Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Asymmetries in the nucleosome core particle at 2.5 Å resolution

Joel M. Harp,[a] B. Leif Hanson,[a] David E. Timm[c] and Gerard J. Bunick[a,b]*

[a]University of Tennessee/Oak Ridge Graduate Program for Genome Sciences and Technology, Oak Ridge National Laboratory, Oak Ridge, TN 37831-8080, USA, [b]Life Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-8080, USA, and [c]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN 46202, USA

Correspondence e-mail: bunick@bio.ornl.gov

Received 23 July 2000
Accepted 29 August 2000

PDB Reference: nucleosome core particle, 1eqz.

The 2.5 Å X-ray crystal structure of the nucleosome core particle presented here provides significant additions to the understanding of the nucleosome, the fundamental unit of chromatin structure. Extensions are made to the structure of the N-terminal histone tails and details are provided on hydration and ion binding. The structure is composed of twofold symmetric molecules, native chicken histone octamer cores and the DNA palindrome, which were expected to form a perfectly twofold symmetric nucleosome core particle. In fact, the result is asymmetric owing to the binding of the DNA to the protein surface and to the packing of the particles in the crystal lattice. An analysis is made of the asymmetries by comparisons both within the nucleosome core particle and to the structure of the histone octamer core of the nucleosome.

1. Introduction

The nucleosome core particle (NCP) is the crystallizable portion of the nucleosome, the fundamental building block of chromatin, and provides the framework on which the processes of genetic modulation occur. In chromatin, structure is function. High-quality models of the molecular interactions in genetic processes are reliant on the quality of the NCP structural model. Heuristic features of the NCP and its role in gene modulation are also enhanced with better quality structural information. Thus, continued improvement of the model for the NCP is a necessary element of molecular chromatin studies. This article describes the structure of the NCP to a resolution of 2.5 Å.

The NCP is a substructure of the canonical nucleosome defined by the DNA protection pattern of histone octamer in nuclease digestion of chromatin. It contains the histone octamer core with two each of the core histones H2A, H2B, H3 and H4 and 145–147 base pairs of DNA wrapped in about 1.75 superhelical turns around the histone core. The NCP lacks the linker DNA and a linker histone such as H1. The DNA and protein constituents of the NCP contribute nearly equal molecular weight to the final 205 kDa complex. The crystal structure of the nucleosome, as well as the histone core of the nucleosome, has been the goal of only a few groups worldwide for more than 20 years. Success has required solution of a number of difficult problems, not the least of which involved engineering NCPs capable of forming crystals that diffract X-rays to a high resolution. A brief recapitulation of the history of NCP structural studies may be useful to place the current structure within the context of a long and often contentious chapter in chromatin research.

The existence of the nucleosome as the defining feature of chromatin is now firmly established. It is instructive to reflect

that before 1971 the possibility of such a structure had not occurred to researchers at the forefront of the field. Before the discovery of the nucleosome, the uniform supercoil model was generally accepted. In this picture of chromatin, the basic histones were seen as uniformly coating the acidic polyelectrolyte structure of the DNA, thereby allowing it to pack tightly into the nucleus. The observations of a ladder of particles of unit sizes in nuclease digestions of chromatin as well as the establishment of the true stoichiometry of the histones led to the realization that the eukaryotic genome was much more than a rococo version of the better known and more tractable bacterial systems. Eukaryotic DNA is organized according to an algorithm for the maintenance and management of genomes whose beauty and complexity can only now be appreciated.

The nucleosome as the repeating packaging motif of chromatin was established by biochemical studies (Kornberg, 1974; Kornberg & Lorch, 1999) and strikingly demonstrated in the electron micrographs of the Olinses and Woodcock (Olins & Olins, 1973, 1974; Woodcock, 1973). The existence of the nucleosome as well as its basic features were accepted by nearly all of the participants of the Cold Spring Harbor Symposium in 1977 (Chambon, 1978). After that initial burst of discovery and excitement, laboratories around the world embarked on a quest for the structure of the nucleosome and its role in packaging DNA into the higher order structures of the chromatin fiber and the condensed metaphase chromosome. Early structural studies resulted in neutron (Bentley *et al.*, 1984) and low-resolution X-ray structures (Finch *et al.*, 1981; Richmond *et al.*, 1984; Uberbacher & Bunick, 1989) based on crystals containing NCPs purified from nuclease digests of soluble chromatin. Those structures provided limited information concerning the dimensions and overall shape of the nucleosome, the path of DNA around the histone core and the existence of protein density extending out between the gyres of DNA. Important details of the structure required examination of higher resolution structures of the histone core and of engineered NCPs.

The quest for a high-resolution structure of the nucleosome has seen three important milestones: the 3.1 Å structure of the histone core of the nucleosome (Arents *et al.*, 1991), the development by this laboratory of the human X chromosome α-satellite palindromic DNA to extend the twofold symmetry to the NCP and avoid potential packing disorder in the crystals (Harp *et al.*, 1996) and the use of recombinant histones to facilitate heavy-atom incorporation in the nucleosome core particle (Luger *et al.*, 1997). The determination of the crystal structure of the nucleosome consumed such a long period of time in a large part because of the packaging function of the structure. The histone octamer core of the nucleosome does not bind DNA in a sequence-specific fashion. The octamer forms a superhelical ramp capable of bending duplex DNA into a compact structure. High-resolution crystallographic studies require that NCPs should possess a degree of uniformity in structure not commonly seen in nature. In those cases where nucleosomes are known to bind at precise locations determined by DNA sequence, those positions are generally involved in gene regulation or formation of nucleosome arrays essential to development of higher order structures.

In addition to the structurally well defined core, the histones contain regions of indeterminate structure in the N-terminal region of each histone and in the C-terminus of H2A. These regions without secondary structure represent about 30% of the amino-acid sequence of the histones. In the 3.1 Å histone octamer structure, 25% of the molecule, all from these tail regions, had no interpretable electron density. These features, in addition to the moderately high solvent content, result in weakly diffracting NCP crystals. As such, atomic position refinement of the NCP was virtually intractable prior to macromolecular cryocrystallography and the advent of synchrotron beamlines. Any feature that would help to structurally stabilize the NCP became the target of investigation as an aid to engineering better NCP crystals. The special contribution of this group was the DNA.

