NOTE TO USERS

DISSERTATION

STRUCTURAL AND BIOCHEMICAL INVESTIGATIONS ON DNA ACCESSIBILITY AND DYNAMICS WITHIN THE NUCLEOSOME

Submitted by

Rajeswari S. Edayathumangalam

Department of Biochemistry and Molecular Biology

In partial fulfillment of the requirements

for the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2005

COLORADO STATE UNIVERSITY

April 1, 2005

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY RAJESWARI S. EDAYATHUMANGALAM ENTITLED "STRUCTURAL AND BIOCHEMICAL INVESTIGATIONS ON DNA ACCESSIBILITY AND DYNAMICS WITHIN THE NUCLEOSOME" BE ACCEPTED AS FULFILLING IN PART THE REQUIREMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

(Oren P. Anderson)

(Jennifer K. Nyborg)

(Olve B. Peersen)

(Laurie A. Stargell)

(Karolin Luger)

(Marvin R. Paule)

ABSTRACT OF DISSERTATION

STRUCTURAL AND BIOCHEMICAL INVESTIGATIONS ON DNA ACCESSIBILITY AND DYNAMICS WITHIN THE NUCLEOSOME

The 'blueprint of life' is organized as highly compacted chromatin in the eukaryotic nucleus. The fundamental building block of chromatin is the nucleosome, which comprises the nucleosome core particle in association with 20 to 80 bp of linker DNA and the linker histone protein. The nucleosome core particle, which is the first step of compaction of naked DNA, consists of 147 bp of DNA wrapped around an approximately equal mass of a proteinaceous histone octamer. The nearly 100,000-fold level of final compaction of DNA achieved in the nucleus places heavy restraints on the cellular machinery. Processes like transcription, DNA replication, DNA repair and recombination, which use DNA as the working platform, must now navigate this very tightly packed scaffold to gain access to specific target sites. Understanding how the cell maintains the fluidity and dynamics of the underlying DNA without compromising the structural integrity of chromatin will provide invaluable insights into the mechanistic details of biological processes like transcription. We approach these questions using the monunucleosome as a model system and pyrrole-imidazole polyamides as small-molecule probes for DNA accessibility and dynamics. First, we perform *in vitro* temperature-induced nucleosome repositioning assays and show that polyamides inhibit the temperature-induced translocation of the

histone octamer on the DNA, thereby suggesting a mechanism for temperature-induced nucleosome repositioning. Second, we report the co-crystal structure of the nucleosome core particle in complex with a 'bivalent' polyamide dimer 'clamp'. The clamp targets a nucleosomal 'supergroove' (two non-contiguous sites that are 80 bp apart on linear DNA) and significantly improves nucleosome crystal diffraction as well as dramatically stabilizes nucleosomes against dilution-induced dissociation *in vitro*. These findings identify 'supergrooves' as unique platforms for molecular recognition of condensed DNA. Third, using x-ray crystallography and footprinting techniques, we demonstrate that nucleosomal DNA exists as a mixture of multiple twist-defect intermediates in solution. Fourth, the structure of a tailless H4-containing nucleosome bound by a polyamide shows that ligand binding leads to unraveling of the ends of the nucleosomal DNA. This also leads to an alternate mode of crystal packing, which does not involve the DNA ends, and thereby reveals other possible interactions between nucleosomes. Taken together, these studies elucidate how the various components resident within the nucleosome modulate nucleosome stability, accessibility, and dynamics. Our results also further our understanding of compacted DNA as a biological substrate.

Rajeswari S. Edayathumangalam
Department of Biochemistry and Molecular Biology
Colorado State University
Fort Collins, CO 80523
Spring 2005

ACKNOWLEDGEMENTS

"கேடில் விழிச்செல்வம் கல்வி யொருவர்க்கு
மாடல்ல மற்ற பிற."

- Thiruvalluvar, ca. 130 B.C.

*"A man's learning is an imperishable and precious wealth.
All other possessions are less golden."*

As I look back, graduate school has been one of the most gratifying experiences in my life so far. There is one faithful emotion I associate with all these years; it is that endless surge of excitement towards science. All this excitement would have stayed a dream had it not been for the wonderful people that have helped make it a dream-come-true.

My love, joy and gratitude know no bounds for the most beautiful people in my universe, my parents, E. N. Balambal and E. V. Sundaram, and my sister, Kalpana Sundaram. I feel fortunate to be raised by model parents whose love, trust and friendship have always been most encouraging. Their utmost respect and passion for education and career have been a crucial driving force in my life. Their fantastic yet most humble attitude towards life and living is a constant source of amazement to my friends and me. I also thank my wonderful sister for her love and friendship through all these years and my brother-in-law, Vivek Ramachandran for his friendship. My parents and my sister are the sole reason I even dare to dream.

Much praise and thanks go to my advisor, Karolin Luger, an amazing friend, guide and philosopher. Karolin welcomed me to her lab with open arms even though I did not

do a rotation in her lab and even when she did not know me. I think we have always worked hand-in-glove and shared a wonderful mentor-student relationship. She has consistently provided me with ample opportunities and allowed me to work unfettered and independently, something I have immensely cherished over the years. At the same time, she has always offered helpful scientific input, and has been the most approachable and available person. I admire her ambition and enthusiasm for science, and her ability to communicate science effectively. I am grateful to her for giving me the opportunity to maintain and repair the x-ray generator, and for allowing me to attend innumerable scientific meetings and workshops over the years. I will also think fondly of all the fun we have had outside the lab and those insane moments of laughter at the synchrotron in the wee hours of the morning.

Each and every one of the past and present Luger lab members have helped in creating a very rich and stimulating environment that has been most conducive to experimental science and learning. All members have brought a unique facet to the lab and I have always appreciated their company, technical help and scientific input. I would like to especially mention Pam Dyer and Uma Muthurajan for their unrelenting help and friendship. I extend my earnest thanks to every member of my PhD committee for their guidance, technical help and critical review of my work. The remaining faculty and staff in our department provided me with the help and resources that were crucial to the completion of my work.

Learning the ropes of crystallography has been a lot of fun thanks to all the help I have received over the years from the innumerable stellar crystallographers I have met at

meetings and over bulletin boards like CCP4BB. The useful scientific advice and help from friends like Gerry McDermott, Ralf Grosse-Kunstleve and Peter Zwart will always be remembered. The generous help offered by Paul Adams and Jeffrey Bonanno was pivotal to the completion of one of my projects. I was always reluctant to use CCP4 programs in the past but not anymore, thanks to Jeffrey Bonanno. I have gained significantly from our numerous scientific discussions and I much respect his help and friendship during the finishing stages.

I would like to thank all my friends in Fort Collins and in the rest of the globe. I affectionately recollect the help and support from Karen and Craig Foltz, and John and Mary Dubler when I moved to Fort Collins. I count myself lucky to have friends like Uma Muthurajan, Venugopal Pujari, Sangita Kalarickal and Nischal Piratla. They have treated me like family and have always offered unflinching support. I am also thankful to Uma for critical reading of my thesis and for her helpful comments, and to Venu and my labmates for their help with my thesis and seminar preparation. Jes Kuruvilla has been an incredible friend throughout this journey. His encouragement and friendship have provided the much-appreciated support system that has sustained me through this entire experience.

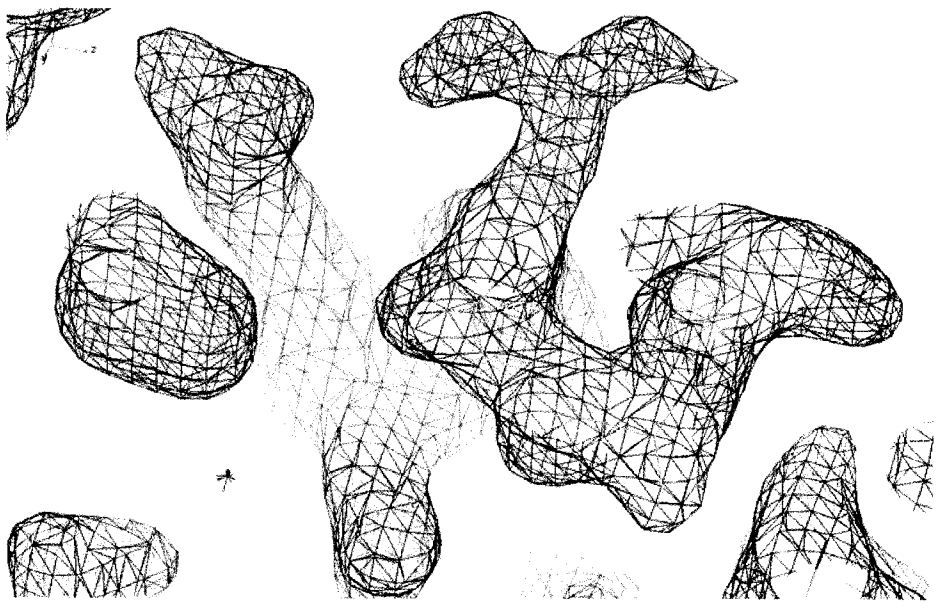
And, Lord Alfred Tennyson once said, "More things are wrought by prayer than this world dreams of".

"கற்க கசடற கற்றவை கற்றபின்
நிற்க அதற்குத் தக."

- Thiruvalluvar, ca. 130 B.C.

"Learn perfectly all that you learn, and thereafter keep your conduct worthy of that learning."

TABLE OF CONTENTS



Chapter 1

Review of the Literature

1.1 Recent advances in chromatin biology

The experimental findings . . . are consistent with the view that histones, by combining with DNA, cause it to assume a condensed state, no longer active as a primer in the synthesis of ribonucleic acid.

- Allfrey and Mirsky, 1964

(Excerpt from Van Holde, 1988)

The past decade or so has seen a plethora of exciting developments in the field of chromatin biology. The elucidation of the first high-resolution crystal structure of the nucleosome core particle (NCP) containing *Xenopus laevis* histone proteins (Luger et al., 1997) followed by the determination of a variety of NCP structures from various organisms (see Luger, 2003 for a comprehensive list of all published nucleosome structures), of NCPs containing histone variants (Suto et al., 2000), and of NCPs bound by small-molecule ligands (Edayathumangalam et al., 2005; Edayathumangalam et al., 2004; Suto et al., 2003) have added significantly to the understanding of DNA structure and compaction. Genetic, biochemical, mutational and structural studies have revealed how some highly specific histone mutations (the SWI/SNF-independent (SIN) and the loss of rDNA silencing (*lrs*) mutants), which occur within the core structured regions of the canonical histone proteins and which are localized at the protein-DNA binding interface formed by histones H3 and H4, modulate nucleosome structure and function (Flaus et al., 2004; Luger, 2003; Muthurajan et al., 2004; Park et al., 2002 and references therein). The identification of histone modifying enzymes and a myriad of histone post-translational modifications soon led to the proposal of a 'histone code' hypothesis to help explain their effects on gene expression and higher-order chromatin structure (Fischle et al., 2003; Strahl and Allis, 2000). Other significant milestones include the characterization of several ATP-

dependent chromatin-remodeling factors and the *in vitro* nucleosome repositioning brought about by these factors (Flaus and Owen-Hughes, 2004; Korber and Horz, 2004), studies on higher-order chromatin assembly and function (Hansen, 2002; Lu and Hansen, 2003), biochemical and structural investigations on histone chaperones (Akey and Luger, 2003), and studies on *in vitro* nucleosome assembly and dynamics, including those of nucleosomes containing histone variants (Bao et al., 2004; Park et al., 2004a; Park et al., 2004b).

Our current view of the nucleosome is no longer one of a 'millstone' in the path of the cellular 'workhorses' and a 'mundane' global repressor of transcription, but of a dynamic structural and functional entity central to regulating gene expression. Although recent advances have underscored the intimate involvement of chromatin at every step in gene expression and transcription regulation, there is a mechanistic void in how all of the above-described facets of chromatin structure and function orchestrate routine cellular processes like transcription, DNA replication, and DNA repair. Many such open questions remain, which will continue to foster exciting developments in chromatin biology in the years to come.

1.2 Chromatin and the nucleosome core particle

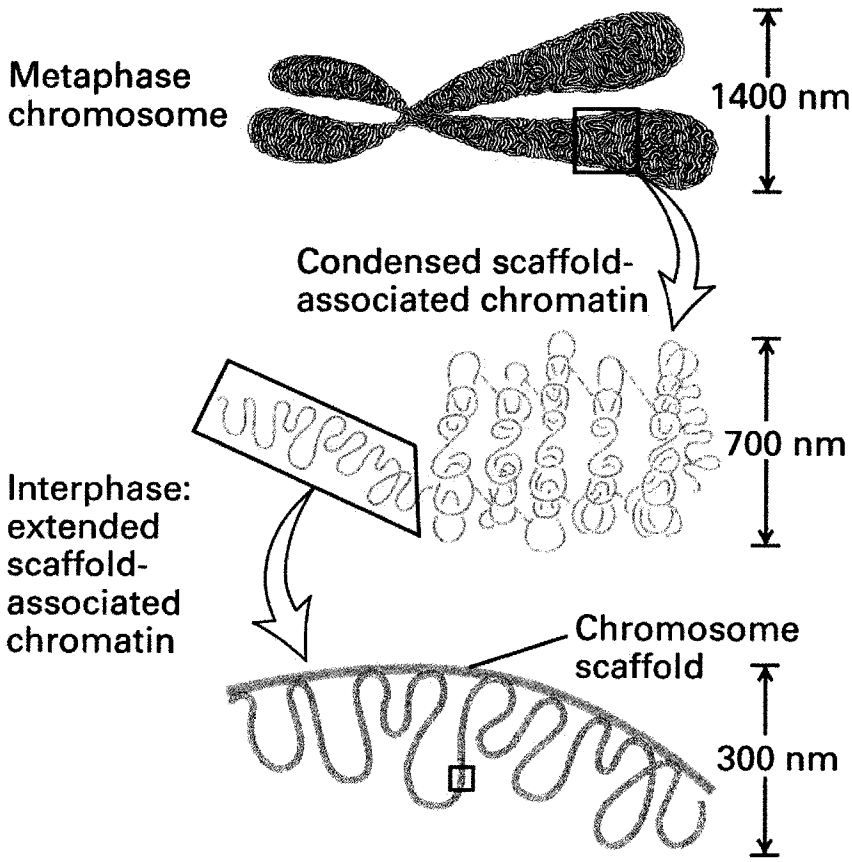
A eukaryotic chromosome made out of self-assembling 70 Å units, which could perhaps be made to crystallize, would necessitate rewriting our basic textbooks on cytology and genetics! I have never read such a naive paper purporting to be of such fundamental significance. Definitely it should not be published anywhere!

- Anonymous review of paper submitted by D. F. L. Woodcock, 1973

(Excerpt from Van Holde, 1988)

The genetic information of a single eukaryotic cell is stored in DNA molecules that are in total over two meters in length, but compacted in the cell's nucleus to nearly 100,000th of this dimension (Figure 1.1; Lodish et al., 2003). This is achieved by increasingly higher orders of compaction into DNA-protein assemblies called chromatin (Kornberg, 1977). Therefore, the substrate for essential biological processes like transcription, replication, recombination, DNA repair and cell division is chromatin, not naked DNA. The nucleosome is the fundamental unit of chromatin, which comprises the NCP and 20-80 bp of linker DNA in association with the linker histone, H1 (Kornberg, 1977). The NCP consists of a histone octamer core containing two molecules each of the four core histone proteins, H2A, H2B, H3, and H4, and a 147 bp stretch of DNA.

The high-resolution structure of the NCP (Figure 1.2; Luger et al., 1997), determined at 2.8 Å resolution, provided the first detailed view of the nature of the protein-DNA interactions that led to the first level of genomic DNA compaction (Luger et al., 1997). In the recent years, high-resolution structures of a variety of NCPs have added to our knowledge of nucleosome structure and function. The results from these structural studies have been discussed in detail elsewhere (Luger, 2003; Muthurajan et al., 2003). The salient features of the structure of the NCP are summarized as follows. The 146 bp DNA fragment wraps around the histone octamer in 1.65 turns of a left-handed superhelix. The DNA structure in the NCP deviates significantly from that of average linear B-form DNA as well as from that of an ideal superhelix (Luger et al., 1997; Richmond and Davey, 2003). The DNA superhelix is not uniformly bent and exhibits regions of high curvature adjacent to regions that are almost straight. The major and minor grooves are



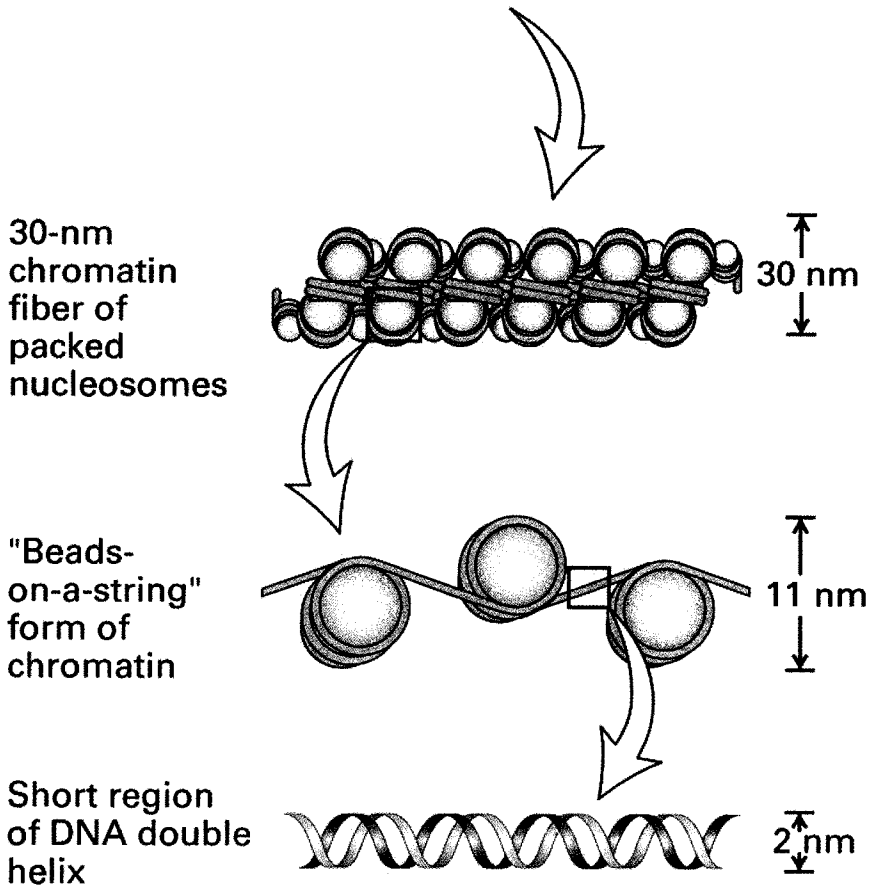


Figure 1.1 Schematic representation of chromatin condensation and chromosomal scaffolding in the nucleus. The histone octamer core is shown as an orange disk in the "beads-on-a-string" representation of the nucleosomes (adapted from Figure 10-24, Lodish et al., 2003).

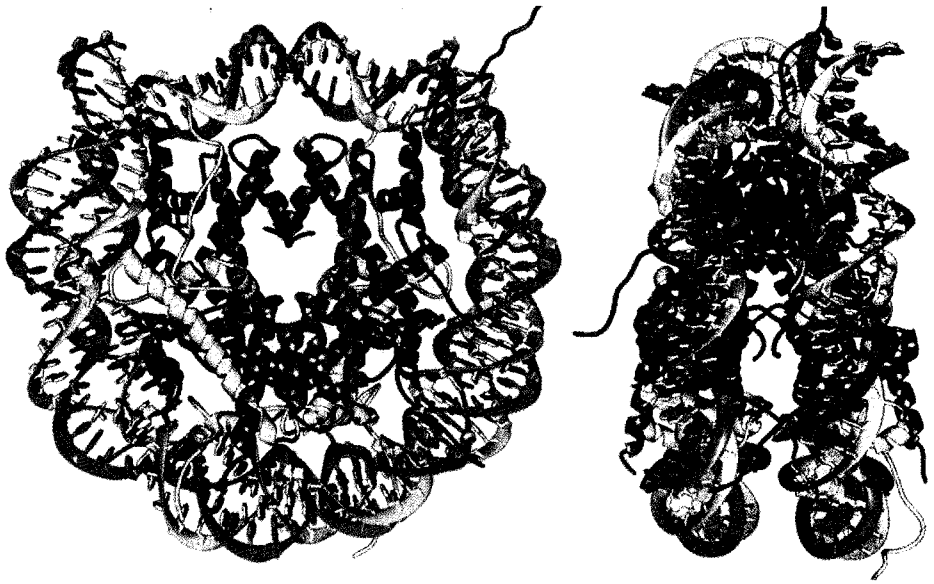


Figure 1.2 Nucleosome core particle: ribbon traces for the 146 bp DNA (brown and turquoise) and the histones (blue: H3; green: H4; yellow: H2A; red: H2B. The views are down the DNA superhelix axis (on the left) and a view perpendicular to it (on the right). For both particles, the pseudo-twofold axis is aligned vertically with the DNA centre at the top (adapted from Luger et al., 1997).

highly compressed as they face the octamer and highly extended when exposed to the solvent. Approximately 75 % of the total DNA surface is solvent exposed. The nucleotides that are facing the solvent are highly dynamic and have two-to-three fold higher crystallographic B-factors as compared to the nucleotides that are in contact with the histones. The four core histones all share a conserved structural motif called the histone fold, which accounts for the organization of 121 bp of DNA. An arginine residue, contributed by the core regions of the four histones, inserts into the DNA minor groove every time the minor groove faces the octamer. The structurally variable regions of the histones include the histone-fold extensions, the flexible histone N-terminal tails, and the C-terminal tail of histone H2A, some of which help organize the terminal 13 bp at both ends of the superhelical DNA. There are ~ 120 direct protein-DNA contacts and an almost equal number of water-mediated protein-DNA interactions that organize the structure of the nucleosome. The total numbers of DNA phosphate-protein main-chain interactions, which remain almost invariant for two NCPs containing different DNA sequences, indicate that the protein main-chain interactions have the ability to override the intrinsic sequence-dependent variability in the structure of DNA.

1.3 Role of histone N-terminal tails in chromatin function

The histone octamer core in the nucleosome contains approximately 75 % structured "histone fold" domains that organize the majority of the 147 bp of nucleosomal DNA. The remainder ~ 25 % of the protein mass is largely unstructured and is too disordered to be fully observed in the crystal structures of nucleosome core particles (NCP) (Davey et al., 2002; Harp et al., 2000; Luger et al., 1997; Muthurajan et al., 2003; Suto et

al., 2000; Suto et al., 2003; White et al., 2001). The N-terminal tails of histones H3 and H2B exit through the aligned minor grooves between the two gyres of the superhelical DNA, in contrast to that of histones H2A and H4, which exit to the exterior of the nucleosome (Luger et al., 1997). Thus, the tails are solvent exposed and are highly accessible to molecular recognition.

The histone tails are the targets for a multitude of post-translational modifications (Cheung et al., 2000; Strahl and Allis, 2000; Van Holde, 1988), which may modulate nucleosome structure and function, in addition to regulating the expression of the underlying genes. Modifications like acetylation and methylation of lysine residues have been linked to transcriptional regulation (Heard et al., 2001; Litt et al., 2001; Noma et al., 2001). Acetylation of histone tails has been shown to restrict the formation of the condensed 30 nm fiber *in vitro* (Garcia Ramirez et al., 1995; Tse et al., 1998). In other studies, the histone N-terminal tails have been shown to be the major targets of the cellular coactivator, p300, and to have a role in transcriptional regulation (Georges et al., 2002). Histone N-terminal tails also serve as platforms for the recruitment of cellular regulatory factors, and have been demonstrated to directly interact with factors like Tup1, Sir3, Sir4, HMG-14, and HP1 (Zheng and Hayes, 2003). Some studies also provide evidence for the interaction of the H4 tail with the linker histone H1 (Baneres et al., 1994; Thiriet and Hayes, 1997). The histone N-terminal tails have been previously implicated in the formation of higher-order chromatin (Dorigo et al., 2003; Hansen, 2002; Lu and Hansen, 2003). Studies using model oligonucleosomes suggest the presence of specific tail-tail interactions within the condensed chromatin fiber (Fletcher and Hansen, 1995). The core histone

tails are essential for the folding of oligonucleosome arrays into intermediate structures and the fully condensed 30 nm chromatin fiber (Allan et al., 1982; Annunziato and Hansen, 2000; Garcia-Ramirez et al., 1992; Schwarz et al., 1996). The crystal structures of NCPs also reveal internucleosomal interactions mediated by the histone H4 N-terminal tail (Davey et al., 2002; Luger et al., 1997; Muthurajan et al., 2003; Suto et al., 2000; Suto et al., 2003). A stretch of amino acids from 14 to 19 in the histone H4 N-terminal tail, which makes contacts with the H2A/H2B dimer of a neighboring particle in the crystals, is also essential for *in vitro* chromatin compaction (Dorigo et al., 2003). More recently, it was demonstrated that the stabilization of a pair of cysteine-replacement mutants (derived from one of the H4 and H2A interaction pairs seen in the crystals) within nucleosomal arrays leads to the formation of a 'two-start' chromatin fiber (Dorigo et al., 2004).

In addition, histone tails are also involved in regulating access to nucleosomal DNA. For example, removal of the H3/H4 tails enhances binding of the transcription factor TFIIIA to nucleosomal DNA (Vitolo et al., 2000). Other studies reveal that Gal4 binding to nucleosomes require acetylation of the histone H4 tail (Vettese-Dadey et al., 1996). Previous studies have also elucidated that the deletion of either H3 or H4 N-terminal tails allows binding of sequence-specific pyrrole-imidazole polyamides to nucleosomal DNA (Gottesfeld et al., 2001). To summarise, although the above-outlined biochemical and structural studies provide ample clues to implicate the histone N-terminal tails in higher-order chromatin organization and in the regulation of gene expression, the mechanistic details of how these tails bring about their biological effects are largely unknown.

1.4 Polyamides as molecular probes of nucleosomal DNA accessibility

The pyrrole-imidazole polyamides (PA) are a class of synthetic small molecules that can be designed to bind predetermined DNA sequences (Dervan and Edelson, 2003; Trauger et al., 1996; Wemmer, 2000; White et al., 1997). The parent molecules for polyamides are natural antibiotics like distamycin and netropsin that also bind DNA in the minor groove. Polyamides contain N-methylpyrrole (Py) and N-methylimidazole (Im) amino acids. Base-pairing specificity depends on side-by-side pairing of the aromatic Py and Im amino acids bound in the minor groove of DNA. Pairing of an Im opposite a Py targets a G•C base pair, whereas a Py opposite an Im targets a C•G base pair (Figure 1.3; Dervan, 2001). The Py/Py pair is degenerate and targets both A•T and T•A base pairs. The aliphatic β-alanine•β-alanine pair also recognizes both A•T and T•A base pairs. These pairing rules are supported by NMR studies (Geierstanger et al., 1994) and by x-ray crystallography (Kielkopf et al., 1998a; Kielkopf et al., 2000; Kielkopf et al., 1998b).

The polyamides bind their target sequences with affinities comparable to those of natural DNA-binding transcriptional regulatory proteins (Trauger et al., 1996), and have been shown to competitively inhibit the TFIIIA-DNA complex as well as several other protein-DNA complexes (Gottesfeld et al., 1997). Polyamides designed to interfere with the binding of well-characterized eukaryotic transcription activators or repressors to their target DNA sites were found to inhibit or derepress transcription, respectively (see Dervan and Edelson, 2003 for various examples). Artificial transcription factors, in which the DNA-binding domain of the activator has been substituted with a specific polyamide

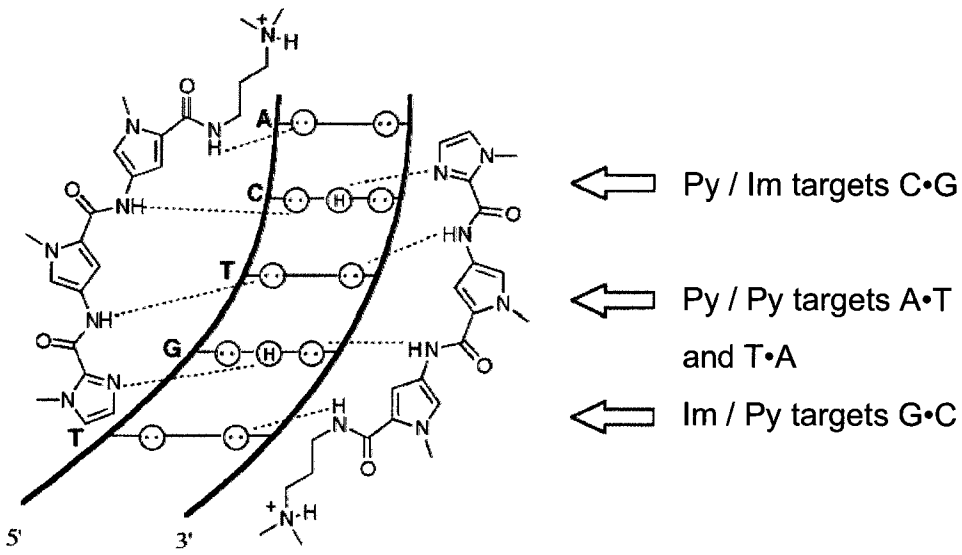


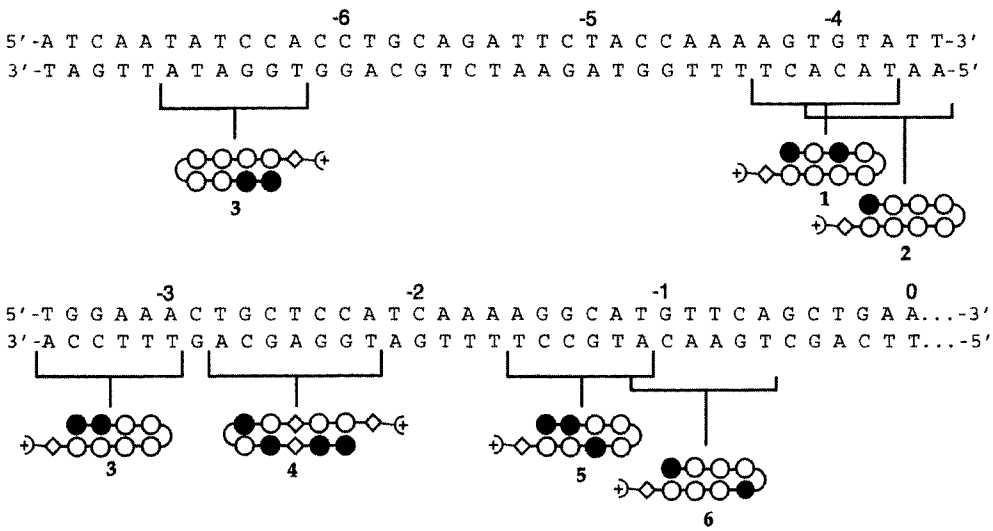
Figure 1.3 Rules for pyrrole-imidazole polyamide recognition of DNA (adapted from Dervan, 2001).

to tether the molecule to the DNA, have been shown to up-regulate specific gene expression *in vitro* (Mapp et al., 2000).

In addition to targeting naked DNA, polyamides can also be designed to bind their target DNA sites in the context of the NCP (Figure 1.4; Gottesfeld et al., 2001). Six different polyamides, polyamides 1 to 6 were synthesized to target specific sites on the 146 bp α-satellite DNA (Gottesfeld et al., 2001). The binding specificity and affinity of the polyamides for the NCP were determined using *in vitro* DNase I and hydroxyl radical footprinting experiments. These studies show that sites on nucleosomal DNA facing away from the histone octamer, or even partially facing the octamer are fully accessible and the nucleosomes remain fully folded upon binding the polyamides. Polyamides only fail to bind where sites are near the N-terminal tails of histones H3 and H4. Removal of the N-terminal tails of either H3 or H4 allows these polyamides (specifically, polyamides 5 and 6) to bind to the NCP (Gottesfeld et al., 2001). These studies suggest that the much of the DNA in the NCP is available for molecular recognition in the minor groove.

The high-resolution x-ray structures of three of these polyamides, polyamides 1, 2, and 3 were subsequently determined in complex with the NCP (Suto et al., 2003). While the histone-DNA interactions with the NCP remain unaffected by polyamide binding in every case, these structures reveal significant changes in DNA parameters at the polyamide-binding site and in adjacent regions. The structure of the NCP-polyamide 1 complex shows that the bound ligand prevents the propagation of DNA twist defects from one end of the nucleosomal DNA to the center. In the case of the NCP-polyamide 3 complex, the DNA distortions resulting from polyamide binding are transmitted from one NCP mole-

A



B

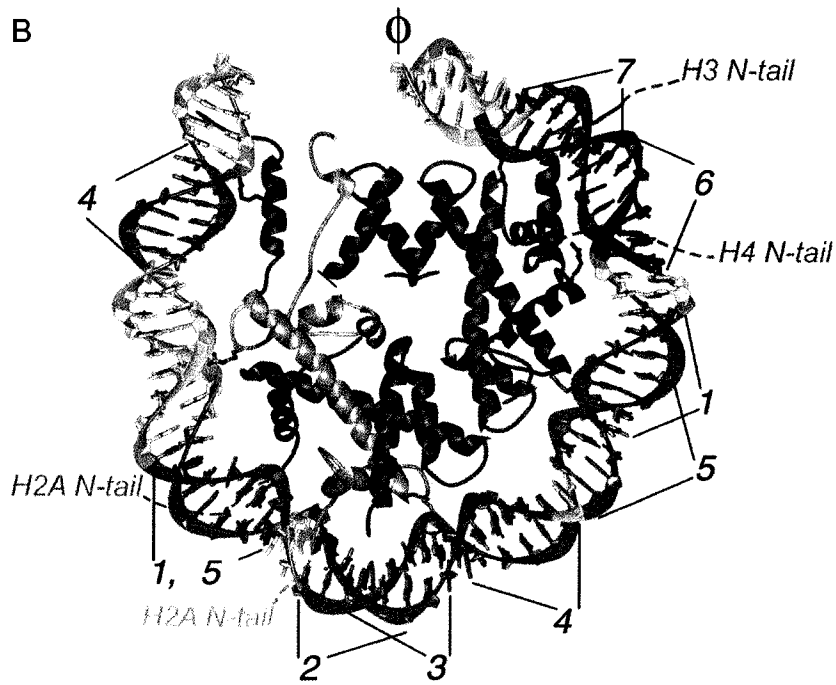


Figure 1.4 Structures and binding sites for polyamides designed to target the 146 bp DNA in the NCP. (A) α-satellite half-site (73 bp) DNA sequence and polyamide-binding models shown along with the location of the target sequences. The filled and open circles represent Imidazole and Pyrrole rings, respectively; the curved line, γ-aminobutyric acid; diamonds, β-alanine; and the half-circle with a plus sign, dimethylaminopropylamide. (B) Ribbon diagram of one half of the nucleosome core particle (73 bp of DNA and associated proteins). H3 is shown in blue, H4 in green, H2A in yellow, H2B in red. Polyamide-binding sites on white DNA are colored in cyan (polyamides 1, 2, and 3), violet (polyamide 4), and magenta (polyamides 5 and 6). The dyad is indicated (ϕ). The location of the N-terminal histone tails is shown. The model (continuous lines) includes H3 residues 38 to 135, H4 residues 24 to 102, H2A residues 14 to 119, and H2B residues 30 to 122 (adapted from Gottesfeld et al., 2001).

cule to the neighboring NCP engaged in crystal contacts. The transmission of DNA distortions from one NCP to another without simultaneous disruption of all histone-DNA contacts argues for the role of twist diffusion in nucleosome translocation.

1.4 Specific aims of the dissertation

Following the sequestration of DNA within the architectural framework of the NCP, the dynamics and accessibility of the nucleosomal DNA are distinct from naked DNA. Consequently, the ability of DNA-binding proteins/transcription factors to recognize their cognate sites in chromatin is restricted by the structure and dynamics of nucleosomal DNA, and by the translational and rotational positioning of the DNA with respect to the histone octamer. This implicates the nucleosomes as key regulators of all cellular processes in which chromatin is the physiological substrate. However, the precise role of nucleosome dynamics in modulating nucleosomal DNA accessibility and in fine-tuning cellular processes is not clearly known.

