2.62	5.0	1.22	0.051
0.025	15	10.00	0.337	2.31	5.0	1.01	0.128
0.03	15	10.00	0.335	1.52	5.0	1.15	0.153

Table 3. Recovered Parameters as a Function of Lower Binding Ratio. The values of different parameters recovered as a function of binding ratio. Time zero was considered to be two ns after the peak of the sum file to avoid inclusion of Raman scattering peak. All parameters were allowed to wander with the exception of N and γ. The value of γ was chosen using global analysis. χ^2 reflects fit from time zero to 170 ns.

Figure 5. <u>Anisotropy Decays at Different Binding Ratios.</u> Anisotropy decays were obtained from data to which a numerical correction for the convolution artifact has been applied. Solutions were prepared at various binding ratios as specified in Table 1. The data indicated that while there is some change with respect to binding ratios at a ratio of less than 0.1, the change is insignificant compared to that between 0.1 and 0.6, where some particles have presumably begun to dissociate.
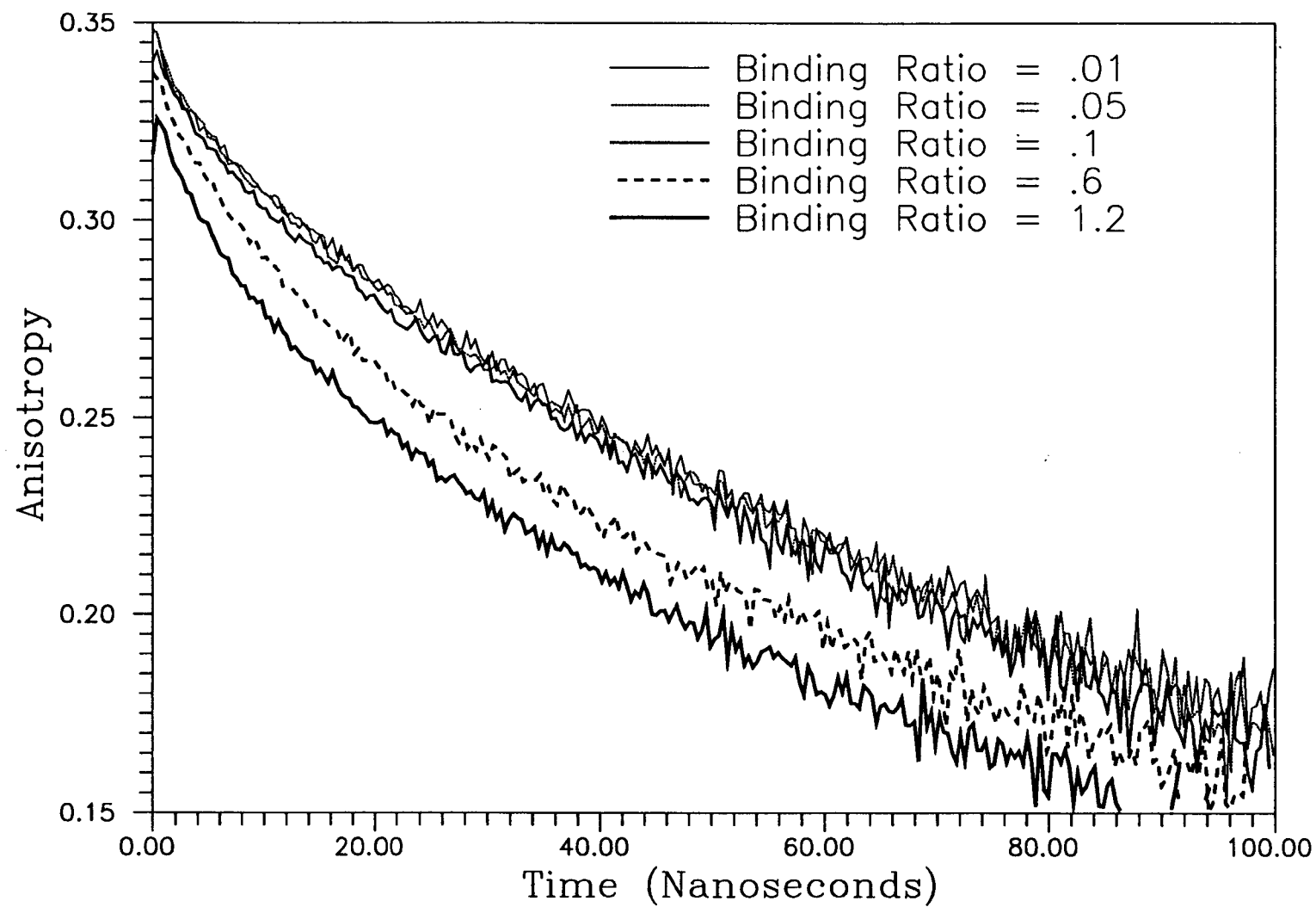
Figure 5

2. Torsional Flexibility of Nucleosome-Bound DNA

There have been numerous studies of the flexibility of free DNA using a model similar to the one used to depict the motion of the DNA on a nucleosome core particle (See Schurr 1989 for a review). The model for free DNA, as does the model for the nucleosome core particle, allows for variation in the torsional coefficient, α; the frictional coefficient, γ; as well the as the number of rigid rods, N and the rotational diffusion coefficient, D_{sph}. Using this method, it is easy to obtain a good fit for any number of values of the parameters α and γ upon the condition that the product of α and γ be equal to some constant which is derived from the anisotropy of DNA under investigation. In order to determine the torsional flexibility, it is assumed that the frictional factor for a segment of DNA can be approximated to be that of a cylinder (Perrin 1934); and can be computed from

$$\gamma = 4\eta\pi a^2 h, \tag{33}$$

where a, h, and η are the radius, length and viscosity, and that the rotational diffusion coefficient of free DNA is known. Because the value of γ is assumed it is possible to determine the value of the torsional coefficient corresponding to cylinders with this value of γ. Thus, it has been shown that the torsional coefficient of free DNA is equal to 3.8 x 10^{-12} dyne cm. (Thomas et al. 1989). Unfortunately, Perrin, who had not heard of DNA, has not provided us with a convenient formula for the frictional coefficient of DNA attached to a nucleosome core particle, nor has the rotational diffusion coefficient been unquestionably determined as a function of pH. These two complications have made the task of determining an exact value of the torsional coefficient for DNA bound to a nucleosome core particle much more difficult.

The torsional and frictional coefficients for DNA bound to a core particle have previously been reported to be 1.5x10^{-12} dyne cm and 9.34 x 10^{-23} dyne cm sec, respectively (Schurr et al. 1985). To obtain these results, N was set at 146 base pairs and the rotational diffusion coefficient was assumed to be 6.12x10^5 s⁻

[1] (adjusted to 20 °C). Such a diffusion coefficient, however, implies a hydrodynamic radius of 62Å, which is inconsistent with the size and shape of the particle as determined by other methods (van Holde, 1988; Brown *et al.* 1990b). In order to confirm that the nucleosome has a smaller hydrodynamic radius than 62.9 Å corresponding to a D_{sph} or 6.12 x 10^5 s^{-1}, we have treated this parameter as variable as well as the α and γ. Additionally, because it is unlikely that DNA is clamped to the nucleosome only at its ends, we have primarily considered values of N which are much lower than 146. For N = 15, at a binding ratio of 0.03, it is possible to obtain a very good fit to the data as shown in Figure 6. Here α is approximately 1.5x10^{-12} dyn cm and γ is about 8 x 10^{-22} dyne cm sec (See Table 3). While the value of α is identical to the value reported by Schurr and Schurr (Schurr *et al.* 1985), there is an important difference. As a result of the choice of N used (146 versus 15) in the Schurr study, 1.5x10^{-12} dyn cm was given as the maximum value of α while here it is close to the minimum.

The reason for this may be rather obscure but can be derived from the twisting correlation functions (2). It has been pointed out (Schurr *et al.* 1985) that the DNA on a core particle has a relaxation time τ which is dependent on the collection conditions. The relaxation time for the longest internal mode is given by

$$\tau = \gamma / 4\alpha \sin^2 \left[\frac{\pi}{2(N+1)} \right]. \tag{34}$$

This means that while we cannot determine either α or γ, we can know the ratio of the two if N is known.

Secondly, at sufficiently long times the twisting correlation function (2) becomes independent of time and reduces to,

$$C_n(\infty) = (n^2 Z)^{1/2} \exp(-n^2 Z) \int^{(n^2 Z)^{1/2}} \exp(y^2) \, dy \tag{35}$$

where

$$Z = k_B T(N+1)/4\alpha. \tag{36}$$

This additional constraint implies that it is possible to know the precise value of α and γ if N is known. Once a set of best fit of parameters has been determined one can calculate a value of τ and Z. Any combination of α, γ and N which results in the same values of τ and Z should also produce a good fit to the data. One can rearrange (35) and (36) to give an approximate expression for the product of α and γ:

$$\alpha\gamma = \frac{\pi\tau k_B T}{8Z^2} \tag{37}$$

Therefore, if τ and Z are known, it is possible to determine the value of α and γ by solving (36) and (37) for any value of N+1 which will give an equally good fit. This phenomena is illustrated in Figure 7. For this analysis a search was made for the best fit value of α, γ and D_{sph} as a function of N. The form of this function (solid) is similar to that (dashed line) which would be expected if α and γ were determined using the method outlined. The values of χ^2 provided little assistance in our efforts to assign N a value. χ^2 was only fractionally lower at smaller values of N. Of course, the linearity of α with respect to N and the dependence of γ on α allows us to extrapolate these values to any value of N. Larger values of N were not presented in the chart because of the excessive time required to fit a dataset; not short for even N = 15, and increasing with respect to N^2.

This indicates that only if the flexing segment of DNA is shorter than 35 base pairs can one say, as has been proposed, that the DNA on a nucleosome core particle is more flexible than DNA free in solution. Over a range of acceptable values of N (from 25 to 60), α is approximately the same as the DNA free in solution. The frictional factor, even at the unlikely N = 80, is at least two fold higher than that proposed for free DNA (7.2×10^{-23} cm s at 20° C).

Our analyses have indicated that the rotational diffusion coefficient is about 10×10^5 s^{-1}. It has been argued that the lifetime of ethidium bromide is too short to allow adequate resolution of this parameter and therefore the use of the methylene blue is more suitable for these studies. Using triplet anisotropy decay of methylene blue on the core particle DNA, the rotational diffusion coefficient has been determined at 5.6×10^5 s^{-1}. Using such a value for our analysis would give a very poor fit to the data at very short time (see Table 3). As can be seen in Figure 6, we are able to achieve an excellent fit at every region of the decay. Therefore we believe that 10×10^5 s^{-1} is a more reasonable value than others reported in the literature (see Brown and Small (1990b) for a comparison and discussion of this subject).