The location of a nucleosome along a length of DNA is known as phasing. It is now known that particular DNA sequence motifs contribute to the ability of a given segment of DNA to phase nucleosome positions. The phasing of nucleosomes is largely dependent on the predisposition of the DNA to bend around the histone octamer core. Statistical analysis of nucleosome phasing sequences has shown a preference for particular dinucleotide pairs flanking points at which the minor groove of the DNA faces in to the histone octamer surface, the minor groove-in position. In particular, AA dinucleotide pairs 5′ to the minor groove-in position tend to narrow the minor groove and may impart a bend to the DNA structure. When these bending regions are distributed in phase with the helical repeat of the duplex DNA, the result is a directional bending of the DNA. This is the basis of the rotational phasing of DNA to the histone octamer. A translational phase is imparted as a consequence of more subtle variations in the distribution of favored dinucleotide steps along the DNA path around the histone core.

A computer algorithm was developed and used to identify candidate sequences (Uberbacher *et al.*, 1988). As a result of that analysis, an especially rich source of nucleosome phasing DNA sequence candidates was found in α-satellite DNA. A clone of a *Bam*HI repeat from the human X chromosome characterized in the laboratory of Dr Barbara Hamkalo (Yang *et al.*, 1982) was obtained through Dr Loren Hauser, who had sequenced the clone as part of his dissertation research (Hauser, 1985). It was known to possess 12 nucleosome phasing regions. The nucleosome positions were successfully predicted by the computer algorithm and were found to be equal in predicted ability to phase nucleosomes.

The concept of using DNA palindromes to engineer NCP crystals was developed in this laboratory to ameliorate the potential twofold packing disorder introduced into the crystal packing when the DNA on either side of the NCP is different. Others have also used this approach (Luger *et al.*, 1997). The use of the DNA palindrome served to extend the twofold symmetry of the histone octamer core to the entire NCP. One nucleosome phasing sequence provided a means for subcloning one half of the sequence as a DNA palindrome.

The half-nucleosome sequence chosen out of 24 possible half-nucleosome candidates in the *Bam*HI repeat was a matter of convenience owing to the presence of an *Alu*I recognition sequence adjacent to the center of the nucleosomal sequence. The palindrome nucleosome phasing sequence was developed prior to widespread availability of PCR technology, so that the construction involved cutting at the *Alu*I site and ligation of a synthetic *Eco*RI linker. The use of a synthetic linker resulted in a change of a single base pair from the native sequence flanking the *Eco*RI recognition sequence.

Development of the α-satellite DNA palindrome was of critical importance to the production of high-quality NCP crystals. Our group was able to obtain well diffracting crystals of the NCP with the α-satellite DNA palindrome prior to 1993 with data that indicated reduced twofold packing disorder and higher resolution. A preliminary diffraction analysis of those crystals was eventually published (Harp *et al.*, 1996) in *Acta Crystallographica Section D: Biological Crystallography* following an unfortunate series of interactions with other journals. During this period, additional opportunities arose which provided the advantage of microgravity for crystal growth and the use of synchrotron radiation to greatly extend diffraction resolution. Both conditions were useful in maximizing diffraction from NCP crystals.

We present here the NCP structure reconstituted from native chicken histones and the α-satellite DNA palindrome. This structure includes numerous well defined solvent molecules and ions previously unreported in atomic models of the NCP. An atomic resolution map of the nucleosome core particle is essential to the investigation of a number of fundamental issues in eukaryotic molecular biology. Mechanisms of regulation of transcription in eukaryotes are tied directly to the necessity of transcribing through DNA bound in nucleosomes, the so-called nucleosome problem of transcription (Studitsky, 1997). Little is known of DNA damage recognition and repair in a chromatin context and the structure of the nucleosome core particle with its details of DNA binding to the histone core provide an excellent basis for such studies. The architecture of the chromosome is also known to depend on nucleosome modifications such as H3-like variants found in nucleosomes in the centromere.

This is a time of great advances in the study of chromatin owing to the emergence of evidence linking chromatin structure as an essential factor in the regulation of chromatin function (Widom, 1998). The structures of the histone octamer and the nucleosome core particle discussed here provide critical insights into structure–function relationships in chromatin. The phasing of DNA on the histone octamer core has been the subject of considerable research, resulting in the prediction that the binding of a twofold symmetric 146 base pair DNA palindrome to the twofold symmetric histone octamer would result in a twofold symmetric nucleosome core particle. The resulting structure is, in fact, asymmetric. There exists asymmetry in the structure of the DNA as well as the structures of the histones. Some of the asymmetries appear to arise from the binding of the DNA to the octamer core of the nucleosome. Others are the result of lattice packing. It would appear *a priori* that the nucleosome core particle could pack into the lattice with an equal probability of a given face of the particle facing a particular direction. The fact that the asymmetric NCPs assume only one orientation suggests that the two faces of the nucleosome are unique within the lattice. An analysis of the structure with regard to the deviations from symmetry has been made by comparison of the NCP structure with the structure of the histone octamer utilizing a low-temperature structure derived from the same histone octamers used for reconstitution of the NCPs.

2. Abbreviations

NCP, nucleosome core particle.

MPD, 2-methyl-2,4-pentanediol.

mGI, minor groove-in position, a point at which the DNA minor groove faces the surface of the histone protein core of the nucleosome.

mH, the median α-helix of the histone fold motif.

CH, the C-terminal α-helix of the histone fold motif.

NH, the N-terminal α-helix of the histone fold motif.

NL, loop connecting the NH and mH α-helices of the histone fold motif.

CL, loop connecting the CH and mH α-helices of the histone fold motif.

CaH, accessory helix C-terminal to the histone fold domain.

NaH, accessory helix N-terminal to the histone fold domain.

Ventral, that side of the NCP containing the DNA half with 72 base pairs lateral to the central base pair.

Dorsal, that side of the NCP containing the DNA half with 73 base pairs lateral to the central base pair.

DCAM, diffusion-controlled crystallization apparatus for microgravity.