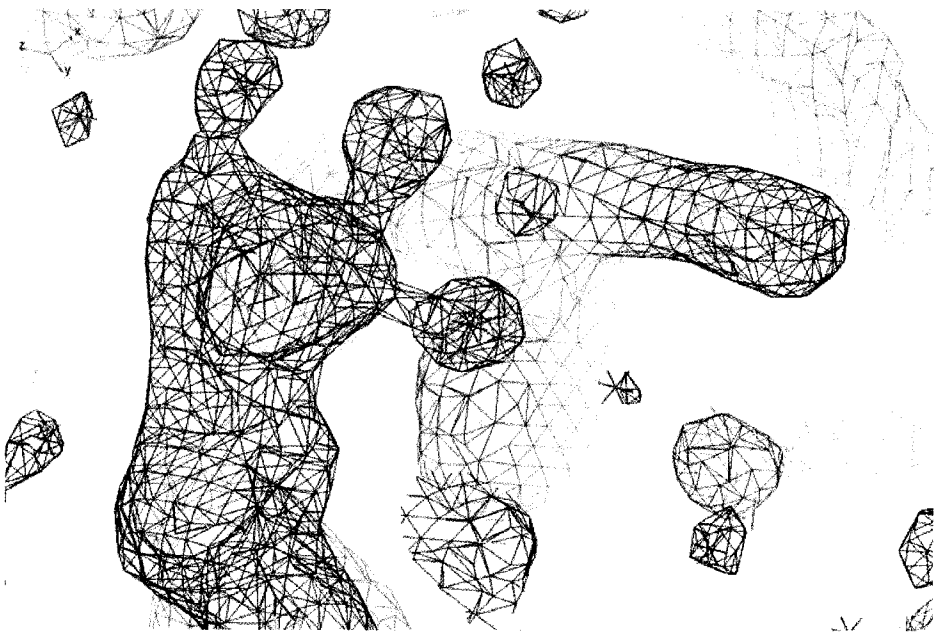
The crystal structure of the nucleosome core particle reveals the multitude of DNA-histone contacts that occur through direct interactions of the histones with the minor grooves of the DNA. In addition, *in vitro* measurements reveal that the octamer-DNA association in the nucleosome is very stable ($K_d \sim 0.5$ nM; Gottesfeld and Luger, 2001). And yet, the nucleosome is surprisingly 'malleable' and accommodates the constant demand to restructure and support all chromatin-dependent cellular processes *in vivo*. The nucleosome is clearly a stable multicomponent assembly stitched together from a myriad of relatively weak non-covalent histone-DNA interactions. Any of the above-mentioned

structural building blocks is a potential 'Achilles' heel' that may perturb the global stability, accessibility and dynamics within the NCP.

What are some of the 'seemingly' modular features of the NCP? These features include the presence of aligned major and minor grooves all along the two gyres of the DNA superhelix, the 'twist diffusion' dynamics of DNA as suggested by crystal packing, and the highly dynamic histone N-terminal tails. I became interested in understanding whether and in what manner the above features affect nucleosomal DNA accessibility and dynamics. The primary goal of my research work is therefore to dissect how the above structural attributes of the NCP impact its function. Understanding how the various components regulate nucleosomal DNA accessibility and dynamics may throw some light on how the transcription machinery functions in the context of highly compacted DNA.

My thesis research is the result of a fruitful collaboration with the groups of Prof. Peter Dervan (California Institute of Technology) and Prof. Joel Gottesfeld (The Scripps Research Institute). The polyamides used in all my studies were provided by the Dervan group. All DNase I and hydroxyl radical footprinting experiments discussed in my thesis were performed by Prof. Joel Gottesfeld. Small-molecule pyrrole-imidazole polyamides (polyamides 1 to 6) were used as sequence-specific DNA-binding probes to study nucleosome dynamics in solution and high-resolution structures of some of the NCP-polyamide complexes were determined to study structural changes imparted to the NCP following ligand binding. The effects of polyamide binding on temperature-induced nucleosome repositioning in the NCPs in solution were also studied. Attempts were made to determine the co-crystal structure of the NCP-polyamide 4 complex. Next, the effects of

the 'bivalent' polyamide 'clamps' (PW12, PW13, and PW14) on the diffraction quality of nucleosome crystals and on nucleosome stability in solution were both examined. The binding of the polyamide clamp (PW12) to the NCP allowed us to study subtle changes in the DNA structure and to reconcile the 'twist defect' phenomenon observed in the NCP crystals with the DNA dynamics within nucleosomes in solution. I also investigated the effects of the removal of either the histone H3 or H4 N-terminal tail on nucleosome structure as well as the effects of polyamide binding to nucleosomes containing either a tailless H3 or tailless H4 histone. In this last study, the effects of ligand binding on the crystallization of the NCP and on crystal packing were examined in detail. A low-resolution crystal structure of the tailless H4-containing NCP in complex with polyamide 6 was also determined. In another cryocrystallography study, I studied the cryocooling protocol for nucleosome crystals in detail to determine the optimal flash-cooling temperature and the conditions for the best diffraction quality.



Chapter 2

Polyamide Binding to the NCP and its Effects on Crystallization and Nucleosome Repositioning

Suto, R. K., Edayathumangalam, S., White, C. L., Melander, C., Gottesfeld, J. M., Dervan, P. B., and Luger K. (2004). Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. J Mol Biol, *326*, 371-380.

This chapter contains results from my experiments that were included in the publication cited above. Other parts of this chapter that include results from nucleosome-polyamide 4 were not included in the above publication.

2.1 Abstract

Several site-specific pyrrole-imidazole polyamides (polyamides 1 to 6) targeted to bind defined sites on the 146 bp nucleosomal DNA were previously characterized and found to bind their target sites with different affinities (Gottesfeld et al., 2001). In order to further characterize the NCP-polyamide complexes, our laboratory has undertaken two approaches: i) determination of the high-resolution x-ray structures of the various NCP-polyamide complexes, and ii) study of the effect of ligand binding on nucleosome repositioning and dynamics in solution. The crystal structures of nucleosomes in complex with polyamides 1, 2, and 3 were recently reported (Suto et al., 2003). In this study, I attempted to determine the structure of the NCP in complex with polyamide 4 and the results are reported here. The effects of polyamide binding on *in vitro* temperature-induced nucleosome repositioning were also studied. We find that polyamides inhibit the temperature-induced translocation of the histone octamer on the DNA thereby suggesting a mechanism for temperature-induced nucleosome repositioning.

2.2 Introduction

Sequence-specific pyrrole-imidazole polyamides have been previously used as molecular probes for DNA accessibility in nucleosomes (Dervan and Edelson, 2003; Gottesfeld et al., 2001; Suto et al., 2003). Unlike many protein transcription factors, polyamides interact exclusively with the minor groove of the DNA in a sequence-specific manner (Dervan and Edelson, 2003); however, they share with protein transcription factors the approximate size of their binding sites (6 bp) and their high (nanomolar) binding affinity. The binding of six different polyamides (polyamides 1 to 6) to the NCP has been previ-

ously characterized. It has been shown that sites on nucleosomal DNA facing away from the histone octamer, or even partially facing the histone octamer, are fully accessible for high-affinity binding (Gottesfeld et al., 2001). Polyamides were only unable to bind to the NCP when the polyamide-binding sites were close to regions of DNA-histone interactions or the histone H3 or H4 N-terminal tails. The deletion of either H3 or H4 N-terminal tail allowed these polyamides (polyamides 5, 6) to bind to the NCP.

The crystal structures of NCPs in complex with three of the six polyamides, polyamides 1, 2, and 3 were determined recently by our group (Suto et al., 2003). The three structures reveal the following structural changes to the NCP as a consequence of polyamide binding: (i) significant distortions are introduced into nucleosomal DNA at the binding site and in adjacent regions, and in one case, these distortions are transmitted between two crystallographically related NCPs; (ii) these distortions do not compromise any of the protein–DNA interactions; and (iii) some bound ligands effectively block the propagation of twist defects from the end of the nucleosomal DNA towards its center.

Structure determination attempts for the NCP-polyamide 4 complex are described in this chapter. Efforts towards the structure determination of NCP-polyamide 5 and NCP-polyamide 6 complexes are detailed in chapter 5. Additionally, the effects of polyamide binding on nucleosome dynamics (nucleosome repositioning) in solution were investigated for polyamides 1 to 6. Nucleosome repositioning involves changes in the translational and rotational positioning of the DNA with respect to the histone octamer. Nucleosome repositioning can be demonstrated *in vitro* as a temperature-dependent process, while ATP-dependent chromatin remodeling factors bring about nucleosome

repositioning/remodeling *in vivo*. It has been shown previously that temperature-dependent nucleosome repositioning can be visualized even on NCPs reconstituted with the 146 bp DNA fragment (Luger et al., 1999; Muthurajan et al., 2003). My experiments show that polyamides inhibit the heat-induced translocation of the histone octamer on DNA. In addition, these findings suggest a mechanism for temperature-induced translocation of the DNA on the histone octamer (Flaus et al., 1996; Luger et al., 1999; Meersseman et al., 1991), and for the effect of bound factors on the dynamic properties of nucleosomes (Pennings et al., 1994).

2.3 Materials and Methods

2.3a Nucleosome reconstitution and formation of NCP-polyamide complex

Previously established protocols were used to reconstitute NCPs from recombinant *Xenopus laevis* histones and a 146 bp DNA fragment derived from human α-satellite DNA and for the crystallization of the NCP-polyamide 4 complex (Dyer et al., 2004). Polyamides were synthesized as described (Baird and Dervan, 1996). The purity and identity of the polyamides were established by analytical HPLC and matrix-assisted laser desorption ionization-time-of-flight mass spectroscopy. Polyamides were purified by preparative HPLC, lyophilized, and stored at 4 °C. For crystallization, polyamide 4 was freshly dissolved in 20 mM potassium cacodylate (pH 6.0) and 1 mM EDTA, and its concentration was determined by measuring UV absorbance and using empirically determined extinction coefficients (ε_{316} = 68,800). NCP (at 40-50 μM) was incubated with a six-to-ten-fold molar excess of polyamide 4 in solution at ambient temperature for 45 min. The integrity of the nucleosome preparation after incubation with the ligand was

checked by electrophoretic mobility-shift assay (Luger et al., 1999). Crystals of NCP-polyamide 4 complex were grown 1-2 weeks in 40-45 mM $MnCl_2$, 35-38 mM KCl, and 20 mM potassium cacodylate (pH 6.0) containing ~ 20 μM (~ 4mg/mL) NCP at 19 °C by using vapor diffusion. Crystals were harvested and flash-cooled as described (Luger et al., 1997).

2.3b Data collection and structure refinement

X-ray data were collected at our in-house Raxis IV detector and at the synchrotron at the Advanced Light Source (ALS), beamline 5.0.2, Berkeley, CA. Data for each of the individual structures were collected on single crystals and processed with Denzo and Scalepack (Otwinowski and Minor, 1997). Molecular replacement (using PDB entry 1AOI as the original search model) and refinement were done with CNS (Brunger et al., 1998).

2.3c Temperature-induced nucleosome repositioning assay

For heat shifting experiments, NCPs and polyamides were prepared in 20 mM Tris–HCl (pH 7.5), 1 mM EDTA, 1 mM DTT. The concentration of each polyamide was determined by measuring UV absorbance and using empirically determined extinction coefficients (ε_{316} = 68,800). NCP (4.56 pmol in 10 μL) was pre-incubated with either a two-fold or a four-fold molar excess of the polyamide, followed by an incubation step at 37 °C for two hours. The NCP-polyamide complexes were then analyzed on a non-denaturing 5% polyacrylamide gel run in 0.2X TBE running buffer at 4 °C. The gel was

first stained with ethidium bromide DNA stain followed by staining with Coomassie brilliant blue.

2.4 Results and Discussion

2.4a Polyamide 4 dissociates from DNA in the crystals of NCP-polyamide 4 complex

NCP-polyamide 4 was crystallized as described in the 'Materials and methods' section. The crystallization conditions did not vary from those for unliganded NCP (Luger et al., 1997) and the crystal morphology is similar to that of unliganded NCP crystals. Several NCP-polyamide 4 crystals were obtained using the 'sitting drop' method. The crystals were harvested by successive soaking in increasing concentrations of the cryoprotectant, 2-methyl-2,4-pentane-diol (MPD), followed by flash-cooling in liquid propane at -120 °C and transfer into a cold nitrogen gas stream (-180°C) for data collection. X-ray diffraction data were obtained for crystals of NCP-polyamide 4 complex both in our in-house X-ray generator and at the synchrotron at ALS. The statistics for data collection are shown in Table 2.1.

Calculation of difference density maps (F_o-F_c as well as $2F_o$-F_c maps) following molecular replacement and initial refinement revealed no density for polyamide 4 in regions where the polyamide would be expected to bind. The lack of polyamide density could be due to low occupancy of the polyamide in its binding site, which could be a consequence of a 10-fold reduced affinity of the polyamide for nucleosomal DNA as compared to free DNA (Gottesfeld et al., 2001). Alternatively, the polyamide may have been displaced

Table 2.1 Data-collection and refinement statistics

Dataset	NCP-polyamide 4 ALS BL5.0.2	NCP-polyamide 4 RAxis IV
Mosaicity	0.6	0.5
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell lengths (*a, b, c* in Å)	105.7 109.7 181.8	106.0 109.7 182.0
Resolution range (Å)†	99-2.80 (2.87-2.80)	45-3.25 (3.33-3.25)
Unique reflections	49205	33939
Completeness (%)	93.0 (78.0)	99.6 (99.6)
Multiplicity	2.9	3.5
Average I/σ(I)	18.3 (2.6)	8.1 (2.7)
R-merge#	0.059 (0.403)	0.091 (0.264)

Numerical values in parentheses denote data for the highest resolution shell.
#R-merge = $\Sigma|I_h - \langle I_h\rangle|/\Sigma I_h$, where $\langle I_h\rangle$ is the mean of the measurements for a single *hkl*.

from its binding site because of crystallization-induced DNA 'stretching' that occurs at the polyamide-binding site. The phenomenon of DNA stretching forms the basis for much of the work described later in chapter 4 and will be only briefly discussed here. The NCP exhibits pseudo twofold symmetry (Luger et al., 1997; Muthurajan et al., 2003). Because DNA binds the histone octamer with the plane of a central base pair coincident with the particle pseudo-twofold axis (dyad axis), and because a 146 bp fragment was chosen for crystallographic studies, the dyad axis divides the 146 bp nucleosomal DNA into a 72 bp- ('short') and a 73 bp- ('long') half, with one bp falling on a dyad. Structural alignment of the two DNA halves shows that a DNA region approximately 1.5 to 2 helical turns away from the molecular dyad axis exhibits significant structural divergence (reviewed in Muthurajan et al., 2003). There, the short half 'stretches' and covers the distance measured by the 12 bp in the long half in only 11 bp. This is caused by the requirement for DNA base stacking between two neighboring nucleosomes in a quasi-continuous helix to drive crystallization. Stretching of DNA leads to a decrease of average minor groove width by more than 1.1 Å, and is accompanied by overwinding of the DNA to an average of 9.66 bp per turn. The binding site for polyamide 4 is coincident with one of the two regions where stretching occurs. This crystallization-induced 'stretching' of DNA may be responsible for the displacement of polyamide 4 from its binding site and may thus explain why there is no density for the polyamide. Although it may be possible for a polyamide that has been displaced to bind again to its target site in solution, it is known that polyamides are insoluble under the crystallization conditions (not shown). Further, precipitation was consistently detected in hanging and sitting drops containing NCP-polyamide 4 complex after a few days of incubation, which may have oc-

curred following dissociation of the polyamide from the nucleosome. This structure was not refined further as we failed to detect any density for the polyamide.

2.4b Polyamides prevent heat-induced octamer repositioning in solution

To examine whether polyamides have an effect on the dynamic behavior of the nucleosome in solution, we tested the effect of binding polyamides 1 to 6 on temperature-induced nucleosome repositioning. After salt-gradient dialysis, the octamer adopts an off-centered position on the 146 bp α-satellite DNA (Luger et al., 1999). After heating reconstituted NCPs for two hours at 37 °C, the octamer transfers to a centered (and thermodynamically more stable) position. These two species are readily distinguished by native gel electrophoresis (Figure 2.1, compare lanes 3 and 4; Luger et al., 1999). Affinity cleavage experiments with Fe[II]-EDTA-coupled polyamides demonstrate that all six polyamides bind to the NCP in the off-centered position (performed by Joel M. Gottesfeld, not shown), consistent with the notion that the off-centered and centered positions on this particular DNA fragment are in the same rotational frame with respect to the octamer and are related by a translocation of a full turn of DNA (~10 bp, Flaus et al., 1996). Newly reconstituted NCPs (before heat treatment, lane 3) were incubated at 37 °C for two hours in the absence (lane 4) or in the presence of either a two- or fourfold molar excess (over binding sites) of polyamides 1 (lanes 5 or 6), 2 (lanes 7 or 8), 3 (lanes 9 or 10), 4 (lanes 1 or 2), 5 (lanes 11 or 12), and 6 (lanes 13 or 14). We found that polyamides 2 and 6 completely inhibit heat-induced repositioning at either concentration (lanes 7 and 8, and 13 and 14, respectively), whereas the other polyamides have very little or no inhibitory effect on repositioning. Ethidium bromide staining of the gel showed that

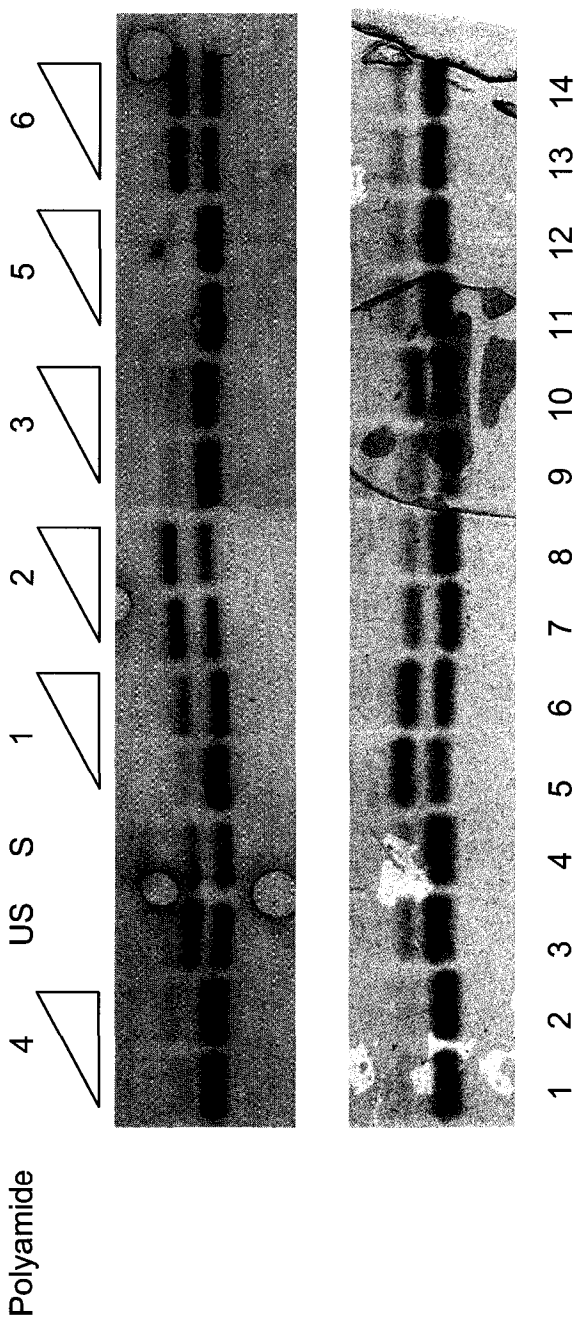


Figure 2.1 The effect of polyamides on heat shifting of nucleosomes. NCPs (4.5 pmol) were run on a native 5% polyacrylamide gel that was first stained with ethidium bromide (upper panel), then with Coomassie brilliant blue (lower panel). Lanes 3 and 4, NCP in the absence of polyamide, before (lane 3) and after (lane 4) a two hour incubation at 37 °C. Lanes 1, 5, 7, 9, 11, and 13, heat treatment in the presence of a two-fold molar excess of polyamide 4, 1, 2, 5, and 6, respectively. Lanes 2, 6, 8, 10, 12, and 14, heat treatment in the presence of a four-fold molar excess of polyamide 4, 1, 2, 5, and 6, respectively.

polyamide binding and heat shifting does not cause dissociation of DNA (Figure 2.1, upper panel). These data indicate that polyamide 2 and 6 may lock the histone octamer in a unique translational position with respect to the underlying DNA sequence, possibly by preventing rotation of the DNA on the surface of the histone octamer.

2.5 Discussion

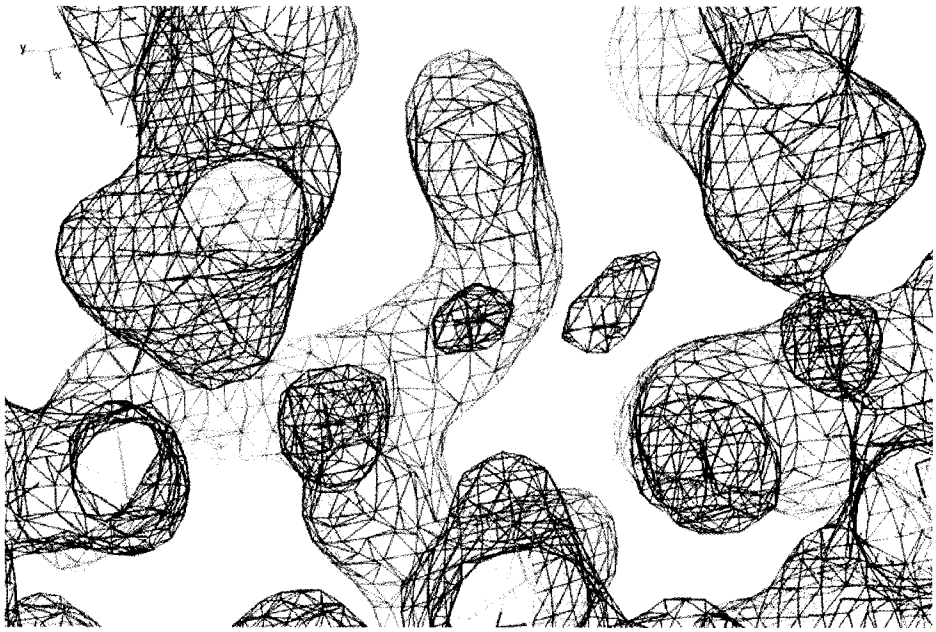
The crystallization studies on the NCP-polyamide 4 complex suggest that the dynamics of DNA within the crystal can influence ligand binding. It is also tempting to speculate that the reported 10-fold reduced affinity of polyamide 4 for the NCP as compared to naked DNA (Gottesfeld et al., 2001) is a result of DNA stretching at the polyamide-binding site. These results also provide indirect evidence for the occurrence of DNA stretching in nucleosomes in solution. The crystallization of polyamide 4 complexed with a nucleosome containing the 147 bp DNA fragment (stretching is absent in NCP containing with 147 bp DNA; (Davey et al., 2002)) may reveal if stretching indeed prevents polyamide 4 binding to the nucleosome. This further suggests that DNA dynamics within the NCP has a significant effect on DNA-ligand interactions thus attributing a regulatory role to DNA dynamics *in vivo*. In fact, additional studies on nucleosome dynamics provide more evidence to corroborate the above interpretation that DNA stretching occurs in nucleosomes in solution (see chapter 4).

The data from the nucleosome repositioning assays suggest that a bound polyamide may lock nucleosomal DNA in one rotational (and translational) orientation with respect to the histone octamer. Indeed, we have found that some, but not all high-affinity-binding polyamides are able to block the temperature-induced histone octamer repositioning by

~10 bp that is observed with the 146 bp DNA sequence. The finding that only two of the six polyamides inhibit this repositioning at lower polyamide-to-NCP ratio can be rationalized by different binding affinities of the polyamides for their binding sites in unshifted NCPs. The phenomenon appears to be general, since it is not limited to the 146 bp α-satellite DNA sequence, but has been found with other DNA sequences (and other polyamides). It has also been shown that polyamides that inhibit sliding also inhibit transcription through a single nucleosome by bacteriophage T7 RNA polymerase (Gottesfeld et al., 2002). Our results show that polyamides lock the histone octamer in a translational position with respect to the underlying DNA sequence and possibly even prevent rotation of the DNA on the surface of the histone octamer. These observations provide a mechanistic rationale for nucleosome repositioning and suggest the involvement of DNA rotational and/or translational rearrangements during nucleosome repositioning.

2.6 Acknowledgements

We thank Stephanie Ebbesen for help with reagents, all members of the Luger lab for technical support and input, and Andrew Flaus for helpful comments. The polyamides used in these studies were synthesized by Christian Melander in the Dervan laboratory. Funding for this work was provided by the Searle Scholars Award to K.L., and by grants from the NIH to K.L., J.M.G., and P.B.D.



Chapter 3

Molecular Recognition of the Nucleosomal "Supergroove"

Edayathumangalam, R. S., Weyermann, P., Gottesfeld, J. M., Dervan, P. B., and Luger, K. (2004) From The Cover: Molecular recognition of the nucleosomal "supergroove". Proc Natl Acad Sci USA, *101*, 6864-6869.

This chapter was published as cited above with our cover image selected for the journal issue. The text (except section 3.7), tables (except Table 3) and figures for this chapter are presented as they appeared in the publication.

3.1 Abstract

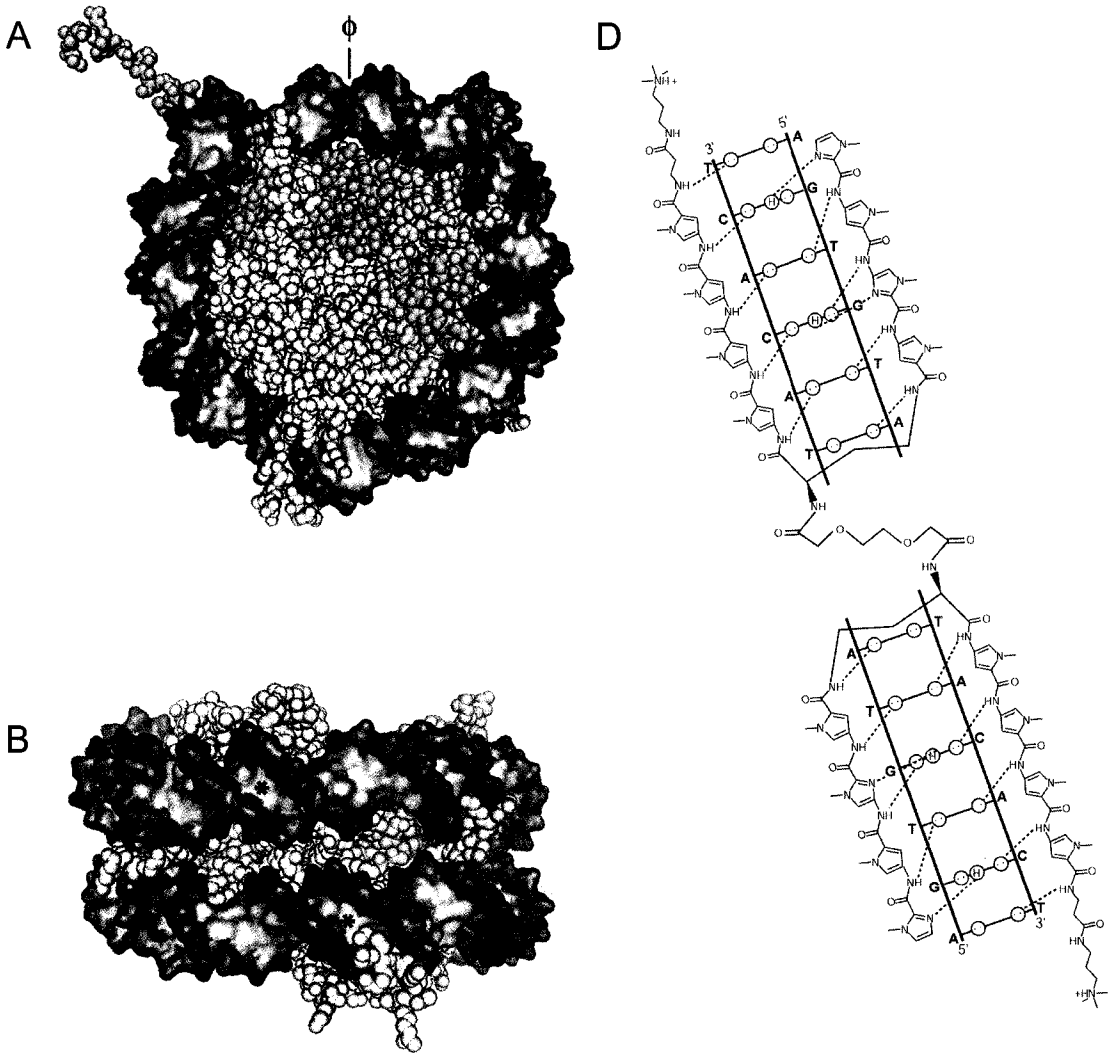
Chromatin is the physiological substrate in all processes involving eukaryotic DNA. By organizing 147 base pairs of DNA into two tight superhelical coils, the nucleosome generates an architecture where DNA regions that are 80 base pairs apart on linear DNA are brought into close proximity, resulting in the formation of DNA "supergrooves." Here, we report the design of a hairpin polyamide dimer that targets one such supergroove. The 2 Å crystal structure of the nucleosome-polyamide complex shows that the bivalent "clamp" effectively crosslinks the two gyres of the DNA superhelix, improves positioning of the DNA on the histone octamer, and stabilizes the nucleosome against dissociation. Our findings identify nucleosomal supergrooves as platforms for molecular recognition of condensed eukaryotic DNA. *In vivo*, supergrooves may foster synergistic protein-protein interactions by bringing two regulatory elements into juxtaposition. Because supergroove formation is independent of the translational position of the DNA on the histone octamer, accurate nucleosome positioning over regulatory elements is not required for supergroove participation in eukaryotic gene regulation.

3.2 Introduction

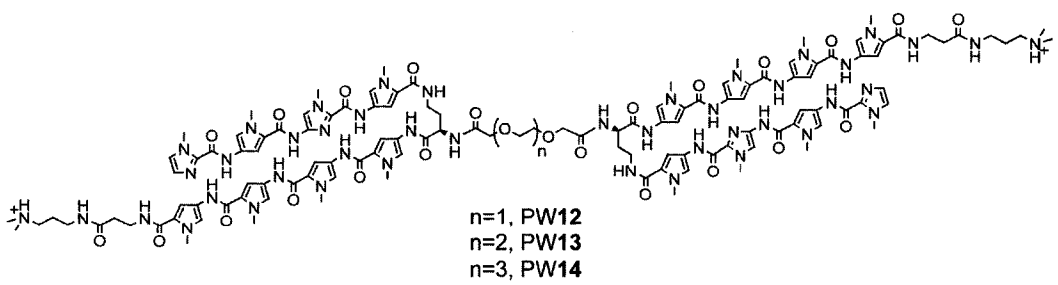
Recent high-resolution structures of nucleosome core particles (NCPs) (Davey et al., 2002; Luger et al., 1997) provide a wealth of details on the intricacies of protein-DNA interactions and provide a view of eukaryotic DNA in a physiologically relevant context (Davey et al., 2002; Luger et al., 1997; Richmond and Davey, 2003; reviewed in Luger, 2003). Two features are of particular relevance. First, due to extensive interactions of the DNA with the core framework of the histone octamer, one face of the DNA is completely

precluded from interactions with sequence-specific DNA binding factors over its entire 147 bp length. Although ~ 70 % of nucleosomal DNA is solvent accessible, transcription factors can access only six to eight consecutive base pairs before being sterically blocked by the histones and by extreme fluctuations in groove width. Thus, a single nucleosome limits access of transcription factors and other cellular factors to the underlying nucleosomal DNA. Second, the modulation of helical parameters within the left-handed DNA superhelix brings the minor (and major) grooves between the two superhelical gyres into alignment throughout almost the entire nucleosome (Figure 3.1A, B). This alignment creates seven minor and six major DNA ''supergrooves'' in which 14-16 bp of DNA would be accessible for site-specific interaction with a single DNA binding ligand. Each supergroove brings sequence elements that are 80 bp apart in linear DNA into close structural proximity (Srinivasan and Olson, 1992). Supergrooves thus provide unique recognition platforms that occur only within the structural context of a nucleosome.

Pyrrole-imidazole polyamides (PAs) bind the DNA minor groove with high affinity and specificity and have been used to modulate gene expression (Dervan and Edelson, 2003). We previously characterized the binding of a small library of eight-ring hairpin PAs to several discrete 6-bp sites on nucleosomes containing 146 bp DNA (NCP146) (Gottesfeld et al., 2001), and determined the crystal structures of NCP146 in complex with three of these hairpin PAs (Suto et al., 2003). Here, we report the design of hairpin PA dimers that target two noncontiguous minor groove sites in the nucleosomal supergroove in a sequence-specific manner, thereby ''clamping'' the two gyres of nucleosomal DNA. Our aim was to generate a sequence-specific bivalent PA clamp that



C



n=1, PW**12**
n=2, PW**13**
n=3, PW**14**

Figure 3.1 Site-specific recognition of nucleosomal DNA by clamp polyamides. (A) NCP146 structure (PDB code 1AOI, Luger et al., 1997) viewed with the superhelical axis along the z-axis. The particle pseudo-twofold dyad axis (ϕ) is shown for orientation. DNA (blue and white) and associated histone proteins (H2A, yellow; H2B, red; H3, blue; H4, green) are shown in sphere or surface representation. (B) 'Supergrooves' in NCP146. A different view of NCP146 with the superhelical axis along the y-axis. Colour scheme is the same as in Figure 1A. One of the DNA supergrooves is indicated by two stars. (C) Chemical structures of clamp polyamides, PW12-14. (D) Hydrogen bonding model of PW12 to its target DNA site. Circles with dots represent lone pairs of N3 of purines and O2 of pyrimidines. Circles containing H represent the N2 hydrogen of guanine. Putative hydrogen bonds are illustrated by dotted lines.

would expand the total size of DNA recognizable in a nucleosomal context, and that could stabilize the histone-DNA complex against dissociation. By using x-ray crystallography, we demonstrate that the clamp binds to NCP146 across the supergroove as designed, and with very high affinity and specificity. In addition, biochemical assays reveal that the clamp has a dramatic effect on the *in vitro* stability of nucleosomes against dilution-induced dissociation. We propose that nucleosomal supergrooves are molecular interaction platforms that can be exploited for synergistic protein-protein interactions, and that supergrooves add an additional level of complexity to eukaryotic gene regulation.

3.3 Materials and Methods

3.3a Co-crystallization of NCP-polyamide complexes

Previously established protocols were used to reconstitute NCP146 from recombinant *Xenopus laevis* histones and a 146 bp DNA fragment derived from human α-satellite DNA (Davey et al., 2002; Dyer et al., 2004; Luger et al., 1999). PAs were synthesized as described (Baird and Dervan, 1996; Weyermann and Dervan, 2002). The purity and identity of the clamps were established by analytical HPLC and matrix-assisted laser desorption ionization-time-of-flight mass spectroscopy. PAs were purified by preparative HPLC, lyophilized, and stored at 4 °C. Each PA was freshly dissolved in 20 mM potassium cacodylate (pH 6.0) and 1 mM EDTA, and its concentration was determined by measuring UV absorbance by using empirically determined extinction coefficients (ε_{316} = 68,800 for PA1; ε_{314} = 139,000 for PW12, PW13, and PW14). NCP146 (at 40-50 μM) was incubated with a 10-fold molar excess of clamp in solution at

ambient temperature for 45 min. The integrity of the nucleosome preparation after incubation with the ligand was checked by electrophoretic mobility-shift assay (Luger et al., 1999). Crystals of NCP146-clamp complexes were grown 1-2 weeks in 40-45 mM $MnCl_2$, 35-38 mM KCl, and 20 mM potassium cacodylate (pH 6.0) containing ~ 20 μM (~ 4mg/mL) NCP146 at 19 °C by using vapor diffusion. Crystals were harvested and flash-cooled as described.