Data#	ns/ch	$\alpha/10^{-12}$ dyn cm	$\gamma/10^{-21}$ dyn cm s	$D_{sph} \times 10^{-5}$ s	N	χ^2
1A	0.376	1.58	0.95	10.1	15	1.11
1A	0.376	1.84	1.66	9.91	15	0.82%
1A	0.376	.318	6.77	6.12&	15	1.60
2A	0.376	1.77	0.94	10.6	15	1.27
2A	0.376	1.83	2.17	9.99	15	0.95%
1B	0.093	1.93	0.73	11.2	15	1.67
1B	0.093	1.76	0.86	10.6	15	1.70*
2B	0.093	1.59	0.30	11.7	15	1.85
2B	0.093	1.41	0.61	10.1	15	1.98*
-	-	.164 $	1.17	6.12	15	-

Table 4. Results from Different Analysis of 4 Datasets. Binding ratio = 0.03. All fits were obtained by fitting the data indirectly. Values of χ^2 indicate quality of fit up to 200 ns for 0.376 ns/channel and up to 85 ns for 0.093 ns/channel. Datasets 1A and 1B, and 2A and 2B were collected from the same sample respectively, but using a different electronic configuration of the instrument.

% χ^2 was determined by fitting the anisotropy data obtained after a numerical correction for the convolution artifact was applied directly to $r_c(t)$.

& D_{sph} was held constant at a value corresponding to a sphere with a hydrodynamic radius of 62.9 Å at 20° C as proposed by Schurr and Schurr (1985).

* D_{sph} was held constant at a value derived from analysis of data collected at .3764 ns/channel. This was done under the presumption that this parameter could not be resolved properly in an 85 ns time span.

$ Parameters reported by Schurr and Schurr (1985) adjusted to 20° C and N = 15.

Figure 6. <u>Fit of the Measured Difference Function</u>. Also shown is the resulting deviation function which indicates the goodness of fit. Binding ratio used was 0.03 ethidium/core particle. 40×10^6 counts were collected at 20 kHz for $F_{||}(t)$ and $F_{\perp}(t)$. For this fit α was $= 1.8 \times 10^{-12}$ dyn cm, $\gamma = 9.5 \times 10^{-22}$ dyn cm s, $D_{sph} = 10.0 \times 10^5$ s^{-1}, $N = 15$ and $\chi^2 = 1.11$
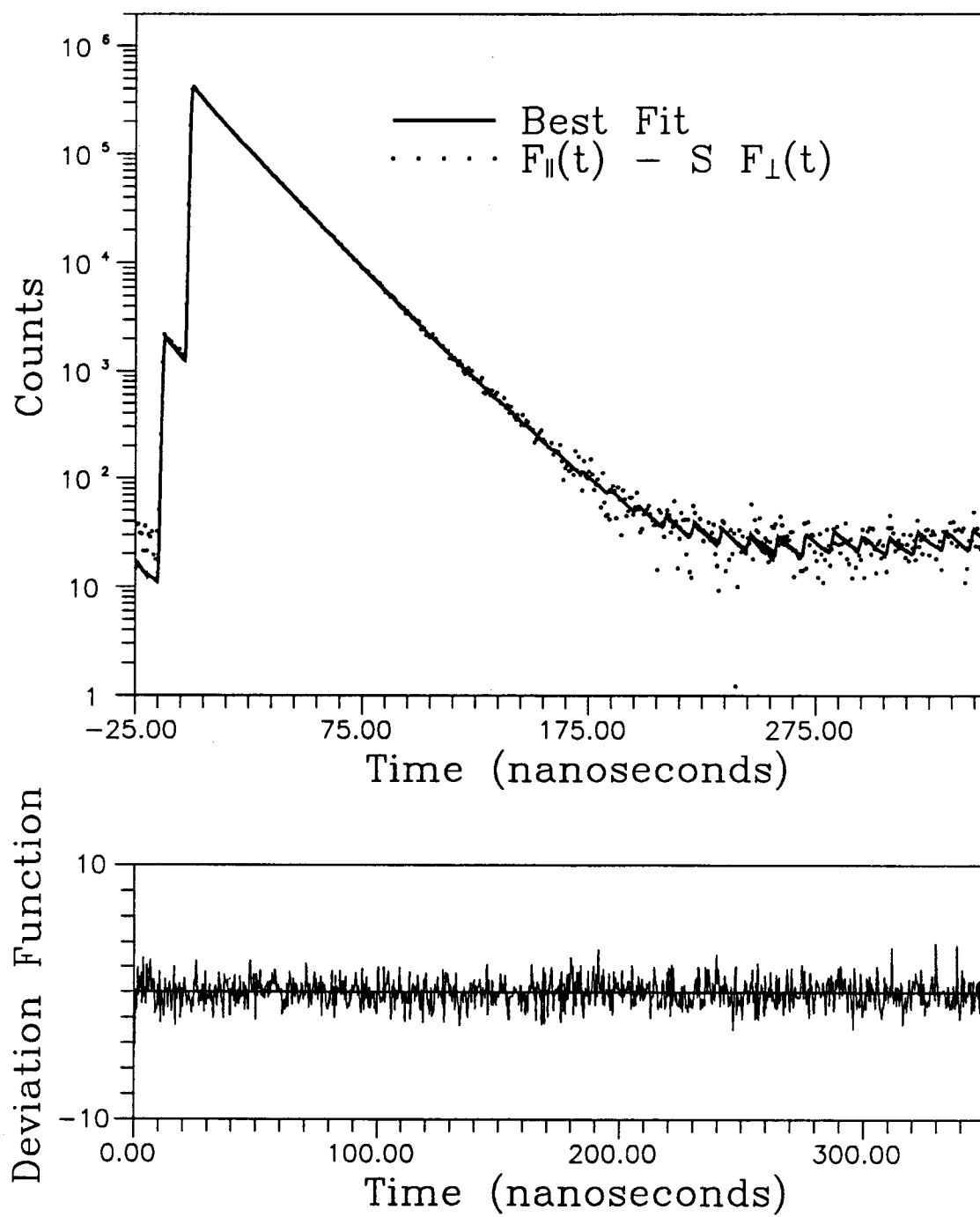
Figure 6

Figure 7. <u>Torsional and Frictional Coefficients as a Function of N.</u> The solid line show the best fit values of α and γ for different N as by obtained iteratively fitting the difference function. Dashed line shows the values of α and γ predicted from the expressions (36) and (37). The data indicate that one can predict the value of either of these parameters for any value of N with some certainty.
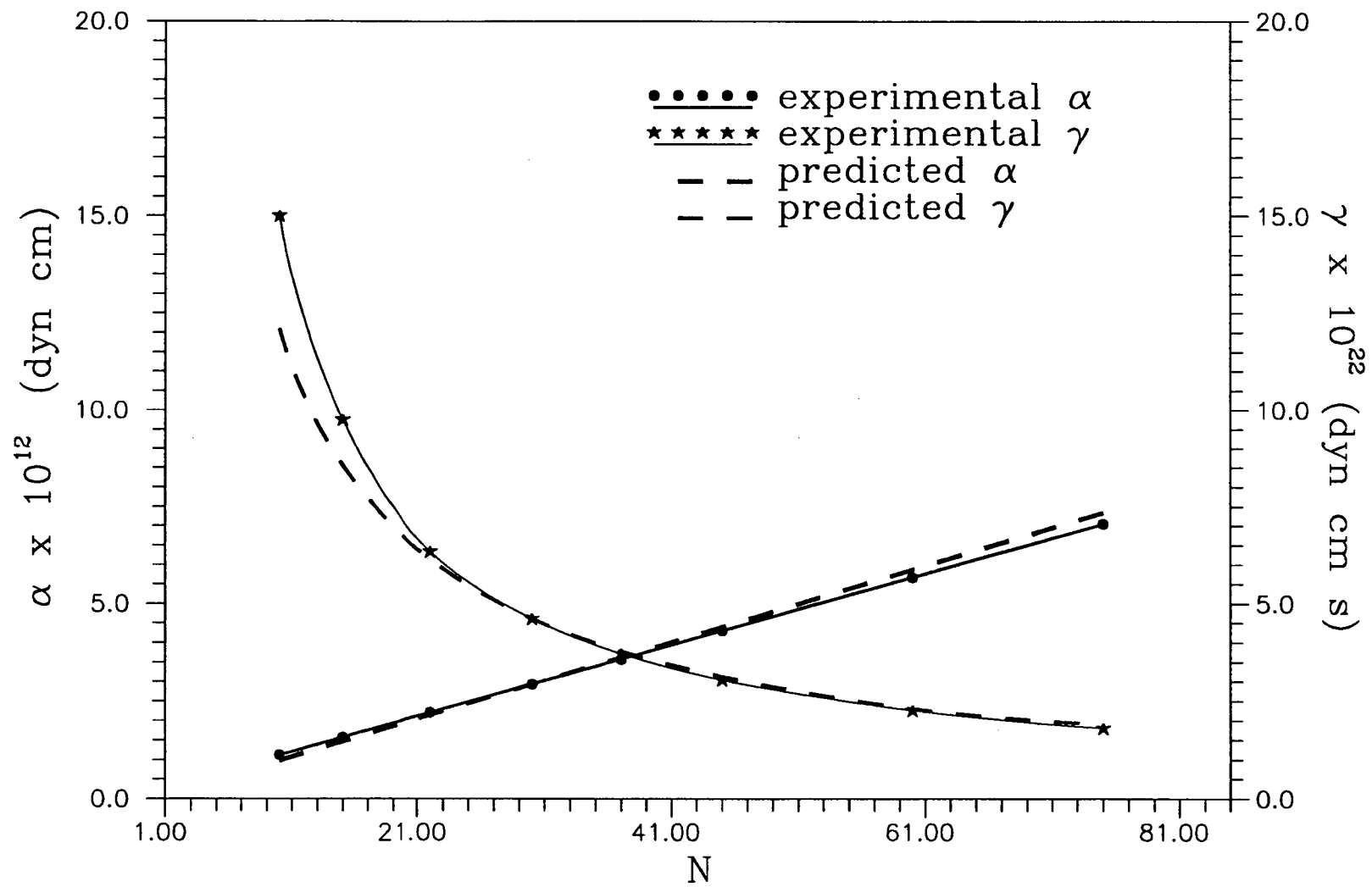
Figure 7

3. Effects of pH on the DNA of a Nucleosome Core Particle

It is readily apparent from viewing the anisotropy decays (Figures 9 and 8) that the core particles undergo some sort of transition as a function of pH. Most notable is the greater curvature, and thus more rapid depolarization, which is introduced as the hydrogen ion concentration increases. Less obvious is whether the rotational diffusion coefficient, and thus the size of particle has changed over the range of pH's under investigation. Analyzing these decays using a sum of exponentials was not as revealing as had been hoped, in part due to the difficulty in achieving a satisfactory fit at very short times. These were the time ranges where the extent of change as a function of pH is most apparent. Secondly, this method only gives insight into the size of the particle and not into the dynamics of the DNA. In order to quantify our initials observations, as well as to determine whether the transition can be attributed to a change in any one physical characteristic, such as torsionally flexibility, we have sought to fit the data to the model developed by Schurr *et al.*

There are two methods available for analyzing the data. The data can be fit directly or indirectly using the convolution routines. The latter has the advantage of producing more accurate results which are more insensitive to the time range used in the analysis. The parameters which are returned when the data are fit directly are highly dependent on the choice of the first channel of data. Time zero is usually chosen to occur at the point where $F_{||}(t) + 2SF_{\perp}(t)$ reaches its maximum, about 1 ns after the peak of the E(t). For indirect fitting, choice of time zero is irrelevant to the analysis and results are reproducible over different analysis ranges. Secondly, with indirect fitting it is possible to analyze out to and fit up to 350 nanoseconds if enough data are available. There is a great deal of temptation to do this, in part because the uncertainty in the parameters is less. The question arises as to whether one gets better or worse results by fitting to a region where there is so little data as to produce artifactual results. Indirect fitting is also very slow and furthermore, possible error is introduced in the fitting of the sum function to a sum of exponentials.