PIP, di-μ-iodobis(ethylenediamine)diplatinum.

3. Methods

3.1. Purification of histones

Histone octamers were prepared from chicken erythrocyte nuclei by gentle salt extraction of histones from soluble chromatin bound to ceramic hydroxyapatite using established protocols (Lutter, 1978; von Holt *et al.*, 1989) and presented in detail in an earlier publication (Harp *et al.*, 1996). Packed erythrocytes were obtained fresh by bleeding chickens immediately prior to the preparation. Erythrocytes were lyzed using a non-ionic detergent to isolate nuclei. Soluble chromatin was prepared by micrococcal nuclease digestion of whole nuclei followed by extraction in low ionic strength buffer. Histones were extracted by first removing linker histones using 0.7 *M* NaCl, 20 m*M* Tris–HCl pH 8.0 and then eluting the core histones with 3 *M* NaCl, 20 m*M* Tris–HCl pH 8.0. The core histones eluted in two peaks corresponding to the H2A–H2B dimer and the $(H3–H4)_2$ tetramer. The peaks were pooled and the octamer reassembled by dialysis against 2 *M* NaCl, 20 m*M* Tris–HCl pH 8.0 and finally purified by gel-filtration chromatography using a Superose 12 column

(Pharmacia-LKB). Care was taken to ensure the correct stoichiometry of the histone subunits in the final histone octamer preparation (Harp *et al.*, 1996).

3.2. Production of 146 base pair DNA palindrome

Large-scale production of the DNA palindrome required using a plasmid containing multiple direct repeats of the half-palindrome fragment. Repeated attempts to clone multiple copies of the full-length 146 bp DNA palindrome were unsuccessful. It was further found that large-scale production using a plasmid containing a single copy of the palindrome was not capable of providing the amounts of DNA required. The DNA used in this study was developed and produced in quantity as previously described (Palmer *et al.*, 1996; Harp *et al.*, 1996). Briefly, the half-palindrome *Eco*RI/*Eco*RV fragment was produced in quantities of tens of milligrams. A large-scale ligation reaction with the half-palindrome fragment followed by recutting with *Eco*RV resulted in the full 146 bp DNA palindrome with the *Eco*RI site at the center of sequence symmetry.

3.3. Reconstitution of nucleosome core particles

Reconstitution of nucleosome core particles from purified histone octamer and α-satellite DNA palindrome was accomplished by slow salt-gradient dialysis as previously described (Harp *et al.*, 1996) using standard techniques (Rhodes *et al.*, 1989). The DNA and protein input ratios were established by small-scale pilot reactions by reconstitution on a microgram scale prior to committing materials to large decigram-scale reactions. Pilot reactions were performed by mixing the protein and DNA components in a 2 µl volume in 1 *M* NaCl, 0.089 *M* Tris–borate buffer pH 8.0, 0.01% Nonidet P40 and diluting the mixture in steps to 200 µl. Reactions were analyzed by non-denaturing polyacrylamide electrophoresis to determine the correct input ratio of protein to DNA.

3.4. Purification of correctly phased nucleosome core particles

Reconstitutions of nucleosome core particles resulted in preparations containing predominantly the correctly phased NCPs with the DNA bound symmetrically to the histone octamer core. However, there were usually some of the incorrectly phased particles in which the DNA was bound asymmetrically. The phasing was generally quantized such that distinct species were visualized in non-denaturing polyacrylamide gels. The incorrectly phased NCPs were separated using preparative non-denaturing polyacrylamide gel electrophoresis to provide strictly homogeneous preparations of symmetrically phased NCP with correct histone stoichiometry (Uberbacher & Bunick, 1986; Harp *et al.*, 1995).

3.5. Crystallization

Crystallization was performed using the diffusion-controlled crystallization apparatus for microgravity (DCAM; Carter, Wright *et al.*, 1999) to take advantage of the microgravity environment in order to improve crystal morphology and quality. NCP preparations were loaded into DCAMs 1–2 d prior to launch for microgravity and transferred to NASA for loading on the Shuttle. The DCAM apparatus possessed two buffer chambers separated by a septum containing an agarose plug to prevent mixing of the buffers while allowing diffusion between the chambers. A 45 µl crystallization button was molded into the end plate sealing one of the chambers. The NCP preparation was placed in the crystallization button and covered with dialysis membrane which was secured using an O-ring. The chamber containing the button was filled with buffer containing 50 m*M* KCl, 10 m*M* potassium cacodylate, 1 m*M* PMSF and a high concentration of $MnCl_2$. The opposite chamber was filled with 50 m*M* KCl, 10 m*M* potassium cacodylate, 1 m*M* PMSF and a low concentration of $MnCl_2$. Because diffusion commenced immediately upon assembly of the DCAM, a major concern in the design of microgravity experiments was the delay between assembly of the apparatus on the ground and its arrival in the microgravity environment. However, since crystallization of the NCP requires lowering of the $MnCl_2$ concentration rather than an increasing precipitant concentration, it was possible to employ a 'time-delay fuse' by preparing the agarose plug separating the two chambers with a very high (200 m*M*) concentration of $MnCl_2$. With the concentration of $MnCl_2$ in the agarose plug much higher than that in either chamber, several days elapsed before an Mn^{2+} concentration gradient began to develop between the two chambers.

3.6. Data collection

Crystals for data collection were harvested from crystallization buttons directly into stabilization buffer containing 30 m*M* KCl, 30 m*M* $MnCl_2$, 10 m*M* potassium cacodylate pH 6.0, 2% 2-methyl-2,4-pentanediol (MPD). The concentration of MPD was increased in steps by addition of buffer containing 30 m*M* KCl, 30 m*M* $MnCl_2$, 10 m*M* potassium cacodylate pH 6.0 and 50% MPD. The amount of buffer containing 50% MPD added at each step was adjusted to provide a 2% increase in MPD concentration and a minimum of 12 h was allowed for equilibration of the crystals before the next MPD addition. The concentration of MPD for data collection was 22.5% adjusted using refractometry to verify the final value. Crystals were mounted on rayon loops in magnetic caps (Hampton Research) and flash-cooled directly in the cold nitrogen-gas stream of a cryostat operating at 100 K. Crystals exhibiting high mosaicity were annealed using the macromolecular crystal-annealing protocol (Harp *et al.*, 1998, 1999).