3.3b Data collection, structure refinement and validation

X-ray data were collected at beamline 5.0.2 at the Advanced Light Source (ALS) in Berkeley, CA. Data for each of the individual structures were collected on single crystals and processed with Denzo and Scalepack (Otwinowski and Minor, 1997). Molecular replacement (using PDB entry 1AOI as the original search model) and refinement were done with CNS (Brunger et al., 1998), and model building with O (Jones et al., 1991). Each model was checked using simulated annealing omit maps during early stages of model building. Polyamides were clearly visible in the original difference density maps calculated following molecular replacement. An initial model for the polyamide was generated using: 1) fragments of previously published polyamide structures (Suto et al., 2003), 2) information from the Uppsala Software Factory HIC-Up server (Kleywegt and Jones, 1998), and 3) Engh and Huber stereochemical parameters (Engh and Huber, 1991). The polyamide was refined along with the rest of the model. The geometry of the final model is excellent (see Table 3.1) with 93.7 % of the residues in the most-favoured regions, 5.6 % in additional allowed regions, 0.7 % in the generously allowed regions, and no residues in the disallowed regions of the Ramachandran plot. The refined

Table 3.1 Crystallographic statistics for NCP146-PW12

Data Set[a]	NCP146-PW12
Data Collection	
Number of crystals merged	1
Space Group	$P2_12_12_1$
Unit Cell Dimensions (Å)	
a	105.4
b	109.7
c	181.8
Resolution range (Å)	2.05 - 100
Unique reflections	124523
Redundancy	4.75
Completeness (%)	94.4 (54.1)
$\langle\langle I\rangle/\sigma(\langle I\rangle)\rangle$	22.99 (2.04)
R_{merge}[b]	0.076 (0.353)
Refinement Statistics	
Reflections in test set	6050
R_{cryst}[c]	0.2194
R_{free}[d]	0.2434
No. amino acids	779[e]
No. DNA base pairs	292
No. water molecules	890
No. Mn^{2+} ions	14
No. Cl^- ions	4
No. polyamides	1
RMSD[f]	
Bonds (Å)	0.008
Angles (°)	1.099
Average B-factors	
Protein	48.5
DNA	98.9
Polyamide	143.7
Solvent	65.1

Numerical values in parentheses are for highest resolution shell. [a]Data collected at ALS BL 5.0.2. [b]$R_{merge} = \sum |I_h - \langle I_h\rangle| / \sum I_h$, where $\langle I_h\rangle$ is the mean of the measurements for a single *hkl*. [c]$R_{cryst} = \sum |F_{obs} - F_{calc}| / \sum F_{obs}$. [d]Calculated using 5 % of the *hkl* data chosen randomly and omitted from the start of refinement. [e]Protein structure also includes a second alternate side chain conformation for 16 histone residues. [f]Root Mean Square Deviation from ideal geometry.

structures were compared with unliganded nucleosome structures using LSQMAN program from the Uppsala Software Factory (Kleywegt, 1996). Several figures in the paper were made using the molecular graphics program, PyMOL (DeLano, W.L. The PyMOL Molecular Graphics System (2002) on World Wide Web http://www.pymol.org).

3.3c Nucleosome dilution assay

The nucleosome dilution experiments were performed as previously described (Gottesfeld and Luger, 2001). Briefly, NCP146 (reconstituted at a molar ratio of 0.6 histone octamer:DNA) was subjected to serial dilution from an initial concentration of 10 nM in steps of 3.3, 1.1, 0.34, 0.11, 0.035, 0.01, 0.004 nM, in a buffer containing 200 mM NaCl, 10 mM Tris-Cl, pH 7.5, 1 mM EDTA, 0.05 % NP-40, 10 % (v/v) glycerol. After incubation at 37 °C for 1 h, the samples were subjected to non-denaturing gel electrophoresis, followed by phosphorimage analysis of the gel.

3.4 Results and Discussion

3.4a Bivalent clamp binds to a nucleosomal supergroove

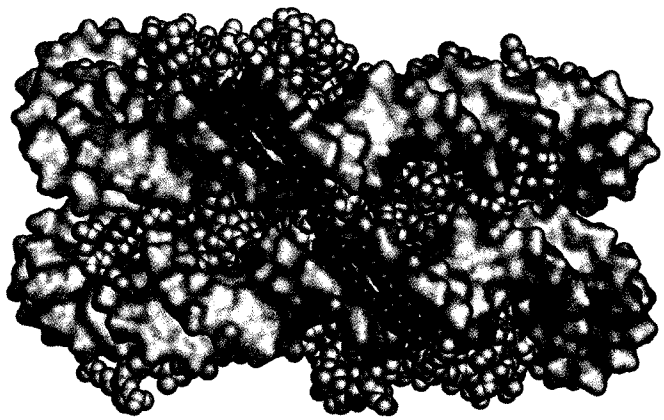
Three bivalent hairpin PA dimers (PW12, PW13, and PW14) were synthesized. A previously characterized eight-ring hairpin, whose binding sites were separated by 80 bp on the palindromic α-satellite DNA fragment used for previous crystallographic studies (PA1) (Gottesfeld et al., 2001; Suto et al., 2003), was bridged in a head-to-head fashion with one, two, or three ethylene glycol units (resulting in distances of ~ 11-18 Å between the two hairpins; Figure 3.1C) to create the bivalent clamps. The predicted hydrogen-bonding pattern of one of the clamps, PW12, to its target DNA sequence is shown in

Figure 3.1D. The NCP146-clamp complexes have the same electrophoretic mobility as unliganded NCP146 and do not show signs of dissociation or aggregation upon binding the clamp (data not shown). Quantitative DNase I footprint titrations (Trauger and Dervan, 2001) (data not shown) demonstrate that the clamps bind NCP146 with affinities in the low nanomolar range. Nonspecific binding of the clamp to NCP146 was not observed.

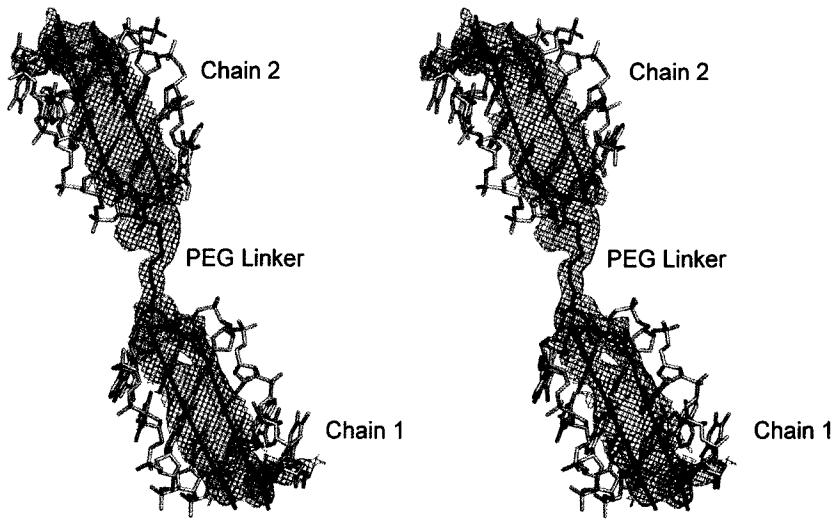
To obtain unambiguous evidence for the mode of interaction of the clamp with nucleosomes, we crystallized NCP146 in complex with each of the three clamps (NCP146-PW12, NCP146-PW13, and NCP146-PW14). Because the biophysical and structural parameters were very similar for the three co-crystal structures, only NCP146-PW12 will be further discussed. Clamp binding increased the size, order, and resolution of NCP146 crystals. High quality x-ray data to 2 Å resolution were obtained for NCP146-PW12 from a single crystal, without detectable radiation damage (Table 3.1). In contrast, nucleosome crystals without bound ligand are highly sensitive to radiation damage, and the recently reported nucleosome structures at similarly high resolution required merging data from as many as 44 crystals (Davey et al., 2002). Because the diffraction qualities of nucleosome crystals depend on the positioning of the DNA with respect to the histone octamer with single bp accuracy (Muthurajan et al., 2003; Richmond and Davey, 2003), our results strongly suggest that the clamp improves the accuracy of DNA positioning with respect to the histone octamer.

We refined the NCP146-PW12 structure to a resolution of 2 Å (Table 3.1 and Figure 3.2A). Electron difference density for PA moieties and linker was clearly visible

A



B



C

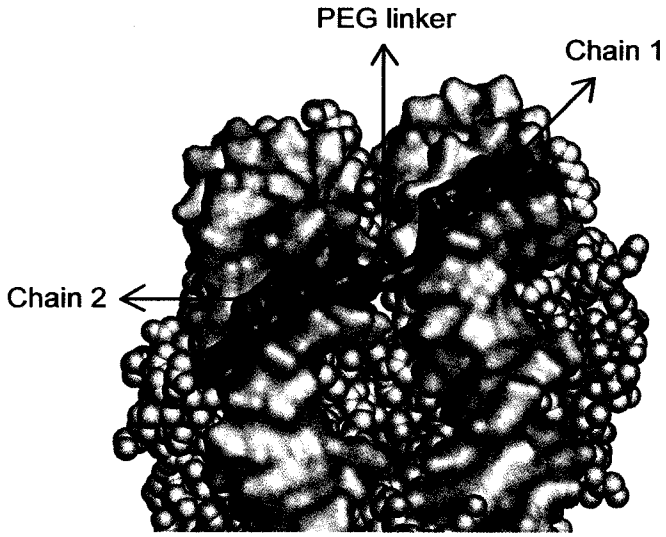


Figure 3.2 Clamp binding in the NCP146-PW12 complex. Color code for histone proteins is as in Figure 1. DNA is shown in green and white, PW12 in green or magenta. (A) Overview of NCP146-PW12 structure, orientation same as in Figure 1B. (B) Stereoview of PW12 bound to its target DNA site. Omit density for the clamp is shown at 2 σ contour level. Chains 1, 2 denote the two hairpin moieties. (C) The linker in the clamp is buried between the two gyres of superhelical DNA. A close-up surface representation of NCP146-PW12 structure is shown.

after molecular replacement and during subsequent refinement cycles (Figure 3.2B). The ethylene glycol linker is buried between the two gyres of the superhelix (Figure 3.2C). Both subunits of the bivalent hairpin ligand bind to their target sites in accordance with previously established pairing rules (Dervan and Edelson, 2003) (Figure 3.1D). Thus, the crystal structure reveals that the clamp indeed targets the supergroove, and that a designed ligand can exploit the special geometry of a DNA supergroove in the nucleosome for high-affinity and high-specificity binding. The structure of the nucleosome and of the hairpin PAs remains virtually unchanged compared with that of NCP146 in complex with the parent PA (PA1) (Suto et al., 2003), demonstrating that no additional structural distortions are imparted to the nucleosome upon clamp binding.

3.4b Clamp binding to a supergroove prevents nucleosome dissociation

Nucleosome dissociation is likely to initiate by the unraveling of the DNA ends from the surface of the histone octamer, followed by the dissociation of one or both (H2A-H2B) dimers (Figure 3.3). Furthermore, unwrapping of the ends of nucleosomal DNA by chromatin remodeling complexes such as SWI/SNF (Kassabov et al., 2003) and Swr1 (Mizuguchi et al., 2004) is thought to expose the DNA-binding surface of the (H2A-H2B) dimer and initiate dissociation and/or exchange of the (H2A-H2B) dimer. The PA clamp may counteract dissociation because it effectively cross-braces the two gyres of nucleosomal DNA and forms a closed DNA circle around the histone octamer (Figure 3.3). Although partial unraveling of the DNA ends may occur in the presence of the clamp, complete dissociation of the DNA is likely to be precluded beyond the clamp

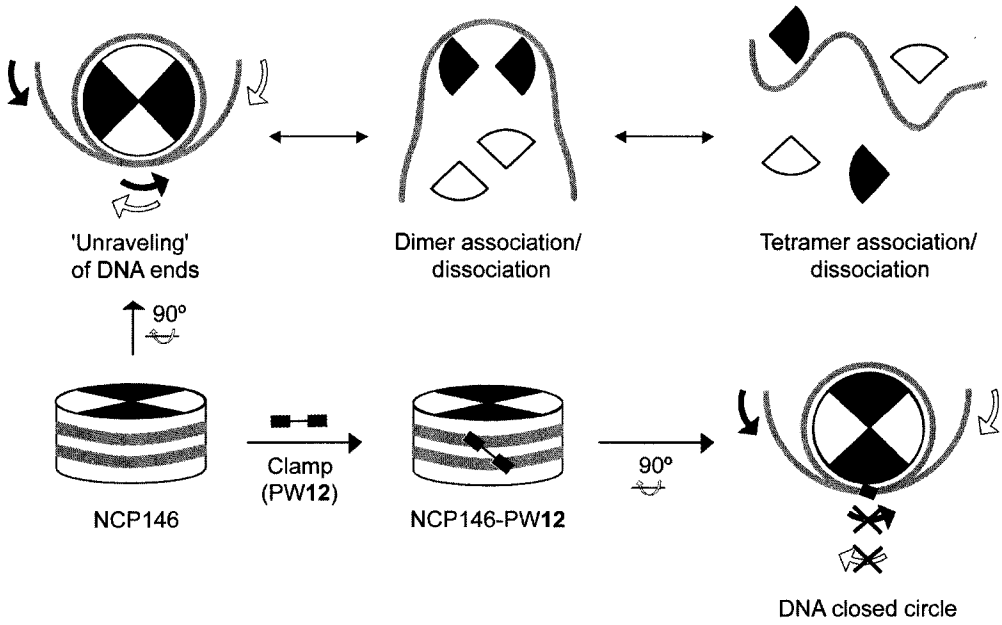


Figure 3.3 Schematic illustration of the predicted effect of polyamide clamps on nucleosome dissociation. In the absence of ligand binding, nucleosome dissociation initiates with unravelling of the DNA ends, followed by dissociation of the (H2A-H2B) dimers, and finally by the dissociation of the $(H3\text{-}H4)_2$ tetramer. Binding clamp to the nucleosomes leads to the formation of a closed ~ 80 bp DNA supercoil that prevents further disassembly of the nucleosomes.

binding sites (Figure 3.3), which would in turn stabilize the nucleosome from dissociation.

The effect of clamp binding on NCP146 stability was investigated by using a simple dilution assay (Ausio et al., 1984; Cotton and Hamkalo, 1981; Gottesfeld and Luger, 2001). This method allows us to compare the kinetic stability of NCP146 in the presence and absence of the clamp. We find that a saturating concentration of PW12 dramatically stabilizes NCP146 against dissociation into free DNA and histones (Figure 3.4A, B). In the absence of PA, we observed 50% dissociation at a NCP146 concentration of 0.33 nM. In the presence of the unlinked parent PA (PA1), 50 % dissociation is observed at ~ 9-fold lower concentration of NCP146 (0.038 nM). In contrast, dissociation is minimal for NCP146-PW12 (compare lane 8 in the three panels in Figure 3.4A, and see Figure 3.4B). The off-rate for PW12 on NCP146 is increased almost by an order of magnitude relative to that of PA1 on NCP146 whereas the equilibrium affinities of both the clamp and PA1 for NCP146 are similar (Table 3.2). Thus, it seems that the kinetic stability (k_{off}) of the clamp on NCP146 is significantly higher than that of PA1 on NCP146. These results are also consistent with our observation that NCP146-clamp complexes form ordered and highly diffracting crystals. PA clamps could be potent reagents to determine whether chromatin remodelers transiently unravel DNA as a step en route to histone exchange. Furthermore, this class of molecules might be used to explore and modulate many aspects of chromatin function: for example, to improve nucleosome positioning and stability on poorly positioned nucleosomal templates, to study the role of nucleosome dissociation

Table 3.2 Dissociation rate constants for PA1 and PW12 with NCP146

Polyamide	PA1	PW12
K_a (M^{-1})[1]	1.1 (± 0.1) X 10^9	0.5 (± 0.05) X 10^9
Half-Life[2]	15 min	180 min
k_{off} (sec^{-1})[3]	7.7 X 10^{-4}	6.4 X 10^{-5}
k_{on} ($M^{-1}sec^{-1}$)[4]	8.5 X10^5	3.2 X 10^4

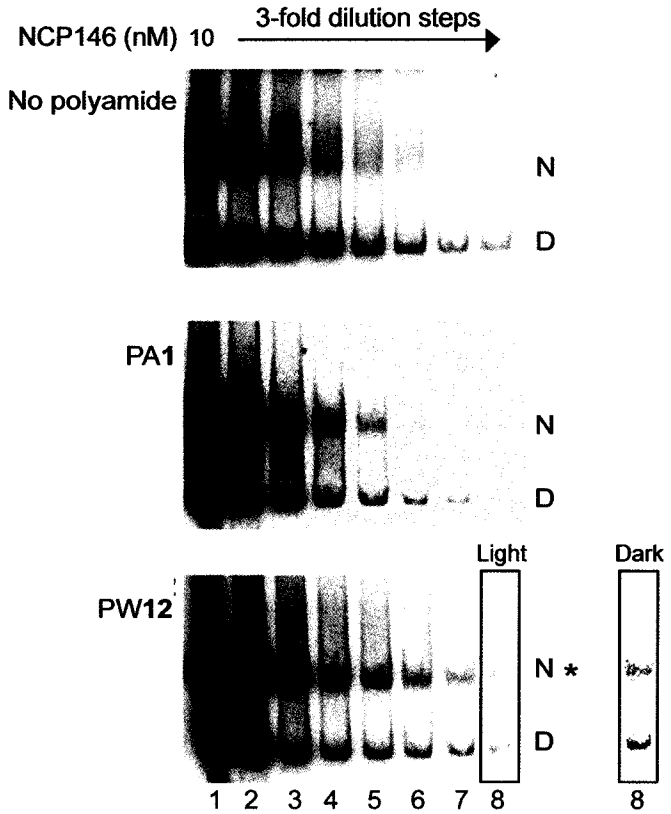
[1] Determined by quantitative DNase I footprinting (data not shown). Values in parentheses denote the standard deviations (SDs). Values reported here are averages and SDs of four independent determinations.

[2] Determined by competition footprinting using unlabelled 146 bp DNA as a competitor (data not shown).

[3] Calculated according to $k_{off} = 0.693/t_{1/2}$.

[4] Calculated from the equation, $k_{on} = (K_a)(k_{off})$.

A



B

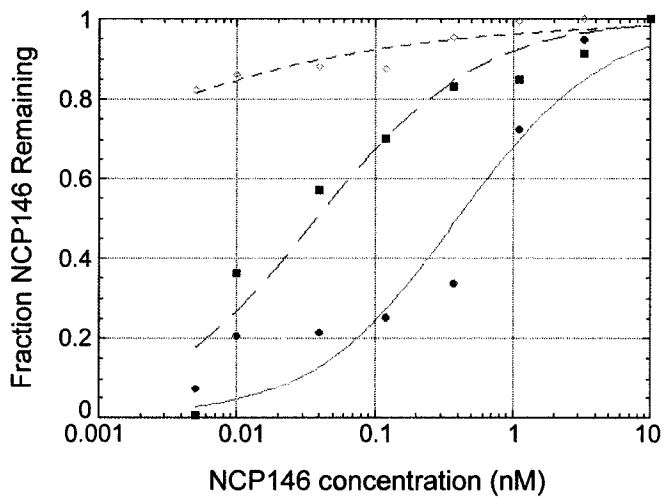


Figure 3.4 Stabilization of NCP146 by the clamp. (A) Non-denaturing gel analysis. NCP146 was subject to serial dilution in a dilution buffer that contained either no polyamide, or 100 nM PA1 or PW12, as indicated. NCP146 samples were then subjected to electrophoresis on a 6 % non-denaturing polyacrylamide gel, and a phosphorimage of the gel is shown. N, NCP146; D, 146 bp DNA; * denotes the stability of the NCP146 upon dilution in the presence of PW12. The side panel shows a darker contrast of lane 8 in the third panel. (B) Graphical analysis of the dissociation data shown in Figure 4A. The fraction of DNA remaining in NCP146, relative to the 10 nM sample, is plotted versus NCP146 concentration. Symbols: ●, no polyamide; ■, PA1; and ◊, PW12.

and dynamics during transcription, replication, and chromatin remodeling, or to target reagents or recruit proteins specifically to particular nucleosomes.

3.4c Nucleosomal supergrooves are platforms of molecular recognition.

Here, we show *in vitro* that DNA architecture embedded within the nucleosome generates interaction platforms or supergrooves that can be targeted by sequence-specific bivalent DNA binding ligands to mediate short-and medium-range interactions, and to modulate nucleosome stability. *In vivo*, nucleosomal supergrooves may mediate medium- and long-range DNA interactions. Indeed, there is evidence that such interactions may regulate transcription of several genes. For example, on the mouse mammary tumor virus long terminal repeat, the precise spacing between two glucocorticoid receptor recognition elements brings about nucleosome-mediated synergistic transcriptional activation (ref. (Collingwood et al., 1999 and references therein). In several other genes, a positioned nucleosome brings two precisely spaced regulatory elements into juxtaposition; examples of this are the *X. laevis* vitellogenin B1 and TRβA genes, the *Drosophila* Adh, hsp26, and hsp27 genes, and the human U6 gene (Collingwood et al., 1999; Stunkel et al., 1997 and references therein). We found numerous examples of tandem or multiple repeat regulatory sequences that are arranged with a regular spacing of ~ 80 bp. Such spacing may have evolved to promote crosstalk between factors bound to noncontiguous regulatory elements in the context of chromatin. Supergrooves occur all along the nucleosomal DNA (seven minor and six major supergrooves per nucleosome), and their occurrence requires only that the two DNA segments be separated by ~ 80 bp and be within the context of a folded nucleosome. Because they do not require precise

translational nucleosome positioning, the presence of nucleosomes over such regulatory elements, and their role in transcription regulation, may often go undetected, and synergistic interactions between regulatory factors may be much more frequent than previously assumed.

3.5 Acknowledgements

We thank Gerry McDermott and Keith Henderson at the Advanced Light Source (Berkeley, CA) for help with data collection; Stephanie Ebbesen and Pamela Dyer for help with reagents; and Laurie Stargell and Clara Kielkopf for helpful comments. The polyamides were synthesized by Philipp Weyermann in the Dervan laboratory, and the nucleosome dilution assays were performed by Joel Gottesfeld. This work was supported by a grant supplement from the National Institutes of Health (to K.L., J.M.G., and P.B.D.) and by a postdoctoral fellowship from the Swiss National Science Foundation (to P.W.).

3.6 Protein Data Bank Coordinates

The atomic coordinates and structure factors for NCP146-PW12 have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 1S32).

3.7 Addendum

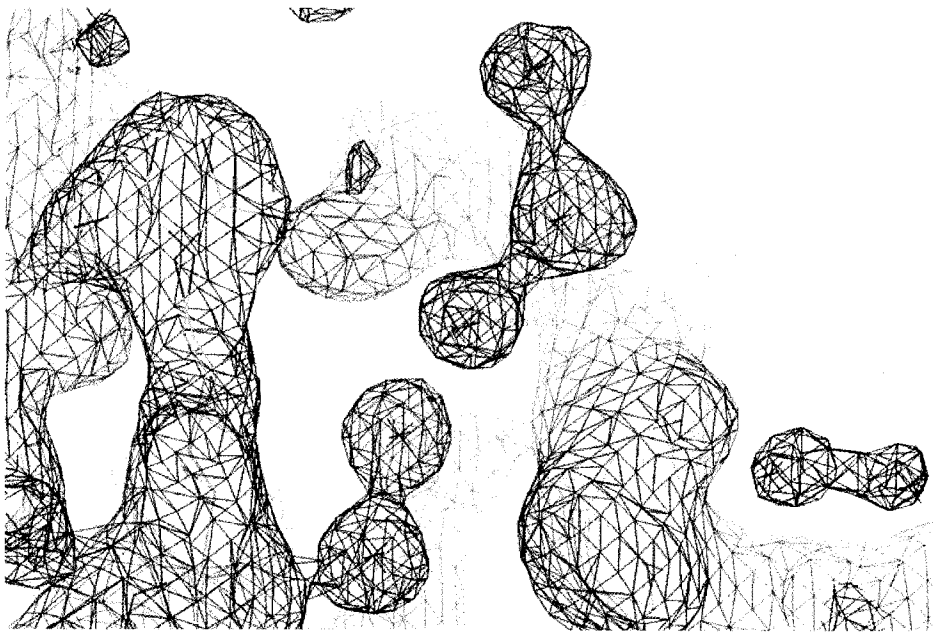
In addition to determining the co-crystal structure of NCP146-PW12 (reported above), I also determined the structures of NCP146-PW13 and NCP146-PW14 (partially refined model only, in the case of NCP-PW14). As compared to the clamp PW12, polyamides PW13 and PW14 contain longer PEG linkers connecting the polyamide

hairpins within the dimers. The crystallographic data collection statistics and refinement statistics for the two complexes are as summarized in Table 3.3. Calculation of difference density maps for NCP146-PW13 and NCP146-PW14 structures reveal density for the two hairpin monomers and density for a part of the linker at the polyamide-linker branching points. However, unlike in the clamp PW12, the density for the PEG linker disappears distal to the polyamide hairpins for both PW13 and PW14 (not shown). The PEG linker may be too flexible in both cases to be visualized. This flexibility could be a result of the increased linker length in these polyamides (~ 14 and ~ 18 Å in PW13 and PW14, respectively) as compared to that in PW12 (~ 11 Å). A least-square superposition of NCP146-PW13 and NCP16-PW14 structures on the structure of NCP146-PW12 reveal no additional differences between the three structures (not shown). We therefore conclude that the polyamide linker length does not influence the structure of the nucleosome.

Table 3.3 Data-collection and refinement statistics

Dataset	NCP-PW13 ALS BL5.0.2	NCP-PW14* RAxis IV
Data Collection		
Mosaicity	0.4	0.6
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell lengths	105.9	106.2
(*a, b, c* in Å)	109.4	109.6
	182.1	181.8
Resolution range (Å)†	99-2.40 (2.46-2.40)	100-2.70 (2.76-2.70)
Unique reflections	82273	57045
Completeness (%)	98.1 (89.8)	97.0 (90.3)
Multiplicity	3.5	3.1
Average $I/\sigma(I)$	26.6 (2.66)	10.4 (1.4)
R-merge#	0.044 (0.397)	0.066 (0.463)
Refinement Statistics		
Reflections in test set	3999	2625
R_{cryst}†	0.2228	0.2308
R_{free}‡	0.2570	0.2805
No. of amino acids	779	764
No. of DNA base pairs	292	292
No. of water molecules	553	-
No. of Mn^{2+} ions	13	-
No. of Cl^{-1} ions	4	-
No. of polyamides	1	-
rmsd		
Bonds (Å)	0.007	0.007
Angles (°)	1.03	1.08
Average B-factors		
Protein	55.5	41.6
DNA	108.0	96.0
Polyamide	145.4	-
Solvent	70.1	-

Numerical values in parentheses denote data for the highest resolution shell.
#R-merge = $\Sigma|I_h - \langle I_h\rangle|/\Sigma I_h$, where $\langle I_h\rangle$ is the mean of the measurements for a single *hkl*.
*Refinement was done for 100-2.76 Å data and the statistics reported are only for a partially refined model.
†$R_{cryst} = \Sigma|F_{obs} - F_{calc}|/ \Sigma F_{obs}$
‡Calculated using 5 % of the *hkl* data chosen randomly and omitted from the start of refinement.
rmsd, rms deviation from ideal geometry.



Chapter 4

Nucleosomes in Solution Exist as a Mixture of Twist-Defect States

Edayathumangalam, R. S., Weyermann, P., Dervan, P. B., Gottesfeld, J. M., and Luger, K. (2004) Nucleosomes in solution exist as a mixture of twist-defect states. J Mol Biol, *345*, 103-114.

This chapter was published as cited above. The text and figures for this chapter are presented as they appeared in the above publication.

Abstract

The 2.0 Å crystal structure of a nucleosome core particle in complex with a bivalent pyrrole-imidazole polyamide reveals that this 'clamp' effectively crossbraces the two gyres of the DNA superhelix, thereby stabilizing the nucleosome against dissociation. Using x-ray crystallography and footprinting techniques, we show that the clamp preferentially binds nucleosomes over free DNA, and that nucleosomal DNA exists as a mixture of multiple twist-defect intermediates in solution. The nucleosomes exist in one of two different conformations in various crystal structures that trap twist-diffusion intermediates, even on a strong positioning sequence. Evidence has been obtained supporting the existence of twist-defect states in nucleosomal DNA in solution that are similar to those obtained in crystal structures. Our results also substantiate the idea that twist diffusion may represent an important means of altering the accessibility of nucleosomal DNA both in the presence and absence of ATP-dependent chromatin-remodelling enzymes.

4.1 Introduction

The three-dimensional structure of the nucleosome core particle (NCP), the fundamental structural unit of chromatin, has previously been determined to high resolution (Davey et al., 2002; Harp et al., 2000; Luger et al., 1997; White et al., 2001). These structures present a detailed view of DNA in a physiologically relevant context (Richmond and Davey, 2003), and provide insights into the intricacies of DNA compaction in the eukaryotic cell. The nucleosome has long been viewed as an 'immovable object' that represents a formidable obstacle to all processes that require access to the DNA substrate (Kornberg and Lorch, 1991). This picture is rapidly changing. Despite over

120 direct contacts (and roughly the same number of water-mediated interactions) made by the histone octamer with the DNA at fourteen independent contact points over its entire length (Davey et al., 2002), nucleosomes are highly dynamic and malleable structures that undergo histone exchange, chromatin remodeling, and reversible posttranslational modifications (Berger, 2002; Cheung et al., 2000; Dutnall, 2003; Flaus and Owen-Hughes, 2003; Flaus and Owen-Hughes, 2004; Jenuwein and Allis, 2001; Korber and Horz, 2004; Langst and Becker, 2004; Narlikar et al., 2002).

Approximately 95 % of genomic DNA sequences do not appear to be capable of pronounced sequence-specific nucleosome positioning (Lowary and Widom, 1997; Richmond and Widom, 2000). Positioning is likely to be governed by a combination of DNA bendability and curvature mainly over the central 80 base pairs of DNA, and is likely to represent a compromise between several competing features that are embedded in the DNA sequence. Whether uniquely, loosely, or not at all positioned, nucleosomes are dynamic and are acted upon by ATP-dependent chromatin-remodeling factors in order to regulate nucleosomal DNA accessibility (Flaus and Owen-Hughes, 2004; Langst and Becker, 2004; Narlikar et al., 2002) and to bring about subsequent gene activation or repression. Although many ATP-dependent chromatin-remodeling factors have been shown to enhance the mobility of nucleosomes and alleviate the repressive effects of chromatin on transcription (Hamiche et al., 1999; Kingston and Narlikar, 1999; Langst et al., 1999; Narlikar et al., 2002; Whitehouse et al., 1999), the molecular mechanism by which nucleosome repositioning occurs is still controversial (Flaus and Owen-Hughes, 2004; Langst and Becker, 2004; Narlikar et al., 2002). Two current models for nu-

cleosome remodeling are based on twist diffusion or propagation of a DNA bulge on the surface of the histone octamer. It has also been recently demonstrated that ATP-dependent chromatin remodeling can exchange or remove histone dimers from nucleosomes (Flaus and Owen-Hughes, 2004). One concept that has emerged from these studies is that the nucleosome is a dynamic entity, and further knowledge concerning DNA conformational flexibility in the nucleosome is key to understanding these vital processes.

Here, we investigate the different possible states of the nucleosome in solution and in the crystal lattice using previously described bivalent pyrrole-imidazole polyamides (Py-Im PAs) that target the DNA 'supergrooves' within the nucleosome. DNA supergrooves have recently been reported as non-contiguous minor grooves (or major grooves) that are brought into structural alignment across the two gyres of superhelical DNA within the NCP (Edayathumangalam et al., 2004). These ligands effectively crossbrace the two gyres of nucleosomal DNA in a site-specific manner, and thus are likely to capture different structural states of the nucleosome in solution. Py-Im PAs bind predetermined DNA sequences with affinities in the nanomolar to the subnanomolar range, and have been shown to modulate specific gene expression (Dervan and Edelson, 2003). Using both DNase I footprinting and affinity cleavage techniques, we have shown previously that Py-Im PAs bind nucleosomes *in vitro*, demonstrating that much of the DNA in the context of the NCP is readily accessible for molecular recognition (Gottesfeld et al., 2001).

X-ray crystallographic studies of NCP-polyamide complexes reveal the nature of structural changes in the nucleosome upon polyamide binding at molecular detail (Suto et al., 2003). These changes occur exclusively in the nucleosomal DNA without altering the underlying protein-DNA interactions. Due to the palindromic nature of the 146 bp human α-satellite DNA used in our studies, and due to the fortuitous location of the binding sites for the parent polyamide PA1 in NCP reconstituted with this particular DNA fragment (NCP146), two molecules of PA1 bind match sites that are immediately opposed across the two gyres of DNA in NCP146 (Suto et al., 2003). Based on these results, we designed and characterized a bivalent Py-Im PA 'clamp' that binds an entire supergroove of nucleosomal DNA in a site-specific manner and effectively crossbraces the two gyres of nucleosomal DNA (Edayathumangalam et al., 2004). The bivalent clamp was synthesized as a PA1 dimer connected by an ethylene glycol (PEG) linker. Using x-ray crystallography and footprinting techniques, we have demonstrated that the clamp binds to NCP146 across the supergroove as designed (Edayathumangalam et al., 2004). The significantly improved diffraction quality of the NCP146-clamp complex, compared to the unliganded NCP146, suggests that the clamp improves DNA positioning with respect to the histone octamer. Biochemical assays reveal that the clamp binds with very high affinity and specificity, and has a dramatic effect on the *in vitro* stability of nucleosomes against dilution-induced dissociation (Edayathumangalam et al., 2004).

Here we present a more detailed analysis of the NCP146-clamp complex, and compare quantitative footprinting data of a clamp bound to nucleosomes reconstituted with two DNA fragments of almost identical sequence but whose lengths differ by one base

pair. We further show that the DNA in NCP146 continuously samples several conformations in solution, and that the binding of the polyamide clamp favors and stabilizes some conformations in preference to others. The multiple conformations observed in solution appear to be 'twist-defect' intermediates, akin to those trapped in the crystal structures. We propose that 'twist diffusion' processes may be involved in the redistribution of twist-defect states of the nucleosomes in solution. The occurrence of such dynamic motions in the nucleosomal DNA could help optimize DNA positioning signals with respect to the octamer, could assist ATP-dependent nucleosome sliding, and regulate DNA accessibility at the level of a mononucleosome. The use of a highly positioned DNA sequence in our studies has allowed us to capture subtle perturbations in the dynamic states of nucleosomal DNA in solution. Our studies suggest that conformational flexibility is an attribute of the nucleosomal DNA regardless of whether or not nucleosomes are uniquely positioned.

4.2 Materials and Methods

4.3a Reconstitution of NCP

Previously established protocols were used to reconstitute NCP146 and NCP147 from recombinant Xenopus laevis histones and a 146 bp or 147 bp palindromic DNA fragment derived from human α-satellite DNA (Davey et al., 2002; Dyer et al., 2004; Luger et al., 1999).

4.3b DNase I, hydroxyl radical footprinting and K_d measurements

Previously published protocols were followed to perform DNase I and hydroxyl radical footprinting experiments, for the measurements of dissociation constants, and the estimation of half-lives of the polyamide PA1 and the clamp polyamide for the 146 bp DNA, NCP146 and NCP147 (Gottesfeld and Luger, 2001; Gottesfeld et al., 2001).

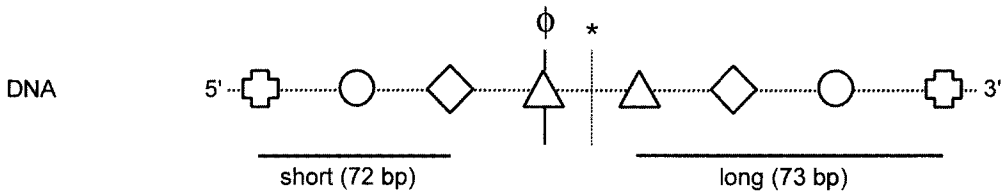
4.4 Results

4.4a The polyamide clamp alters the dynamics of DNA 'stretching' in NCP146

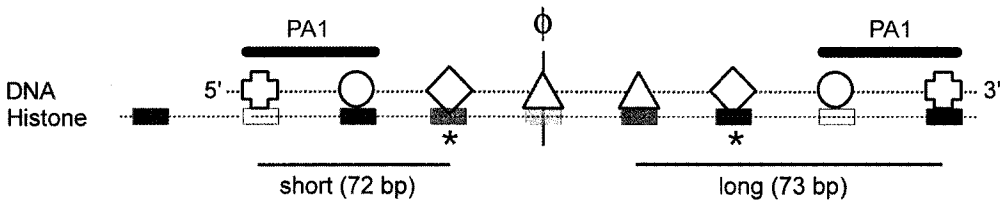
The 2.0 Å crystal structure of NCP146 in complex with the bivalent clamp polyamide PW12 (NCP146-clamp, PDB accession code 1S32) has been recently reported (Edayathumangalam et al., 2004). A comparison with the structure of NCP146 in complex with the unlinked parent molecule, PA1 (NCP146-PA1 (Suto et al., 2003); PDB accession code 1M18) by least square superposition reveals that the structure of the polyamide itself remains largely unaffected by the presence of the PEG linker. Local DNA distortions due to ligand binding, as described earlier for NCP146-PA1 (Suto et al., 2003), are also very similar between the two structures (data not shown).