As it is difficult to know which is the best way to analyze the data, results from both methods will be presented.

Some of the values of the parameter obtained using an indirect fit for high salt are given in Table 5. The high salt data contains about 50 percent more data than the lower salt data and so the long time fit was better. We are able to obtain excellent results up to 350 nanoseconds here. Plots of α versus pH and D_{sph} versus pH, both at N = 15, for the two salt concentrations are given in Figures 12 and 13. It is obvious that there is a transition with respect to α. The figures show that the data may be roughly divided into 2 or 3 groups and that this observation is substantiated by being reproducible at a higher and more physiological salt concentration. The data suggest a a two-fold decrease in α. The transition appears to be shifted toward a higher acidity at the lower salt concentration. It is less clear what the effect of pH is on D_{sph}. The 100 mM ionic strength data do not indicate any dependence while a transition is revealed at the lower ionic strength. Secondly, if data is analyzed using the direct fitting approach, the results of which are shown in Figures 10 and 11, both the high salt and low salt data show that D_{sph} decreases with pH. It should be noted that we were unable to achieve a satisfactory fit at times greater than 200 ns in all the cases where the data showed a pH dependent shape transition. It remains unclear as to whether one actually exists. The differences between the data collected at two different ionic strengths may, in part, be due to differences in the viscosity of the solution, a product of the 10 fold higher salt concentration used in the 100 mM measurements. It is may also be due to artifacts in the analysis resulting from the different relative amounts of data collected, or the loosening of the core particle structure in preparation for the unfolding which occurs at 600 mM.

While it is also unclear as to whether there is any transition with respect to γ, it is readily apparent from the data that there is a decrease in α with pH. However it is more likely that N has changed and the torsional coefficient has remained the same. Just as was demonstrated above, it is possible to show that α increases linearly with increasing N. See Figure 14. Here a line has been drawn at $\alpha = 3.5 \times 10^{-12}$ dyn cm sec to illustrate that *it is possible to hold*

this value constant and achieve the same fits by only changing N. J. M. Schurr (personal communication) has indicated that there is no significant change in the flexibility of free DNA over this range of pH. Therefore, it is appears that the flexing segment length increases as the conditions become more alkaline.

pH	Cation conc. mM	$\alpha/10^{-12}$ dyn cm	$\gamma/10^{-21}$ dyn cm s	r_0	D_{sph} x 10^{-5} s	$\chi^{2\ \&}$	$\chi^{2\ *}$
4.79	100	2.10	2.42	0.352	8.26	1.17	1.39
5.49	100	2.69	1.92	0.352	9.51	0.97	1.14
5.99	100	2.57	1.79	0.351	9.27	0.94	1.04
6.40	100	2.29	1.30	0.350	9.58	1.09	1.18
6.88	100	1.95	1.04	0.352	9.76	1.16	1.17
7.64	100	1.48	1.48	0.350	8.99	1.04	1.08
8.31	100	1.37	1.55	0.346	9.55	1.06	1.34
8.80	100	1.34	1.23	0.360	9.19	1.13	1.26
9.26	100	1.35	0.89	0.362	9.58	1.12	1.28
4.85	10	1.71	3.43	0.351	8.47	1.22	1.64
5.48	10	1.96	1.55	0.355	10.0	1.39	1.50
5.99	10	2.15	0.91	0.353	10.5	1.40	1.76
6.85	10	1.53	1.27	0.351	9.94	1.45	1.78
7.50	10	1.37	1.30	0.358	9.60	1.30	1.62
8.30	10	1.19	1.69	0.355	8.97	1.26	1.53
8.57	10	1.19	1.27	0.356	8.98	1.30	1.50

Table 5 Best Fit Parameters for Different pH. Values were obtained by fitting the difference function using an iterative convolution routine as a function of ionic strength and pH. Binding ratios were 0.05 and 0.03 for the 10 mM and 100 mM data respectively.

& χ^2 describes the fit obtained beginning at the peak of the sum file up to 200 ns.

* χ^2 describes the fit obtained beginning at the peak of the sum file up to 350 ns.

Figure 8. <u>Anisotropy Decays as a Function of pH (100 mM Salt).</u> Anistropy decays obtained from data to which a numerical correction for the convolution artifact has been applied. Solutions had a final ionic strength of 100 mM and an ethidium/core particle binding ratio of 0.03. A total of 25×10^6 counts were collected for $F_{||}(t)$ and $F_{\perp}(t)$.
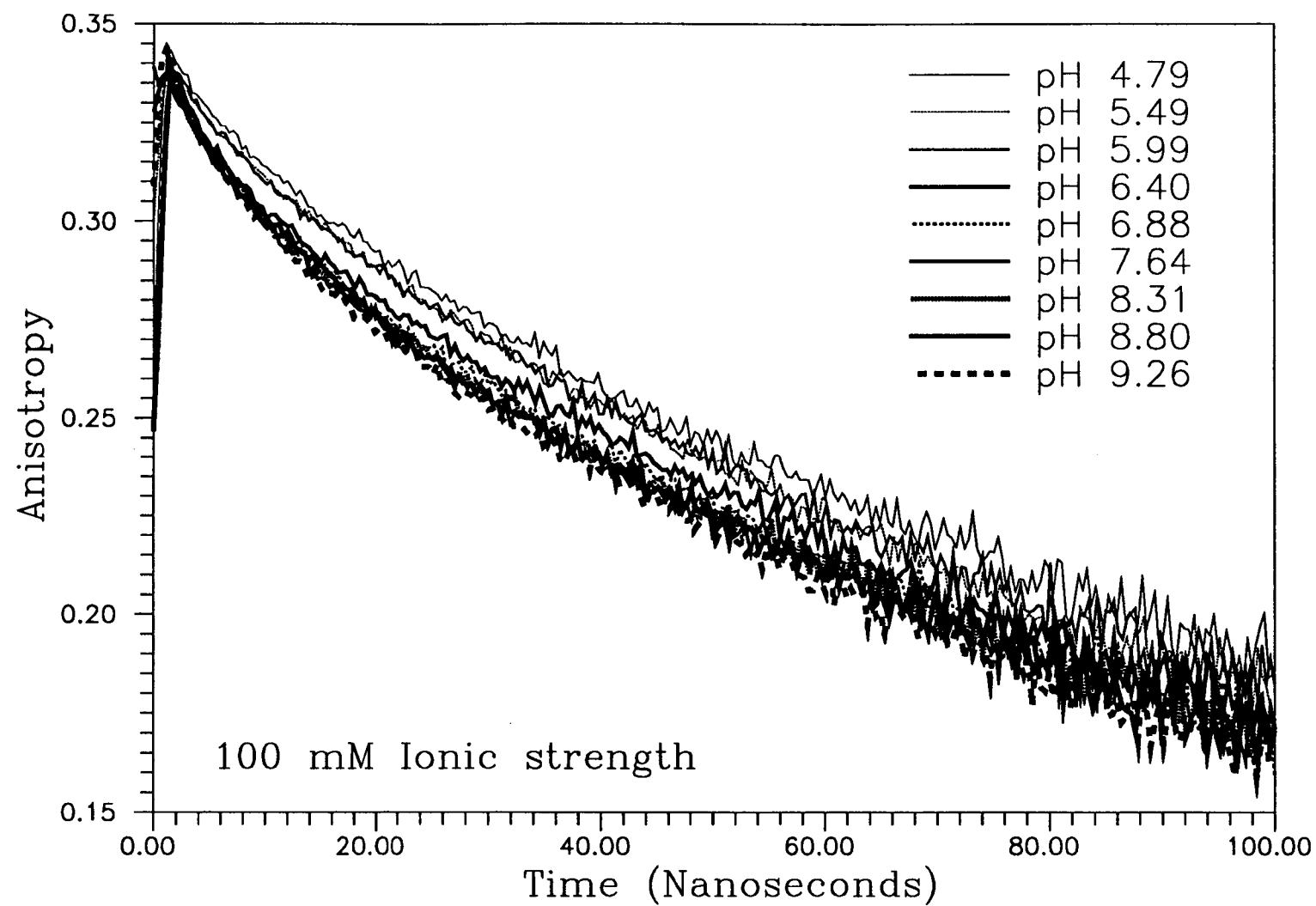
Figure 8