A heavy-atom derivative was prepared by soaking an NCP crystal in a solution of di-μ-iodobis(ethylenediamine)-diplatinum (PIP; O'Halloran *et al.*, 1987) in thoroughly degassed buffer in the dark.

3.7. Phasing

Phases were obtained from three sources. Experimental phases were obtained by single isomorphous phasing using data from a PIP-derivative NCP crystal. The SIR method

(Wang, 1985) was used as implemented in the *PHASES* program package (Furey, 1997) and the heavy-atom position refined and phasing statistics calculated using *MLPHARE* (Otwinowski & Minor, 1997). Molecular replacement was performed using *CNS* (Brunger *et al.*, 1998). Models for molecular replacement were the histone octamer (PDB code 2hio; Arents *et al.*, 1991) and the nucleosome core particle (PDB code 1aoi; Luger *et al.*, 1997).

3.8. Model fitting and refinement

Model fitting was performed using *TURBO-FRODO* and *XtalView* (McRee, 1999) using σ_A-weighted electron-density maps calculated using either *XtalView* or *CNS* with Fourier coefficients $2mF_o - DF_c$ or $mF_o - F_c$. SIR and MR phases were combined using the *SIGMAA* program (Read, 1986) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Refinements were performed using *CNS* with cross-validation throughout (Brünger, 1992). The simulated-annealing protocol used the maximum-likelihood target (Adams *et al.*, 1997) and bulk-solvent correction along with individual temperature-factor refinement between each cycle of fitting. Water peaks were identified and evaluated using the *DDQ* program (van den Akker & Hol, 1999). Water molecules which refined to an individual temperature factor significantly lower than the surrounding protein or DNA atoms were replaced as ions by consideration of chemical environment and evaluated by refinement to ascertain changes in temperature-factor and real-space *R* factor. Secondary structure was evaluated using the program *DSSP* (Kabsch & Sander, 1983) and by manual inspection of hydrogen-bonding patterns and interatomic distances.

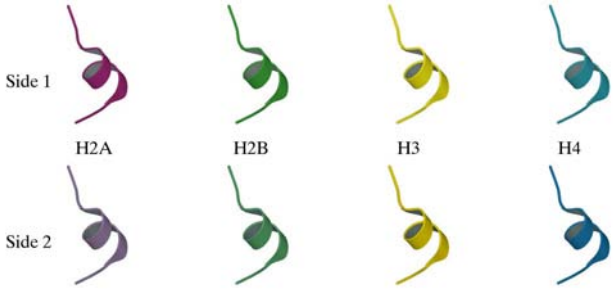

Figure 1
Color code for identification of histone subunits in the figures of this paper. Unless otherwise indicated, all figures conform to this code. Note the change in hue denoting the two sides of the histone octamer.

3.9. Preparation of figures

Figures were prepared using *TURBO-FRODO* (Roussel & Cambillau, 1991), *GRASP* (Nicholls *et al.*, 1991), *MOLSCRIPT* (Kraulis, 1991), *Raster3D* (Merritt & Murphy, 1994), *Rasmol* (Sayle & Milner-White, 1995), and the *Raster3D* interface from the *xfit* program in the *XtalView* package (McRee, 1999). A color code was prepared to aid recognition of specific histone subunits (Fig. 1).

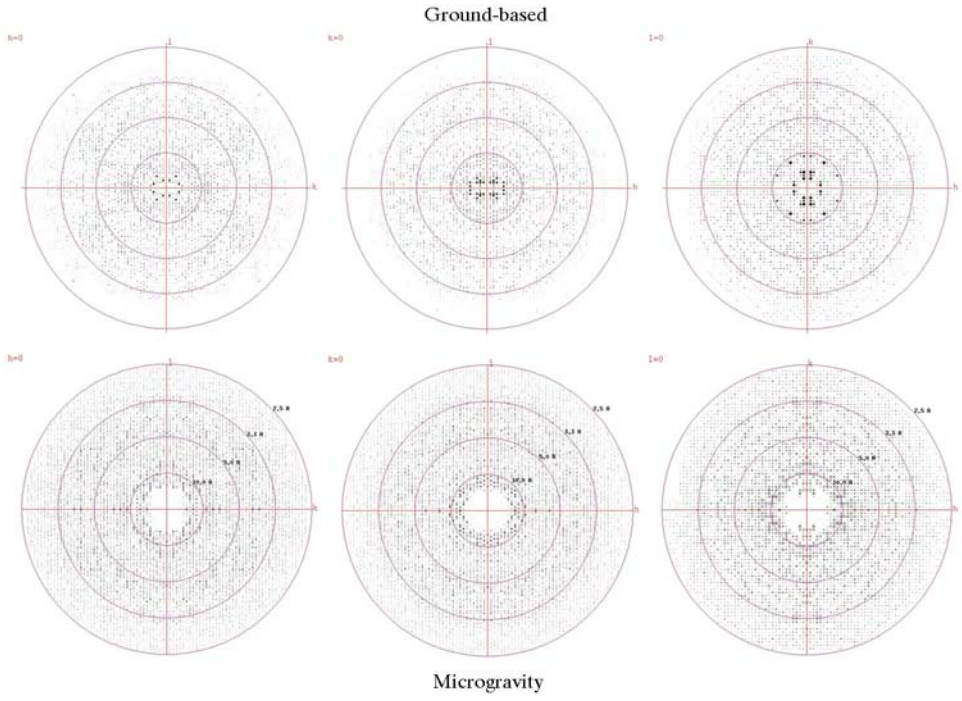

Figure 2
A comparison of microgravity *versus* ground-based diffraction data. Pseudoprecession images, generated by the *HKLview* program from the *CCP*4 program suite, illustrate completeness and degree of anisotropy in the diffraction from both crystals. Note the higher resolution of the microgravity data. The microgravity data were from an NCP crystal grown on the USML-2 mission and were collected at beamline X12-C at NSLS. The ground-based data are representative of crystals grown on the ground and were collected using the same synchrotron beamline.