The most striking structural difference between NCP146-PA1 and NCP146-clamp is found in the manner in which the two nucleosomes accommodate the 'twist defect' that results from the use of a 146 bp (instead of a 147 bp) DNA fragment for crystallographic purposes. As noted earlier, the palindromic DNA fragment in NCP146 crystals is not bound symmetrically by the histone octamer since the non-crystallographic 'dyad' axis of symmetry (ϕ in Figure 4.1A) that bisects the NCP passes through the central base pair of

A Observed location of dyad (ϕ) in NCP146 structures



B Asymmetric location of PA1 sites in NCP146-PA1 structure



C Predicted symmetric location of PA1 sites in NCP147

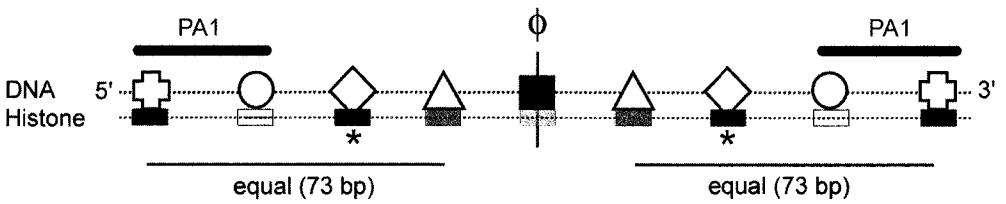


Figure 4.1 Schematic to describe the relative positions of polyamide binding sites on different nucleosomal DNA sequences. Only a few bases and histone contacts are depicted for illustration. Open, black symbols represent individual DNA bases; coloured rectangles, underlying histone contacts; ϕ, pseudo-twofold dyad axis; ■, PA1 binding sites; *, symmetry axis of the palindrome; and *, an example of altered histone-DNA contacts between NCP146 and NCP147 structures. (A) Observed location of the structural dyad axis in relation to the palindrome axis of symmetry in NCP146 structures. (B) Asymmetric location of PA1 binding sites in NCP146-PA1 structure. (C) Predicted symmetric location of PA1 sites in NCP147.

the DNA (Muthurajan et al., 2003). Since the dyad axis is not coincident with the symmetry axis of the palindromic DNA fragment (* in Figure 4.1A), equivalent base pairs encounter non-equivalent histone contacts in the two halves of the palindrome, as shown schematically in Figure 4.1B (for example, see *). As a second consequence, structural alignment of the two halves around the central base pair reveals that the palindrome is bisected into a short (72 bp) and a long (73 bp) half.

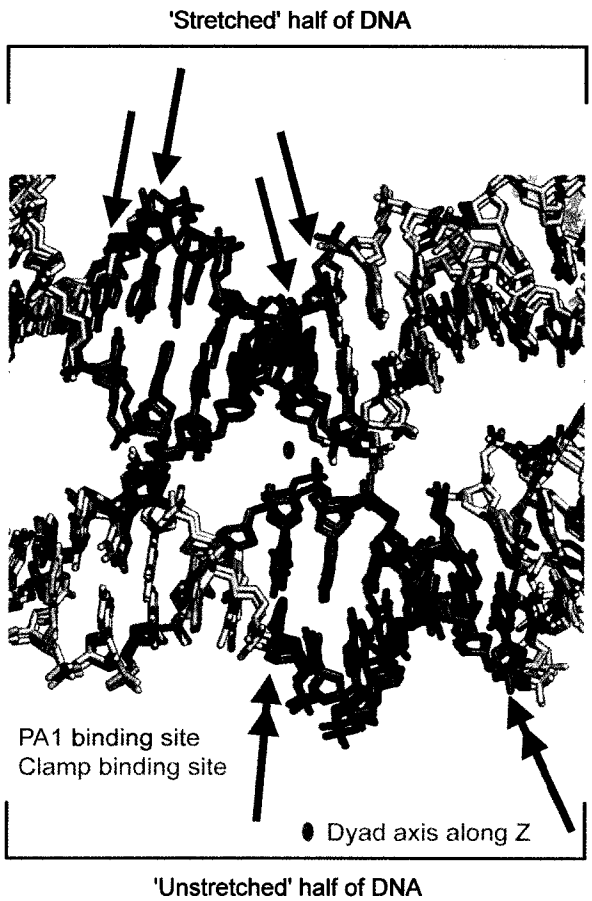
Crystal packing requires that both ends of the DNA be engaged in base-stacking interactions with a neighboring nucleosome. Consequently, the 72-bp half is twisted and pulled to accommodate the one-bp length difference with respect to the 73-bp half, thereby generating a 'twist defect' (Muthurajan et al., 2003). This twist defect diffuses along the length of the DNA and is usually dissipated at a location ~ 20 bp from the nucleosomal dyad (by simultaneous DNA 'stretching' and overwinding). Thus, the discrepancy between sequence and structural symmetry remains only in the 20 base pairs between the dyad and the 'stretch site'. The preference for twist diffusion to this particular location has been studied earlier and is observed in most unliganded NCP146 structures determined to date (Luger, 2003). In contrast, DNA stretching is absent in the structure of a NCP reconstituted with a 147 bp DNA fragment (Davey et al., 2002; NCP147), suggesting that the incorporation of one additional bp in the 146 bp DNA (shown by ■ in Figure 4.1C alleviates the requirement for stretching prior to crystallization.

We have shown earlier that twist diffusion cannot proceed past the polyamide binding site in NCP146-PA1 (Muthurajan et al., 2003; Suto et al., 2003), and that a different region, which is closer to the end of the DNA (~ 45 bp from the dyad) and situated be-

tween the DNA end and the bound polyamide, is chosen to dissipate the twist defect. Interestingly, this ~ 45 bp region is identical to the location where stretching was previously reported in nucleosomes containing a different 146 bp DNA sequence (Davey et al., 2002). The alternative stretch site in NCP146-PA1 places the two binding sites for PA1 in different structural contexts with respect to the underlying histones (Figure 4.1B). For NCP147, we predict that PA1 would occupy structurally equivalent sites since both sequence and structural symmetry exist on either side of the dyad in NCP147 (Figure 4.1C, see * for example). Unexpectedly, stretching in NCP146-clamp reverts to the location ~ 20 bp near the dyad, which is different from that in NCP146-PA1, but identical to that observed in previously determined NCP146 structures (Luger, 2003). This difference in stretch sites places one of the two polyamide binding sites in a different structural context in NCP146-PA1 compared to NCP146-clamp, as shown in Figure 4.2A (indicated by purple and green arrows). Thus, while the two sites for PA1 in NCP146-PA1 are asymmetrically placed about the octamer, binding sites for both polyamide modules of the clamp in NCP146-clamp are present in structurally identical contexts. Consequently, the polyamide modules bound in the 'unstretched' halves in the two structures are perfectly superimposed, while they are offset by one aromatic ring unit in the 'stretched' half (Figure 4.2B, indicated by purple and green arrows). Identical results were obtained with two other bivalent polyamides that exhibited longer linker lengths (not shown).

What causes this difference in the stretching sites between two nucleosomes reconstituted with identical DNA fragments and complexed with identical polyamide modules?

A



B

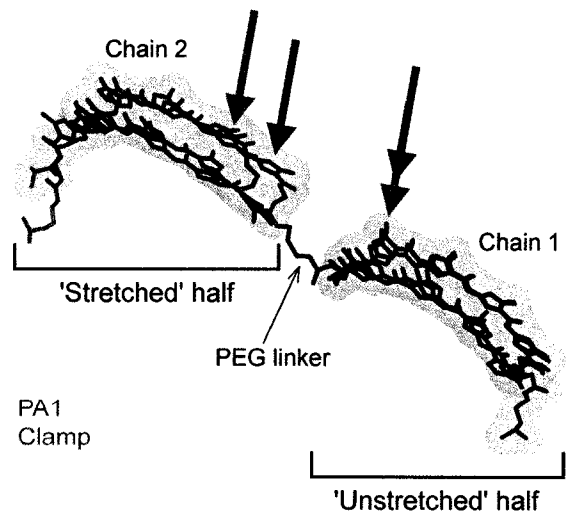


Figure 4.2 Comparison of NCP146-PA1 and NCP146-clamp structures. A superposition of NCP146-PA1 (PDB entry code 1M18) and NCP146-clamp (PDB entry code 1S32) structures based on alignment of histone octamer is shown. The nucleosome is viewed with the superhelical axis along the y-axis. (A) Comparison of the structural context of the recognition sites for PA1 and the clamp. Only the DNA in NCP146 around the two binding sites is shown. (B) Comparison of the structures of PA1 and the clamp. Polyamides alone are shown for clarity.

The only difference between PA1 and the clamp is the PEG linker that connects the two PA1 modules in the latter. Structural modeling and grafting of the PEG linker into the NCP146-PA1 structure reveals that the PEG linker would have to adopt an unfavorably kinked conformation in order to link the two molecules of PA1 in NCP146-PA1 (data not shown). Therefore, as a result of the conformational constraints of the PEG linker, the clamp adopts a different conformation and in turn also selects for a different DNA conformation than that observed in NCP146-PA1. Since the clamp binds to NCP146 in solution, this hypothesis stipulates that DNA stretching occurs not only during crystallization, but also may occur in solution. The following scenarios are possible. First, positioning determinants embedded in the 146 bp DNA sequence are strong enough to favor a single stretched state in solution for NCP146 at all times (such that the structures of NCP146 and NCP147 in solution are in fact identical over most of the DNA sequence). Second, the nucleosomes exist as a mixture of various stretched (and perhaps unstretched) states in solution. Third, stretching is absent in NCP146 in solution and clamp binding alone drives DNA into a conformation similar to that in NCP147.

4.4b The clamp prefers NCP147 over NCP146 and free DNA

In order to distinguish between the above possibilities, we compared the relative affinity of the clamp for NCP146 and NCP147. If NCP146 exist as one single stretched species that favors clamp binding in solution, then the affinity of the clamp for NCP146 and NCP147 would be expected to be indistinguishable. However, if NCP146 were distributed in solution between states, only a sub-population of which is optimally bound by the clamp, the overall on-rate of the clamp for NCP146 would be slower than that for

NCP147 and the overall affinity of the clamp for NCP146 would be expected to be lower than that for NCP147. Similarly, if stretching is absent in solution, and clamp binding drives stretching, the energetic cost of this conformational change would predict a lower affinity of the clamp for NCP146 compared to NCP147. This energetic penalty would likely be reflected in a slower on-rate as well. The second and third scenarios might be distinguished by comparing the structures of NCP146 and NCP147 in solution using footprinting methods (see below). We measured the binding affinities (K_a) of the clamp for both NCP146 and NCP147 by quantitative DNase I footprint titrations, and found that the clamp has a four-fold higher affinity for NCP147 compared to NCP146 (Table 4.1). We also performed footprint competition experiments to directly measure the half-lives ($t_{1/2}$) of the clamp on 146 bp DNA (Figure 4.3), NCP146 and NCP147 (Figure 4.4). We formed DNA- or NCP-polyamide complexes and then challenged these complexes with a large excess of unlabelled DNA, to provide a "sink" for dissociating clamp molecules. After various incubation times, we subjected the samples to DNase I digestion and gel analysis. Off-rates (k_{off}) were then calculated using the $t_{1/2}$ values from the half-life experiment, as shown in Table 4.1. We also calculated on-rates (k_{on}) from K_a and k_{off} values, as described in Table 4.1. A comparison of the on-rates reveals that the clamp binds to NCP147 with a calculated k_{on} value that is almost an order of magnitude higher than that for NCP146, while the k_{off} values are similar for both NCP146 and NCP147. Therefore, the clamp forms a more efficient complex with NCP147, where no stretching is required for optimal access to the clamp binding sites, suggesting that NCP147 is the optimal binding substrate for the clamp in solution. However, once bound to the NCP, the

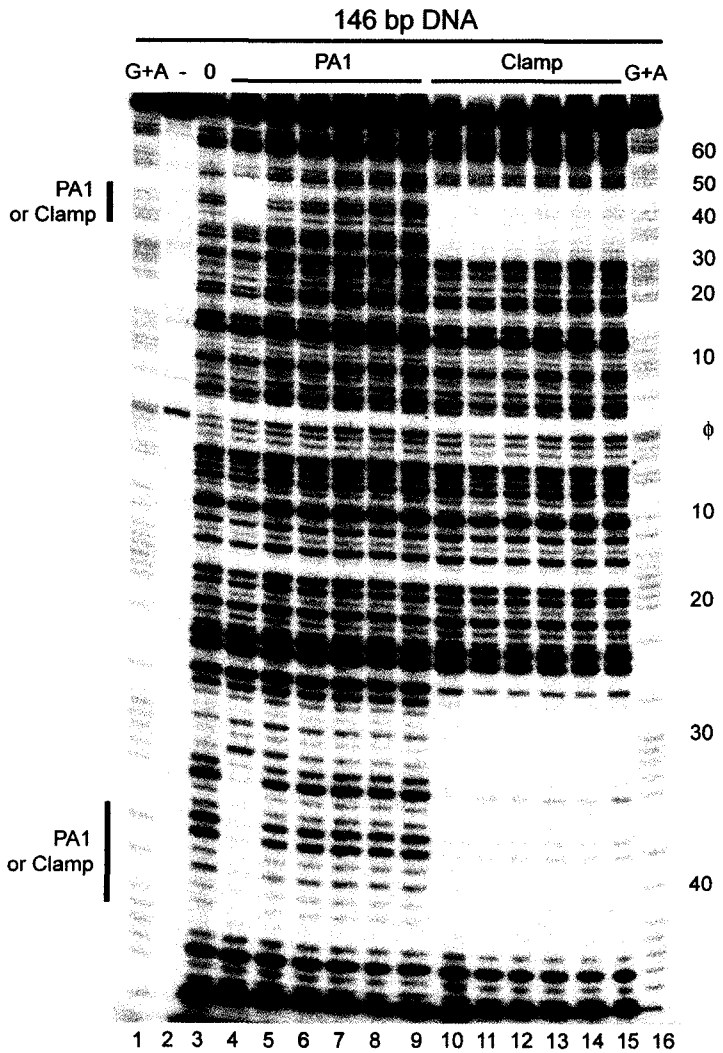


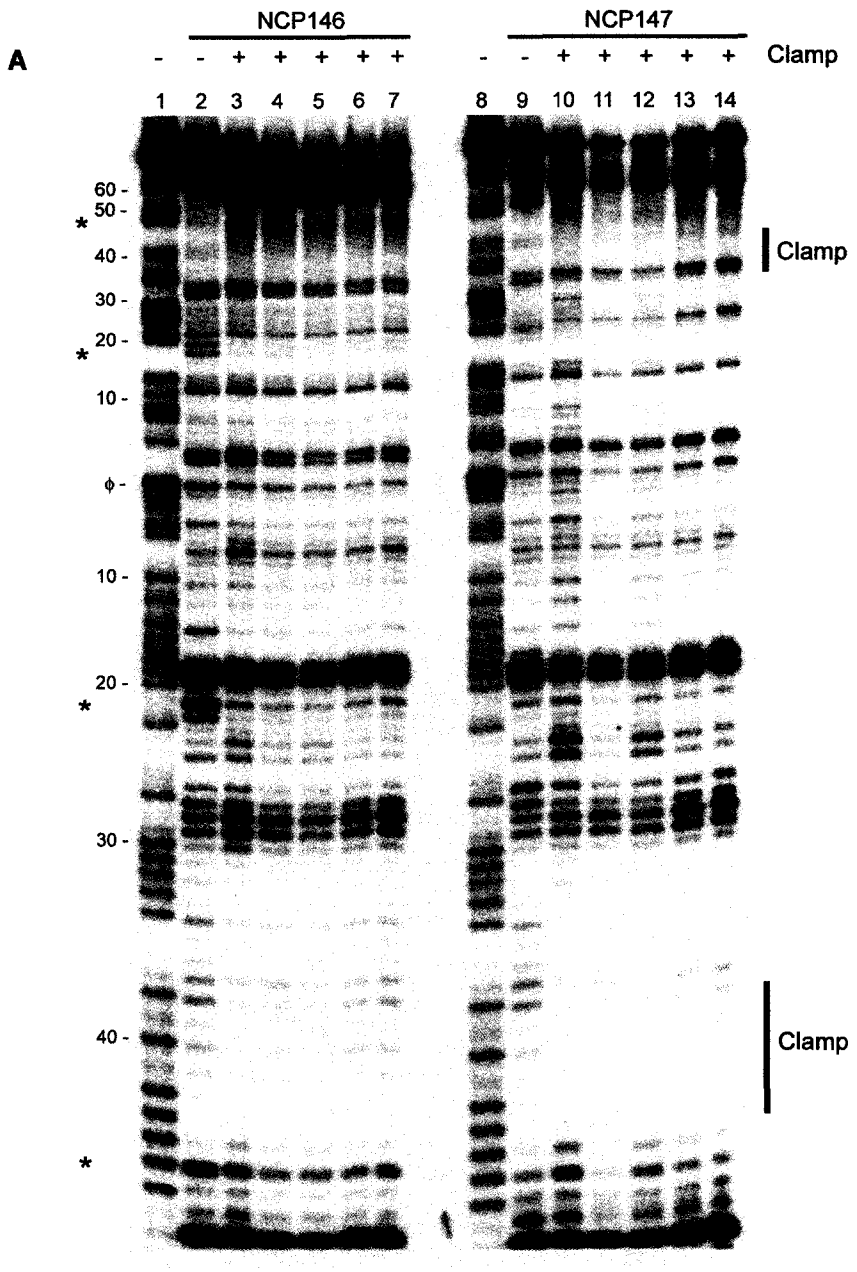
Figure 4.3 Estimation of polyamide half-lives for 146 bp DNA by DNase I footprint competition. Each reaction contained radiolabelled DNA at a final concentration of 1.3 nM in a total volume of 200 µl. Lanes 1 and 16, G+A sequencing reactions of 146 bp DNA; lane 2, no DNase; lane 3, no polyamide; lanes 4-9, 100 nM PA1 as a parent polyamide control; lanes 10-15, 100 nM clamp. Polyamide-146 bp DNA complexes were allowed to form for 16 h at room temperature. In lanes 5-9 and 11-15, complexes were challenged with 100 nM unlabeled 146 bp DNA for various times prior to DNase digestion: 15 min, 30 min, 1 h, 1.5 h, and 2h, respectively for PA1; 15 min, 30 min, 1 h, 2 h, and 4 h, respectively for the clamp. PA1 or clamp binding sites and nucleotide positions from the dyad (ϕ) are shown alongside the gel. The band migrating around the dyad region in lane 2 corresponds to remaining unligated 73 bp DNA that has not been purified away from the 146 bp DNA.

interactions of the clamp with both NCP146 and NCP147 are identical as determined by identical off-rates of the clamp for both NCP146 and NCP147.

Inspection of the K_a values also reveals that the affinity of the clamp for free 146 bp DNA is similar to that for NCP146, both being four-fold lower than that for NCP147 (Table 4.1). This suggests that the clamp can discriminate nucleosomes over free DNA if the 'correct' design of the DNA is chosen (namely, 147 bp DNA instead of 146 bp DNA), and that the NCP146 population is heterogeneous in solution. Only a sub-population of NCP146 (likely identical to NCP147) among the heterogeneous species of NCP146 may be optimal for clamp binding, thereby reducing the overall k_{on} and affinity of the clamp for NCP146 (Table 4.1). The above-mentioned results therefore argue against the idea that NCP146 already exist in a single stretched conformation similar to NCP147 in solution.

4.4c Structural differences between NCP146 and NCP147 suggest the presence of twist-diffusion intermediates in solution

The lowered affinity of the clamp for NCP146 could lead to two possibilities, as outlined above. To distinguish between the above two possibilities, and to investigate the structural differences between NCP146 and NCP147 in solution, we compared the DNase I footprints for NCP146, NCP147, NCP146-clamp and NCP147-clamp (Figure 4.4). It is important to note that the footprint obtained for the nucleosomes is an average of the two strands of DNA (since the DNA is palindromic, 5' end labeling of the DNA allows a direct readout of both strands). Figure 4.4A shows a typical experiment. If stretching were absent in nucleosomes in solution, the addition of one base pair in 147 bp DNA would



B

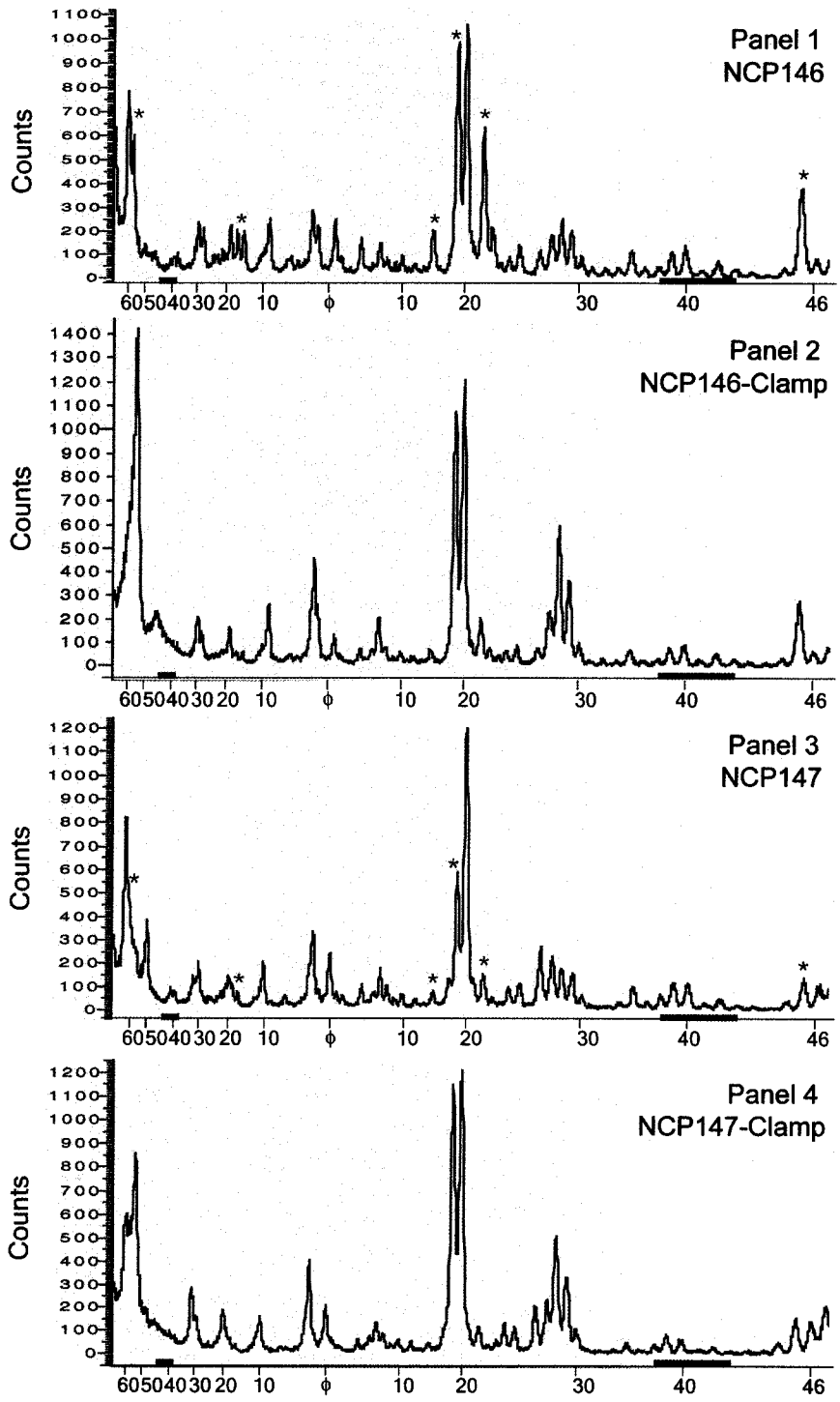


Figure 4.4 (A) Polyamide half-lives determined for NCP146 and NCP147 by DNase I footprint competition, in the presence and absence of the clamp. A phosphorimage is shown with nucleotide positions from the NCP dyad (ϕ) indicated alongside the gel. G+A sequencing reactions of the 146 bp and 147 bp DNA are shown in lanes 1 and 8, respectively. Lanes 2-7 and 9-14 represent digests of NCP146 and NCP147, respectively. Binding reactions were performed for 16 h prior to competition with unlabeled DNA, and contained 250 pM NCP, and 10 nM clamp. Low salt buffer conditions in this experiment (10 mM Tris-Cl, pH 7.5, 10 mM NaCl) allow the use of low NCP concentrations without any observable NCP dissociation (Gottesfeld and Luger, 2001). The complexes were challenged with 200 nM unlabeled 146 bp or 147 bp DNA for various times: no competitor DNA (lanes 3, 10), 30 min (lanes 4, 11), 1 h (lanes 5, 12), 2 h (lanes 6, 13), and 4 h (lanes 7, 14) prior to DNase I digestion. Lanes 2, 9 did not contain any polyamides or competitor DNA. The "match" binding sites for the clamp are indicated alongside the gel. The prominent band in lanes 3-7 and lanes 10-14 (at position ~55 bases from the dyad) is a gel artefact and represents the full-length input DNA. (B) Quantification of the bands in lanes 2, 7, 9, 14 are shown in panels 1, 2, 3, 4, respectively, as determined using ImageQuant software (Molecular Dynamics). The clamp binding sites are shown in the scans (▬).

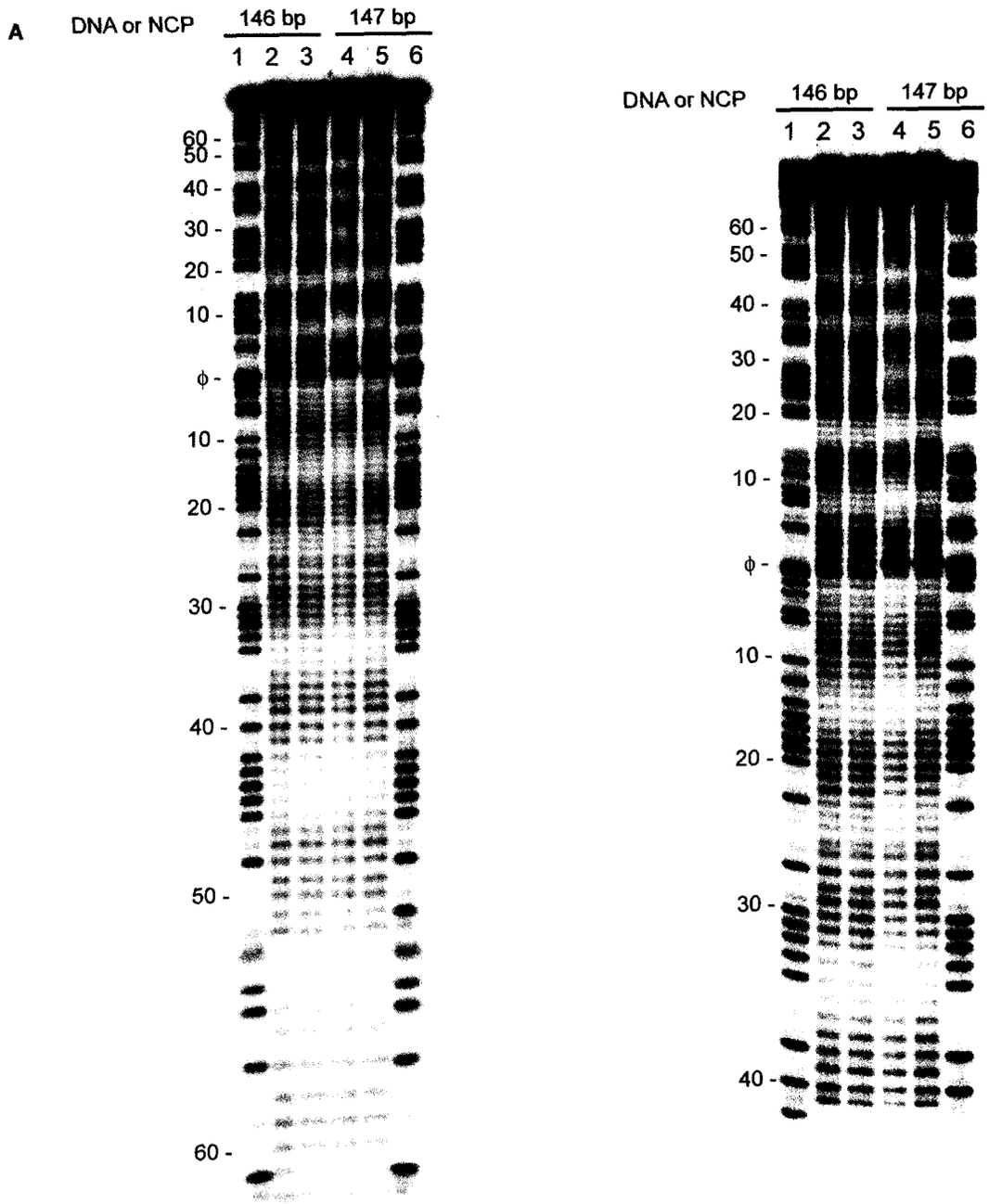
lead to subtle differences between NCP146 and NCP147 throughout the entire length of the footprint. DNA stretching in localized regions of the nucleosome would be expected to cause prominent differences over and above the subtle differences expected above.

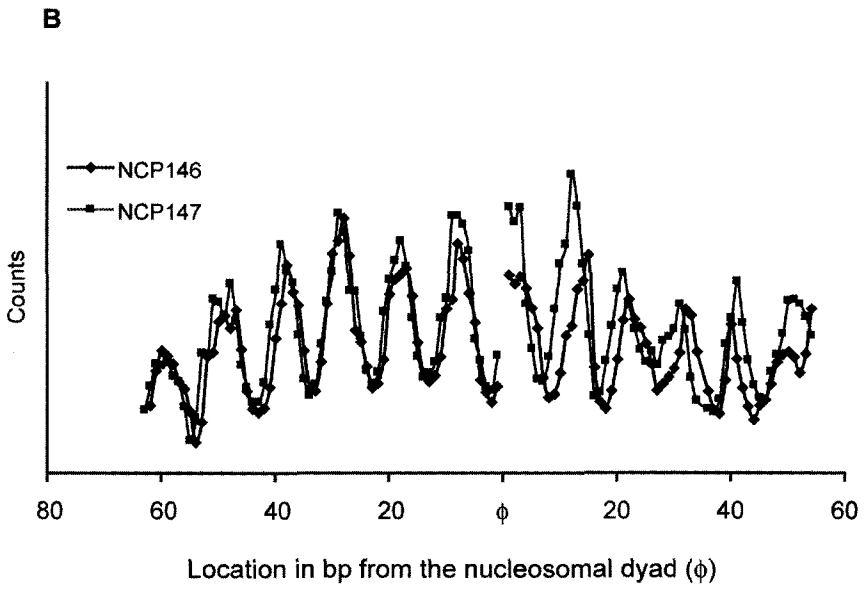
As expected, we find some subtle differences in the footprints for NCP146 and NCP147, which can be attributed to strand averaging. Importantly, prominent differences occur in the DNase I footprinting patterns between NCP146 and NCP147 (Figure 4.4A, see *, compare lanes 2 and 9; and Figure 4.4B, compare * in panels 1 and 3) despite virtually identical DNA sequences and three-dimensional structures. The most dramatic differences occur ~ 45 bp and ~ 20 bp on either side of the dyad, indicating structural differences between NCP146 and NCP147 in these two regions. We also see additional differences at ~ 5 bp and ~ 60 bp. The resolution of the footprints do not allow for the detection of differences beyond ~ 45 bp on one side of the gel. The differences at ~ 5 bp from the dyad could be attributed to the insertion of an additional bp in the 147 bp fragment. These results thereby suggest heterogeneity in the location of the twist defects in nucleosomes in solution, which is in contrast to the homogeneous diffusion of the twist defect to a single location in the crystal structures. The heterogeneity in solution is also suggestive of the highly dynamic nature of nucleosomal DNA. Intriguingly, the major footprinting differences seen at ~ 20 bp and ~ 45 bp are identical to the locations where stretching is observed in the crystals of NCP146 (Muthurajan et al., 2003). This indicates that at least a fraction of the NCP146 population in solution exists in a stretched state, some with the stretch site centered ~ 20 bp, and others ~ 45 bp from the nucleosomal dyad. Therefore, we can rule out the possibility that stretching is absent in solution. If

stretching were completely absent in NCP146 in solution and clamp binding drives NCP146 to a conformation identical to that of NCP147, we would not expect to see the above-mentioned footprinting differences at multiple discrete locations between NCP146 and NCP147.

We also compared the effect of clamp binding on both NCP146 and NCP147. In contrast to the changes induced to the DNA structure following clamp binding to NCP146 (Figure 4.4B, compare panel 1 with 2), ligand binding to NCP147 brings about very little changes to the DNA structure, and the footprints for NCP147 and NCP147-clamp look very similar (see Figure 4.4B, compare panels 3 with 4). The most striking observation from these studies is that the footprints of clamp-bound NCP146 and clamp-bound NCP147 very closely resemble one another, with both being similar to that of NCP147 without the ligand (Figure 4.4B, compare panels 2-4). This suggests that NCP146-clamp and NCP147-clamp complexes exhibit a highly similar DNA structure in solution, despite the pronounced footprinting differences between NCP146 and NCP147 in the absence of the ligand. Therefore, while the DNA in NCP146 exhibits conformational heterogeneity in the absence of the clamp, the DNA structure in the NCP146-clamp complex appears to be selected by the ligand to be one which is similar to NCP147.