Figure 9. <u>Anisotropy Decays as a Function of pH (10 mM Salt)</u>. Anisotropy decays obtained from data to which a numerical correction for the convolution artifact has been applied. Solutions were diluted to a final ionic strength of 10 mM, and had a binding ratio of 0.05 ethidium/core particle. A total of 25×10^6 counts were collected for $F_{||}(t)$ and $F_{\perp}(t)$.
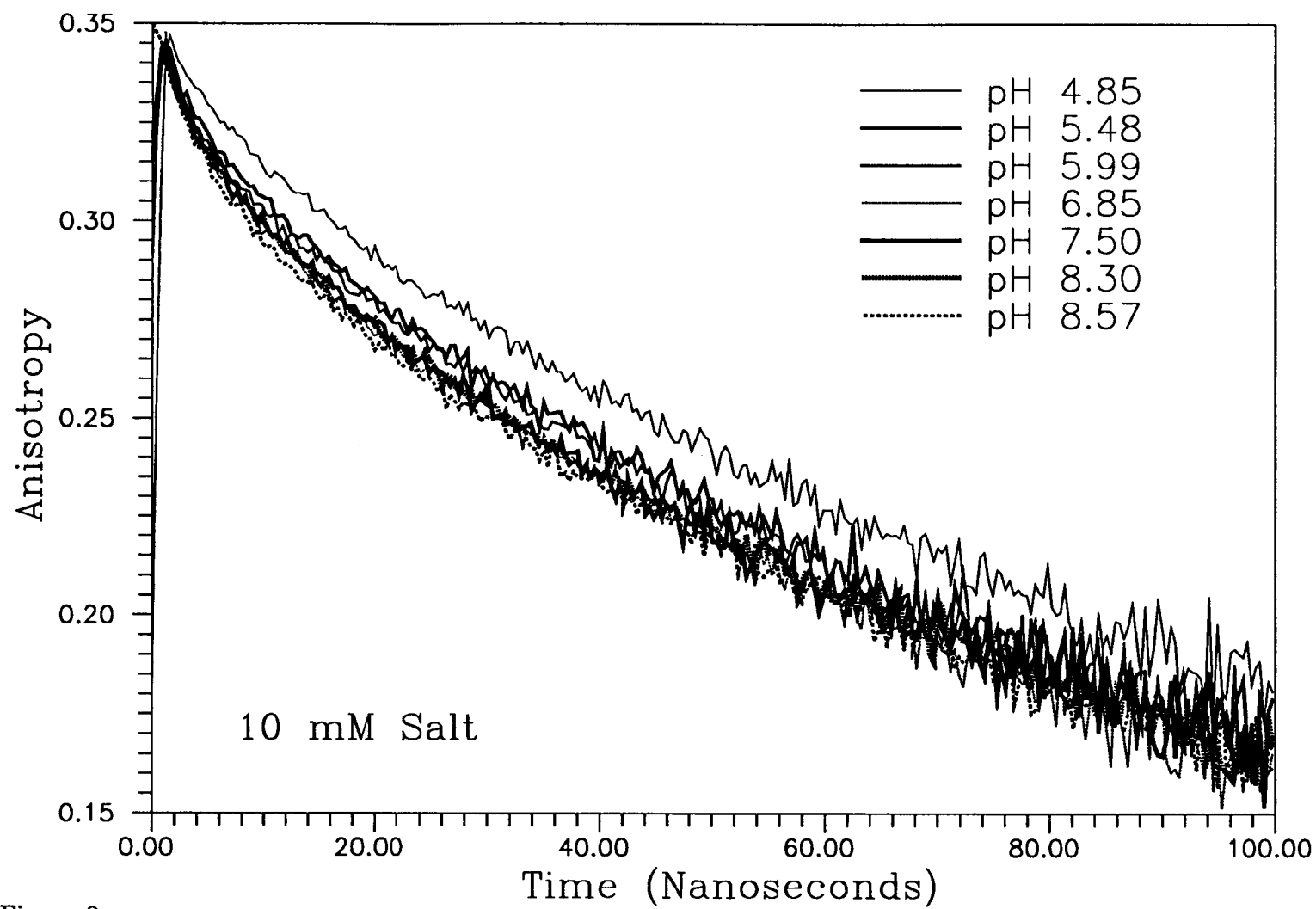
Figure 9

Figure 10. <u>Torsional Coefficient as a Function of pH (Direct Fit)</u>. The best fit values of α are plotted as a function of pH for two salt concentrations. The parameters were obtained by fitting the corrected anisotropy data directly to expression (1). Time zero was uniformly considered to occur 6 ns after the peak of the excitation function. Data were fit at every point up to 170 ns. 0.05 and 0.03 ethidium/core particle binding ratios were used for the for the 100 mM and 10 mM data, respectively. Sample solutions contained 2.55 $\times 10^{-7}$ M EtBr, 5.125$\times 10^{-6}$ M core particles with an OD = 205, and 10 mM salt buffer whose identity varied with pH. For pH<5 a Sodium acetate buffer was used; for 5<pH<7, MES; ph> 7, Tris HCl. pH was determined following the measurement. All other parameters, including r_o, Dsp, and γ were considered as variable.
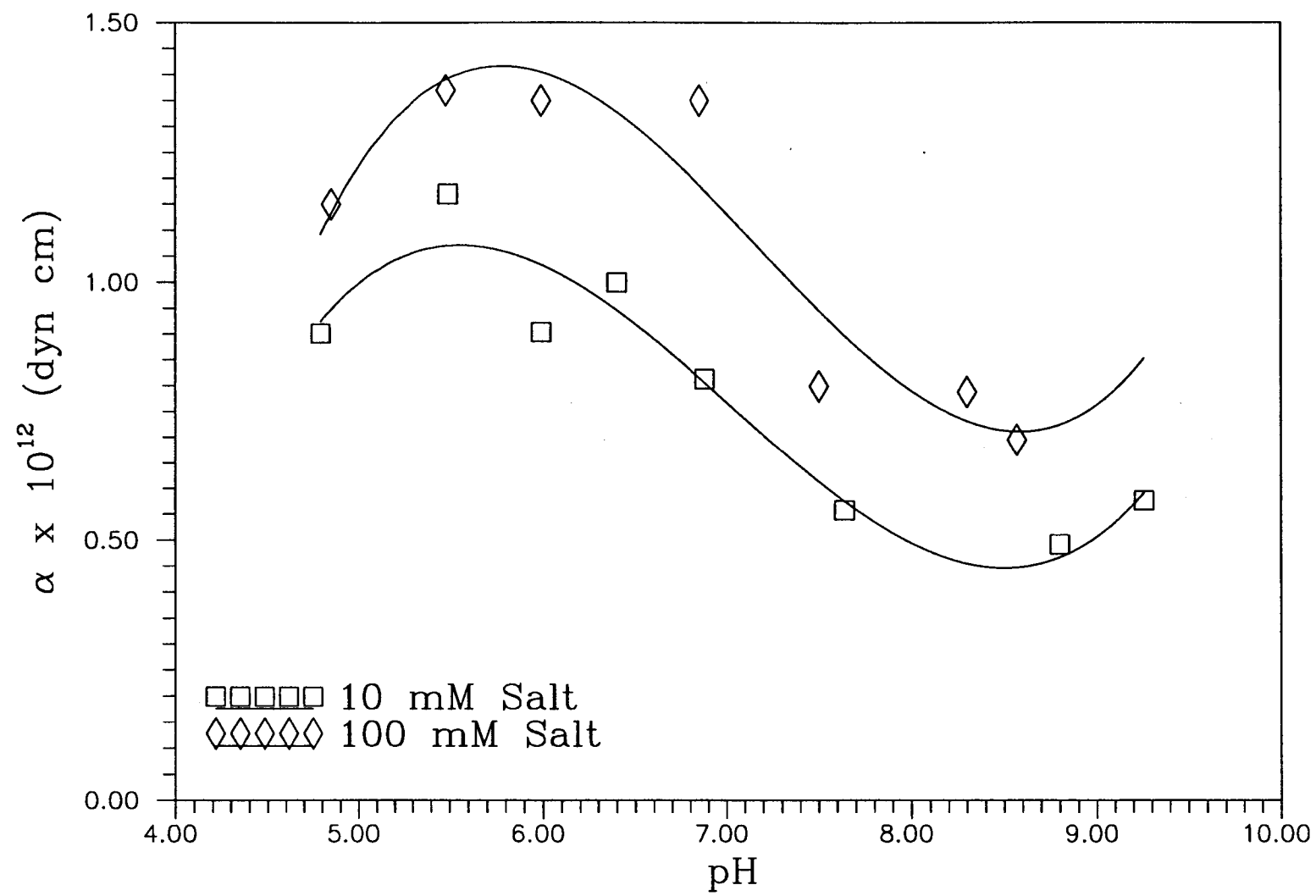
Figure 10

Figure 11. <u>Rotational Diffusion Coefficient as Function of pH (Direct Fit)</u>. The best fit values of D_{sph} are plotted as a function of pH for two salt concentrations. Data analysis was performed as described for Figure 10.
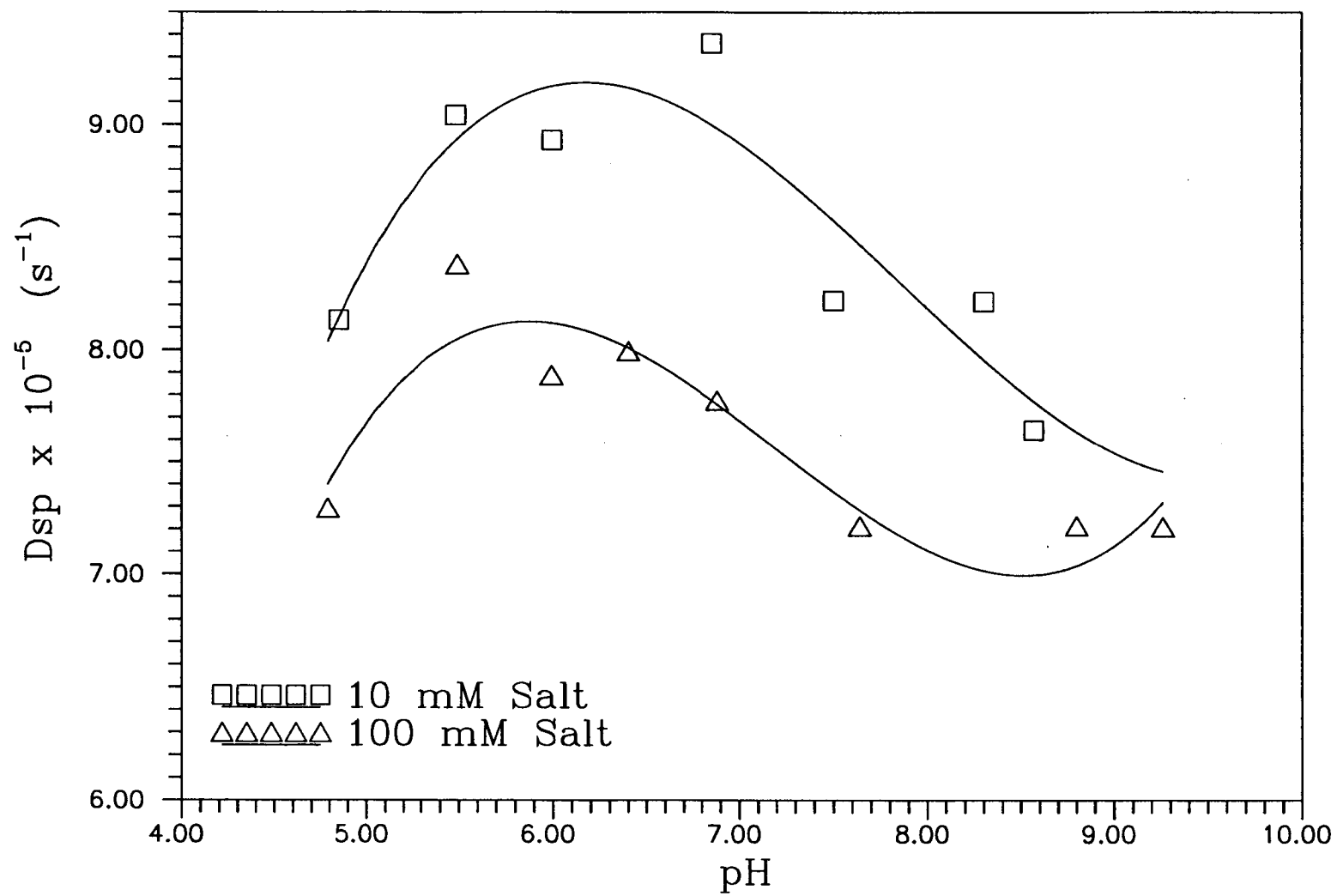
Figure 11