4. Results and discussion

4.1. Crystallization and crystal quality

A crystal grown on NASA Shuttle mission USML-2 was used to collect native data on beamline X12C at the National Synchrotron Light Source. The crystal grew in the orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 105.28$, $b = 109.71$, $c = 181.12$ Å. The data were collected to a resolution of 2.5 Å. With limited synchrotron beam time, the 2.5 Å resolution limit used in this study was selected, although reflections were seen well beyond 2.5 Å resolution. No ground-based crystals provided data of the same quality, primarily because of anisotropy in ground-based crystals but also in overall diffraction quality. Fig. 2 provides a graphical illustration of the difference in microgravity *versus* ground-based crystals of the nucleosome core particle using pseudoprecession

photographs. The ground-based crystal was grown using a DCAM in the laboratory and provided the best quality data recorded from ground-based crystals during these studies. It was flash-cooled and annealed (Harp *et al.*, 1998, 1999) and data were collected using beamline X12C at the National Synchrotron Light Source. As is evident in Fig. 2, the microgravity crystal does not exhibit the degree of anisotropy typically seen in ground-based crystals of the NCP. The microgravity crystal used in this study flash-cooled quite well and was used before the phenomenon of macromolecular crystal annealing had been discovered. As will be discussed later in relation to the solvent environment of the crystals, the NCP crystals were soaked in 22.5% MPD to improve diffraction quality. This level of MPD provides a marginal cryoprotection for flash-cooling so that, prior to the development of macromolecular crystal annealing, many crystals were required to obtain an adequate flash-cooling without a significant increase in mosaicity.

Diffraction statistics for the microgravity crystal and for the PIP derivative are provided in Table 1. The PIP derivative crystal data were collected in-house using a Rigaku HU-200 X-ray generator and an 18 cm MAR Research image-plate area detector. The crystal was flash-cooled to 100 K using a Molecular Structure Corporation (The Woodlands, Texas, USA) transfer-tube cryostat.

Table 1
Crystallographic diffraction and phasing statistics.

	Native	PIP
Observations	605950	379171
Unique reflections	72200	32814
Completeness (%)	98.7	93.7
Resolution (Å)	25–2.5	40–3.3
Linear *R* factor	0.045	0.095
Cullis *R* factor		0.80
Phasing power		2.00

4.2. Data reduction and model refinement

The experimental phases obtained from SIR and *MLPHARE* refinement were combined with molecular-replacement phases using coordinates from the histone octamer core of the nucleosome (PDB code 2hio; Arents *et al.*, 1991) and the resulting map was used to verify the placement of residues in the MR map. Additional phasing information for fitting of the DNA was provided by molecular replacement using coordinates from the nucleosome core particle structure (PDB code 1aoi). Solvent molecules were placed using the program *DDQ* (van den Akker & Hol, 1999) to locate positive difference density peaks in an $mF_o - F_c$ electron-density map calculated using the model with waters intentionally removed. A σ cutoff of 3 was used and the program utilized hydrogen-bonding distance restraints to differentiate putative solvent peaks from shift peaks produced by errors in placement of atoms in the model or from errors in temperature-factor refinement. Each of the putative water molecules were examined in both the hydrated $mF_o - F_c$ and $2mF_o - DF_c$ maps. Subsequent iterations of the protocol using $mF_o - F_c$ maps calculated with the accepted water molecules located additional water molecules. Refinement in *CNS* of the entire structure with cross-validation and maximum-likelihood refinement targets was completed after each round of interactive fitting using Powell and simulated-annealing energy-minimization protocols. In total, 349 ordered waters have been placed in the NCP model. The ions, represented by both anions and cations, include 15 Mn^{2+}, seven K^+, two Cl^- and one cacodylate ion. Final refinement statistics for the NCP model are presented in Table 2.

4.3. General features of the molecular structure

The model presented here represents a significant extension of the atomic position data for the NCP. We are able to model 152 more amino acids than were seen in the previously published 2.8 Å model (PDB code 1aoi) and the overall *B* factors, especially for the DNA, are significantly lower. All DNA was fitted into observed density and no modeling of ideal structure was performed to complete the structure. The DNA fragment used in this study includes terminal phosphates, making this the longest fragment of DNA for which atomic position data exists. Ions and water molecules are seen in this structure and their modeled positions provide insight into the ordering of larger chromatin arrays.

The engineering of the NCP crystal used all of the constituent features to optimize symmetry within the NCP. Gentle salt extraction of native histone octamers from the mostly inactive

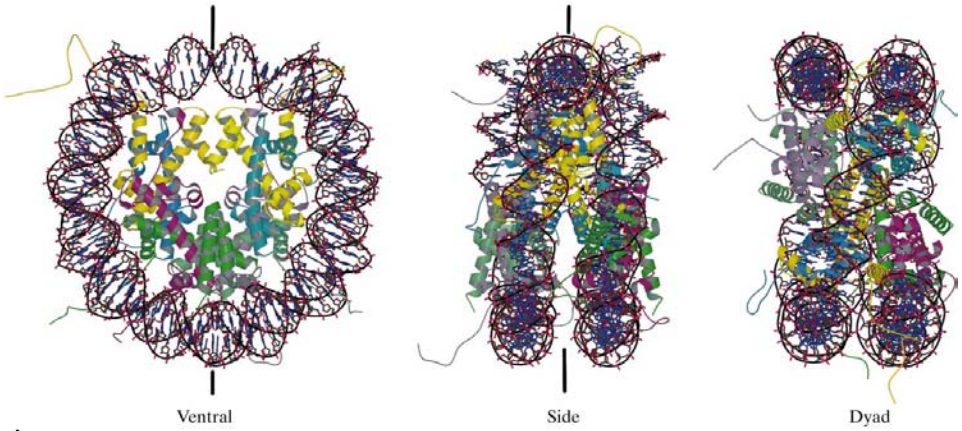

Figure 3
Views of the nucleosome core particle with the protein represented by ribbon models of the secondary-structure elements and DNA indicating base pairing between complementary strands. The pseudo-dyad axis is represented by vertical bars for both the ventral and side view. The pseudo-dyad axis passes through the center of the dyad view orthogonal to the plane of the page.