Next, we performed hydroxyl radical footprinting for NCP146 and NCP147 to further substantiate our DNase I footprinting results (Figure 4.5A). As discussed above, if stretching did not occur in NCP146 in solution, the difference between the 146 bp and 147 bp DNA fragments would lead to subtle overall differences between NCP146 and NCP147. In contrast, if twist-defect intermediates were present for NCP146 in solution,





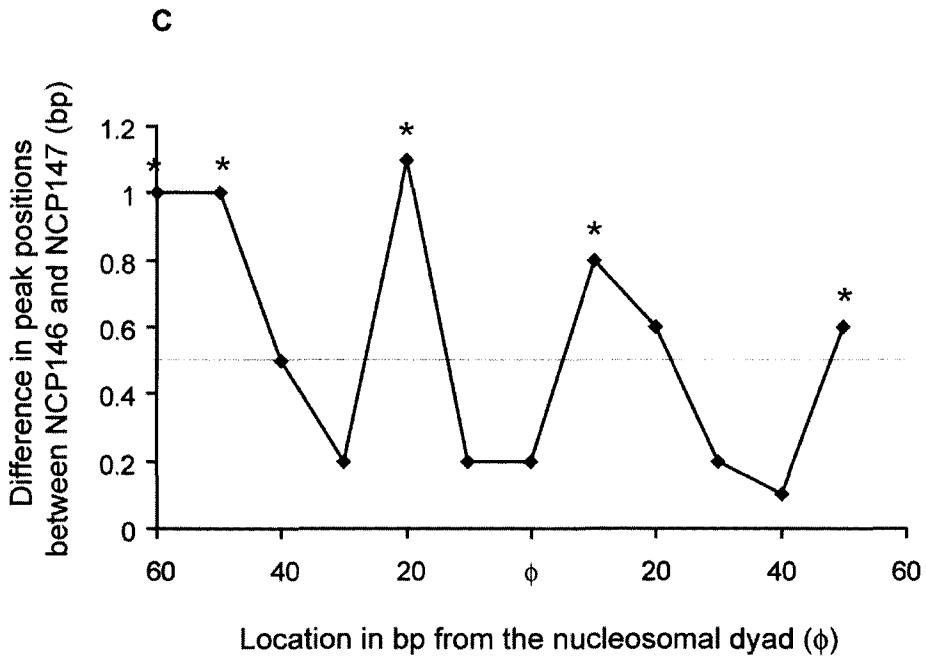


Figure 4.5 (A) Hydroxyl radical footprint comparing NCP146 and NCP147. Lane 1 corresponds to 146 bp DNA; lanes 2, 3, NCP146; lanes 4, 5, NCP147; and lane 6, 147 bp DNA. NCP were reconstituted at 0 °C (lanes 2, 4) by stepwise dilution, which results in an off-centred position octamer, and then incubated at 37° C (lanes 3, 5) for 2 h prior to hydroxyl radical treatment to allow the octamer to adopt a central position on the DNA. G+A sequencing ladders for the 146 bp and 147 bp DNA fragments are shown in lanes 1 and 6, respectively. The products were electrophoresed on an 8 % sequencing gel for 1.75 h (left panel) and subsequently for 2.75 h (right panel), and phosphorimages are shown. The nucleotide positions from the dyad (ϕ) are shown alongside the gel. (B) The quantification of the bands for NCP146 (lane 3) and NCP147 (lane 5) is plotted, derived from the phosphorimage using ImageQuant. (C) The differences in location of the corresponding peaks of hydroxyl radical cleavages between NCP146 and NCP147 are shown for both sides of the dyad bp (maxima indicated by *). The major band at the dyad corresponds to a small amount of singly nicked molecules (those that have not been ligated on both strands). Thus, these bands differ by one base between NCP146 and NCP147, and are not included in the densitometric traces or in our analysis of the data.

this would give rise to some prominent footprinting differences between NCP146 and NCP147. Since the 146 bp DNA and 147 bp DNA differ in length by 1bp, the hydroxyl radical footprints for NCP146 and NCP147 are expected to have a one-bp offset beyond the dyad. It is difficult to interpret the data by simply looking at the gel and therefore, the data were mathematically analyzed by densitometry analysis of the hydroxyl radical footprints for NCP146 and NCP147 in Figure 4.5B, after adjusting for the one-bp offset. As reported earlier (Gottesfeld et al., 2001), the scan reveals the 10 base cleavage periodicity that is characteristic of a rotationally positioned nucleosome on a unique DNA sequence for both NCP146 and NCP147 (Figure 4.5B). Likewise, to see the regions of maximal differences between NCP146 and NCP147, we plotted the differences in location of the corresponding peaks of hydroxyl radical cleavages between NCP146 and NCP147 in Figure 4.5C after adjusting for the one-bp offset. In the regions where the NCP146 and NCP147 structures are identical, the hydroxyl radical peak locations for the two structures will coincide and the differences in peak locations will be zero. In regions where there are structural differences over and above the one-bp offset, as shown in Figure 4.5C, the peak positions will not coincide and the differences will be non-zero. As seen before with the DNase I footprints, the largest differences in the relative positions of the peak centers are found ~ 20 bp and ~ 45-50 bp on either side of the dyad (see * in Figure 4.5C), supporting the notion that several populations of NCP146 exist in solution, which differ from NCP147 mostly in the 20 bp and 45 bp regions. Again, the distribution of NCP146 population is heterogeneous in solution as revealed by smaller differences in additional regions of the nucleosomal DNA. Taken together, the affinity measurements described above and the footprinting results rule out the possibility that stretching is ab-

sent in NCP146 in solution. Our results also provide evidence supporting a heterogeneous mixture of various stretched or twist-defect states of NCP146 in solution.

4.5 Discussion

The energetic stability and position of a nucleosome depend on how well individual nucleosome positioning signals in the DNA sequence are accommodated on the surface of the histone octamer (Thastrom et al., 2004). With the exception of simple repeating DNA sequences, unique nucleosome positioning is likely to be a compromise to accommodate several 'conflicting interests', since it is unlikely that any given nucleosome position would place all positioning signals in a perfect orientation. Here, we provide evidence by x-ray crystallography and DNA footprinting experiments that even nucleosomes reconstituted with a strong positioning sequence exist in several twist-defect states in solution. These states differ by the presence and location of a so-called twist defect that allows, within limits, the optimization of the location of DNA positioning signals. Furthermore, DNase I footprinting reveals that a bivalent Py-Im PA clamp (which effectively crossbraces two gyres of nucleosomal DNA) favors binding to nucleosomes compared to free DNA when the optimization of positioning signals generates a DNA conformation that is favorable for clamp binding.

The recently described polyamide clamp (Edayathumangalam et al., 2004) has provided unexpected insight into the dynamic behavior of nucleosomes in solution. First, the crystal structure of NCP146-clamp as well as affinity measurements (increased k_{on} and K_a of the clamp for NCP147 compared to NCP146) suggest that 'stretching' of the DNA by one base pair is a prerequisite for optimal binding of the clamp to NCP146 in solution,

since it places the PEG linker that connects the two polyamide modules in the clamp in a sterically allowed conformation. Second, DNase I and hydroxyl radical footprints of NCP146 and NCP147 (in the absence of ligand) reveal major differences at the two locations (~ 20 bp and ~ 45 bp from the dyad) where a twist defect is trapped in the various NCP146 crystal structures (Muthurajan et al., 2003). These differences and changes in additional locations likely stem from a mixture of various twist-defect states of NCP146 in solution. Therefore, in addition to indicating structural differences between NCP146 and NCP147 in solution, our results strongly suggest positional heterogeneity within the population of NCP146 in solution. Lastly, binding of the clamp to NCP146 selects and stabilizes a conformation of nucleosomes that is very similar to NCP147, resulting in similar DNase footprints for NCP147 and liganded NCP146. In contrast, clamp binding to NCP147 in solution has little effect on DNA structure. The driving force for the existence of twist-defect states in solution is likely to be the need to optimize the location of the DNA sequence-specific positioning signals with respect to the surface of the histone octamer. This may partially counterbalance the energetic penalty that must exist for overwinding and stretching the DNA. In comparison to the crystal structures where a discrete twist-defect state is captured, the footprinting data suggest that twist-defect states may not be homogeneously distributed in nucleosomes in solution. DNA parameters that are influenced mostly by the conformation of nucleosomal DNA, as determined by the structural context of the histone octamer, predispose the regions around 20 and 45 base pairs removed from the nucleosomal dyad for stretching in the crystal structures (Muthurajan et al., 2003). The ~ 20 bp stretching location is the only site observed in all unliganded NCP structures containing the identical canonical 146 bp sequence. In con-

trast, stretching in the ~ 45 bp region is observed in NCP146-PA1 (Suto et al., 2003) and in the structure of a nucleosome containing a 146 bp sequence different from that in NCP146 (Davey et al., 2002), suggesting that this is an alternative region for the dissipation of twist diffusion. In cases where the ~ 20 bp region is chosen as the stretch site, it appears that the system is driven towards minimizing the region in which sequence and structure are out of register. Interestingly, it was recently shown that one of the two major contacts made by the ISW2 chromatin-remodeling complex with the nucleosome also occur at ~ 20 bp from the nucleosomal dyad (Kagalwala et al., 2004). It is tempting to speculate that perhaps the specific properties imparted on the DNA in this particular region by the histone octamer may be pertinent to ISW2 function (Muthurajan et al., 2003).

Based on the crystallographic and footprinting results, we propose a model for a dynamic equilibrium between the multiple twist-defect states of NCP146 in solution (Figure 4.6). In contrast to the heterogeneous distribution of twist-defect states in solution, only one discrete state is observed or 'trapped' in the various crystal structures (Muthurajan et al., 2003). For example, in state II, stretching close to the dyad (indicated by ◄) positions clamp sites optimally. In our model, stabilization of state II upon clamp binding in solution shifts the equilibrium towards state II allowing all NCP146 molecules to be eventually bound by the clamp in solution. Consequently, state II is observed in the crystal structure of NCP146-clamp. Binding of the parent polyamide PA1 to nucleosomes in solution precludes migration of the twist defect to the most favorable site (indicated by ►), resulting in state I in the NCP146-PA1 crystal (Figure 4.6) (Suto et al., 2003). Twist diffusion mechanisms may allow spontaneous interconversion of one state to the other, as

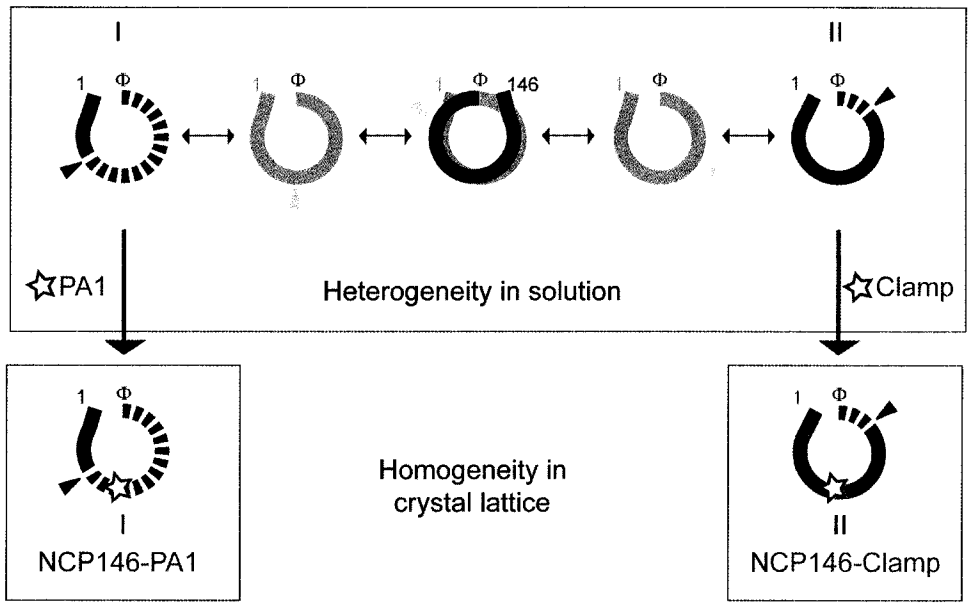


Figure 4.6 Schematic of nucleosome twist-defect intermediates observed in solution and in crystal structures. Only one half of the DNA is shown in most cases for clarity. DNA ends are numbered and the nucleosomal dyad axis is indicated by ϕ. The histone octamer has been omitted. (Top panel) NCP146 exists in solution as a mixture of different twist-defect conformations. The various locations of the twist defects are indicated schematically by ◀, ◀, and ◀. States I and II are observed in solution and in the crystal structures. Others are hypothetical states in solution. (Bottom panel) Only one of the many conformations in solution is captured in the crystal lattice. States I and II are observed in the structures of NCP146-PA1 and NCP146-clamp, respectively.

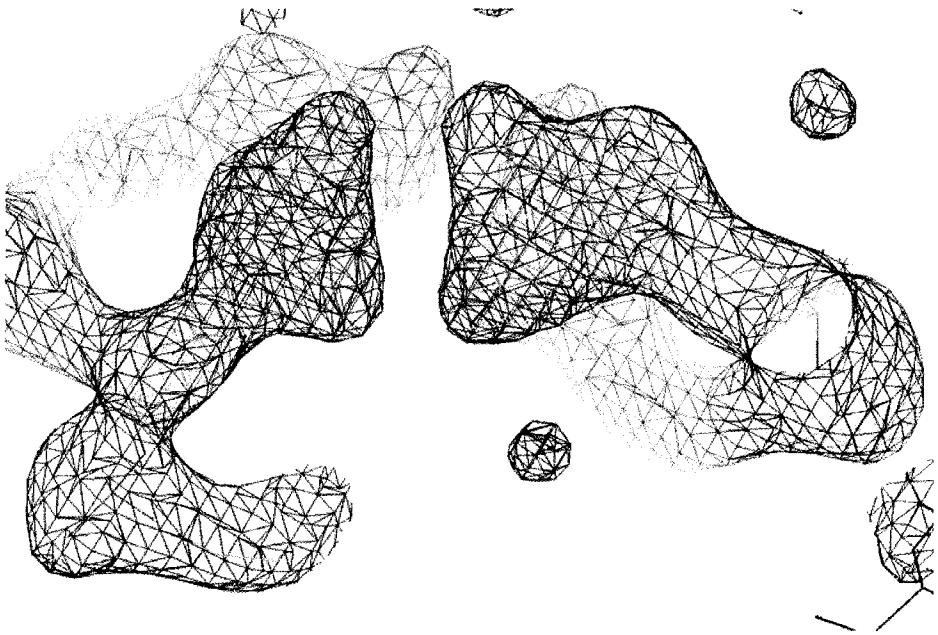
outlined in Figure 4.6 (top panel). If DNA stretching or twist diffusion results in a dynamic equilibrium in solution, it may be a means to altering nucleosomal DNA accessibility in the absence of ATP hydrolysis and/or chromatin-remodeling factors.

In conclusion, our results are relevant to the biological function of nucleosomes and indicate the occurrence of a heterogeneous population of twist-defect nucleosomal states in solution. Our results also suggest twist diffusion as a plausible mechanism for establishing a dynamic equilibrium between the various twist-defect states in solution. Additionally, twist diffusion mechanisms may also allow nucleosomes to adjust to a certain extent to optimize binding to suboptimal positioning sequences (or to better accommodate sequence-specific protein factors). Whether this is indeed the case remains to be seen. This intrinsic conformational flexibility of nucleosomal DNA may facilitate molecular recognition *in vivo* to allow transcription factor binding, chromatin remodeling, nucleosome positioning, and chromatin compaction, and may thus play an important role in chromatin structure and function.

4.6 Acknowledgements

We thank Kevin Sullivan for joining us in developing the original idea of the nucleosome clamp; Pamela Dyer for help with reagents; and all Luger lab members for help and discussion. The polyamides were synthesized by Philipp Weyermann in the Dervan laboratory and the footprinting experiments were done by Joel Gottesfeld. We also thank Andrew Flaus for critical reading of the manuscript and for helpful suggestions. This work was supported by a grant supplement GM57148 from the US

National Institutes of Health (NIH) to K.L., J.M.G., and P.B.D, and by NIH grant GM61909 to K.L.



Chapter 5

Ligand Binding to Tailless H4 NCP Leads to Unwrapping of the DNA Ends

Edayathumangalam, R. S., Bonanno, J. B., Melander, C., Dervan, P. B., Adams, P. D., and Luger, K.

This chapter is work in progress and a manuscript in preparation.

5.1 Abstract

The histone N-terminal tails constitute about 28 % of the total protein mass in the nucleosome. Over two-thirds of the histone tails is disordered in the nucleosome crystals. In spite of numerous examples showing the involvement of histone N-terminal tails in gene regulation and higher-order chromatin organization, the precise mechanism is unknown. The role of the histone H4 N-terminal tail in regulating nucleosome accessibility and dynamics was investigated. The structure of the tailless H4 nucleosome does not reveal any differences *per se* when compared to that of the nucleosome containing full-length histone H4. Next, tailless H4 nucleosomes were crystallized in complex with polyamide 6, which has been shown to bind nucleosomes only when either the H3 or H4 tail is removed. The co-crystal structure of the tailless H4 nucleosome-polyamide 6 complex reveals that the DNA ends are unraveled in order to accommodate the bound polyamide. Our results strongly suggest that H4 tail removal makes the nucleosome more amenable to structural rearrangements, which in turn expose previously inaccessible nucleosomal sites for ligand binding.

5.2 Introduction

The x-ray structure of the nucleosome core particle shows that the histone N-terminal tails are predominantly unstructured and extend outside the confines of the nucleosome core (Luger et al., 1997a). Although the histone sequences are highly conserved throughout evolution, the histone H3 and H4 have redundant functions in chromatin assembly and are individually dispensable for viability in the yeast, *Saccharomyces cerevisiae* (Kayne et al., 1988; Ling et al., 1996; Wells and Brown, 1991). The H2A and H2B

N-terminals tails, as with H3 and H4, are essential but with redundant functions in yeast (Schuster et al., 1986). The exact mechanism by which the flexible histone N-terminal tail domains function in gene regulation is not known although multiple roles have been attributed to the tail domains. The tail domains are targets for a series of posttranslational modifications, which affect the recruitment of non-histone regulatory factors that establish either active or repressive chromatin regions. The reader is referred to a review by Strahl and Allis to examine in detail the effects of histone tail acetylation, phosphorylation, methylation, ADP-ribosylation, and ubiquitination on gene expression (Strahl and Allis, 2000). Again, the mechanistic details of how tail modifications impact its function are poorly understood.

The flexible tails have been known to function in stabilizing higher-order chromatin folding (Allan et al., 1982), presumably by binding to DNA (or histones) on neighboring nucleosomes and/or by interacting with the linker histone within the chromatin fiber. Although the precise structure of the 30 nm chromatin fiber is the subject of continued debate, it is becoming increasingly evident that the histone tails play a role in controlling the dynamics of the chromatin fiber (reviewed extensively in Hansen, 2002). Numerous solution studies show that the histone tails participate in interactions with the DNA of nucleosome core particles under low salt and physiological salt conditions. For example, model studies show that the H4 tail peptide has a strong interaction with DNA, with an apparent K_d of $\sim 10^{-9}$ M in physiological salts. This interaction is dramatically reduced by several orders of magnitude when the same peptide is acetylated at the lysine residues (Hong et al., 1993). The relevance of this observation to tail function still remains to be

elucidated. Several groups have demonstrated that tailless nucleosomal arrays are unable to form the intramolecular maximally folded 55S structure or the intermolecular oligomers in solution ((Hansen, 2002) and references therein), even in several hundred millimolar divalent cation concentrations. This indicates that the tails influence chromatin dynamics in part through non-electrostatic mechanisms and that unknown functions of the histone tail domains mediate chromatin condensation. Richmond and colleagues recently showed that the H4 N-terminal tail is critically required for full fiber compaction (Dorigo et al., 2003). The data from the above studies support the notion that the H4 tail is the most important for chromatin compaction in the absence of any other factors, whereas the H2A, H2B, and H3 perhaps contribute to some other processes that impact chromatin biology.

At the nucleosome level, the tails affect nucleosome dynamics, which in turn affect DNA accessibility within individual nucleosomes. Recent studies have found that deletion of the tail domains leads to a 1.5- to 14-fold increase in the equilibrium accessibility of DNA target sites buried inside nucleosomes (Polach et al., 2000). Three different assays were used: equilibrium binding by the site-specific DNA-binding protein Gal4, and dynamic DNA accessibility to restriction enzymes and to endonucleases. The effects of deletion of the histone N-terminal tails were studied within the framework of the 'site exposure' model. In the site exposure model, localized transient dissociation and rebinding of the nucleosomal DNA to the histone octamer are thought to occur. The equilibrium constants for these local DNA unwrapping events and transient site exposure are ~ 10^{-2} and ~ 10^{-5} for sites located near the periphery and at the nucleosomal dyad, respectively.

These studies show that a modest increase (1.5- to 14-fold) in the time-averaged accessibility of nucleosomal sites was a result of changes in the distribution of nucleosome positions upon tail deletion. In another study examining the role of histone tails in determining access to nucleosomal target sites, Vitolo and colleagues reported that the H3 and H4 N-terminal tails modulate the binding of transcription factors to nucleosomal DNA (Vitolo et al., 2000). The authors showed that reconstitution of the 5S DNA fragment into a nucleosome restricts the binding of transcription factor, TFIIIA to its target nucleosomal site. Removal of all histone N-terminal tails relieved this inhibition and allowed high-affinity binding of TFIIIA to nucleosomal DNA. Further characterization showed that the H3 and H4 tails, but not H2A and/or H2B, directly modulate the ability of TFIIIA to bind nucleosomal DNA.

The effect of the H3 and H4 tails on polyamide binding to the NCP has been studied before (Gottesfeld et al., 2001). The binding sites for polyamides 5 and 6, which are located near the nucleosomal dyad, are inaccessible in nucleosomes containing full-length H3 and H4. Interestingly, removal of either tail allows polyamide binding. Since removal of either tail (H3 or H4) allows ligand binding, the effect of tail removal is expected to be indirect rather than from a direct competition between the histone tails and the polyamides. These studies also showed that binding of polyamide 5 to tailless H3 or tailless nucleosomes led to a dramatic change in the DNase I digestion profile. These results strongly suggest that removal of either histone tail allows the DNA to adopt alternative rotational positions, which in turn enables ligand binding. The crystal structure of the NCP reveals that the polyamide-binding site faces the histone octamer, and the minor

groove is occluded by a short 3_{10} helix of H4. The structure also suggests that the binding site might participate in interactions with the H4 N-terminal tail.

The footprinting results (Gottesfeld et al., 2001), in accordance with similar reports from other groups (Polach et al., 2000; Vitolo et al., 2000), strongly implicate the histone H3/H4 tails in regulating DNA accessibility within nucleosomes. However, the structural and mechanistic basis for these observations is currently unclear. The current studies focus on elucidating the structural basis for the role of the histone H3 and H4 N-terminal tails as determinants of nucleosomal DNA accessibility. Nucleosomes containing either tailless H3 (residues 27 to 135) or tailless H4 (residues 20 to 102) were crystallized in the presence and absence of ligand (polyamides 5 or 6). Tailless H3 nucleosomes (both in the absence and presence of ligand) did not yield well diffracting crystals. Hence, our focus shifted to characterizing tailless H4 nucleosomes in the absence and presence of bound polyamides.

We report that ligand binding leads to the capture of a nucleosome conformation in the crystal with unraveled DNA ends in order to accommodate the bound ligand. Although we do not see structural changes in the case of the tailless H4-containing NCP in the absence of the ligand, we hypothesize that tail removal in itself leads to unraveling of DNA ends. We suggest that these dynamic changes are not stabilized in the absence of a bound ligand. On the other hand, ligand binding appears to stabilize this unraveled state of the NCP in solution and subsequently in the crystal. Our results show that the H4 tail plays an important role in regulating DNA dynamics and accessibility within the nucleosome.

5.3 Materials and Methods

5.3a Co-crystallization of NCP-polyamide complexes

Previously established protocols were used to reconstitute NCP from recombinant *Xenopus laevis* full-length and tailless histones (tailless H3 (residues 27 to 135) or tailless H4 (residues 20 to 102)) and a 146 bp DNA fragment derived from human α-satellite DNA (Davey et al., 2002; Dyer et al., 2004; Luger et al., 1999b). The tailless histones were initially defined by trypsin digestion of full-length histones (Luger et al., 1997b and references therein). The expression of recombinant tailless histones has been previously described (Luger et al., 1997b; Luger et al., 1999a). Polyamides were synthesized as described (Baird and Dervan, 1996; Weyermann and Dervan, 2002). The purity and identity of the polyamides were established by analytical HPLC and matrix-assisted laser desorption ionization-time-of-flight mass spectroscopy. Polyamides were purified by preparative HPLC, lyophilized, and stored at 4 °C. Each polyamide was freshly dissolved in 20 mM potassium cacodylate (pH 6.0) and 1 mM EDTA, and its concentration was determined by measuring UV absorbance and using empirically determined extinction coefficients (ε_{316} = 68,800). Tailless H3 or tailless H4 NCP (at 40-50 μM) was incubated with a 10-fold molar excess of the polyamide in solution at ambient temperature for 45 min. The integrity of the nucleosome preparation after incubation with the ligand was checked by electrophoretic mobility-shift assay (Luger et al., 1999b). Crystals of tailless H3 NCP-polyamide and tailless H4 NCP-polyamide complexes were grown in 1-2 weeks in 40-45 mM $MnCl_2$, 35-38 mM KCl, and 20 mM potassium cacodylate (pH 6.0) containing ~ 20 μM (~ 4mg/mL) NCP146 at 19 °C by using vapor diffusion. Crystals for

tailless H4 NCP (in the absence and presence of ligand) were harvested and flash-cooled as previously described (Luger et al., 1997a).

5.3b Solution labeling of tailless H4 NCP with tetrakis(acetoxymercuri)-methane (TAMM)

In order to label tailless H4 NCP in solution with TAMM, a cysteine replacement mutant of H2B, H2B T112C (site for covalent attachment of TAMM) was chosen. The residue C110 in histone H3 has been previously shown to be inaccessible for modification following soaking for a few hours in NCP in solution (Luger et al., 1997a). A histone mutant (H2B T112C) was used for covalent modification of NCP in solution. The tailless H4 NCP used for solution labeling with TAMM contained the following recombinant *Xenopus* histones: H2A, H2B T112C, H3 C110A (to rule out any problems with partial modification of this site), and tailless H4. Large-scale preparation of the tailless H4 NCP was done using previously established methods and the purified NCP in TCS buffer (20 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM DTT) was harvested for labeling at this point. The NCP was loaded onto 1 mL Sephadex G-25 (medium) spun columns (~ 0.4 mg NCP per column) and centrifuged for 10 min. at 2000 rpm in order to remove EDTA and DTT prior to TAMM addition. A stock solution of 4 mM TAMM was prepared in labeling buffer (20 mM Tris-acetate, pH 7.5). The NCP was incubated with a 6-fold molar excess of TAMM in the dark for one to two hours. The NCP labeled with TAMM was dialyzed in CCS storage buffer (20 mM K-cacodylate, pH 6.0, 1 mM EDTA) prior to incubation with polyamide and crystallization. Conditions for incubation

with polyamide and crystallization conditions for TAMM-labeled NCP were identical to that of unlabeled NCP.

5.3c Heavy atom (TAMM) soaking procedure for tailless H4 NCP crystals

The crystals of tailless H4 NCP-polyamide 6 complex were harvested in mother liquor and solid TAMM was added to the mother liquor containing the crystals. The TAMM in the mother liquor was allowed to soak into the crystals over a period of about two weeks. In this particular case, no histone point mutants are required. It has been shown earlier that although the histone H3 C110 residue is inaccessible for modification in NCP in solution, it is fairly accessible in nucleosomes when the local concentration of TAMM is kept high and upon prolonged soaking (as above) into the crystal lattice. TAMM soaking of NCP crystals leads to covalent modification of the H3 C110 residue (Luger et al., 1997a). Following the heavy atom soak, the crystals were harvested in cryoprotectant and flash-cooled in a manner identical to that for regular NCP crystals (Luger et al., 1997a).

5.3d Data collection and structure refinement

X-ray data for several crystals were collected both in-house and at beamline 8.2.2 at the ALS in Berkeley, CA. Data were collected on single crystals and processed with the HKL suite (Otwinowski and Minor, 1997). Molecular replacement (using PDB entry 1AOI as the original search model) and refinement were done with CNS (Brunger et al., 1998). Alternatively, single wavelength anomalous dispersion (SAD) phasing was done using SHARP (La Fortelle and Bricogne, 1997) followed by density modification with

non-crystallographic symmetry (NCS) averaging with DM in the CCP4 suite (Collaborative Computational Project, 1994). SAD phasing followed by density modification was also done using CNS. Initial Fourier synthesis maps were calculated both in CNS and using FFT from the CCP4 suite. Details of the above procedures are discussed below.

5.4 Results and Discussion

5.4a Crystals of tailless H3 NCP and tailless H4 NCP

Both tailless H3 NCP and tailless H4 NCP were crystallized in the unliganded form as well as in complex with polyamides 5 and 6. The six types of NCP combinations yielded crystals under the same conditions as for regular NCP. However, the crystal sizes and morphologies were very different from that of regular NCP and from each other (Figure 5.1). Of all these complexes, the crystals of tailless H4 and tailless H4-polyamide 6 complex alone diffracted better than 4 Å (Figure 5.2) and therefore structure determination was pursued in these two cases.

5.4b The crystal structure of tailless H4 NCP is identical to regular NCP

High-resolution data to 2.3 Å were obtained for tailless H4 NCP crystals at beam-line 5.0.2 at ALS, Berkeley, CA. The tailless H4 NCP crystallized in the same space group with the same unit cell parameters as regular NCP (Table 5.1). The structure was phased using molecular replacement and the refinement statistics for the final model are shown in Table 5.1. The crystal packing interactions in tailless H4 NCP are identical to that of regular NCP. The refined structure of tailless H4 NCP was compared with the

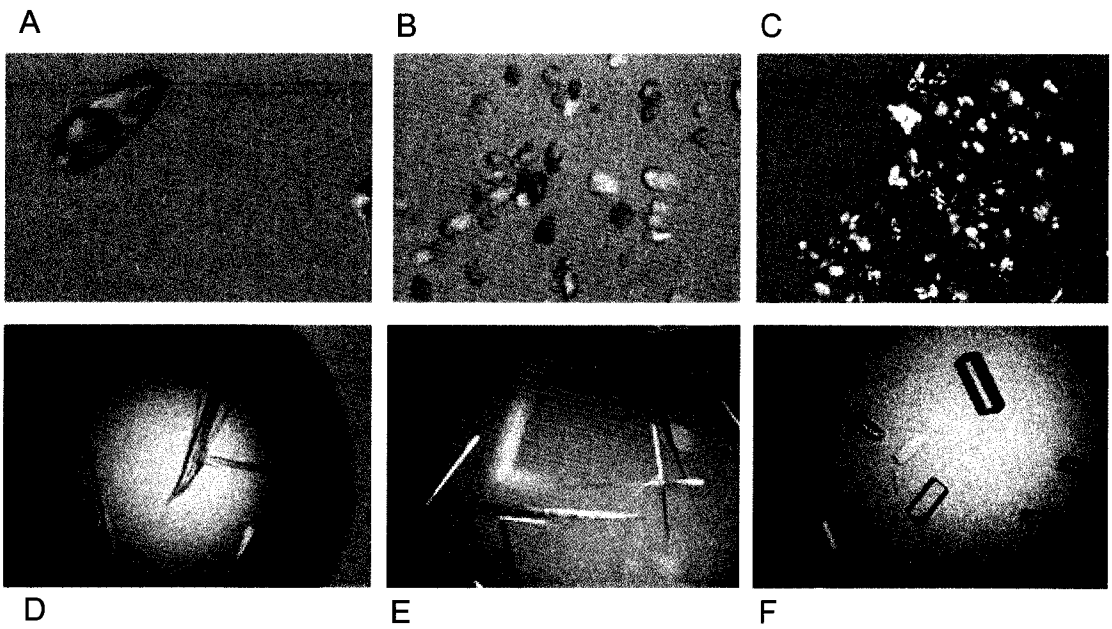
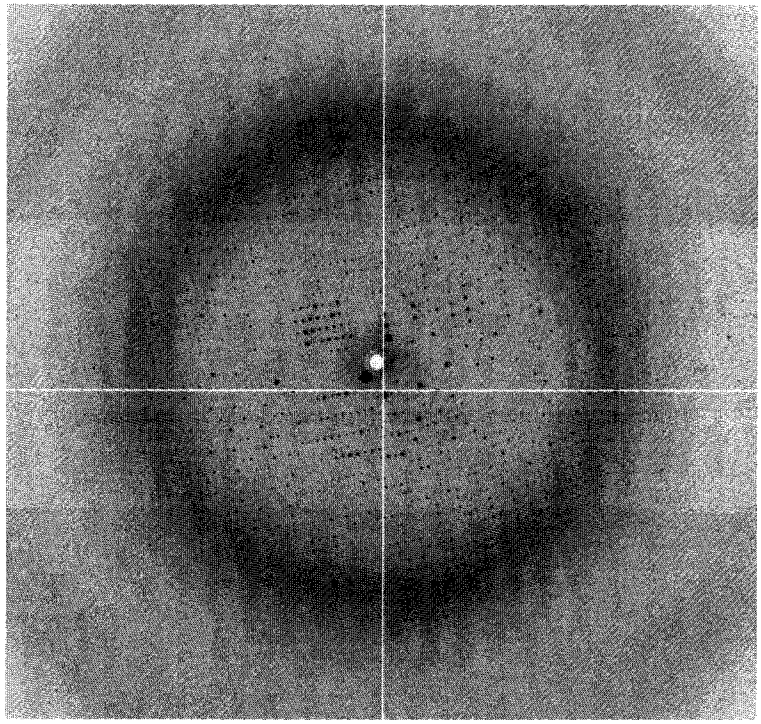


Figure 5.1 Crystals of NCP containing tailless histones. (A-C) Crystals of tailless H3 NCP. (D-F) Crystals of tailless H4 NCP. Crystals contain either NCP alone (A, D), NCP with polyamide 5 (B, E), or NCP with polyamide 6 (C, F).

A



B

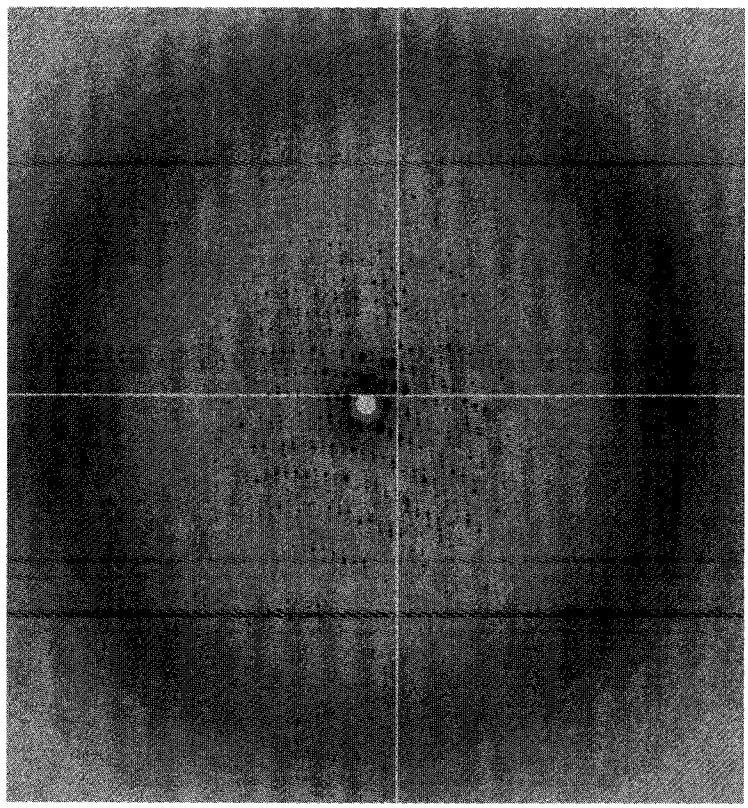


Figure 5.2 X-ray diffraction patterns collected from crystals of tailless H4 NCP without (A) and with (B) polyamide 6. X-ray data collected at -180°C with an oscillation width of 0.5 ° per frame. Resolution limit at the edge is ~ 2.2 Å (A) and ~ 3.5 Å (B). Data was collected at beam-line 5.0.2 at the Advanced Light Source (Berkeley, CA).

regular NCP structure following least-squares superposition of the two structures using LSQMAN from the Uppsala Software Factory (Kleywegt, 1996). The comparison revealed no significant differences between the two structures (data not shown), indicating that changes to the nucleosome structure following H4 tail removal, if any, are not stabilized in the nucleosome crystals (also see Section 5.4f).

5.4c The crystals of tailless H4 NCP-polyamide 6 crack in cryoprotectant containing 2-methyl-2,4-pentanediol (MPD)

The crystals of tailless H4 NCP-polyamide 6 complex were grown in conditions similar to that of the regular NCP and tailless H4 NCP (without ligand). The crystals belong to a different space group ($P2_1$) as compared to that of regular NCP and tailless H4 NCP ($P2_12_12_1$). Also, the loss of symmetry is accompanied by a doubling of one of the cell axes (*b*-axis of tailless H4 NCP-polyamide 6 is twice *a*-axis of tailless H4 NCP, compare Tables 5.1 and 5.3) leading to a doubling of the cell volume. Matthews coefficient calculations (Matthews, 1968; Table 5.2) indicated the presence of four molecules in the asymmetric unit (ASU). These crystals routinely diffract to ~ 8 Å and one in every 30-40 crystals diffracts to ~ 3.5 Å. However, all crystals of the tailless NCP-polyamide 6 complex develop visible cracks following ~ 4 days of soaking in the cryoprotectant containing 24 %(*w*/*v*) final MPD (Figure 5.3). Many alternative approaches were tried to prevent crystal cracking: 1) addition of different concentrations of MPD to the crystallization drop and the reservoir during crystallization, 2) varying the final concentrations of MPD in the cryoprotectant and soaking times, and 3) screening for alternative cryoprotectants (including various oils, PEGs and sugars like sucrose). All

Table 5.1 Data-collection and refinement statistics

Dataset	Tailless H4 NCP ALS BL5.0.2
Data Collection	
Mosaicity	0.5
Space group	$P2_12_12_1$
Unit-cell lengths	106.4
(*a, b, c* in Å)	109.9
	182.1
Resolution range (Å)†	99-2.30
Unique reflections	82273
Completeness (%)	85.9 (75.2)
Multiplicity	5.1
Average I/σ(I)	16.6 (2.3)
R-merge#	0.078 (0.405)
Refinement Statistics	
Reflections in test set	3977
R_{cryst}†	0.2112
R_{free}‡	0.2454
No. of amino acids	767
No. of DNA base pairs	292
No. of water molecules	820
No. of Mn^{2+} ions	14
No. of Cl^{-1} ions	4
rmsd	
Bonds (Å)	0.005
Angles (°)	0.992
Average B-factors	
Protein	44.7
DNA	94.2
Solvent	62.4

Numerical values in parentheses denote data for the highest resolution shell.
#R-merge = $\Sigma|I_h - \langle I_h\rangle|/\Sigma I_h$, where $\langle I_h\rangle$ is the mean of the measurements for a single *hkl*.
†$R_{cryst} = \Sigma|F_{obs} - F_{calc}|/ \Sigma F_{obs}$
‡Calculated using 5 % of the *hkl* data chosen randomly and omitted from the start of refinement.
rmsd, rms deviation from ideal geometry.