Figure 12. <u>Torsional Coefficient as a Function of pH (Indirect Fit)</u>. The best fit values of α are plotted as a function of pH for solutions with 10 and 100 mM ionic strengths. Values were obtained by iteratively fitting the difference function.
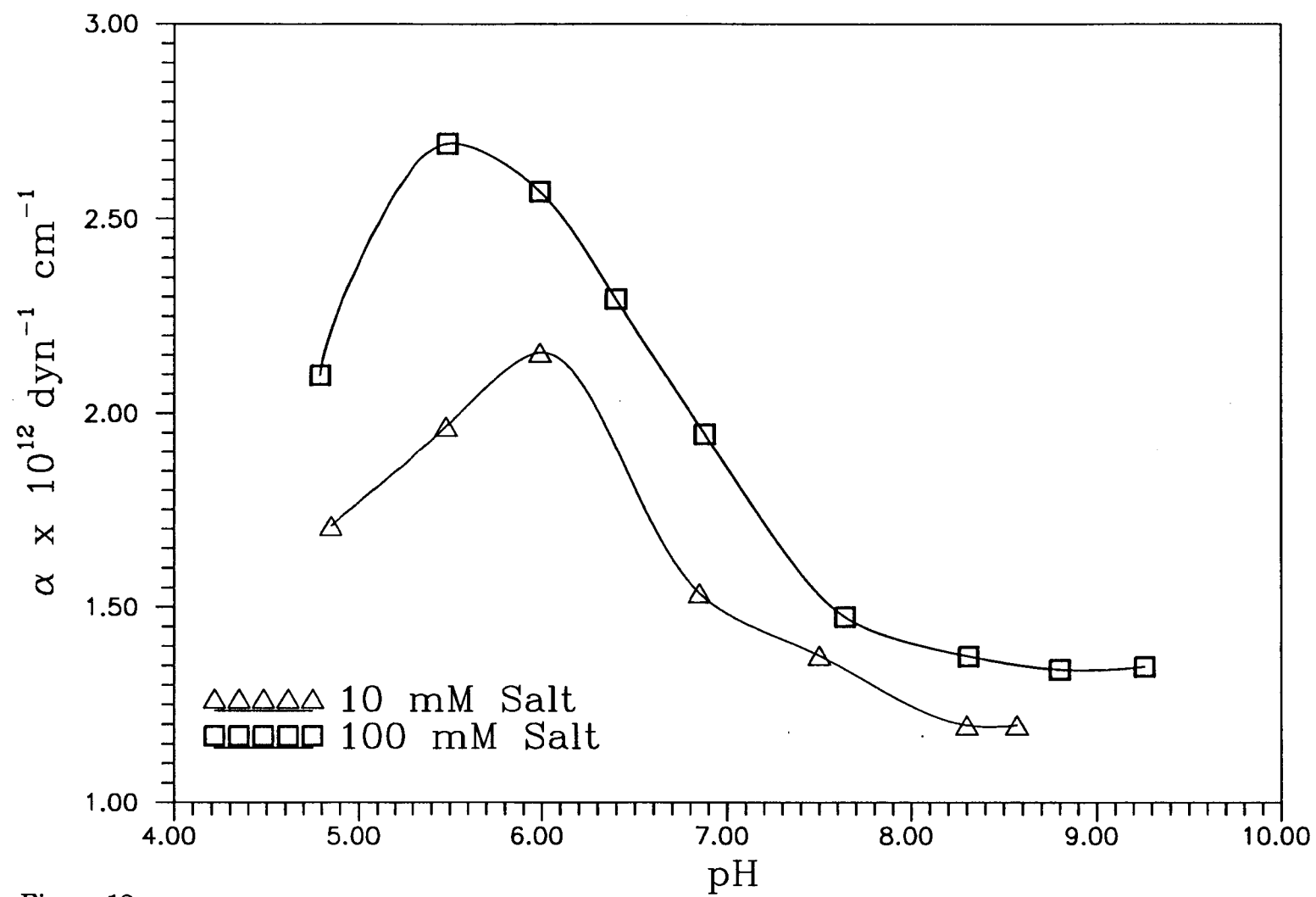
Figure 12

Figure 13. <u>Rotational Diffusion Coefficient as Function of pH (Indirect Fit).</u> Best fit values of D_{sph} as a function of pH for two solutions with a 10 and 100 mM ionic strengths, respectively. Values were obtained by iteratively fitting the difference function.
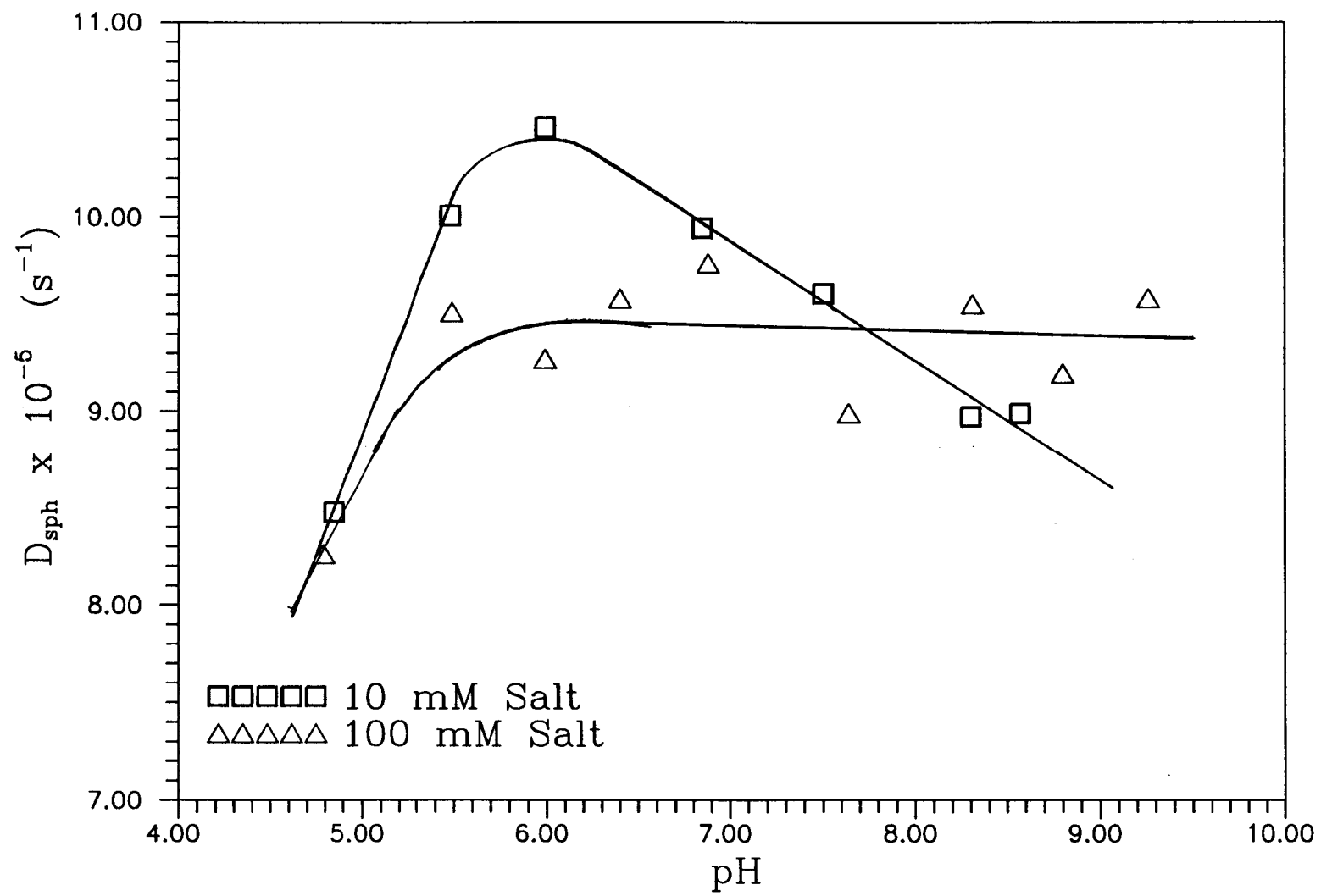
Figure 13