Table 2
Refinement statistics.

Resolution range (Å)	25–2.5
Observations	72200
Observations in test set	7303
Test-set size (%)	10.1
R value	0.219
Free *R* value	0.282
Estimated error in free *R* value	0.003

nuclei of avian mature erythrocytes has provided a homogeneous population of unmodified histones with unaltered and thus physiologically authentic folding. The use of an engineered DNA palindrome based on a sequence designed by nature to possess strong phasing signals, α-satellite DNA, was intended as a way to extend the twofold symmetry of the histone octamer and result in nucleosome core particle with no dorsal–ventral asymmetry. Possibly because of the necessity of mutation at the dyad to introduce a restriction site, this prediction was off by 0.5 base pairs.

Arbitrary anatomical designations (see Fig. 3) have been made to facilitate discussion of the structure of the NCP. The portion of the molecule where the DNA enters and exits the NCP is referred to as the top of the structure. The dyad axis of the histone core of the NCP divides the octamer into halves, with each half, designated as side 1 and side 2, possessing one each of the histone proteins. Thus, the two subunits of histone H2A are designated as 1H2A or 2H2A depending on whether the subunit is placed on side 1 or side 2, respectively. The molecule is obviously asymmetric, with one side having greater order in the tail regions. Asymmetry is also seen as a consequence of the dyad intersecting DNA base 73. This results in positioning 72 base pairs to the side of the dyad on one face (the ventral face) and 73 base pairs to the other face (the dorsal face). Remarkably, the asymmetries did not result in twofold packing disorder. The fortunate alteration of dorsal and ventral orientations of NCPs in the lattice followed the continuous path of DNA base stacking between NCPs. The asymmetry of the DNA binding to the octamer core in this structure has provided insights into the plasticity of the protein–DNA interactions in the nucleosome, allowing formation of nucleosomes on random sequences of DNA throughout the nucleus. Binding of DNA to the superhelical ramp of the histone octamer is primarily mediated through a series of binding sites constructed around an arginine residue whose side chain inserts into the minor groove of the DNA. These positions mark the closest approach of the minor groove to the protein surface and are here defined as minor groove-in positions or mGIs.

Although the histone octamer structure displays a twofold symmetry about a central axis, the structures of the individual histones in the NCP deviate from symmetry in a number of positions. Least-squares fitting of the octamer moiety of the NCP onto the histone octamer structure indicates clearly that the disposition of the protein subunits deviates from symmetry upon binding with DNA. Fig. 4 shows the displacement of the

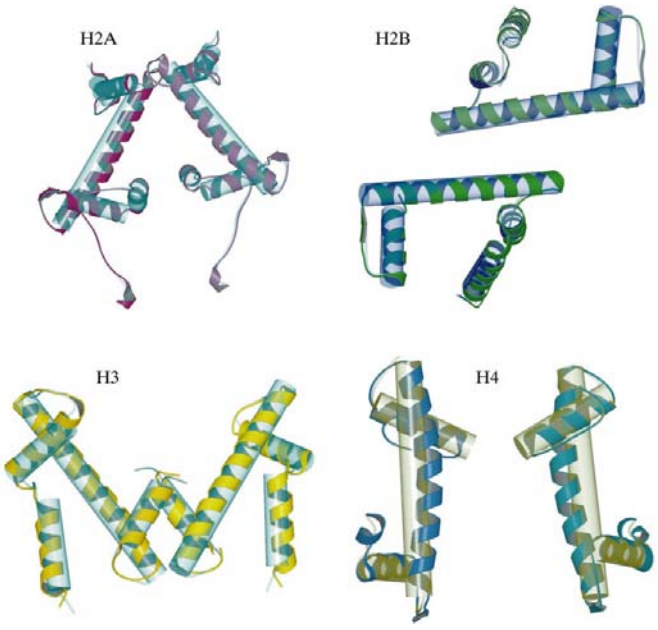

Figure 4
Ribbon models of histone C^{α} atoms showing least-squares fitting of individual histone subunits from the octamer of the NCP compared with the same subunits in the histone octamer without DNA bound. Colored models represent the histone structure within the NCP, while the transparent cylinders represent the boundaries of the low-temperature histone octamer structure. Deviations from symmetry induced by binding of DNA are apparent from the differences in the disposition of helices.

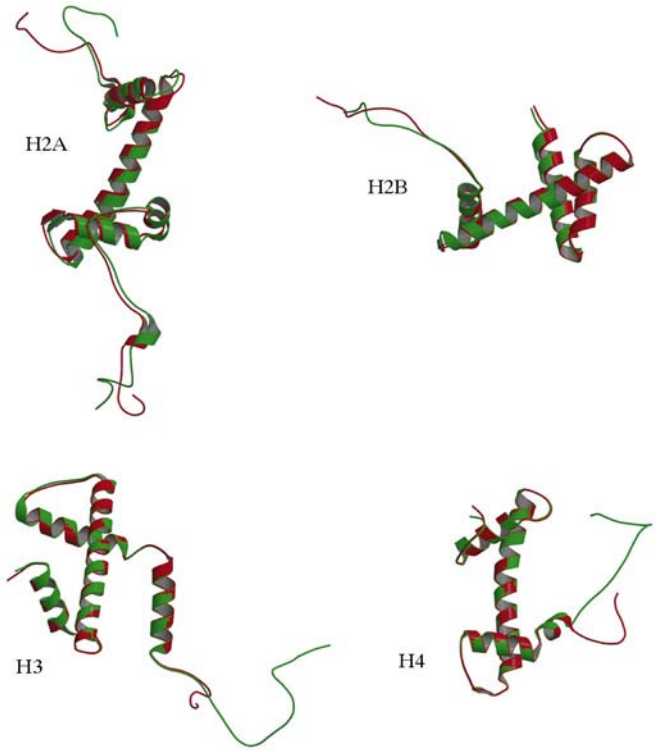

Figure 5
Ribbon models representing a least-squares fitting of homologous histone subunits in the NCP to reveal locations of asymmetry between the two halves of the histone octamer core of the NCP structure.

helices of individual histone subunits from the symmetrical histone octamer. This figure demonstrates some remodeling of the histone structure with DNA bound, particularly with respect to histone H2A and H4. Fig. 5 similarly represents a least-squares fitting of the coordinates of homologous subunits except that only the histones with the NCP structure are used. This more detailed rendering of the structures shows that H3 and H4 are least impacted by DNA binding and that the dimer histones show the largest degree of remodeling. However, as will be discussed later, DNA binding is not the only factor influencing the remodeling seen here. Stacking interactions between the histone faces of adjacent NCPs are largely responsible for the asymmetry in the histone subunit structures.