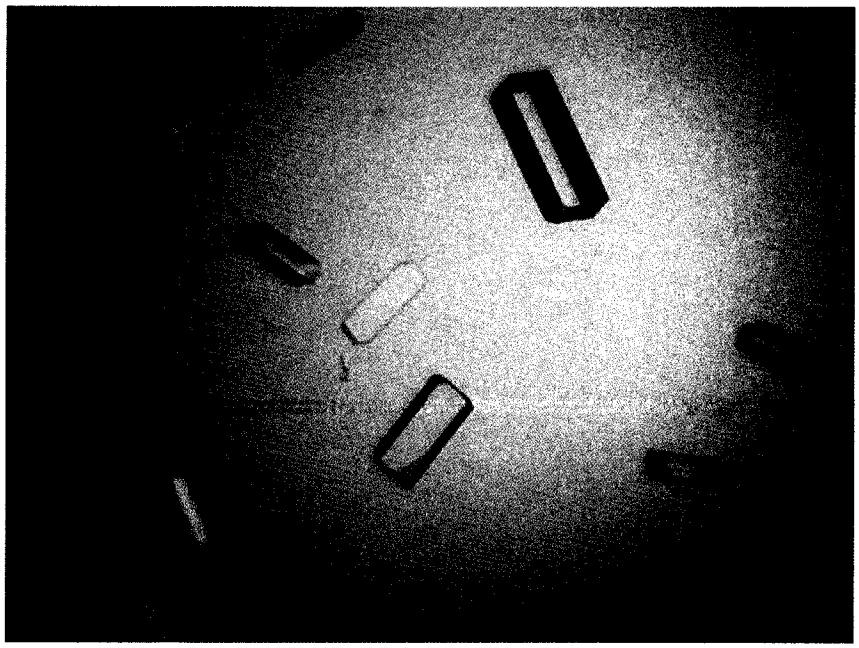
Table 5.2 Matthews coefficient calculations*

For estimated protein molecular weight 210,000		
Nmol/asym	Matthews Coeff	% solvent
1	10.3	87.9
2	5.1	75.9
3	3.4	63.8
4	2.6	51.8
5	2.1	39.7
6	1.7	27.7
7	1.5	15.6
8	1.3	3.6

*Results obtained from the CCP4 suite (Collaborative Computational Project, 1994).

these methods only worsened the diffraction limit of the crystals (not shown). Several datasets were obtained both in-house and at the synchrotron for the cracked crystals. Molecular replacement searches were done in both CNS and Molrep (CCP4 suite). However, none of these datasets yielded a molecular replacement solution in either program (not shown). Next, in the absence of any molecular replacement phases, attempts were made to obtain experimental phase information from heavy atom derivatives of these crystals. The nucleosomes were either labeled in solution with TAMM prior to crystallization, or first grown and subsequently soaked with TAMM. Both procedures have been previously described for nucleosome crystals (Luger et al., 1997a). Nucleosomes labeled in solution with TAMM yielded crystals under similar conditions to that of regular NCP crystals. However, the diffraction limit was severely reduced for these crystals (10 Å or worse). Crystals soaked with TAMM were similar to crystals without TAMM in terms of their diffraction limit. A serendipitous result was that soaking these crystals in TAMM prevented crystals from cracking when harvested in the cryoprotectant containing MPD. Also, while the crystals that did not contain TAMM suffered severe radiation damage during data collection making it difficult to collect complete datasets, the crystals soaked in TAMM had much reduced radiation damage. Although it is a common occurrence that heavy atom soaks have adverse effects on the diffraction obtained from macromolecular crystals, it is also observed, albeit less frequently, that heavy atoms can enhance crystal stability and/or diffraction. To date, it has not been possible to obtain good diffraction data for the tailless H4 NCP-polyamide 6 crystals without soaking these crystals in TAMM.

A



B

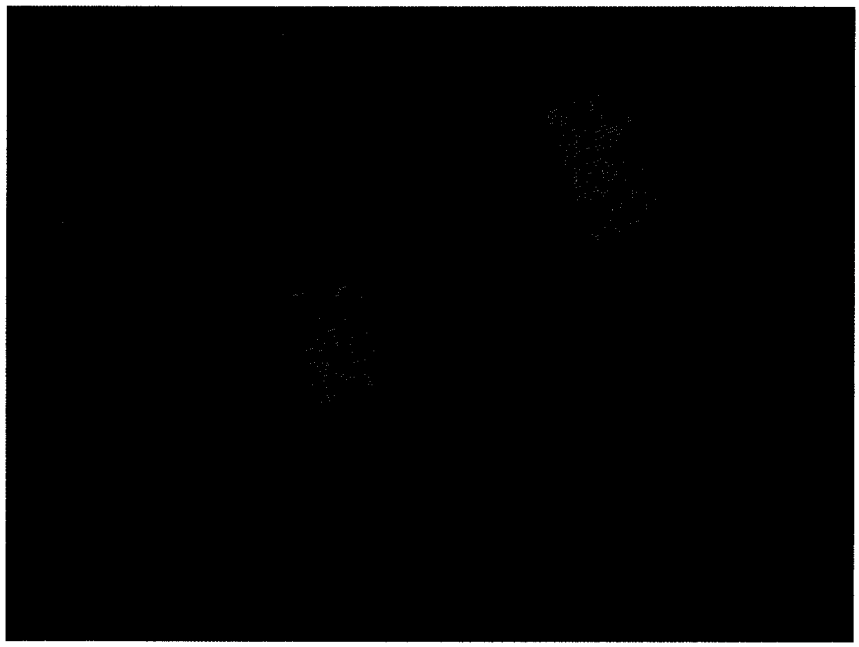


Figure 5.3 Crystals of tailless H4 NCP-polyamide 6 develop visible cracks after ~ 4 days of harvesting in cryoprotectant containing MPD. Crystals before harvesting (A) and after soaking in MPD (B).

5.4d Structure determination and calculation of minimally-biased electron density maps

Several crystals of tailless H4 NCP-polyamide 6 soaked in TAMM were screened and data were obtained at the ALS beamline 8.2.2. The data collection statistics for one of the crystals are reported in Table 5.3. Molecular replacement searches in CNS as well as Molrep led to one outstanding solution with four molecules in the ASU, as expected (Figure 5.4). In addition to obtaining a molecular replacement solution, attempts were also made to derive experimental phases from this dataset. The synchrotron data were collected at a wavelength of 1.0000 Å, which corresponds to the L-III absorption edge of the Hg atom. Attempts to locate the TAMM peak from the experimental data using the HySS module of the Phenix package (Grosse-Kunstleve and Adams, 2003) failed. However, anomalous difference Fourier calculated with the FFT program (CCP4 suite) using anomalous differences from the experimental data and the model phases (from the above molecular replacement solution) led to the identification of the positions of the anomalous scatterers in the nucleosome. Four TAMM peaks were located in the ASU with one TAMM cluster bound per NCP, which also confirms the validity of the molecular replacement solution. Since TAMM is a heavy atom cluster containing four Hg atoms, the TAMM peak was not treated as a single peak from one heavy 'superatom' but as a peak containing four Hg atoms to obtain the best initial phases. Often, heavy atom clusters can adopt many orientations in heavy atom peaks and this problem has to be resolved before meaningful phases can be obtained. In our case, fairly reliable positions of the four Hg atoms in the TAMM peak were easily determined as TAMM is involved in

Table 5.3 Data-collection and refinement statistics

Dataset	Tailless H4 NCP-polyamide 6 ALS BL8.2.2
Data Collection	
Mosaicity	0.6
Space group	$P2_1$
Unit-cell lengths	112.6
(*a*, *b*, *c* in Å)	212.8
	183.1
Resolution range (Å)†	50-3.40
Unique reflections	116510
Completeness (%)	99.6 (98.3)
Multiplicity	5.2
Average I/σ(I)	18.3 (1.8)
R-merge#	0.084 (0.720)
Refinement Statistics	
Reflections in test set	5286
R_{cryst}†	0.3197
R_{free}‡	0.3384
No. of amino acids	753
No. of DNA base pairs	272
Overall B-factor	120

Numerical values in parentheses denote data for the highest resolution shell.
#R-merge = $\Sigma|I_h - \langle I_h\rangle|/\Sigma I_h$, where $\langle I_h\rangle$ is the mean of the measurements for a single *hkl*.
†$R_{cryst} = \Sigma|F_{obs} - F_{calc}|/ \Sigma F_{obs}$
‡Calculated using 5 % of the *hkl* data chosen randomly and omitted from the start of refinement.

Figure 5.4 The crystals of tailless H4 NCP-polyamide 6 have four NCP molecules in the ASU. DNA and histone proteins are shown in stick and ribbon representation, respectively. Color code is as follow – DNA, grey; histone H2A, yellow; H2B, red; H3, blue; and H4, green.

covalent interactions with four neighboring sulphur atoms in the histone proteins (residues C110 and H113 from both histone H3 chains). The various strategies for phasing and calculation of initial Fourier synthesis maps that were attempted are summarized in Table 5.4. Two observations suggest that the experimental phases obtained from the TAMM soaks are very weak: 1) the TAMM positions in the structure could only be identified upon calculation of anomalous difference Fourier using the model phases, 2) both SAD-phased and random-phased density modification procedures yielded very similar maps. The phases obtained from the random-phased density modification were used in subsequent refinement as described below.

The following procedure (also see Table 5.4), similar to ones described previously (Chaudhry et al., 2004), was finally used to calculate minimally biased electron density maps. First, a mask was calculated around one of the four NCP molecules in the ASU. The mask was expanded and smoothed using the MAMA program from the Uppsala Software Factory (Kleywegt, 1996). NCS operators were calculated from the atomic co-ordinates of the protomers in the ASU. Random phases were then calculated for all the experimental observations and figures-of-merit (FOM) were all assigned a constant value of 0.1. The experimental amplitudes, random phases and FOMs were then used in the density modification procedure in CNS, along with the protomer mask generated from MAMA and the NCS operators to average the electron density. The protomer mask was also used to define the solvent region for solvent flipping. The basis for the above procedure is that the NCS relationship has a strong influence on the amplitudes when there is high-enough NCS in the crystal (4-fold NCS in our case). Therefore, there is essentially

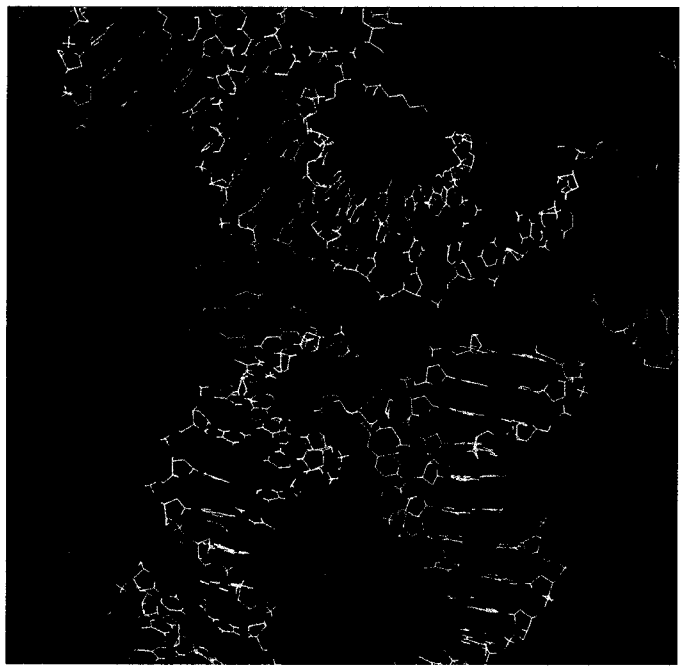
Table 5.4 Strategies for phasing of tailless H4 NCP-polyamide 6

Phasing Strategy	Outcome
1. Molecular replacement (CNS version 1.1)	Uninterpretable maps full of random noise
2. SAD phasing and density modification with four-fold NCS averaging an solvent flipping (CNS)	Interpretable initial maps
3. Random phases and density modification with four-fold NCS averaging an solvent flipping (CNS)	Maps very similar to that from strategy 2
4. SAD phasing in SHARP and density modification with four-fold NCS averaging in DM	Maps very similar to scenarios 2 and 3

only one phase set that gives rise to the observed amplitudes, given the constraints of NCS and solvent flatness. The improved phases after density modification were then used to obtain the initial set of electron density maps.

Upon visualization of the maps generated from the various phasing strategies described in Table 5.4, it was consistently observed that ~ 10 bp on either ends of the nucleosomal DNA are disordered (Figure 5.5). Also, the nucleosomal DNA near the dyad region was moved relative to the input model (Figure 5.5). Hence, 10 bp were deleted from either ends of the nucleosomal DNA such that the input model for refinement only contained 126 bp of DNA instead of the regular 146 bp. Also, residual density is visible at both ligand-binding sites (for polyamide 6) on the nucleosomal DNA (Figure 5.6). However, given the number of artifacts that are commonly encountered in low-resolution maps, the observed density cannot be interpreted in detail. Furthermore, the low-resolution maps did not permit any manual model building. Most protein and DNA backbone features of the model fit reasonably well into the map, as ascertained by visual inspection of the fit of the model to the map. Next, positional refinement and grouped B-factor refinement of the modified input model were performed using 'strict' NCS restraints (as defined in CNS), DNA-base-pairing restraints and the maximum likelihood refinement target using the density-modified phases. The refinement protocol did not make any global or local changes to the model. The refinement statistics at this point are reported in Table 5.3. It must be noted that the overall B-factor reported for the structure perhaps reflects the inherent disorder in the crystal. No further refinement could be carried out for this structure in CNS.

A



B

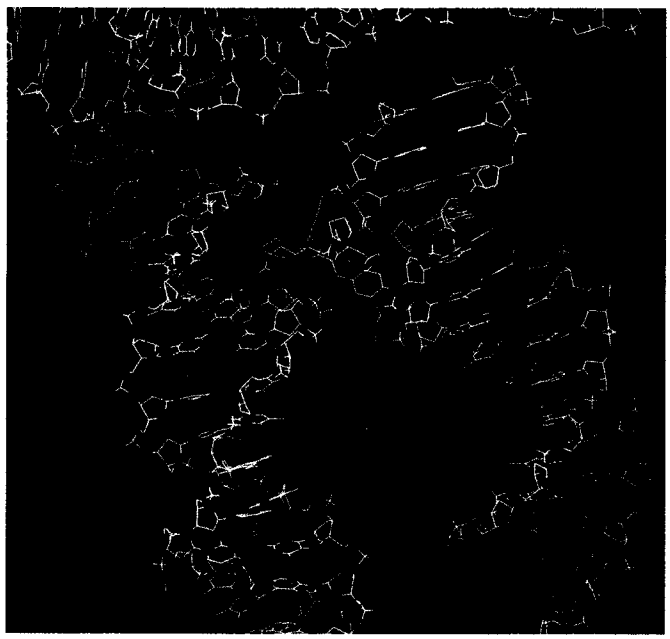
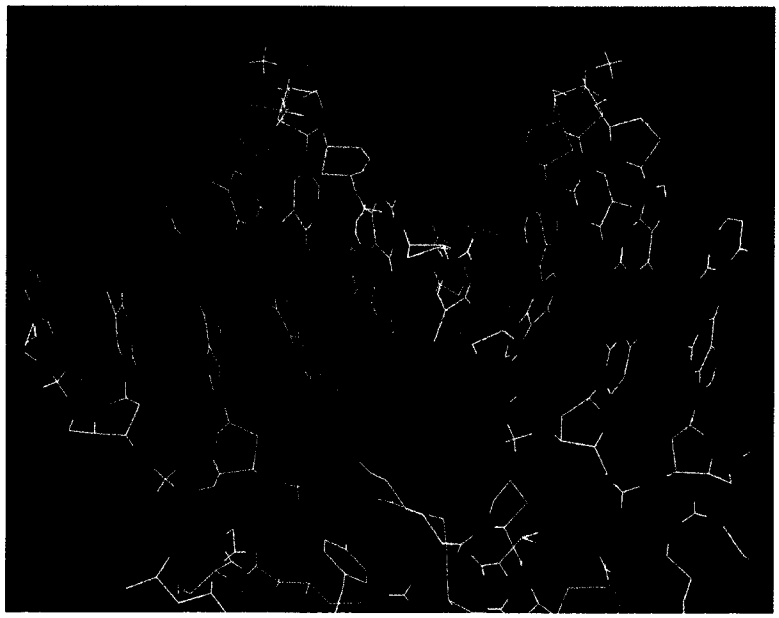


Figure 5.5 The DNA ends are disordered in the tailless H4 NCP-polyamide 6 complex. Fourier synthesis maps following random-phased density modification are shown. (A) and (B) show the two different ends of the nucleosomal DNA. Maps are contoured at 1 σ level.

A



B

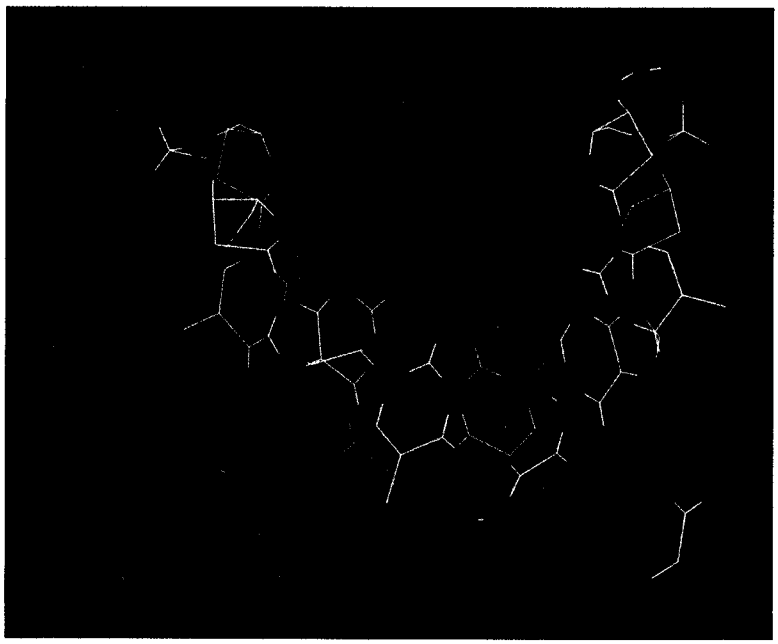


Figure 5.6 Evidence for ligand binding in the tailless H4 NCP-polyamide 6 complex. Fourier synthesis maps following random-phased density modification are shown. (A) and (B) show the two polyamide 6 binding sites on the nucleosomal DNA. Maps are contoured at 1 σ level.

As an additional test, the density-modified phases obtained following strategy 4 in Table 5.4 were used in TLS refinement in Refmac along with tight NCS restraints (Winn et al., 2003). Simply put, TLS parameters describe the translation, libration and screw-rotation of a rigid body defined by a group of atoms. TLS motion models anisotropy against x-ray data using far fewer parameters and is therefore possible even at intermediate resolutions. This is in contrast to the situation with high-resolution data, where treatment of individual atoms anisotropically using individual anisotropic displacement parameters (U's) requires a much larger number of refinement parameters. TLS refinement has been shown in several instances to improve B-factor estimates and lead to significantly improved electron density maps (Chaudhry et al., 2004). Each of the histone chains in the NCP and the two DNA chains were treated as separate rigid TLS groups. Cycles of TLS refinement were carried out using all experimental data and the amplitude-based maximum likelihood refinement target. Five cycles of TLS refinement cycles were required for convergence. After TLS refinement, positional and isotropic B-factor refinement was carried out. Again, as seen with refinement in CNS, there were no global or local changes to the model. Also, the quality of the electron density maps did not improve following TLS refinement. No further model refinement was attempted in Refmac.

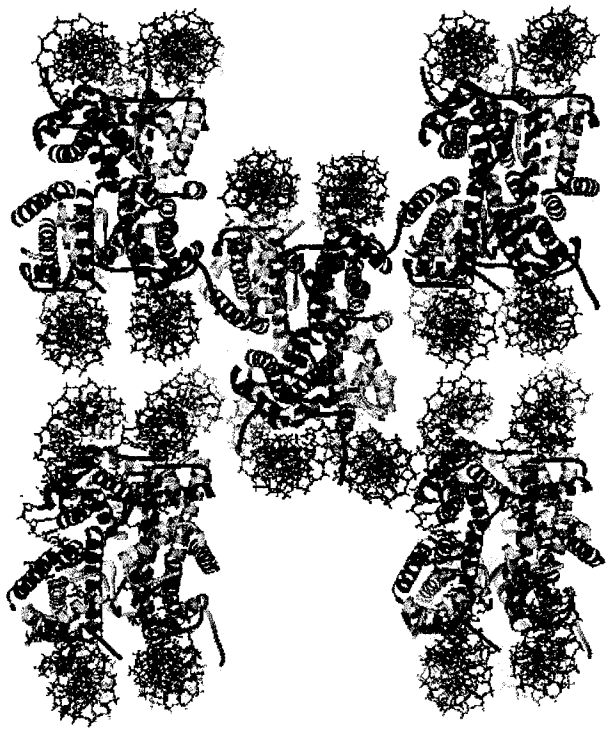
In the absence of reliable experimental phases, and given the inherent disorder and the poor diffraction quality of the crystals (partly explained by the disorder in the ends of the DNA), the current model obtained may be the best possible model. Although high-resolution features of the structure cannot be gleaned from the current model, we discuss some valid global features of the structure in the following sections.

5.4e Inter-particle interactions in tailless H4 NCP-polyamide 6 crystals

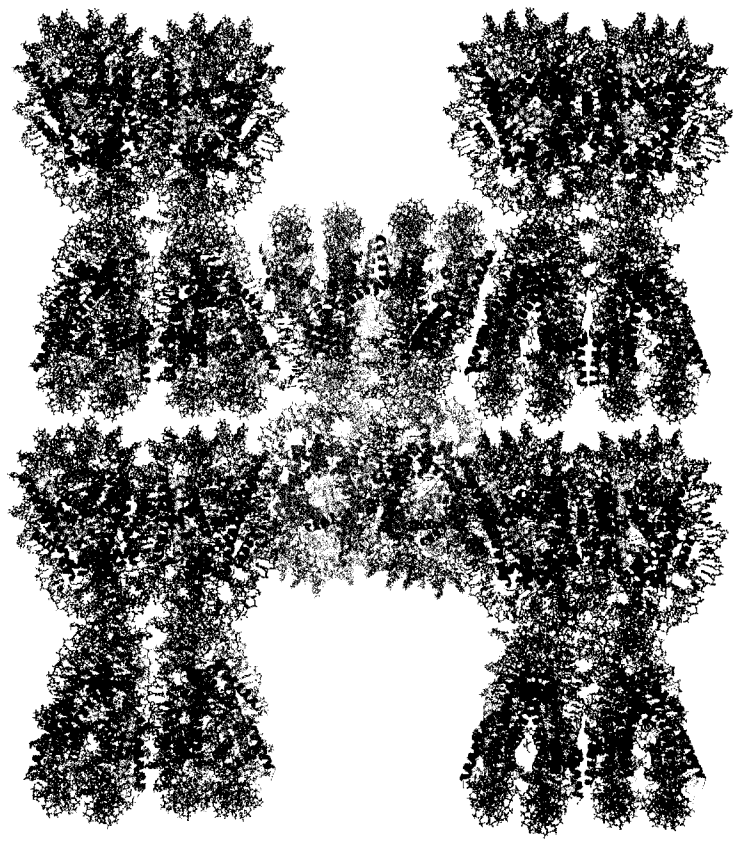
Crystal packing in regular NCP crystals utilize protein-protein and DNA-DNA interactions (Luger et al., 1997a). Tailless H4 NCP crystals exhibit an identical set of interactions as that in the regular NCP (Figure 5.7A, C; Figure 5.8 A, B). The N-terminal tail of histone H4 usually makes crystal contacts with an acidic patch in the H2A/H2B dimer of a neighboring NCP (Figure 5.7A, C). More specifically, the H4 region (basic residues between 16 to 25; K16, K19, K20, R23) make multiple hydrogen bonds and salt bridges with an 'acidic patch' formed by acidic residues from H2A (E56, E61, E64, D90, E91, E92) and H2B (E110). Also, in the case of every NCP structure published to date, the DNA ends of one NCP molecule engage in quasi-helical stacking with those from neighboring NCPs (Figure 5.8A, B). In contrast, an alternative set of protein-protein and DNA-DNA interactions drive the formation of the ASU as well as that of the crystal lattice in the tailless H4 NCP-polyamide 6 complex.

In examining the set of protein-protein interactions that drive crystallization of the tailless H4 NCP-polyamide 6 structure, several interesting observations are made. First, one set of H4 tail and (H2A/H2B) dimer interactions occur on every interface between neighboring nucleosomes (Figure 5.7C), while two such sets of interactions occur in the interface regions between two adjacent nucleosomes in the ASU (Figure 5.7D). Second, the H4 N-terminal tail interacts with the acidic patch formed by the α2 helix of H2A and αC helix of H2B in regular NCP. These interactions are replaced by a different subset of interactions in the tailless H4 NCP-polyamide 6 complex: the H3 α1 helix and the N-terminal residues of the tailless H4 of one NCP appear to make crystal contacts with

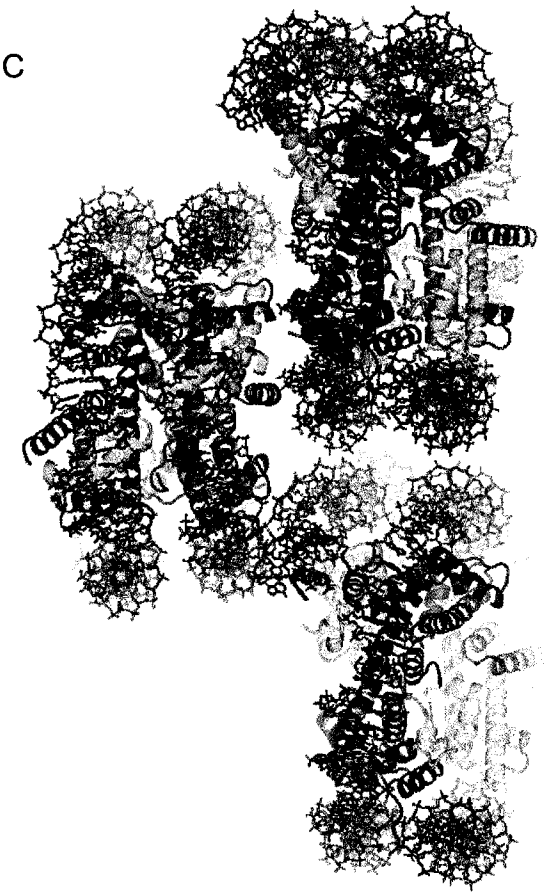
A



B



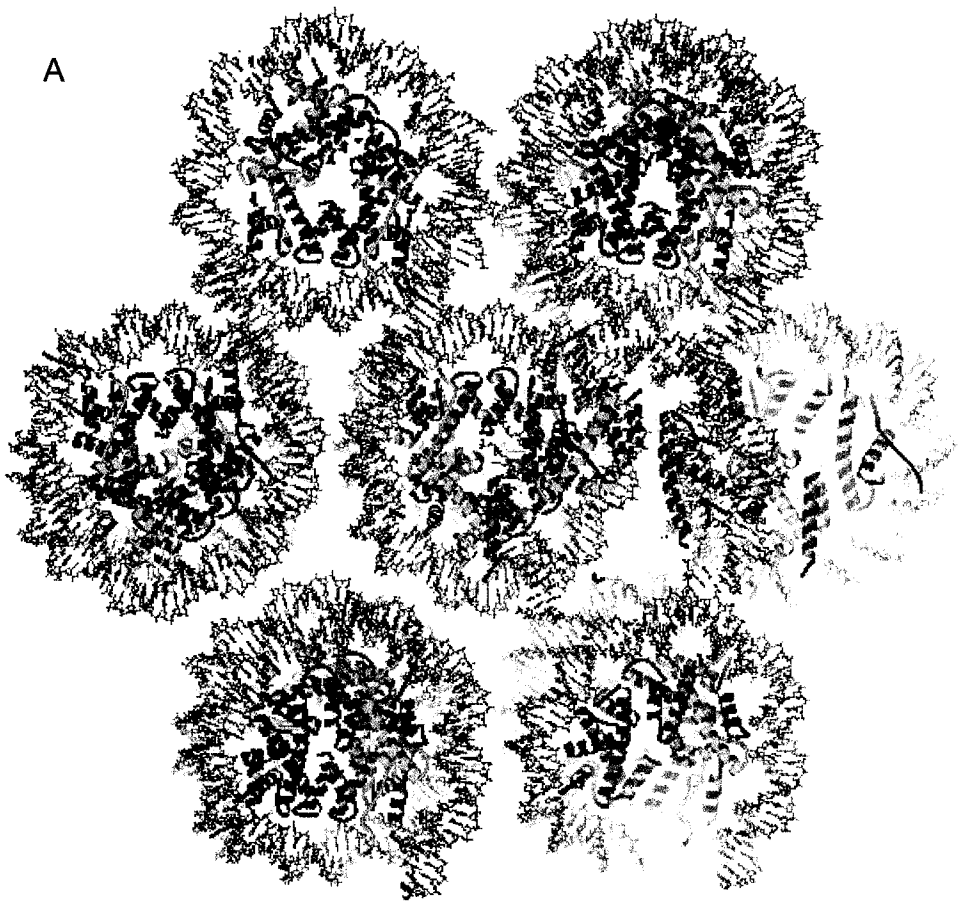
C



D

Figure 5.7 Protein-protein inter-particle contacts are altered in tailless H4 NCP-polyamide 6 crystals. Histone-histone contacts in tailless H4 NCP (A; close-up view in C) and in tailless H4 NCP-polyamide 6 (B; close-up view in D).

A



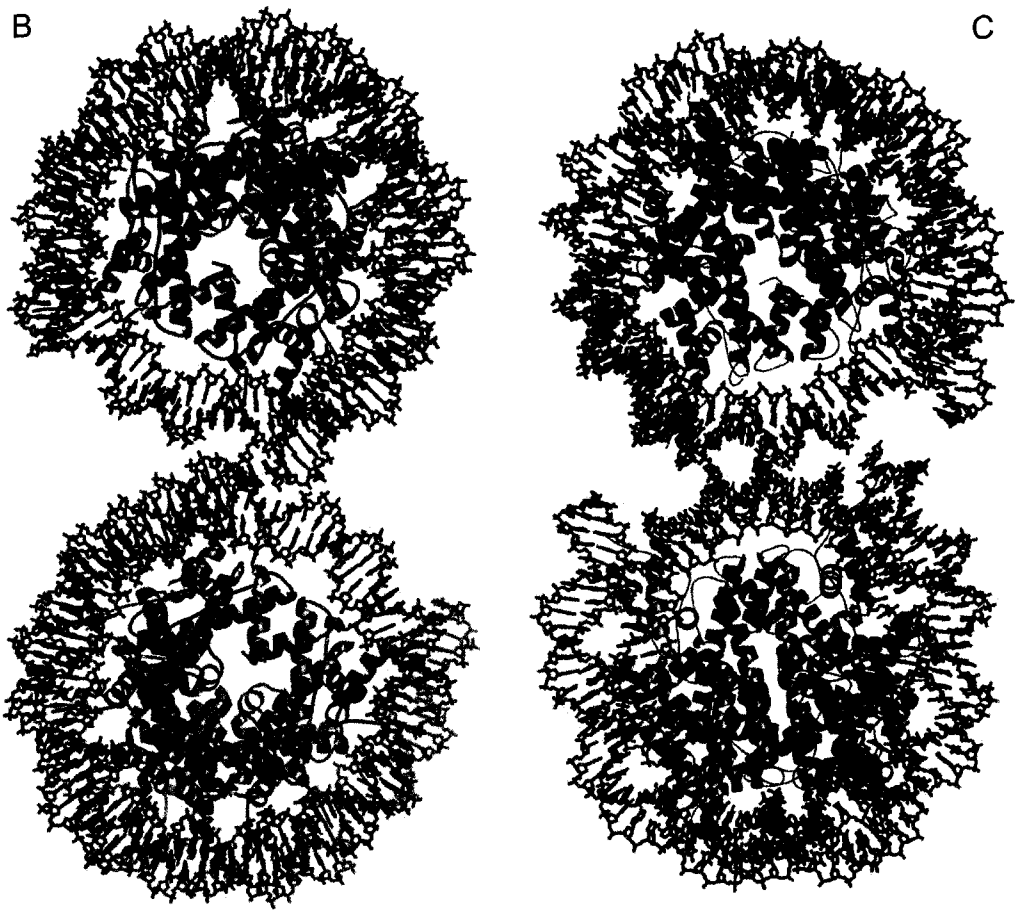


Figure 5.8 End-to-end DNA interactions are absent in the crystals of tailless H4 NCP-polyamide 6. (A, B) DNA end-to-end quasi-helical stacking in tailless H4 NCP crystals. A close-up view is shown in B. (C) Helix-in-helix stacking at the nucleosomal dyads in tailless H4 NCP-polyamide 6 crystals.

residues from the α2 helix of H2A and αC helix of H2B in the neighboring NCP. Third, the path of the residues 20 to 22 of histone H4 is highly disordered in the ligand-bound structure.

While protein-protein interactions assemble the crystal lattice along the *a*- and *b*-unit cell axes in the case of regular NCP and the tailless H4 NCP, and DNA-DNA quasi-helical base-stacking drives lattice formation along the *c*-axis, they are no longer the same for ligand-bound tailless H4 NCP. Protein-protein interactions occur along the *b*-axis while helix-in-helix stacking interactions dictate lattice formation along the other two lattice axes (for example, see Figure 5.8C). The ends (~ 10 bp) of the nucleosomal DNA are highly flexible and no longer engage in crystal contacts (Figure 5.8C). Instead, the DNA at the nucleosomal dyad of one NCP has moved and engages in a helix-in-helix stacking interaction with that from a neighboring NCP (Figure 5.8C). We report the first structure of a nucleosome with altered crystal packing arrangements that do not utilize the ends of the nucleosomal DNA. Our results suggest that the movement and unwrapping of the DNA ends is a direct consequence of histone H4 tail removal, as discussed below. Interestingly, the dyad-to-dyad internucleosomal interactions that we observe in our crystals are a common feature of the various structural models proposed to date for the 30 nm fiber, and in the recent elegant studies that showed that the chromatin fiber comprises two stacks of nucleosomes in accord with the two-start model (Dorigo et al., 2004; Van Holde, 1988). In the two-start model for the chromatin fiber, the minimal building block of the helical fiber is a row of two nucleosomes that run along the fiber long axis. The model suggests that the nucleosome 'dimers' are in close proximity to one

another at their respective nucleosomal dyads. For the first time, our structure may allow for the analysis of these proposed 30 nm fiber structures, provide us with the possibility of placing individual nucleosomes within the fiber, and to glean higher-resolution information than currently afforded by the models of the 30 nm fiber. Such analyses will make it possible to visualize how NCP molecules containing longer stretches of DNA might be even able to interact with one another and at the same time accommodate the extranucleosomal DNA within the confines of the chromatin fiber.