Figure 14. <u>Best Fit of Torsional Coefficient for Different N and pH</u>. The solid lines show the best fit values of α for different N as by obtained iteratively fitting the difference function of pH data. The solid horizontal line at approximately the value of free DNA is included to emphasize that the observed change in α is actually due to a change N.
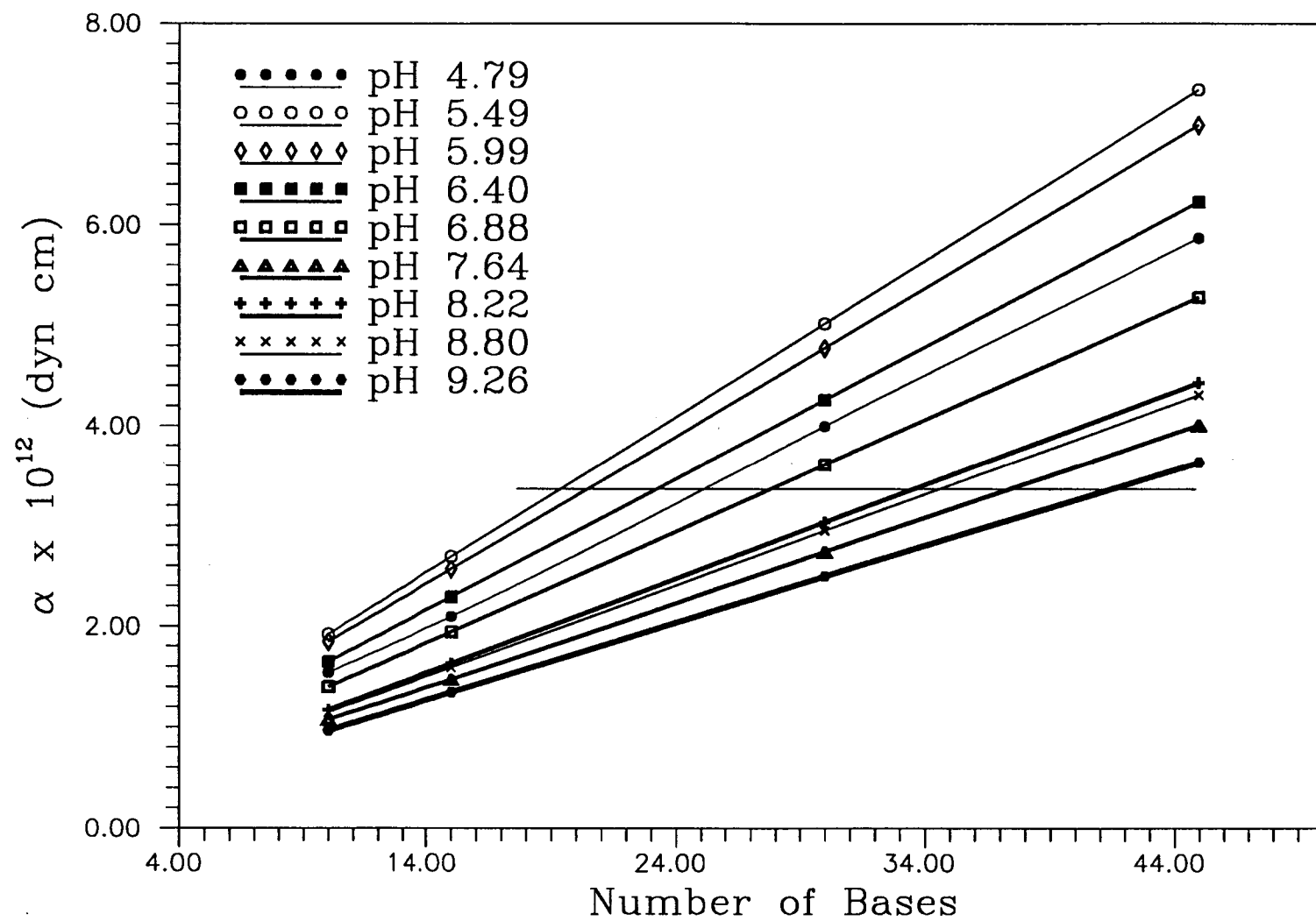
Figure 14

DISCUSSION

We have assumed that nucleosomal DNA can be modeled as a series of rigid rods connected to one another by Hookean springs, and that the rods are constrained to girdle the axis of sphere and are bound to it at both ends to this sphere. Some may disagree with this model; however, we have attained a beautiful fit to the data both at short and long times using this model and are thus very satisfied with it. Given our successes with this model we have sought to fully explore its potential and determine exactly what it is able to tell us about the physical constants or conformation of DNA bound to a nucleosome, as well as if it can be used as an indicator of the effects of ethidium binding and hydrogen ion concentration on the DNA. There is, however, a question which first must be answered.

This question is the choice of N, the number of base pairs between points of attachment. It is more likely than not that the DNA is attached to the nucleosome core particle at a number of places between the two ends. DNase I digestion of nucleosome core particles has been shown to be inhibited at sites approximately 30, 60, 80 and 110 base pairs from the ends of the DNA (Lutter, 1978) suggesting possible contacts with the DNA at these points. Thermal denaturation studies (McMurray *et al.*,1985) as well as studies using NMR, and micrococcal nuclease digestion have indicated that DNA on a core particle melts from the ends inward and is clearly biphasic. These studies suggest attachment points at 20-25 base pairs from either end. Studies calculating the number of phosphate groups which are neutralized suggest that the DNA contacts the histones about once every 10 base pairs, equivalent to one turn of the double helix (Record *et al.*, 1978). For a review of recent research on this field see van Holde (1988). Given this information, it is unlikely that N is as large as 145, and we have therefore primarily considered much smaller values of this parameter. However, we are able to obtain equally good fits for any number of values of N though χ^2 is fractionally better for lower N. If the number of base pairs between points of attachment is any number greater than about 30,

which is very likely, then the Schurr model implies that the torsional rigidity of the DNA on a core particle is higher than or equal to that of free DNA.

One must also consider the possibility that N is less than 30 and that the torsional rigidity of nucleosomal DNA is less than free DNA. This seems illogical because the DNA is obviously subject to a number of constraints in its association with the core particle. Secondly, the number of base pairs per turn for DNA on a nucleosome core particle is thought to be 10.0, in contrast with 10.5 for free DNA (van Holde, 1988). This implies that the DNA on a core particle is more tightly wound than free DNA and suggests that the torsional flexibility of core particle DNA should be impeded, and not encouraged. However, given the similarity between the number we attain for the torsional coefficient, and the one observed for free DNA, it is perhaps most plausible to assume that the torsional rigidity of DNA on a nucleosome core particle is the same as free DNA. This implies that if the Schurr model is correct, that ethidium binds to DNA on a strand which is about 30 base pairs long and which is bound at either end.

It should also be noted that the binding studies also indicate that the recovered values of α (or N) are dependent on binding ratio. While we feel that we have used a low enough dye to core particle ratio that we have avoided the possible consequences of multiple binding and resulting dissociation of the core particle, other researchers may have not. The torsional coefficient reported here is 10 times higher than that previously reported (Schurr and Schurr, 1985) when adjusted for differing values of N. There appear to be no obvious discrepancies in the method of analysis so the difference must result from the data used. The data used in the Schurr and Schurr (1985) study was obtained by Wang *et al.*(1982) using triplet anistropy decays of methylene blue and *a binding ratio of 2.0 dye per core particle*. While it is questionable whether binding of methylene blue to core particles can be compared to ethidium, though both are planar intercalators, we have shown that at binding ratios as low as 1.2 dye per core particle there is evidence that the dynamics of the DNA has been severely affected.

While there is some question as to what is the correct choice of N, the value of γ with respect to that of free DNA is much easier to determine. It is clear from our studies that it must be higher than free DNA. Furthermore, there is little uncertainty about the value of the rotational diffusion coefficient, D_{sph}. The lowest point in the χ^2 surface is approximately 10.0 $x10^5s^{-1}$ regardless of choice of N. We are unable to fit the data as well using a lower value of this parameter. This would suggest that the values we recover are accurate or the model is inaccurate. Discussion of some of the values obtained in other laboratories is given by Brown *et al.* (1990b).

We observe that the data can be separated into two or three pH dependent groups. The first is observed at about pH 5. Decays from samples whose pH was 4.79 appear to have a lower rotational diffusion than the rest. The effects on the torsional coefficient are less obvious. It is questionable as two whether this region can be considered separate from the second whose range is from pH 5 to 7. In this region, (pH 5-7) regardless of the approach to analysis, our data indicate that the DNA has maximal torsional rigidity, or that there are the most frequent contacts between DNA and the histone core. In this region, the torsional rigidity decreases steadily with pH up to pH 7. At pH 7 and beyond there is little change in the torsional rigidity (or distance between histone-DNA interactions). In the region beyond neutrality, either torsional rigidity of the DNA decreases or the number of base pairs between points of attachment increase between pH 5.5 and pH 7.5. The effect of pH on the rotational diffusion coefficient cannot be determined with certainty. Some methods of analysis suggest that this parameter decreases under alkaline conditions. The pH dependent transition with respect to either α or N occurs at a higher acidity at 10 mM salt than it does at 100 mM. The possible low pH transition is not observed at 10 mM ionic strength. This could result from this shift, or it could mean that the transition observed at the higher ionic strength is artifactual.

There have been a great number of studies of the low salt transition. There have been relatively few which have investigated the possibility of a pH transition. A number of those looked specifically at transitions only involving the histones. The studies of the effects of pH on naked histone complexes

(Kawashima *et al.* 1982) are not as relevant to our research as the studies involving DNA. A few early studies (Gordon *et al.* 1979; Zama *et al.* 1978a) have detected a transition between pH 4 and 5 but not around pH 7. Zama *et al.* (1978a) used circular dichroism, laser Raman spectroscopy and the fluorescence of N-(3-pyrene)maleimide covalently bound to histone H-3 to investigate the effects of pH on a nucleosome core particle. In all cases they observed no changes around pH 7, but did detect small changes in the CD at 281 nm (a region sensitive to DNA) and relatively smaller changes at 223 nanometers (a wavelength sensitive to changes in the alpha helicity of the histone residues) between pH 4 and 5. Gordon *et al.*, (1979) have made similar measurements. These researchers showed that at 1 mM ionic strength, the CD at 285 nm was significantly higher at pH 7 than at 4.8. Unfortunately, CD data for pH greater than 7 were not given in this work. Their data also indicated that there was little change in the sedimentation coefficient between pH 5.5 and 9 at ionic strengths between 1 and 11 mM. However, a relatively large change was observed in the sedimentation coefficient, indicating aggregation was observed between pH 4.5 and 5.5 in solutions with an ionic strength of less than 11 mM. In both of these cases the results may be somewhat compromised by the use of nucleosomes which may have been contaminated with histone H1 as well as linker DNA.

More recent studies have detected transitions around neutrality (Libertini and Small 1982; Libertini and Small 1984a; Muller *et al.* 1985;) as well as at pH 4 (Libertini and Small 1982). The transition observed by Libertini and Small, (1982) at pH 4, was shown to result from precipitation using dynamic light scattering measurements at 400 nm. These researchers also noted that proton binding to the core particle can prevent the transition from the high salt form to the low salt from. Libertini and Small (1984a) detected small changes in the CD at 284 nm centered at about pH 7 using core particles at 100 mM ionic strength, but were unable to see any significant changes when using 260 nm, where conformational changes in the histone residues might have been detected. Using circular dichroism, Muller *et al.* (1985) also demonstrated changes in the ellipticity of DNA at 283.5 nm centered at pH 6.65 using a 20 mM phosphate buffer. Libertini and Small

(1984a) also detected a small change in the intrinsic tyrosine steady-state fluorescence anisotropy centered around pH 7, and were able to show that this transition was blocked by cross-linking the protein core with dimethyl-suberimidate. Very little change in the sedimentation coefficient was observed at 5<pH<8.