4.4. The histone fold and the handshake motif

The four core histones each possess a similar central structural design, the histone fold (Arents *et al.*, 1991), possibly produced by the symmetrical duplication of a helix–turn–helix motif (Arents & Moudrianakis, 1995; Ramakrishnan, 1995). The strongly conserved architecture of the histone fold motif consists of a long central helix with shorter helices attached by loop structures at the N-terminal and C-terminal ends. The histone fold is illustrated in Fig. 6. The elements of the histone fold structure therefore consist of three helices: a short helix on the N-terminal side of the symmetry center of the fold (NH) the long median helix (mH) and a short helix on the C-terminal side (CH). The helices are joined by loop NL between the NH helix and the mH helix, and loop CL between the median helix and the CH helix. The short terminal helices are folded back and rotated over the median helix. This structure is seen in a number of DNA-binding proteins belonging to the histone fold family (Baxevanis *et al.*, 1995) which were identified after the first description of the histone fold motif (Arents *et al.*, 1991). The core histones assemble as heterodimers, (H2A–H2B) and (H3–H4), by interdigitation of the terminal helices and overlapping of the central helices as shown in Fig. 7. The heterotetramer, $(H3–H4)_2$, is assembled by formation of a four-helix bundle composed of the mH and CH helices of the two H3s. The tight association of histones in the handshake motif is characterized by extensive hydrophobic contacts between the helices and by a short β-bridge structure formed between the CL loop of H2A and the NL loop of H2B in the (H2A–H2B) heterodimer and between the CL loop of H3 and the NL loop of H4 in the (H3–H4) heterodimer. The β-bridges form part of the primary DNA docking sites on the histone surface which are described later. The close contact of the histones in the heterodimers was described as a handshake motif by Arents *et al.* (1991).

4.5. Structure of the histone subunits

In addition to the basic architecture of the core histones, individual histones may possess accessory helices N-terminal or C-terminal to the histone fold motif. Each of the four core histones possesses an unstructured N-terminal tail which extends beyond the boundaries of the histone octamer structure to make contacts with intraparticle DNA as well as numerous interparticle contacts. An unusual feature of the

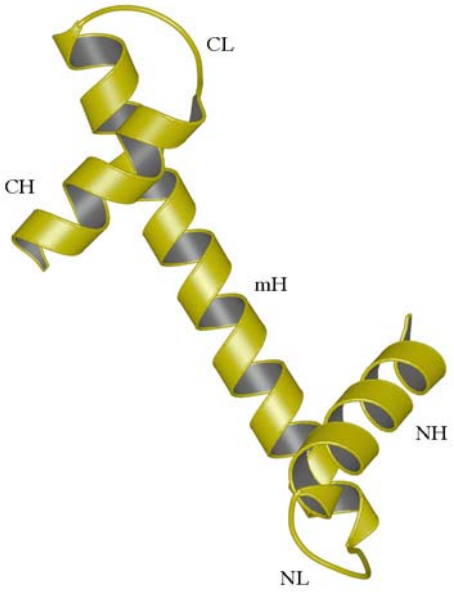

Figure 6
Structure of the histone fold motif using a ribbon C^{α} model of H3 as an exemplar. The histone fold is constructed from three helices: the N-terminal helix (NH), the central or median helix (mH) and the C-terminal helix (CH) joined by loops: the N-terminal loop (NL) and the C-terminal loop (CL).

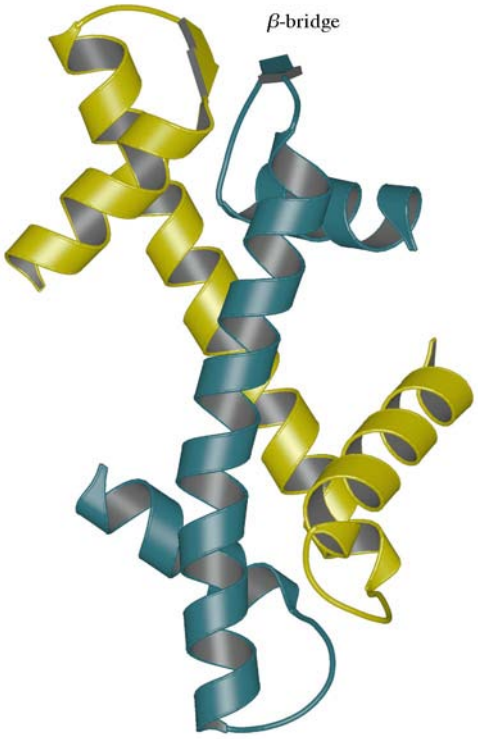

Figure 7
Ribbon C^{α} model of H3–H4 heterodimer showing the two histones interdigitated in a handshake motif. The CL loop of one subunit is involved with the NL loop of the other in a short β-bridge.

NCP crystals is the extent and variety of interparticle contacts made by the N-terminal unstructured tails. Histone H2A also displays an extensive C-terminal tail which, although random coil, is reasonably well ordered within the overall structure of the histone octamer and does not appear to be involved in interparticle interactions.

Alignments of the secondary structures of the individual histone subunits with their respective amino-acid sequences are presented in Figs. 8, 9, 10 and 11. These figures show the extent of the chains observed in the octamer structure 2hio (Arents *et al.*, 1991) which defines the boundaries of the unstructured tail regions of the protein in the absence of stabilizing interactions with bound DNA. The extent of the chains observed in the NCP structure are also indicated. Residues forming the histone fold of each subunit are shown within solid lines and residues within accessory helices are shown inside dashed lines. Each of the eight histones subunits is shown in the figure, although the secondary structure is essentially the same on either side of the pseudodyad of the histone octamer. This is to present and to clarify the observed differences.