5.4f Polyamide binding to tailless H4 NCP leads to unraveling of DNA ends

The experimental maps show that the DNA is displaced at the nucleosomal dyad and the ends (~ 10 bp) of the DNA are highly flexible and unraveled (Figure 5.5). The results from the altered crystal packing interactions strongly suggest that the unraveling of the DNA ends may be a consequence of the H4 tail removal and that this altered nucleosome conformation, already captured in solution by the bound ligand, may in fact drive the alternative mode of crystal packing. Recent footprinting and fluorescence studies demonstrate that binding of the transcription factor, Amt1 to NCPs near the dyad leads to partial dissociation of the DNA ends (White and Luger, 2004). In another study, it was shown that binding of the LexA protein to NCPs leads to an unwrapped, accessible state, simultaneously allowing LexA binding (Li and Widom, 2004). Our results, along with the above studies and the previous report that ligand binding requires tail removal (Gottesfeld et al., 2001), suggest that the H4 N-terminal tail regulates the dynamics of the nucleosomal DNA, and in particular that of the ends of the DNA. The histone H4 N-terminal tail regulates access to nucleosomal DNA (especially to sites near the

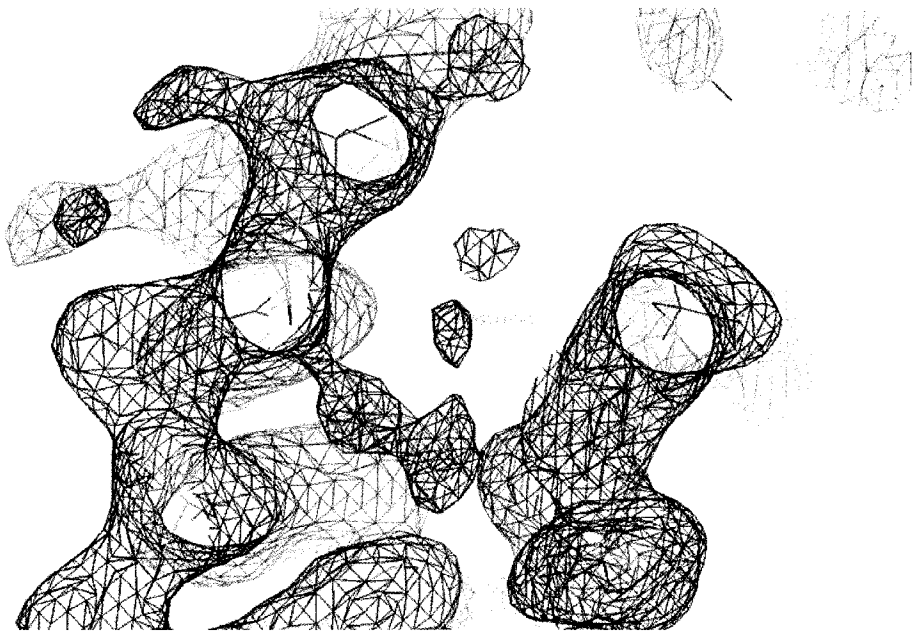
nucleosomal dyad) and removal of the H4 tail therefore allows 'unrestricted entry' to previously inaccessible sites in the NCP. This implies that the DNA ends are dynamic in the tailless H4 NCP even in the absence of the ligand. Our results are also supported by kinetic studies using quantitative site exposure assays (also discussed earlier in Section 5.2), which show that DNA sites are constantly and transiently exposed on the nucleosomal surface (Polach et al., 2000). It was also noted that the equilibrium constants for site exposure are small, from ~ 10^{-2} at the nucleosomal periphery to ~ 10^{-5} at the dyad. The removal of the tail domains brought about only a moderate but a significant 1.5- to 14-fold increase in the position-dependent equilibrium constants for site exposure. At any given time, the abundance of a defined exposed site in a population as well as the lifetime of this transiently exposed site is highly limiting. Hence, it is not surprising that the transient changes following H4 tail removal are not globally captured in the tailless H4 NCP crystals. However, in the case of the tailless H4 NCP-polyamide 6 complex, when the local concentration of the ligand is high (which may not be the case for regulatory factors *in vivo*), the equilibrium is shifted towards stabilization of the transient species, thereby capturing an altered DNA conformation in the crystal lattice.

It would be insightful to study how dynamic the DNA ends are in the tailless H4 NCPs in solution, both in the absence and presence of the ligand. First, fluorescence spectroscopic experiments designed to test the above would corroborate the crystallographic findings and provide additional biophysical evidence for our hypothesis. Second, the use of shorter DNA lengths (~ 126 bp) in the crystals may also reduce the disorder arising from unraveled DNA ends. A combination of fluorescence and

crystallographic studies will provide unprecedented insight into how tail removal and ligand binding influence the dynamics and accessibility of the nucleosomal DNA.

5.5 Acknowledgements

We thank Gerry McDermott, Corie Ralston and Christine Trame at the ALS for help with data collection and Pamela Dyer for help with reagents. The polyamides were synthesized by Christian Melander in the Dervan laboratory. We thank Oren Anderson for providing the atomic coordinates for TAMM. The technical input and help offered by Paul Adams and Jeffrey Bonanno were critical to data analysis of the tailless H4 NCP-polyamide 6 crystals. We also thank Ralf W. Grosse-Kunstleve, Zbigniew Dauter, Peter Zwart, Eleanor Dodson, K. R. Rajashankar, and several program authors for helpful discussions.



Chapter 6

The Temperature of Flash-Cooling has Dramatic Effects on the Diffraction Quality of Nucleosome Crystals

Edayathumangalam, R. S., and Luger, K. (*Acta Crystallogr D Biol Crystallogr* (2005), *in press*).

The above-cited manuscript is currently in press. The text and figures for this chapter are presented as they appear in the manuscript.

6.1 Abstract

Nucleosome core particle crystals are routinely flash-cooled in liquid propane at temperatures ~ 153 K, followed by transfer into a cold nitrogen gas stream (~ 93 K). Analysis of diffraction data from crystals flash-cooled at different temperatures shows that the optimal temperature is ~ 153 K. The data quality worsens with a concomitant reduction in the diffraction limit at temperatures both higher and lower than 153 K. With some batches of crystals, significant shrinkage of the unit cell volume is also observed at temperatures 138 K and lower. The lattice shrinkage is always restricted to the *c*-axis, concurrent with closer packing of two nucleosomes. Direct plunge-cooling of crystals in liquid nitrogen leads to loss of diffraction quality and resolution limit. Thus, in cases where flash-cooling into liquid nitrogen is detrimental to diffraction, optimising cooling protocols at higher temperatures using liquid propane or other cryogens with similar properties may lead to dramatically improved results. In a related study, we show that a nucleosome crystal transported under 'cryocooled' conditions has higher mosaicity and yields inferior data quality in comparison to a crystal cryocooled at the synchrotron. For fragile crystals, transport in mother liquor and/or cryoprotectant followed by subsequent flash-cooling at the synchrotron may be the best procedure.

6.2 Introduction

In macromolecular crystallography, it is now standard practice to collect diffraction data under cryogenic temperatures (~ 100 K) (Garman & Schneider, 1997; Pflugrath, 2004). Cryocrystallography dramatically reduces radiation damage to crystals leading to diffraction data of better quality and higher resolution. It also facilitates the handling,

storage and transport of crystals. The prolonged lifetime of cryocooled crystals has made collection of multi-wavelength anomalous dispersion (MAD) data from single crystals a very common technique (Garman & Schneider, 1997; Hendrickson, 1991; Hendrickson, 1999; Pflugrath, 2004). The crystals are usually harvested and stabilised in cryoprotectant solution prior to cryocooling. The various cryocooling procedures include cryogens such as liquid nitrogen, liquid propane, liquid ethane and tetrafluoromethane (CF_4). Flash-cooling directly in the cold nitrogen gas stream is also common. While liquid nitrogen is commonly used to flash-cool crystals, there are few examples for large macromolecular complexes where flash-cooling in cryogens such as liquid propane (at temperatures higher than that of liquid nitrogen) is the method of choice (for example, see Hope et al., 1989; Luger et al., 1997; Sargent & Richmond, 2004). Some previous studies comparing the cooling rates for liquid nitrogen and liquid propane show that liquid propane has significantly higher cooling rates (Kriminski et al., 2003; Teng & Moffat, 1998), while another study reported that the cooling rates for liquid nitrogen and liquid propane were similar (Walker et al., 1998). These various studies involve intricate experimental setups and the results are not directly comparable. Also, thermocouples were used as samples to obtain the measurements reported in these studies, which cannot be easily or directly extrapolated to macromolecular crystals. While the boiling point (231 K) and melting point (83 K) for liquid propane are well separated, the boiling point (77 K) and melting point (63 K) for liquid nitrogen are very close. Teng and Moffat reported that this narrow spread between the boiling and melting points for liquid nitrogen leads to 'film boiling' between 250 K and 150 K whereby boiling of liquid nitrogen leads to the formation of an insulating film during flash-cooling. This 'film boiling', in turn, reduces the efficiency of

heat transfer in liquid nitrogen (Teng & Moffat, 1998). On the other hand, liquid propane cooling is dominated by 'nucleate boiling' between 260 K and 90 K, during which the coolant makes direct contact with the crystal. This leads to efficient heat transfer and faster cooling rates over a much broader temperature range in liquid propane. Efficient heat exchange occurs with liquid nitrogen only when the temperature falls below 150 K (Teng & Moffat, 1998). Liquid ethane is also an effective coolant. Ethane has been historically used in cryo-electron microscopy, and shares very similar properties to that of propane except that ethane is more difficult to liquefy than propane (ethane boiling point (184.5 K) is lower than propane). Additionally, while the use of liquid ethane is allowed at most facilities, the use of liquid propane at synchrotron sources is restricted due to safety issues and often necessitates the cryocooling of crystals in the home laboratories. Other cryogens include tetrafluoromethane (CF_4). CF_4 has been routinely used for several years as a cryogen in the Cryo-Xe-Siter device (Rigaku/MSC, Inc.) used in the preparation of xenon derivatives (Jim W. Pflugrath, personal communications). The boiling and melting points for CF_4 are 145 K and 86 K, respectively.

As seen above, liquid propane, ethane and CF_4 have a larger spread between boiling and melting points than liquid nitrogen, with liquid propane having the largest dynamic spread. This allows for the optimisation of cryo-cooling protocols at temperatures higher than liquid nitrogen temperatures. Optimisation of flash-cooling temperatures may be more critical than currently appreciated. The crystals of nucleosome core particles (nucleosome) are one such example. The crystals in cryoprotectant are initially flash-cooled in liquid propane (at ~ 153 K) followed by subsequent exposure to the cold nitrogen gas stream. This method is regularly used to obtain well diffracting nucleosome crystals

(Luger et al., 1997; Sargent & Richmond, 2004), although the mechanism underlying this observation remained largely unknown.

Here, we perform a systematic study to examine the effects of various propane temperatures on the diffraction quality of nucleosome crystals. We also compare the results obtained from crystals flash-cooled directly in liquid nitrogen with those obtained from our propane cooling procedure at intermediate temperatures. Our results can be summarized as follows. First, we do not see ice formation for the entire temperature range studied here (113 K to 193 K). The lack of ice nucleation during all our measurements suggests that the cryoprotectant is presumably in a 'vitrified' or 'glass state' in the temperature range tested, although the precise 'glass-transition' temperature (see Angell, 1995 for a discussion of the phenomenon of glass formation in liquids), which characterizes the vitrification properties for the cryoprotectant, has not been measured. This also suggests that the solvent, which is confined to narrow channels and cavities in the crystals (see Section 2.1 for a discussion of the solvent in the nucleosome crystals), remains in a glass state and does not crystallize within the studied temperature range. There have been similar reports on the absence of crystallization of solvent in narrow channels and cavities in protein crystals (Weik et al., 2001). Second, the best diffraction data are obtained when the propane temperature is maintained at ~ 153 K during the flash-cooling step. Flash-cooling crystals at temperatures both above and below 153 K leads to significant deterioration in data quality with a concomitant reduction in the diffraction limit. We also observe that around and below 138 K, some crystals undergo a defined transition during which the *c*-axis shrinks by 3.3 %. Third, we find that direct plunge-cooling of nucleosome crystals into liquid nitrogen yields irreproducible results in terms of the crystal

unit cell volume, diffraction quality and data resolution. Therefore, for fragile crystals, for crystals with high solvent content, or for larger crystals with a tendency for lattice and/or mosaicity changes, optimizing cooling protocols at intermediate temperatures in cryogens like liquid propane (or other cryogens with a dynamic temperature spread) may result in spectacularly improved results. Thus, cryogen, flash-cooling temperature, and flash-cooling protocol need to be a variable in screening for optimal diffraction qualities, in addition to screening for crystallization conditions and cryoprotectant.

We also studied the effect of shipping a crystal in the 'frozen' state versus shipping it in mother liquor containing cryoprotectant. From this experiment, it appears that fragile crystals may be better preserved when transported under liquid conditions prior to cooling.

6.3 Materials and Methods

6.3a Crystals

Preparation and crystallization of nucleosomes containing a palindromic 146 bp DNA fragment derived from human α-satellite DNA and the four full-length *Xenopus laevis* histone proteins have been previously described (Dyer et al., 2004 and references therein). The nucleosome crystals were grown by sitting-drop vapour diffusion. The crystals were grown in 1-2 weeks at 292 K in 40-45 mM $MnCl_2$, 35-38 mM KCl and 20 mM K-cacodylate (pH 6.0) containing ~ 20 μM nucleosome core particle complex. The sitting-drop technique yielded several crystals with dimensions of ~ 0.3 × 0.3 × 0.7 mm. The crystals belong to space group $P2_12_12_1$, with unit cell parameters as described in Ta-

ble 6.1 and Table 6.2. The crystals have a solvent content of 55 %. The orthorhombic nucleosome crystals have narrow intermolecular solvent channels of ~ 15 Å in their longest dimension. The nucleosome core particle also has only a narrow central hole of ~ 10 Å along the superhelical axis, which is not completely hollow and is occupied by some of the histone protein side-chains (Luger et al., 1997).

6.3b Flash-cooling procedure

The nucleosome crystals were harvested by stepwise soaking (~ 1-2 min. each) of the crystals in mother liquor containing increasing concentrations of 2-methyl-2,4-pentanediol (MPD). The MPD concentration was changed gradually in 3 %(*w*/*v*) increments 'in situ'. The 'in situ' harvesting procedure prevents repeated manipulation of these fragile crystals. The final cryoprotectant solution contains 24 %(*w*/*v*) MPD and 5 %(*w*/*v*) trehalose (often used in cryoprotectants as it reduces osmotic stress and prevents excessive crystal dehydration). The crystals were stable after soaking overnight in the final cryoprotectant solution.

Nucleosome crystals are routinely cryocooled in two steps. First, liquid propane (propane tank with a diptube, instrument grade 99.5% purity, Scott Speciality Gases) is collected and pre-cooled in liquid nitrogen for a few minutes. An empty cryovial containing a stirbar and surrounded by a layer of insulating material is cooled in liquid nitrogen. The pre-cooled liquid propane is poured into the cryovial and allowed to gradually warm to 153 K or the desired final temperature, as shown in Figure 6.1A, B. Precooling the empty cryovial in liquid nitrogen ensures a slow and constant rate of increase in propane temperature (not shown). The slow kinetics of propane temperature

Table 6.1 Effect of flash-cooling temperature on data quality

Dataset*	193 K	173 K	153 K	143 K	138 K	133 K	113 K
Mosaicity (°)	0.7	0.6	0.8	0.6	0.7	0.9	0.8
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell lengths	106.0	106.0	106.3	106.1	105.3	105.2	104.9
(*a, b, c* in Å)	110.1	109.6	109.8	109.6	109.3	109.4	109.2
	181.9	181.9	181.9	182.0	175.6	175.4	175.1
Resolution range (Å)†	50-2.88	50-2.88	50-2.88	50-2.89	50-3.41	50-3.41	50-3.62
	(2.98-2.88)	(2.98-2.88)	(2.98-2.88)	(2.99-2.89)	(3.62-3.41)	(3.62-3.41)	(3.90-3.62)
No. of unique reflections	43345	45523	44890	45425	27946	27927	20825
Completeness (%)	88.7 (77.9)	93.7 (88.3)	91.7 (78.6)	94.0 (83.7)	97.9 (93.5)	98.5 (96.6)	87.6 (90.4)
Multiplicity	3.0	3.0	3.1	3.1	3.2	3.1	4.0
Average I/σ(I)	9.6 (2.2)	19.3 (3.6)	27.9 (6.5)	15.1 (2.4)	13.9 (2.8)	11.5 (2.7)	12.3 (2.8)
I/σ(I) 3.91-3.62 Å bin	6.5	15.4	26.2	12.3	4.6	4.0	2.8
I/σ(I) 2.98-2.88 Å bin	2.2	3.6	6.5	2.4	-	-	-
R-merge#	0.095	0.048	0.036	0.067	0.08	0.081	0.099
	(0.517)	(0.303)	(0.181)	(0.448)	(0.391)	(0.49)	(0.639)
R-merge 3.91-3.62 Å bin	0.063	0.147	0.041	0.088	0.244	0.265	0.639
R-merge 2.98-2.88 Å bin	0.517	0.303	0.181	0.448	-	-	-

* Datasets were collected in-house at the indicated propane temperature during the flash-cooling step.

† Effective high-resolution limit reported is equal to the d_{edge} (d_{edge} = 2.89 Å) for datasets 193 K, 173 K, 153 K and 143 K. For datasets 138 K, 133 K and 113 K, the effective high-resolution limit was determined using a criterion of I/σ(I) ≥ 2.

R-merge = $\Sigma|I_h - \langle I_h\rangle|/\Sigma I_h$, where $\langle I_h\rangle$ is the mean of the measurements for a single *hkl*.

Table 6.2 Effect of mode of crystal transport on data quality

Dataset*	1 (flash-cooled in-house)	2 (flash-cooled at synchrotron)
Mosaicity (°)	0.7	0.3
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell lengths	106.2	106.2
(*a, b, c* in Å)	109.5	109.6
	181.8	181.8
Resolution range (Å) †	50-2.60 (2.70-2.60)	50-2.45 (2.54-2.45)
No. of unique	64298	77169
reflections	97.3 (96.5)	99.0 (99.3)
Completeness (%)	4.1	4.1
Multiplicity	14.2 (3.1)	21.2 (2.7)
Average $I/\sigma(I)$	3.1 (2.70-2.60 Å bin)	5.1 (2.76-2.54 Å bin)
	0.083 (0.344)	0.064 (0.546)
R-merge#	0.344 (2.70-2.60 Å bin)	0.327 (2.76-2.54 Å bin)

* Datasets were collected at Advanced Light Source beamline 8.2.1.
† Effective high-resolution limit was determined using a criterion of $I/\sigma(I) \geq 2$.
# R-merge = $\Sigma|I_h - \langle I_h \rangle|/\Sigma I_h$, where $\langle I_h \rangle$ is the mean of the measurements for a single *hkl*.

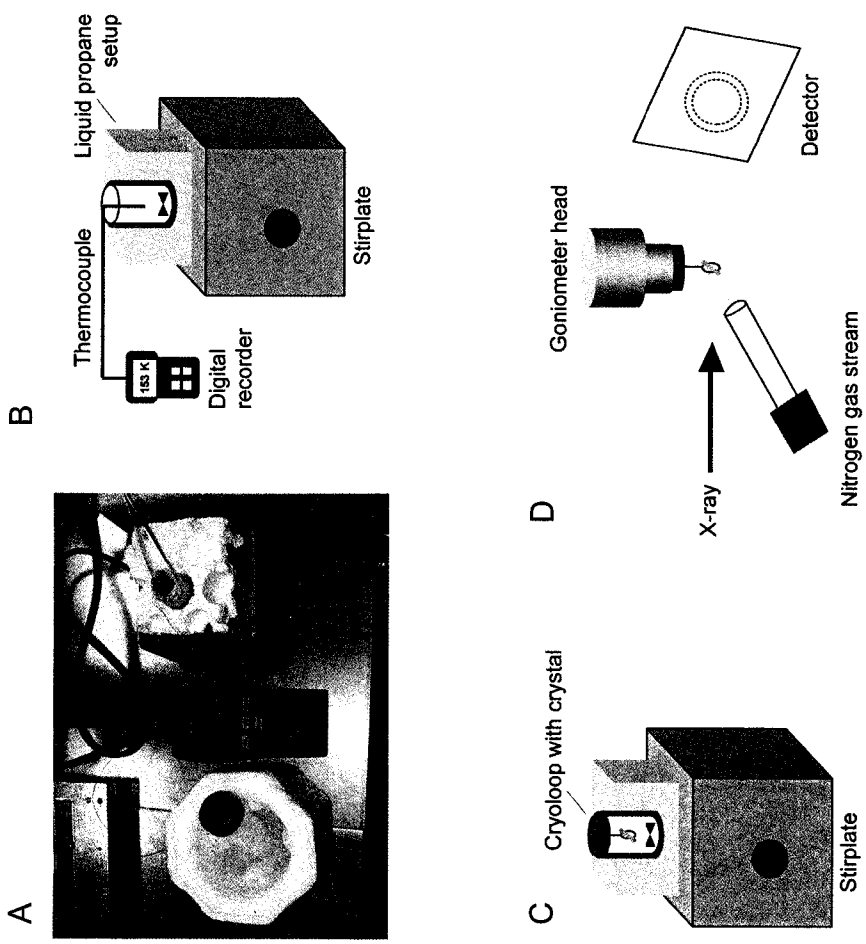


Figure 6.1 Experimental set up for flash-cooling nucleosome crystals in liquid propane. (A) Photograph of the propane flash-cooling set up. (B) Liquid propane (pre-cooled in liquid nitrogen) is allowed to slowly thaw to around 153 K under continuous stirring. (C) The crystal is mounted in a cryoloop and flash-cooled by immersion in the cryovial when the propane temperature is ~ 153 K. (D) The flash-cooled crystal is quickly transferred to a cold nitrogen gas stream (93 K).

increase have been previously studied (Vernede & Fontecilla-Camps, 1999). The nucleosome crystal cooling procedure takes advantage of the slow kinetics of propane for the precise control of propane temperature during flash-cooling, and we exploit the same property in this study (as discussed below) for testing a wide range of propane temperatures during flash-cooling. The temperature of the propane in the vial is monitored using a thermocouple attached to a digital recorder. The length of the thermocouple inserted in the vial is identical to the length from the centre of the cryoloop to the base of the pin (Figure 6.1 A, B). This monitors the propane temperature at the height of the crystal in the vial. The crystal is flash-cooled by immersing in propane at ~ 153 K (or other desired final temperatures, see Table 6.1), as in Figure 6.1 A, C. The crystal is then immediately transferred to the cold nitrogen gas stream (93 K) for data collection, as in Figure 6.1D. The flash-cooled crystal can also be stored in liquid nitrogen for long-term storage, transport or subsequent data collection.

6.3c Data collection and refinement

Diffraction data from similar-sized crystals (~ 0.3 × 0.3 × 0.7 mm) that were flash-cooled at different propane temperatures (Table 6.1) were collected using our in-house rotating anode RU-H3R generator and RAxis IV detector (Rigaku/MSC, Inc.) at a wavelength of 1.5418 Å. Data comparing the effect of transport on cryocooled crystals (Table 6.2) were obtained at the Advanced Light Source, beamline 8.2.1, Berkeley, USA. Single measurements at each temperature are shown in Table 6.1, although data were collected from three crystals for each temperature and similar results were obtained from repeated measurements (not shown). Data were collected with identical data-collection

strategies and parameters. Similarly, the two datasets reported in Table 6.2 were collected from two similar-sized fragments of one single long crystal. Diffraction data were integrated, scaled and merged with HKL2000 (Otwinowski & Minor, 1997). Mosaicity is defined in HKL2000 as the smallest angle through which the crystal can rotate, about any axis or combination of axes, while a reflection is still observed. Mosaicity estimates from the program also include contributions from the x-ray bandwidth and beam crossfire. An error model, as implemented in HKL2000, was used to allow $\chi 2$ values (goodness-of-fit estimate) to converge to ~ 1.0 for all datasets prior to calculation of the intensity statistics reported in Table 6.1 and Table 6.2. Molecular replacement (using Protein Data Bank ID code 1AOI as the original search model) and subsequent partial structure refinement for two of the structures (datasets 153 K and 138 K in Table 6.1) was carried out using CNS (Brunger et al., 1998). Steps in refinement included rigid body refinement, simulated annealing, positional minimization and grouped B-factor refinement. Crystal packing analysis and structure comparisons were done using the program LSQMAN from the Uppsala Software Factory (Kleywegt, 1996).

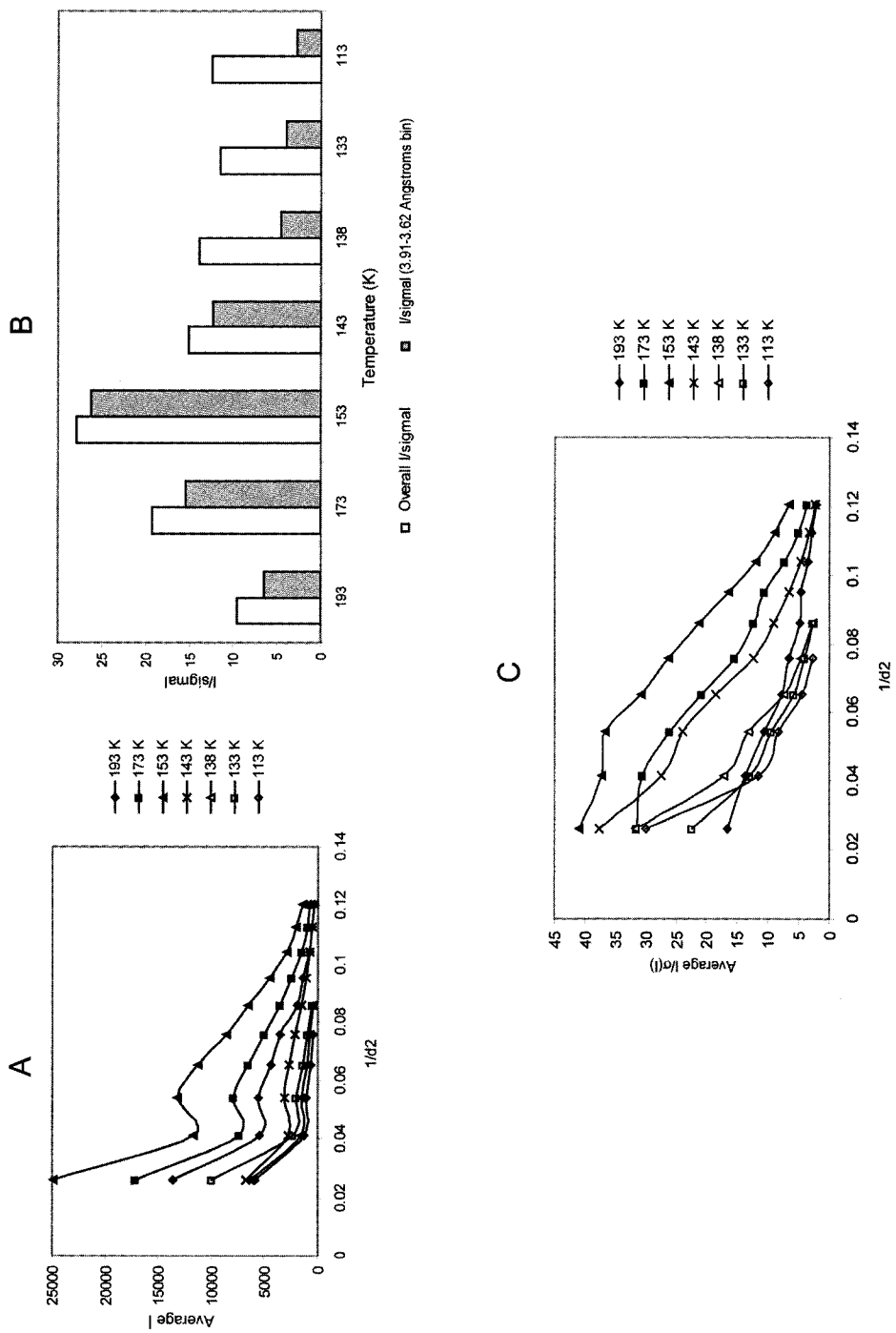
6.4 Results and Discussion

6.4a Optimal propane flash-cooling temperature

The temperature for the propane flash-cooling step for nucleosome crystals has been previously reported by our group and others to be ~ 153 K (Luger et al., 1997; Muthurajan et al., 2003; Sargent & Richmond, 2004). We performed a systematic and detailed analysis to investigate the effect of vitrification temperatures on the diffraction quality of nucleosome crystals. Three nucleosome crystals were flash-cooled at each of the various

liquid-propane temperatures and diffraction data were obtained using an identical data-collection strategy for each crystal. All crystals were of uniform size (~ 0.3 × 0.3 × 0.7 mm). Representative data from only one crystal for each temperature is shown in Table 6.1, although the observed data statistics were reproducible for the multiple measurements at each temperature.

A comparison of the various datasets reveals several temperature-dependent effects on the data quality, as shown in Table 6.1 and Figure 6.2. The overall completeness and completeness in the highest resolution bin are both lower at extreme temperatures (193 K, 113 K) because of a higher percentage of reflections rejected following data reduction. A comparison of the average diffraction intensities and the R-merge values for all datasets reveals that flash-cooling crystals at a propane temperature of 153 K leads to the highest diffraction intensities with lowest R-merge values (Figure 6.2). We observe a two-to-five-fold increase in the average diffraction intensities (Table 6.1, Figure 6.2A) and ~ 3-fold increase in the average $I/\sigma(I)$ (Table 6.1, Figure 6.2B, C) at 153 K compared to both higher and lower temperatures. The R-merge values also reflect this trend with data at 153 K yielding R-merge values which are ~ 3-fold lower than those at 193 K and 113 K (Table 6.1, Figure 6.2D, E). These differences are significant considering that the comparisons are drawn between crystals of similar sizes, with identical data collection parameters, space groups and data redundancies. Further characterization of crystals at different temperatures with different detector distances showed a concomitant reduction in the diffraction limit at temperatures both above and below 153 K (not shown), even though the data shown in Table 6.1 were obtained only to d_{edge} = 2.89 Å in each case. Routinely, the observable diffraction limit deteriorates from ~ 2.2 Å at 153 K to ~ 3.2 Å



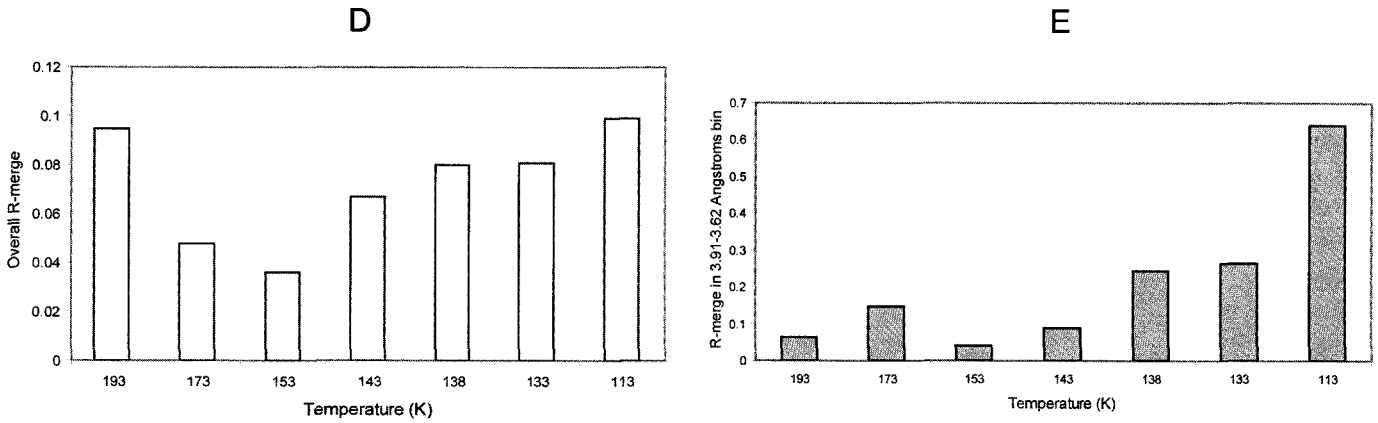
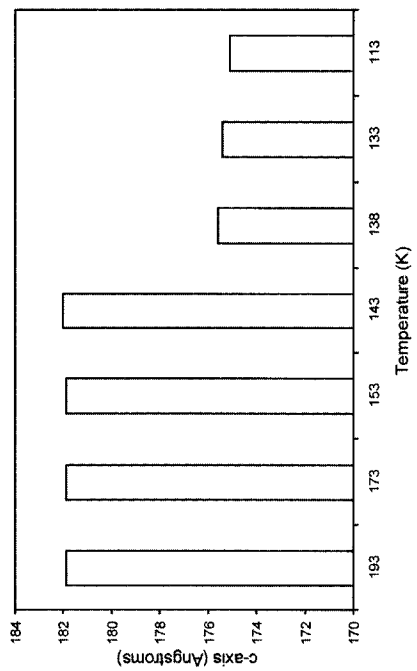


Figure 6.2 Plots showing the dependence of the data statistics (from Table 1) on propane flash-cooling temperature. (A) Plot of average intensities from all reflections as a function of resolution. (B) $I/\sigma(I)$ values for the various datasets. (C) Average $I/\sigma(I)$ as a function of resolution. Overall R-merge (D) and R-merge in the indicated resolution bin (E) for the various datasets.

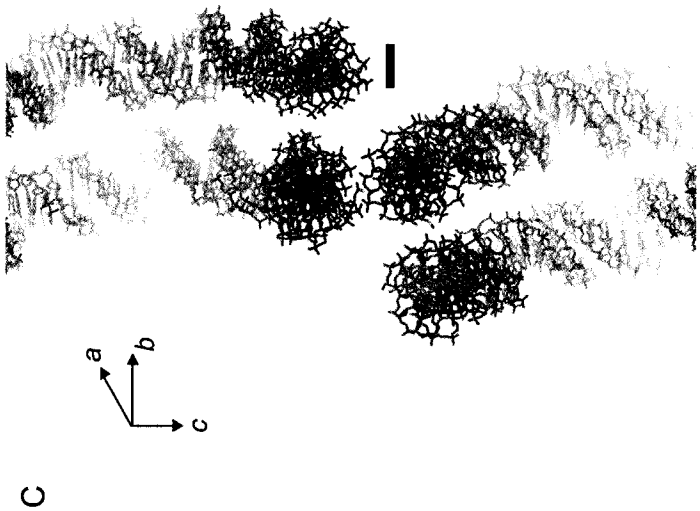
at the extreme propane temperatures tested (not shown). Hence, our results reveal that ~ 153 K is the optimal cooling temperature for nucleosome crystals and that a temperature fluctuation of a few tens of degrees above and below the optimal temperature has a detrimental effect on diffraction quality.

6.4b Crystal packing analysis

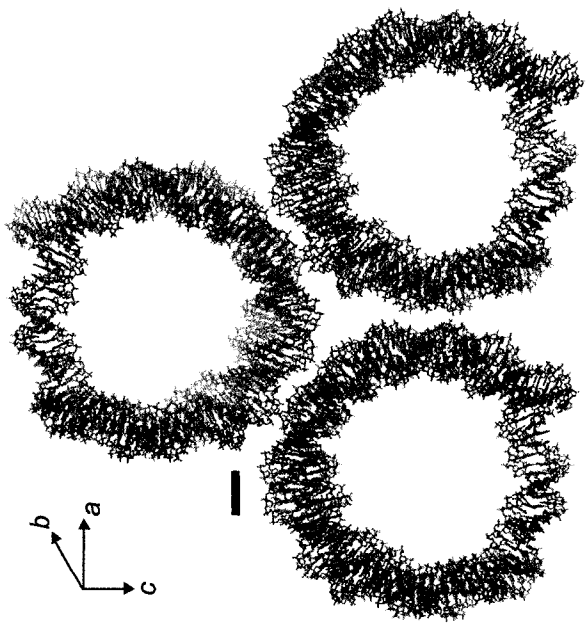
In the course of these experiments, we also observed that the lattice spacing along the *c*-axis changes in a temperature-dependent manner, while the *a*- and *b*- axes remain unchanged (Table 6.1, Figure 6.3A). The unit cell length along the *c*-axis is shortened by ~ 3.3 % (from ~ 182 Å to ~ 175 Å). It must be noted that lattice changes are only observed with some batches of nucleosome crystals tested. If they do occur, they are always observed at temperatures lower than 153 K. Although the basis for the lattice change is not clear, it may be related to the kinetics of nucleosome crystal flash-cooling and to the extent of thermomechanical stress induced during flash-cooling. It is a fairly common observation that mild dehydration of crystals prior to or during flash-cooling often improves diffraction. The effects of lattice changes on the diffraction quality, the crystal packing and the domain conformation of HIV-1 reverse transcriptase have been examined in detail and up to 18 % reduction in lattice volume following induced dehydration was reported (Esnouf et al., 1998). We show here that even ~ 3.3 % reduction of the lattice volume (Table 6.1, Figure 6.3A) during flash-cooling in our case correlates with a worsening of diffraction data quality (Table 6.1, Figure 6.2). This reduction in length from ~ 182 Å to ~ 175 Å stays constant with no further decrease beyond this end-point at any of the lower temperatures tested.