The results presented here confirm that there is some change in the conformation of the core particle centered near neutrality. The difference between the midpoint of the transition observed in the Libertini and Small experiment (1984a), and the Muller *et al.* (1985) experiment may be due to the different ionic strength used. It is possible that the ionic strength of the solution dictates to some degree the point at which the pH transition will occur- -the lower the ionic strength, the lower the pH needed to achieve the transition. If this is true, the results presented in these experiments, where the transition appears to occur at higher hydrogen ion concentrations at lower ionic strengths, suggests that the same change is being observed here as in the Muller *et al.* and Libertini and Small experiments.

The transition observed between pH 6 and 7 could result from a number of physical phenomena. While most (Libertini and Small 1984a; Eickbush *et al.* 1978; Gordon *et al.* 1978) suggest titration of a histidine residue to be responsible for the observed transitions, there are other possibilities -- protonation of base pairs for one. A new pH dependent conformation of DNA, dubbed the H form, have been reported in the recent years (Lyamichev *et al.*,1987). This conformation which occurs mostly in homopyrimidine tracts is thought to be dependent on the protonation of the cytosine residues. The transition to this form, unlike the transition to the Z form is highly dependent on pH and occurs at a higher frequency at pH less than neutrality. Of course the idea that the changes in the DNA flexibility on the nucleosome core particle may be due to loss of protonation in pyrimidine residues is entirely speculative.

In conclusion, our results indicate that the flexibility of DNA on a nucleosome core particle is similar to that of free DNA (Thomas *et al.* 1980) and the amount of friction experienced by the DNA on a nucleosome core particle is

much higher than that experienced by free DNA. We also believe that value of the rotational diffusion coefficient is higher than has previously been reported (Harrington 1981; Wang *et al.* 1982; Schurr and Schurr 1985). Secondly, our results indicate that there is a very pH-dependent significant change in the fluorescent anisotropy decay of DNA bound to a nucleosome core particle. The fluorescence is depolarized more rapidly at high pH suggesting that the DNA has more freedom of movement. The most plausible explanation for the observed changes is that the number of places the DNA makes contact with the histone core is decreased as the hydrogen ion concentration is decreased. This may result from titration of a histidine residue as this amino acid has an isoelectric point in the vicinity of the transition observed.

BIBLIOGRAPHY

Allison, S. A. & Schurr, J. M. (1979) Chem. Phys. 41, 35-59.

Andrews, D. F., Bickel, P.J., Hampet, F. R., Huber P. J., Rogers, W. H., Tukey, J. W. (1972) Robust Estimates of Location, Princeton University Press.

Ashikawa, I., Kinosita, K., Jr., Ikegami, A., Nishimura, Y., Tsuboi, M., Watanabe, K. & Iso, K. (1983) J. Biochem. 93, 665-668.

Barkley, M. D. & Zimm, B. H. (1979) J. Chem. Phys. 70, 2991-3007.

Beechem, Joseph M. & Gratton, Enrico (1988) in Time-Resolved Laser Spectoroscopy in Biochemistry (Lakowicz, J. R., ed.), pp. 70-81, SPIE Proceedings, Vol. 909, Bellington, WA. .

Beechem, J. M., Knutson, J. R., Brand, L. (1986) Biochem. Soc. Trans. 14, 832-835.

Bevington, Philip R. (1969) Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill Book Company, New York.

Belford, G. G., Belford, R. L. & Weber, G. (1972) Proc. Nat. Acad. Sci. USA 69, 1392-1398.

Brown, D. W., Libertini, L. J. & Small, E. W. (1990a) in Time-Resolved Laser Spectroscopy in Biochemistry II (Lakowicz, J. R., ed.), SPIE Proceedings Vol. 1204, Bellingham, WA.

Brown, D. W., Libertini, L. J. & Small, E. W. (1990b) in preparation.

Cantor, C. R. & Schimmel, P. R. (1980) Biophysical Chemistry, W. H. Greeman and Company, San Francisco.

Chambard, J. C., & Pouryssegur, J. (1986) Exp. Cell. Res., 164(2), 284-94.

Chuang, T. J. & Eisenthal, K. B. (1972) J. Chem. Phys. 57, 5094-5097.

Crothers, D. M., Dattagupta, N., Hogan, M., Klevan, L. & Lee, K. S. (1978) Biochemistry 17, 4525-4533.

Davie, J. R., Saunders, C. A., Walsh, J. M. & Weber, S. C. (1981) Nucleic Acids Res. 9, 3205-3216.

Ehrenberg, M. & Rigler, R. (1972) Chem. Phys. Lett. 14, 539-544.

Eickbush, T. H. & Moudrianakis, E. N. (1978) Biochemistry 17, 4955-4964.

Eisenfeld, J. & Ford, C. C. (1979) Biophys. J. 26, 73-83.

Erard, M., Das, G. D., de Murcia, G., Mazen, A. Bouyet, J., Champagne, M., Daune, M. (1979) Nuelaic Acids Res., 3231-3253.

Genest, D. & Wahl, P., (1981) Biochimie, 63, 561-564.

Genest, D., Wahl, P., Erard, M., Champagne, M. & Daune, M. (1982) Biochimie 64, 419-427.

Gillies, Robert J. (1982) in Intracellular pH, (Nuccitelli, Richard & Deamer, David W., eds.) 341-359, Alan R. Liss, Inc., New York.

Gordon, V. C., Knobler, C. M., Olins, D. E. & Schumaker, V. N. (1978) Proc. Nat. Acad. Sci. USA 75, 660-663.

Gordon, V. C., Schumaker, V. N., Olins, D. E., Knobler, C. M. & Horwitz, J. (1979) Nucleic Acids Res. 6, 3845-3858.

Grandin, N. & Charbonneau, M. (1989) Exp. Cell. Res., 185(2), 436-452.

Harrington, R. E. (1981) Biopolymers 20, 719-752.

Hutchings, J. J. & Small, E. W. (1990) in Time-Resolved Laser Spectroscopy in Biochemistry II (Lakowicz, J. R., ed.), SPIE Proceedings Vol. 1204, Bellingham, WA.

Kawashima, S. & Imahori, K. (1982) J. Biochemistry 91, 959-966.

Libertini, L. J. & Small, E. W. (1980) Nucleic Acids Res. 8, 3517-3534.

Libertini, L. J. & Small, E. W. (1982) Biochemistry 21, 3327-3334.

Libertini, L. J. & Small, E. W. (1983) Rev. Sci. Instrum. 54, 1458-1466.

Libertini, L. J. & Small, E. W. (1984a) Nucleic Acids Res. 10, 4351-4359.

Libertini, L. J. & Small, E. W. (1984b) Anal. Biochem. 138, 314-318.

Libertini, L. J. & Small, E. W. (1987a) Anal. Biochem. 163, 500-505.

Libertini, L. J. & Small, E. W. (1987b) Nucleic Acids Res. 15, 6655-6664.

Libertini, L. J. & Small, E. W. (1989) Biophys. Chem. 34, 269-282.

Lutter, L. C. (1978) J. Mol. Biol. 124, 391-420.

Lyamichev, V. I., Mirkin, M. S., Frank-Kamenetskii M. D. (1987) J. Biomol Structure & Dynamics, 5, 275-282.

Marquardt, D. W. (1963) J. Soc. Indust. Appl. Math. 11, 431-441.

McGhee, J. D. & von Hippel, P. H. (1974) J. Mol. Biol., 86, 469-489.

McMurray, C. & van Holde, K. (1986) Proc. Nat. Acad. Sci. USA 83, 8472-8476.

McMurray, C. T., van Holde, K. E., Jones, R. L. & Wilson, W. D. (1985) Biochemistry 24, 7037-7044.

McMurray, C. T., Small, E. W. & van Holde, K. E. (1990a) in preparation.

McMurray, C. T. & van Holde, K. E. (1990b) Biochemistry submitted for publication,.

Muller, S., Bertrand, E., Erard, M. & van Regenmortel, M. H. V. (1985) Int. J. Biol. Macromol. 7, 113-119.

Nuccitelli, Richard & Heiple, Jeanne M. (1982) in Intracellular pH, (Nuccitelli, R. & Deamer, D. W., eds.) 567-586, Alan R. Liss, Inc., New York.

Ober, S. S. and Pardee, A. B. (1987) Proc. Natl. Acad. Sci., 84(9), 2766-2770.

O'Connor, D. V. & Phillips, D. (1984) Time-correlated Single Photon Counting, Academic Press, London.

Olins, D. E., Bryan, P. N., Harrington, R. E., Hill, W. E. & Olins, A. L. (1977) Nucleic Acids Res. 4, 1911-1931.

Perrin, P. F. (1934) Le J. de Physique et le Radium (Serie VII) 5, 487-511.

Record, M. T., Anderson, C. F. & Lohman, T. M. (1978) Quart. Rev. Biophys. 11, 103-178.

Ricka, J., Amsler, K. & Binkert, T. (1983) Biopolymers 22, 1301-1318.

Schurr, J. M. (1984) Chem. Phys. 84, 71-96.

Schurr, J. M. & Schurr, R. L. (1985) Biopolymers 24, 1931-1940.

Schurr, J. Michael, Fujimoto, Bryant S., Wu, Pengguang, & Song, Lu (1989) in Modern Fluorescence Spectroscopy, Vol. 1: Principlies and Techniques (Lakowicz, J. R., ed.), Plenum Publishing Corpartion, New York.

Schuyler, R. & Isenberg, I. (1971) Rev. Sci. Instrum. 42, 813-817.

Small, E. W., Fanconi, B. & Peticolas, W. L. (1970) J. Chem. Phys. 52, 4369-4379.

Small, E. W., Libertini, L. J. & Isenberg, I. (1984) Rev. Sci. Instrum. 55, 879-885.

Small, E. W. & Anderson, S. R. (1988a) Biochemistry 27, 419-428.

Small, E. W., Libertini, L. J. & Small, J. R. (1988b) in Time-Resolved Laser Spectroscopy in Biochemistry (Lakowicz, J. R., ed.), pp. 97-107, SPIE Proceedings Vol. 909, Bellingham, WA.

Small, E. W., Libertini, L. J., Brown, D. W. & Small, J. R. (1989a) in Fluorescence Detection III (Menzel, E. R., ed.), pp. 36-53, SPIE Proceedings Vol. 1054, Bellingham, WA.

Small, E. W. (1989b) in Modern Fluorescence Spectroscopy, Volume 1: Principles and Techniques (Lakowicz, J. R., ed.), Plenum Publishing Corporation, New York.

Small, E. W., Libertini, L. J., Brown, D. W. & Small, J. R. (1990) Optical Engineering, in press.

Snare, M. J., Tan, K. L. & Treloar, F. E. (1982) J. Macromol. Sci. Chem. 17, 189-201.

Thomas, J. C., Allison, S. A., Appellof, C. J. & Schurr, J. M. (1980) Biophys. Chem. 12, 177-188.

van Holde, K. E. (1988) Chromatin, Springer-Verlag, New York .

Wahl, P., Tawada, K. & Auchet, J. C. (1978) Eur. J. Biochem. 88, 421-424.

Wang, J., Hogan, M. & Austin, R. H. (1982) Proc. Nat. Acad. Sci. USA 79, 5896-5900.

Webb, Dennis J. & Nuccitelli, Richard (1982) in Intracellular pH (Nuccitelli, Richard and Deamer, David W., eds.), 294-324, Alan R. Liss, Inc., New York.

Wu, H. -M., Dattagupta, N., Hogan, M. & Crothers, D. M. (1979) Biochemistry 18, 3960-3965.

Zama, M., Olins, D. E., Prescott, B. & Thomas, G. J., Jr. (1978a) Nucleic Acids Res. 5, 3881-3897.

Zama, M., Olins, D. E., Wilkinson-Singley, E. & Olins, A. L. (1978b) Biochem. Biophys. Res. Comm. 85, 1446-1452.

APPENDIX

APPENDIX

As discussed above, most of the software used to analyze the data in these experiments was developed exclusively for this project and has not been previously catalogued. The following description of the programs written by the author is provided for the individual who may use or modify these programs. A chart giving the possible pathways through which data can flow is given in Figure 15. Files containing data are represented by a polygons (a reminder of the days in which punch cards were used to carry data) and programs are represented by rectangles. Programs developed by the author are denoted with an *. It is these with which this discussion will be principally concerned.

GLOBAL

This program will directly fit an anisotropy decay which has been corrected for the convolution artifact to the Schurr expression for the anisotropy (equation 1). There are five parameters required by the Schurr expression, α, γ, D_{sph}, r_o, or N. Each of these five parameters can be specified as either variable or constant. While this program is capable of fitting N, it is not recommended-- the program has difficulty understanding that N can only take integer values and becomes confused. This program can be used to analyze one or more datasets either separately (as described in DATA ANALYSIS: General Case) or globally (as described in DATA ANALYSIS: Global Analysis). For each dataset to be fit, there is an input file, *parm-file(n)*. Here *(n)* indicates the order in which the constant-files are read into the program. This file contains information such as the starting value for each of the five parameters, α, γ, D_{sph}, N or r_o; and a *linkgroup* for each of the five parameter. *Linkgroup* = -1 means that the parameter should be considered constant; *linkgroup* = 0 means that the parameter is variable and independent; and *linkgroup* > 0 means that the parmeter is variable and its identity is the same as that of the corresponding parameter from *parm-file(linkgroup)*. Thus when *linkgroup* > 0 for a parameter, it means that information from other datasets will be used to

determine the best value of that parameter. *Parm-file* also contains information about the desired analysis range, the name of the FF file (laboratory standard IBM or Microsoft format) containing the anisotropy data to be fit, and the value of certain constants, such as temperature. The input file later becomes the receptacle for program output. At the end of the program, the final value of the parameters, as well as the value of χ^2 are written to *parm-file*. *Parm-file* will also serve as input to GENDECAY, a program which will generate a best fit file and deviation function from the parameters provided.

When the program is run, it will first ask you whether you wish to print the data. Generally you respond with a logical "false" (or "n"), unless you are debugging the program, (or like to see evidence that the program hasn't crashed--something about which the novice will often wonder as this program is very slow). The next information requested is an integer value for *step-number*. Enough memory is allocated to fit up to 512 points from 5 datasets simultaneously. To fit a longer time range, including more than 512 datapoints you may specify that only every other, or every fourth point be considered in the fitting procedure by entering 2, 4 *etc.* respectively at this point. This option can also be used to speed the fitting procedure, though it does waste data. You will next be requested to provide the number of files which you wish to analyze. The program will then demand the names of that number of *parm-files*. These files should be located in the same directory from which the program is being called. For each *parm-file* an interactive session will be established in which the contents of *parm-file* are displayed and the user is allowed to enter changes. During this time you will be allowed to change the starting values of any of the five parameters, their value of *linkgroup*, the name of the dataset to be analyzed, or the analysis range. The *first-channel-of -data* specifies where time zero should be. The calculated anisotropy will have a value of r_0 at this point. *First-chi-squared* names the channel where the calculation of χ^2 will start and *last-channel-of-data* indicates where this calculation will end. The difference between *last-channel-of-data* and *first-channel-of-data* should not exceed 512 unless *step-number* > 1. If no changes to the contents of *parm-file* are

necessary, or if editing is complete, you may type "g" to go on to the next *parm-file* or begin the fitting procedure.

As mentioned previously, the best-fit values of the parameters determined in the analysis will be written to *parm-file* once χ^2 ceases to diminish, usually in three or four iterations. The results of the analyses can be visualized by running GENDECAY as described below.

CONVOL

This program determines values of α, γ, D_{sph}, and/or r_0 by iteratively fitting the difference function as described in DATA ANALYSIS: Deconvolution. Prior to running this program, one must fit the sum function to a sum of exponentials. The values of amplitudes (a_i) and lifetimes (τ_i) will be needed as input. Also needed is the value of **S** which was used to generate the sum function. This program uses an EFF type file as input. It is essentials to have a good E file with this program. Also, this program works best when there is a large number of counts in the data.

While much of the code is the same, this program is not an sophisticated as GLOBAL. The user is limited in the number of options available. First, there is no provision for fitting the parameter N. This parameter is assigned a constant value at the start of the program. Secondly, the user does not have the liberty of specifying that any of the four remaining parameters should be considered constant or variable in any combination. The user is only allowed to enter the *number-of-parameters-to-be-fit*. If the number is four, then α and γ, r_0 and D_{sph} will all be considered variable. If this number is three then only α, γ, and r_0 will be considered variable; for two only α and γ. While theoretically this program could process two or more datasets simultaneously, as does GLOBAL, the memory limitations of the Microsoft FORTRAN compiler we use have made the implementation of such a feature all but impossible. An intrepid programmer might wish to investigate this matter further.

The first information requested by the program is whether the user desires to run the job in *batch mode*. A logical "true" (or "y") allows the user to consecutively process any number of datasets without being present. Since this program is VERY SLOW, this is one of the program's most useful features. To do this, the sequence of commands which would be used in a normal run of the program are stored in a file called "INFILE.DAT." This file must reside in the same directory from which the program is being called. A negative to *batch-mode* will cause input to be read from the keyboard.

Before entering the batch mode (if so desired) the program will request the user to provide the *number-of-datasets* to be processed. The following sequence will then be replayed for each of the datasets. In between, however, the program will perform a fitting routine which can take hours. It is for this reason that the batch mode is suggested unless the *number-of-datasets* = 1.

For each dataset, the program will first enter an interactive session where the values of the parameters used the last time the program was executed are displayed and edited much as described above. This program uses a numerical convolution recipe. For it to be successful the user must specify the *first-channel-of-data* be such that integral of E(t) up to this channel will contain less than 1% of the data. The *last-channel-of-data* can be any number up to the very last channel. The program has sufficient memory allocated to process up to 1024 points of data and is entirely capable of fitting the entire decay, including the noise at the end. One can calculate the value of χ^2 from any point between the first and last channels of data up to the last channel of data. This point, *first-chi-squared,* is typically chosen to be the channel where $F_{\parallel}(t)$ + $2F_{\perp}(t)$ is at its maximum. This program currently does not allow the user the option of specifying that only alternate points should be fit. (Upon examining the source code one might note that this feature has been installed as evidenced by frequent appearance of the variable *istep*. However, it is not a feature which is currently available as it has not been adequately tested. The same intrepid user might also wish to investigate implementing this option while modifying to program to work globally).

Next the program will request the *number-of-exponentials*, and then that number of pairs of values of amplitudes (a_i) and lifetimes (τ_i). A value for the sensitivity correction and the starting values of α and γ will also be requested at this point and finally the *number-of-parameters-to-be-fit*.

Following the completion of the analysis of one dataset, the program will ask if the user wishes to write the results to disk. After specifying "y" or "true", the user will be asked to supply a file name. Three datasets will be then written to this file using the laboratory standard format. The first will contian the function

$$
E(t)*\left\{ r(t) \sum_{i=1}^{n} a_i e^{-t/\tau} \right\},
$$

where r(t) is calculated from the Schurr expression using the parameters determined in the analysis. The second dataset will contain $F_{\parallel}(t) - S2F_{\perp}(t)$.

The third dataset will contain the deviation function comparing the two. The program PLOTS can now be used to view the results by plotting dataset 1 on top of dataset 3.

Additionally, best fit values of the parameters as well as the variances for each parameters, and the final values of χ^2 will be recorded in a file called "OUTFILE.DAT" which will reside in the directory from which the program is being called.

GENDECAY

This program uses the same *parm-file* as input as is used by GLOBAL. Its purpose is to allow the user to select the best analysis results from GLOBAL and put them in a format whereby they can be visualized. Before running this program, values of *linkgroup* in *parm-file* should be set to either -1 or 0. As mentioned above, *parm-file* contains all the information needed to generate an anisotropy decay function from the Schurr expression. Additionally, the *parm-*

file will contain the name of a file containing a dataset to which the calculated anisotropy can be compared. As output this program will produce a file to be named by the user in the laboratory standard format. The first dataset in the file will contain the calculated anisotropy function. The third dataset will contain the first dataset from the file named in *parm-file*. Usually this will be the experimental corrected anisotropy decay from which the best fit parameters in *parm-file* were derived, (so to allow the user to easily compare the two using PLOTS), but it does not necessarily need to be so. The second dataset will contain the deviation function resulting from the comparison of the first to the third datasets.

AMT

This simple program is used to calculate a sine function M estimate, as described in DATA COLLECTION, from a series of sensitivity correction data. The input consists of the contents of a file named "SC.DAT." This file should reside in the directory from which the program is being called and contain multiple groups of 16 numbers separated by some identifier (usually a four character file name). The program will read these numbers, 16 at a time, and determine the mean after discarding any numbers which lie outside of some range. The output will be displayed on the screen, and the program will stop when it runs out of numbers. The number 16 is written in the source code. Changing it involves simply re-linking and compiling the program.

Figure 15. <u>Data Analysis Flow Chart.</u> This chart depicts the different pathways through which raw data recieved from the pulse height analyzer can flow. Programs (rectangles) developed for this project are denoted with an * and are documented in this section. Pre-existing programs are documented elsewhere.