A



C



B

Figure 6.3 Lattice changes in nucleosome crystals. (*a*) Reduction in *c*-axis length at lower propane temperatures. (*b*-*e*) Comparison of crystal packing between crystals flash-cooled at propane temperatures of 153 K (green) and 138 K (blue). DNA alone is shown. Histone proteins have been omitted for clarity. (*b*) Crystal packing with view down the unit cell *b*-axis. Superposition of the two structures is shown. The scale bar is 15 Å. (*c*) Crystal packing with view down the *a*-axis. The scale bar is 9 Å. (*d*) A magnified view of Fig. 3*c* is shown. Top panel shows a superposition of the two structures; bottom panel shows the individual structures. (*e*) Magnified view of Fig. 3*d* (top panel) is shown.

In order to study how the shrinkage of the unit cell *c*-axis affects nucleosome structure and/or crystal packing, the structures obtained at 153 K and 138 K (datasets in Table 6.1) were compared in detail. Both datasets were phased and the structures were partially refined as outlined in the 'Materials and methods' section. Molecular replacement (using PDB code 1AOI as the original model) was used for initial phasing followed by partial refinement of the model, as described in the experimental procedures. The R-factor and R-free values following partial refinement are 22.0 % and 26.9 %, respectively, for dataset 153 K, and 24.7 % and 32.2 %, respectively, for the dataset 138 K. As shown in Table 6.1, while the unit cell *a*- and *b*-axes are similar for the two datasets, the *c*-axis is shorter by ~ 3.3 % for dataset 138 K, resulting in a decrease in the solvent content from ~ 55 % to ~ 52 % for dataset 138 K.

We compared the two structures following a least-squares superposition of the nucleosome structure at 138 K on that obtained from the dataset 153 K using the LSQMAN program from the Uppsala Software Factory (Kleywegt, 1996). The two structures superimposed with an overall root-mean-square deviation of 0.33 Å and an inspection of the two structures revealed no global differences between them (not shown). Next, we compared the crystal packing between the two structures, which revealed some interesting changes. Figure 6.3 summarises our findings. The packing is altered along the *c*-axis and neighbouring molecules pack much closer to each other along the *c*-axis at ~ 138 K (Figure 6.3B-E). More specifically, the outer strand phosphates on the nucleosomal DNA between neighbouring molecules are in close proximity as a consequence of the lattice change (for example, see Figure 6.3E). The negatively charged phosphates in close contact may cause repulsion and increased disorder within the crystal, which may partly ex-

plain the deterioration in diffraction in crystals with the reduced *c*-axis (Table 6.1). Also, the molecules cannot pack any closer along the *c*-axis due to steric constraints. This explains why we do not observe further reduction beyond the *c*-axis length of ~ 175 Å at any of those temperatures.

Nucleosome crystals are highly unstable, susceptible to radiation damage and are not amenable to data collection at room temperature. It is therefore not possible to collect complete datasets and obtain the mosaicity value and other data statistics for crystals at room temperature. We however tested how nucleosome crystals diffracted when flash-cooled in liquid nitrogen. Five crystals were cooled by direct plunging into liquid nitrogen (data not shown). Again, as in the case of flash-cooling in liquid propane, flash-cooling in liquid nitrogen did not lead to ice nucleation in the crystals. All crystals were of similar size (~ 0.3 × 0.3 × 0.7 mm). Flash-cooling the crystals by plunging into liquid nitrogen leads to irreproducible results. In some cases, a precise cell shrinkage identical to that discussed above (shrinkage of the *c*-axis from ~ 182 Å to ~ 175 Å) is observed. A severe deterioration of the data quality and a decrease in the diffraction limit is also observed in most cases in liquid nitrogen as compared to that from crystals flash-cooled in liquid propane at 153 K. However, because of the lack of reproducibility, we made no further use of the data.

6.4c Effect of transport on cryocooled crystals

We have previously observed that the mosaicity estimates obtained in-house for a freshly cooled nucleosome crystal are often lower than those obtained following transport of the same cryocooled crystal to the synchrotron. Since the estimate of mosaicity

includes contributions from both the mosaic spread of the crystal and beam divergence, it is not possible to separate these factors when comparing data collected in-house and at the synchrotron for the same crystal. In order to study the effect of the mode of crystal transport on the diffraction quality, diffraction data from two similar-sized fragments from one large crystal were compared. Crystal 1 was flash-cooled in-house and transported to the synchrotron in liquid nitrogen while crystal 2 was transported in a harvesting tray in cryoprotectant solution and then mounted in a cryoloop and flash-cooled at the synchrotron (Table 6.2). The datasets were collected at the same beamline using identical data collection strategies and processed with the same program (outlined in the experimental procedures). As shown in Table 6.2, the two datasets are isomorphous. Most noticeably, the mosaicity of the crystal flash-cooled on-site is much lower (mosaicity 0.3) as compared to that flash-cooled in-house (mosaicity 0.7). The diffraction limit, the diffraction intensities (~ 1.5-fold higher) and the overall data quality (Table 6.2) were significantly better for the crystal cooled on-site, although the diffracting volumes for the two crystals were almost identical. Both crystals were flash-cooled in an identical manner in liquid propane at 153 K using our regular cryocooling protocol (as discussed before). Our results suggest that better data quality and a higher diffraction limit are obtained when the crystal is transported in cryoprotectant and flash-cooled at the synchrotron. Dehydration or other lattice changes do not contribute to the differences as such changes have not been previously observed for nucleosome crystals upon prolonged incubation in the cryoprotectant (also seen by the almost identical unit cell values for the two datasets in Table 6.2). Only a single measurement in each case is reported in Table 6.2 and therefore several datasets will have to be obtained to analyse

these effects in detail and to rule out contributions of any other variables to the observed differences. Therefore, for fragile crystals that are susceptible to mosaicity changes, it may be worthwhile to transport the crystal under liquid conditions (either in mother liquor and/or cryoprotectant solution) and subsequently flash-cool at the synchrotron.

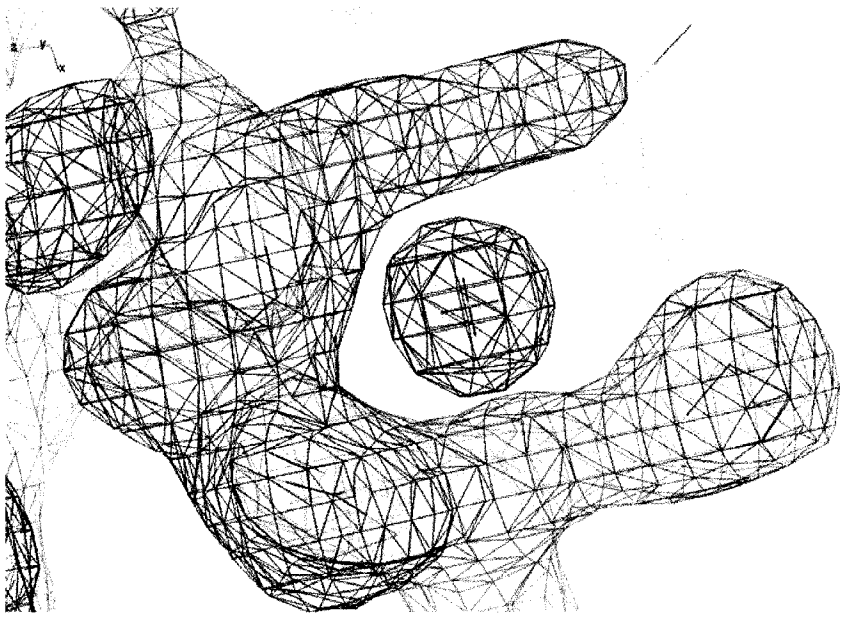
6.4d Conclusions

In conclusion, our results show that plunge-cooling in liquid nitrogen is irreproducible, can cause significant damage to the nucleosome crystals and that limiting the magnitude of the temperature jump by plunge-cooling in liquid propane at intermediate temperatures (~ 153 K) prevents or at least minimizes this damage. We routinely observe in our laboratory that plunge-cooling in liquid nitrogen works well for crystals of other macromolecules. In contrast, nucleosome crystals are fragile and have a tendency for high mosaicity and lattice changes during flash-cooling. Flash-cooling these crystals in liquid propane at an intermediary temperature of 153 K prior to data collection at 93 K is essential for optimal data quality. Dehydration during soaking and/or flash-cooling may be a common phenomenon with macromolecular crystals, and thus careful observation of lattice changes may be helpful since it may either be advantageous or detrimental to diffraction. We conclude that the lack of 'film boiling' of liquid propane (which increases the cooling rates), the slow kinetics of propane temperature increase (which allows for a precise control of the temperature during cooling) and a delicate balance between the cooling rates and the thermomechanical stress at ~ 153 K determine the optimal cryocooling temperature for nucleosome crystals harvested in the MPD-containing cryoprotectant.

In the second study, we find that the mode of crystal transport has an effect on the diffraction quality of nucleosome crystals. The crystal transported in cryoprotectant produced better diffraction data with a higher resolution limit as compared to the crystal transported under flash-cooled conditions. Hence, it may be worthwhile to transport fragile crystals in cryoprotectant in order to minimize stress during transport.

6.5 Acknowledgements

We thank Vidya Subramanian, Jayanth Chodaparambil, Trey Toombs and Srinivas Chakravarthy for providing nucleosome crystals or technical help; Pamela N. Dyer for help with reagents; Corey Ralston and Gerry McDermott for help with data collection at the Advanced Light Source beamline 8.2.1; and Eleanor J. Dodson, Jim W. Pflugrath and Seth A. Darst for helpful discussions. This work was supported by a grant GM61909 from the U. S. National Institutes of Health (NIH) to K. L.



Chapter 7

Summary and Future Considerations

The results obtained during the course of my thesis work support the conclusion that the NCP is a very dynamic entity and provide evidence to demonstrate that the architecture of the nucleosome provides multiple ways to regulating its stability, accessibility and dynamics. Our results also show that the DNA is very dynamic in spite of the fact that our studies were performed with the 146 bp human α-satellite DNA fragment, which is a highly positioned nucleosome positioning sequence.

The nucleosome repositioning assays show that polyamides inhibit *in vitro* nucleosome repositioning and indicate that perhaps some bound ligands inhibit the rotational and translational rearrangements of the DNA with respect to the octamer, thereby inhibiting nucleosome repositioning. Since this phenomenon appears to be more general with similar reports for other polyamides and other DNA sequences, it would be interesting to test whether polyamides can be used to inhibit ATP-dependent chromatin remodeling events *in vitro* and to capture remodeling intermediates. Such intermediates would help decipher the events during nucleosome remodeling by chromatin remodeling complexes.

Studies using the polyamide clamp show that targeting the nucleosomal supergroove leads to enhanced crystallization properties for the nucleosome crystals. Also, the clamp prevents *in vitro* nucleosome dissociation. Nucleosome dissociation *in vitro* as well as *in vivo* is a multiple step process. Our results allow for the rational design of clamps that target other supergrooves in the NCP. For example, clamps designed to target the ends of nucleosomal DNA would be useful in studies of histone exchange mediated by histone chaperones like Nap-1 and in the co-crystallization of the NCP in complex with the linker

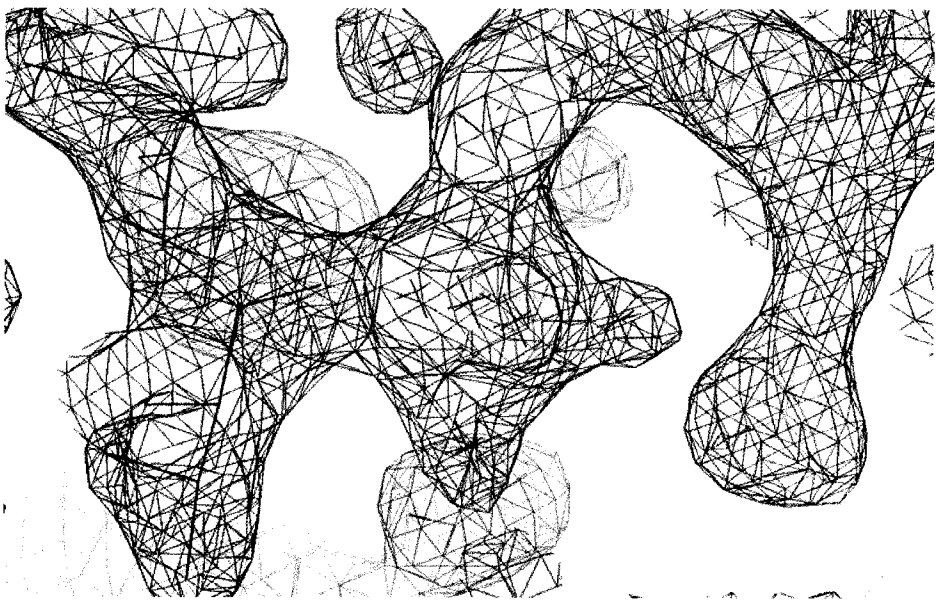
histone, H1. Clamps designed to target nucleosomes on longer DNA templates can also be used to drive nucleosome positioning towards certain nucleosome positions and to capture chromatin-remodeling intermediates.

Our studies comparing NCPs differing only by one-bp length of DNA (146 bp versus 147 bp DNA) indicate the occurrence of a heterogeneous population of twist-defect nucleosomal states in solution. We provide evidence to suggest twist diffusion as a plausible mechanism for establishing a dynamic equilibrium between the various twist-defect states in nucleosomes in solution. The intrinsic conformational flexibility of nucleosomal DNA may modulate molecular recognition *in vivo*, especially in regions of the nucleosome with high conformational flexibility.

Results from the co-crystal structure of the tailless H4-containing NCP in complex with polyamide 6 suggest that tail removal makes the NCP more amenable to structural rearrangements. The structure reveals that the DNA ends are disordered in order to accommodate the bound polyamide. It would be interesting to test how dynamic the ends of the DNA are in the tailless H4-containing nucleosome in solution, both in the absence and presence of the ligand. Fluorescence resonance energy transfer (FRET) experiments designed to test the above would corroborate the crystallographic findings and provide additional biophysical evidence for our hypothesis. We also report the first structure of a nucleosome with altered crystal packing arrangements that do not utilize the ends of the nucleosomal DNA. It may be worthwhile to attempt to improve the order within the crystals with the use shorter DNA fragments (~ 126 bp) since ~ 10 bp of DNA are disordered on both ends. A combination of the FRET studies and the crystallographic studies will

provide unprecedented insight into how tail removal and ligand binding influence the dynamics and accessibility of nucleosomal DNA.

Studies on flash-cooling procedures for nucleosome crystals clearly show a dramatic difference in diffraction quality as a function of the temperature during the flash-cooling step. Our studies suggest that optimization of cryogen, temperature and flash-cooling protocol are critical steps for best diffraction and need to be a variable in screening for optimal diffraction qualities, in addition to screening for crystallization conditions and cryoprotectant, especially when conventional methods for cryocooling compromise diffraction.



REFERENCES

Akey, C. W., and Luger, K. (2003). Histone chaperones and nucleosome assembly. Curr Opin Struct Biol *13*, 6-14.

Allan, J., Harborne, N., Rau, D. C., and Gould, H. (1982). Participation of core histone "tails" in the stabilization of the chromatin solenoid. J Cell Biol *93*, 285-297.

Angell, C. A. (1995). Formation of glasses from liquids and biopolymers. Science *267*, 1924-1935.

Annunziato, A. T., and Hansen, J. C. (2000). Role of histone acetylation in the assembly and modulation of chromatin structures. Gene Expr *9*, 37-61.

Ausio, J., Seger, D., and Eisenberg, H. (1984). Nucleosome core particle stability and conformational change. Effect of temperature, particle and NaCl concentrations, and crosslinking of histone H3 sulfhydryl groups. J Mol Biol *176*, 77-104.

Baird, E. E., and Dervan, P. B. (1996). Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. J Am Chem Soc *118*, 6141-6146.

Baneres, J. L., Essalouh, L., Jariel-Encontre, I., Mesnier, D., Garrod, S., and Parello, J. (1994). Evidence indicating proximity in the nucleosome between the histone H4 N termini and the globular domain of histone H1. J Mol Biol *243*, 48-59.

Bao, Y., Konesky, K., Park, Y. J., Rosu, S., Dyer, P. N., Rangasamy, D., Tremethick, D. J., Laybourn, P. J., and Luger, K. (2004). Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. Embo J *23*, 3314-3324.

Berger, S. L. (2002). Histone modifications in transcriptional regulation. Curr Opin Genet Dev *12*, 142-148.

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al.* (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr *54*, 905-921.

Chaudhry, C., Horwich, A. L., Brunger, A. T., and Adams, P. D. (2004). Exploring the structural dynamics of the E.coli chaperonin GroEL using translation-libration-screw crystallographic refinement of intermediate states. J Mol Biol *342*, 229-245.

Cheung, P., Allis, C. D., and Sassone-Corsi, P. (2000). Signaling to chromatin through histone modifications. Cell *103*, 263-271.

Collaborative Computational Project, N. (1994). The CCP4 Suite: Programs for Protein Crystallography. Acta Crystallogr D Biol Crystallogr *D50*, 760-763.

Collingwood, T. N., Urnov, F. D., and Wolffe, A. P. (1999). Nuclear receptors: coactivators, corepressors and chromatin remodeling in the control of transcription. J Mol Endocrinol *23*, 255-275.

Cotton, R. W., and Hamkalo, B. A. (1981). Nucleosome dissociation at physiological ionic strengths. Nucleic Acids Res *9*, 445-457.

Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W., and Richmond, T. J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. J Mol Biol *319*, 1097-1113.

Dervan, P. B. (2001). Molecular recognition of DNA by small molecules. Bioorg Med Chem *9*, 2215-2235.

Dervan, P. B., and Edelson, B. S. (2003). Recognition of the DNA minor groove by pyrrole-imidazole polyamides. Curr Opin Struct Biol *13*, 284-299.

Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T. J. (2003). Chromatin fiber folding: requirement for the histone H4 N-terminal tail. J Mol Biol *327*, 85-96.

Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R. R., and Richmond, T. J. (2004). Nucleosome arrays reveal the two-start organization of the chromatin fiber. Science *306*, 1571-1573.

Dutnall, R. N. (2003). Cracking the histone code: one, two, three methyls, you're out! Mol Cell *12*, 3-4.

Dyer, P. N., Edayathumangalam, R. S., White, C. L., Bao, Y., Chakravarthy, S., Muthurajan, U. M., and Luger, K. (2004). Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods Enzymol *375*, 23-44.

Edayathumangalam, R. S., Weyermann, P., Dervan, P. B., Gottesfeld, J. M., and Luger, K. (2005). Nucleosomes in solution exist as a mixture of twist-defect states. J Mol Biol *345*, 103-114.

Edayathumangalam, R. S., Weyermann, P., Gottesfeld, J. M., Dervan, P. B., and Luger, K. (2004). Molecular recognition of the nucleosomal "supergroove". Proc Natl Acad Sci U S A *101*, 6864-6869.

Engh, R. A., and Huber, R. (1991). Accurate bond and angle parameters for X-ray protein-structure refinement. Acta Crystallog sect *A 47*, 392-400.

Esnouf, R. M., Ren, J., Garman, E. F., Somers, D. O., Ross, C. K., Jones, E. Y., Stammers, D. K., and Stuart, D. I. (1998). Continuous and discontinuous changes in the unit cell of HIV-1 reverse transcriptase crystals on dehydration. Acta Crystallogr D Biol Crystallogr *54 (Pt 5)*, 938-953.

Fischle, W., Wang, Y., and Allis, C. D. (2003). Binary switches and modification cassettes in histone biology and beyond. Nature *425*, 475-479.

Flaus, A., and Owen-Hughes, T. (2004). Mechanisms for ATP-dependent chromatin remodelling: farewell to the tuna-can octamer? Curr Opin Genet Dev *14*, 165-173.

Flaus, A., and Owen-Hughes, T. (2003). Mechanisms for nucleosome mobilization. Biopolymers *68*, 563-578.

Flaus, A., Luger, K., Tan, S., and Richmond, T. J. (1996). Mapping nucleosome position at single base-pair resolution by using site-directed hydroxyl radicals. Proc Natl Acad Sci U S A *93*, 1370-1375.

Flaus, A., Rencurel, C., Ferreira, H., Wiechens, N., and Owen-Hughes, T. (2004). Sin mutations alter inherent nucleosome mobility. Embo J *23*, 343-353.

Fletcher, T. M., and Hansen, J. C. (1995). Core histone tail domains mediate oligonucleosome folding and nucleosomal DNA organization through distinct molecular mechanisms. J Biol Chem *270*, 25359-25362.

Garcia-Ramirez, M., Dong, F., and Ausio, J. (1992). Role of the histone "tails" in the folding of oligonucleosomes depleted of histone H1. J Biol Chem *267*, 19587-19595.

Garcia Ramirez, M., Rocchini, C., and Ausio, J. (1995). Modulation of chromatin folding by histone acetylation. J Biol Chem *270*, 17923-17928.

Garman, E. F., and Schneider, T. R. (1997). Macromolecular Crystallography. J Appl Cryst *30*, 211-237.

Geierstanger, B. H., Mrksich, M., Dervan, P. B., and Wemmer, D. E. (1994). Design of a G.C-specific DNA minor groove-binding peptide. Science *266*, 646-650.

Georges, S. A., Kraus, W. L., Luger, K., Nyborg, J. K., and Laybourn, P. J. (2002). p300-mediated tax transactivation from recombinant chromatin: histone tail deletion mimics coactivator function. Mol Cell Biol *22*, 127-137.

Gottesfeld, J. M., Belitsky, J. M., Melander, C., Dervan, P. B., and Luger, K. (2002). Blocking transcription through a nucleosome with synthetic DNA ligands. J Mol Biol *321*, 249-263.

Gottesfeld, J. M., and Luger, K. (2001). Energetics and affinity of the histone octamer for defined DNA sequences. Biochemistry *40*, 10927-10933.

Gottesfeld, J. M., Melander, C., Suto, R. K., Raviol, H., Luger, K., and Dervan, P. B. (2001). Sequence-specific recognition of DNA in the nucleosome by pyrrole- imidazole polyamides. J Mol Biol *309*, 615-629.

Gottesfeld, J. M., Neely, L., Trauger, J. W., Baird, E. E., and Dervan, P. B. (1997). Regulation of gene expression by small molecules. Nature *387*, 202-205.

Grosse-Kunstleve, R. W., and Adams, P. D. (2003). Substructure search procedures for macromolecular structures. Acta Crystallogr D Biol Crystallogr *59*, 1966-1973.

Hamiche, A., Sandaltzopoulos, R., Gdula, D. A., and Wu, C. (1999). ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF. Cell *97*, 833-842.

Hansen, J. C. (2002). Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. Annu Rev Biophys Biomol Struct *31*, 361-392.

Harp, J. M., Hanson, B. L., Timm, D. E., and Bunick, G. J. (2000). Asymmetries in the nucleosome core particle at 2.5 A resolution. Acta Crystallogr D Biol Crystallogr *56 Pt 12*, 1513-1534.

Heard, E., Rougeulle, C., Arnaud, D., Avner, P., Allis, C. D., and Spector, D. L. (2001). Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. Cell *107*, 727-738.

Hendrickson, W. A. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. Science *254*, 51-58.

Hendrickson, W. A. (1999). Maturation of MAD phasing for the determination of macromolecular structures. J Synchrotron Rad *6*, 845-851.

Hong, L., Schroth, G. P., Matthews, H. R., Yau, P., and Bradbury, E. M. (1993). Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. Biol Chem *268*, 305-314.

Hope, H., Frolow, F., von Bohlen, K., Makowski, I., Kratky, C., Halfon, Y., Danz, H., Webster, P., Bartels, K. S., Wittmann, H. G., and et al. (1989). Cryocrystallography of ribosomal particles. Acta Crystallogr B *45 (Pt 2)*, 190-199.

Jenuwein, T., and Allis, C. D. (2001). Translating the histone code. Science *293*, 1074-1080.

Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjelgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Cryst *A47*, 110-119.

Kagalwala, M. N., Glaus, B. J., Dang, W., Zofall, M., and Bartholomew, B. (2004). Topography of the ISW2-nucleosome complex: insights into nucleosome spacing and chromatin remodeling. Embo J *23*, 2092-2104.

Kassabov, S. R., Zhang, B., Persinger, J., and Bartholomew, B. (2003). SWI/SNF unwraps, slides, and rewraps the nucleosome. Mol Cell *11*, 391-403.

Kayne, P. S., Kim, U. J., Han, M., Mullen, J. R., Yoshizaki, F., and Grunstein, M. (1988). Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast. Cell *55*, 27-39.

Kielkopf, C. L., Baird, E. E., Dervan, P. B., and Rees, D. C. (1998a). Structural basis for G.C recognition in the DNA minor groove. Nat Struct Biol *5*, 104-109.

Kielkopf, C. L., Bremer, R. E., White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., Dervan, P. B., and Rees, D. C. (2000). Structural effects of DNA sequence on T.A

recognition by hydroxypyrrole/pyrrole pairs in the minor groove. J Mol Biol *295*, 557-567.

Kielkopf, C. L., White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., Dervan, P. B., and Rees, D. C. (1998b). A structural basis for recognition of A.T and T.A base pairs in the minor groove of B-DNA. Science *282*, 111-115.

Kingston, R. E., and Narlikar, G. J. (1999). ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. Genes Dev *13*, 2339-2352.

Kleywegt, G. J. (1996). Use of non-crystallographic symmetry in protein structure refinement. Acta Crystallog sect *D 52*, 842-857.

Kleywegt, G. J., and Jones, T. A. (1998). Databases in protein crystallography. Acta Crystallogr D Biol Crystallogr *54*, 1119-1131.

Korber, P., and Horz, W. (2004). SWRred not shaken; mixing the histones. Cell *117*, 5-7.

Kornberg, R. D. (1977). Structure of Chromatin. Annu Rev Biochem *46*, 931-964.

Litt, M. D., Simpson, M., Gaszner, M., Allis, C. D., and Felsenfeld, G. (2001). Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. Science *293*, 2453-2455.

Kornberg, R. D., and Lorch, Y. (1991). Irresistible force meets immovable object: transcription and the nucleosome. Cell *67*, 833-836.

Kriminski, S., Kazmierczak, M., and Thorne, R. E. (2003). Heat transfer from protein crystals: implications for flash-cooling and X-ray beam heating. Acta Crystallogr D Biol Crystallogr *59*, 697-708.

La Fortelle, E. d., and Bricogne, G. (1997). SHARP: A maximum-likelihood heavy-atom parameter refinement program for the MIR and MAD methods. Methods Enzymol *276*, 472-494.

Langst, G., and Becker, P. B. (2004). Nucleosome remodeling: one mechanism, many phenomena? Biochim Biophys Acta *1677*, 58-63.

Langst, G., Bonte, E. J., Corona, D. F., and Becker, P. B. (1999). Nucleosome movement by CHRAC and ISWI without disruption or trans- displacement of the histone octamer. Cell *97*, 843-852.

Li, G., and Widom, J. (2004). Nucleosomes facilitate their own invasion. Nat Struct Mol Biol *11*, 763-769.

Ling, X., Harkness, T. A., Schultz, M. C., Fisher-Adams, G., and Grunstein, M. (1996). Yeast histone H3 and H4 amino termini are important for nucleosome assembly in vivo and in vitro: redundant and position-independent functions in assembly but not in gene regulation. Genes Dev *10*, 686-699.

Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Krieger, M., Scott, M. P., Zipursky, L., and Darnell, J. (2003). Molecular Cell Biology, fifth edn (New York: W. H. Freeman and Company).

Lowary, P. T., and Widom, J. (1997). Nucleosome packaging and nucleosome positioning of genomic DNA. Proc Natl Acad Sci U S A *94*, 1183-1188.

Lu, X., and Hansen, J. C. (2003). Revisiting the structure and functions of the linker histone C-terminal tail domain. Biochem Cell Biol *81*, 173-176.

Luger, K. (2003). Structure and dynamic behavior of nucleosomes. Curr Opin Genet Dev *13*, 127-135.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997a). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Luger, K., Rechsteiner, T., and Richmond, T. J. (1997b). Characterization of nucleosome core particles containing histone proteins made in bacteria. JMB *272*, 301-311.

Luger, K., Rechsteiner, T. J., and Richmond, T. J. (1999a). Expression and purification of recombinant histones and nucleosome reconstitution. Methods Mol Biol *119*, 1-16.

Luger, K., Rechsteiner, T. J., and Richmond, T. J. (1999b). Preparation of nucleosome core particle from recombinant histones. Methods Enzymol *304*, 3-19.

Mapp, A. K., Ansari, A. Z., Ptashne, M., and Dervan, P. B. (2000). Activation of gene expression by small molecule transcription factors. Proc Natl Acad Sci U S A *97*, 3930-3935.

Matthews, B. W. (1968). Solvent content of protein crystals. J Mol Biol *33*, 491-497.

Meersseman, G., Pennings, S., and Bradbury, E. M. (1991). Chromatosome positioning on assembled long chromatin. Linker histones affect nucleosome placement on 5 S rDNA. J Mol Biol *220*, 89-100.

Mizuguchi, G., Shen, X., Landry, J., Wu, W. H., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science *303*, 343-348.

Muthurajan, U. M., Bao, Y., Forsberg, L. J., Edayathumangalam, R. S., Dyer, P. N., White, C. L., and Luger, K. (2004). Crystal structures of histone Sin mutant nucleosomes reveal altered protein-DNA interactions. Embo J *23*, 260-271.

Muthurajan, U. M., Park, Y. J., Edayathumangalam, R. S., Suto, R. K., Chakravarthy, S., Dyer, P. N., and Luger, K. (2003). Structure and dynamics of nucleosomal DNA. Biopolymers *68*, 547-556.

Narlikar, G. J., Fan, H. Y., and Kingston, R. E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. Cell *108*, 475-487.

Noma, K., Allis, C. D., and Grewal, S. I. (2001). Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. Science *293*, 1150-1155.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode, Vol 276, Macromolecular Crystallography, part A (New York: Academic Press).

Park, Y. J., Chodaparambil, J. V., Bao, Y., McBryant, S. J., and Luger, K. (2004a). Nucleosome assembly protein 1 exchanges histone H2A-H2B dimers and assists nucleosome sliding. J Biol Chem.

Park, J. H., Cosgrove, M. S., Youngman, E., Wolberger, C., and Boeke, J. D. (2002). A core nucleosome surface crucial for transcriptional silencing. Nat Genet *32*, 273-279.

Park, Y. J., Dyer, P. N., Tremethick, D. J., and Luger, K. (2004b). A new FRET approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome. J Biol Chem.

Pennings, S., Meersseman, G., and Bradbury, E. M. (1994). Linker histones H1 and H5 prevent the mobility of positioned nucleosomes. Proc Natl Acad Sci U S A *91*, 10275-10279.

Pflugrath, J. W. (2004). Macromolecular cryocrystallography-methods for cooling and mounting protein crystals at cryogenic temperatures. Methods *34*, 415-423.

Polach, K. J., Lowary, P. T., and Widom, J. (2000). Effects of core histone tail domains on the equilibrium constants for dynamic DNA site accessibility in nucleosomes. J Mol Biol *298*, 211-223.

Richmond, T. J., and Davey, C. A. (2003). The structure of DNA in the nucleosome core. Nature *423*, 145-150.

Richmond, T. J., and Widom, J. (2000). Nucleosome and chromatin structure, In Chromatin structure and gene expression, S. C. R. Elgin, and J. L. Workman, eds. (New York: Oxford University Press Inc.), pp. 1-23.

Sargent, D. F., and Richmond, T. J. (2004). Comments on flash-cooling of macromolecular crystals to overcome increased mosaicity by Yao, Yasutake & Tanaka (2004). Acta Crystallogr D Biol Crystallogr *60*, 616.

Schuster, T., Han, M., and Grunstein, M. (1986). Yeast histone H2A and H2B amino termini have interchangeable functions. Cell *45*, 445-451.

Schwarz, P. M., Felthauser, A., Fletcher, T. M., and Hansen, J. C. (1996). Reversible oligonucleosome self-association: dependence on divalent cations and core histone tail domains. Biochemistry *35*, 4009-4015.

Strahl, B. D., and Allis, C. D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Srinivasan, A. R., and Olson, W. K. (1992). DNA associations: packing calculations in A-, B-, and Z-DNA structures. Biophys Chem *43*, 279-310.

Stunkel, W., Kober, I., and Seifart, K. H. (1997). A nucleosome positioned in the distal promoter region activates transcription of the human U6 gene. Mol Cell Biol *17*, 4397-4405.

Suto, R. K., Clarkson, M. J., Tremethick, D. J., and Luger, K. (2000). Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. Nat Struct Biol *7*, 1121-1124.

Suto, R. K., Edayathumangalam, R. S., White, C. L., Melander, C., Gottesfeld, J. M., Dervan, P. B., and Luger, K. (2003). Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. J Mol Biol *326*, 371-380.

Teng, T.-Y., and Moffat, K. (1998). Cooling rates during flash cooling. J Appl Cryst *31*, 252-257.

Thastrom, A., Bingham, L. M., and Widom, J. (2004). Nucleosomal locations of dominant DNA sequence motifs for histone-DNA interactions and nucleosome positioning. J Mol Biol *338*, 695-709.

Thiriet, C., and Hayes, J. J. (1997). Antisera directed against anti-histone H4 antibodies recognize linker histones. Novel immunological probes to detect histone interactions. J Biol Chem *272*, 18740-18745.

Trauger, J. W., Baird, E. E., and Dervan, P. B. (1996). Recognition of DNA by designed ligands at subnanomolar concentrations. Nature *382*, 559-561.

Trauger, J. W., and Dervan, P. B. (2001). Footprinting methods for analysis of pyrrole-imidazole polyamide/DNA complexes. Methods Enzymol *340*, 450-466.

Tse, C., Sera, T., Wolffe, A. P., and Hansen, J. C. (1998). Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. Mol Cell Biol *18*, 4629-4638.

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

Vernede, X., and Fontecilla-Camps, J. C. (1999). A method to stabilize reduced and/or gas-treated protein crystals by flash-cooling under a controlled atmosphere. J Appl Cryst *32*, 505-509.

Vettese-Dadey, M., Grant, P. A., Hebbes, T. R., Crane- Robinson, C., Allis, C. D., and Workman, J. L. (1996). Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. EMBO J *15*, 2508-2518.

Vitolo, J. M., Thiriet, C., and Hayes, J. J. (2000). The H3-H4 N-terminal tail domains are the primary mediators of transcription factor IIIA access to 5S DNA within a nucleosome. Mol Cell Biol *20*, 2167-2175.

Wells, D., and Brown, D. (1991). Histone and histone gene compilation and alignment update. Nucleic Acids Res *25*, 2173-2188.

Wemmer, D. E. (2000). Designed sequence-specific minor groove ligands. Annu Rev Biophys Biomol Struct *29*, 439-461.

Weyermann, P., and Dervan, P. B. (2002). Recognition of ten base pairs of DNA by head-to-head hairpin dimers. J Am Chem Soc *124*, 6872-6878.

White, C. L., and Luger, K. (2004). Defined structural changes occur in a nucleosome upon Amt1 transcription factor binding. J Mol Biol *342*, 1391-1402.

White, C. L., Suto, R. K., and Luger, K. (2001). Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. Embo J *20*, 5207-5218.

White, S., Baird, E. E., and Dervan, P. B. (1997). On the pairing rules for recognition in the minor groove of DNA by pyrrole-imidazole polyamides. Chem Biol *4*, 569-578.

Whitehouse, I., Flaus, A., Cairns, B. R., White, M. F., Workman, J. L., and Owen-Hughes, T. (1999). Nucleosome mobilization catalysed by the yeast SWI/SNF complex. Nature *400*, 784-787.

Winn, M. D., Murshudov, G. N., and Papiz, M. Z. (2003). Macromolecular TLS refinement in REFMAC at moderate resolutions. Methods Enzymol *374*, 300-321.

Zheng, C., and Hayes, J. J. (2003). Structures and interactions of the core histone tail domains. Biopolymers *68*, 539-546.