The examination of the structure of the histones in the NCP was facilitated by comparisons with the histone octamer structure (PDB code 2hio) and this laboratory's low-temperature histone octamer structure, which will be presented in a separate publication. The low-temperature structure is characterized by higher resolution and details of ion binding. References to the histone structure refer to both the 2hio structure and to our 2.7 Å low-temperature structure.

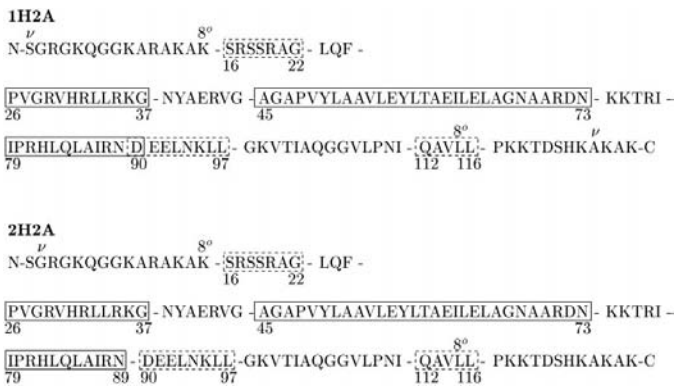

Figure 8
Secondary-structure alignment with the amino-acid sequence of H2A. Helices belonging to the histone fold are within solid lines and accessory helices are within dashed lines. The extent of the chain visible in the NCP structure is denoted by ν over the final residue and the extent of the chain visible in the octamer structure is denoted by 8° over the terminal residue.

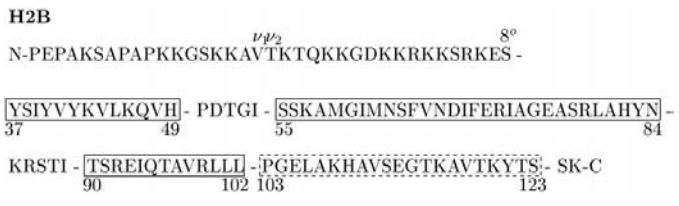

Figure 9
Secondary-structure alignment with amino-acid sequence of histone H2B. Helices belonging to the histone fold are within solid lines and accessory helices are within dashed lines. The N-terminal extent of the chain visible in the NCP structure is denoted by ν_1 for 1H2B and ν_2 for 2H2B. The N-terminal extent of the chain visible in the octamer structure is denoted by 8° over the terminal residue. The C-terminus of the chain is visible in both structures.

H2A possesses a 14-residue unstructured N-terminal tail beyond Lys15 extending around the DNA gyres between mGI 4 and mGI 5 and interacting with both intraparticle DNA and with the N-terminal tail of H2A from the opposite side of a neighboring NCP. That is, 1H2A interacts with the N-terminal tail of 2H2A of a symmetry neighbor within the crystal lattice (see Figs. 12 and 13). The entire tail is visible in the NCP electron-density map for 1H2A and to Gly2 in 2H2A. At the base of the tail an N-terminal accessory helix, NaH, interacts directly with the DNA phosphate backbone.

H2A extends well beyond the histone fold as seen in Fig. 14 and is the only one of the four core histones with a significant C-terminal tail. This tail is visible only to Leu115 in the 2hio histone structure, leaving the 13 C-terminal residues disordered. Histone 1H2A is the only one of the eight protein subunits for which all of the C-terminal residues are not apparent in the NCP electron-density map, with the three C-terminal residues not visible. The H2A chain turns immediately at the C-terminus of the CH helix into a 3_{10} accessory helix, CaH, which then extends across the face of the octamer and into a random-coil structure which serves to bind the (H2A–H2B) dimer to the $(H3–H4)_2$ tetramer. The significance of a 3_{10} accessory helix is unclear but may reflect the fact that the C-terminal tail of H2A is not a disordered region in the same sense as the N-terminal tail domain.

The N-terminal disordered tail of histone H2B extends for 38 residues, with only a few residues N-terminal to helix NH

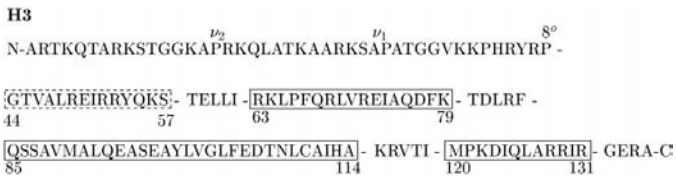

Figure 10
Secondary-structure alignment with amino-acid sequence of H3. Helices belonging to the histone fold are within solid lines and accessory helices are within dashed lines. The N-terminal extent of the chain visible in the NCP structure is denoted by ν_1 for 1H3 and by ν_2 for 2H3. The N-terminal extent of the chain visible in the octamer structure is denoted by 8°. The C-terminus is visible in both structures.

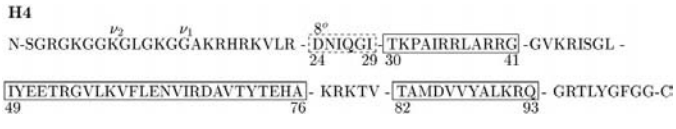

Figure 11
Secondary-structure alignment with amino-acid sequence of H4. Helices belonging to the histone fold are within solid lines and accessory helices are within dashed lines. The N-terminal extent of the chain visible in the NCP structure is denoted by ν_1 for 1H4 and by ν_2 for 2H4. The N-terminal extent of the chain visible in the octamer structure is denoted by 8°. The C-terminus is visible in both structures.

Table 3
Tabulation of minor groove-in, mGI, positions.

Each mGI is identified by a number from 1 to 6 counting outward from the pseudo-dyad and with a letter, *v* for ventral or *d* for dorsal. Each mGI is also identified by the arginine residue inserted into the minor groove of the DNA and by the type of binding motif, α or β, as discussed in the text. The DNA phosphates listed are those positioned closest to the guanidino group of the inserted arginine residue and the base position is the base located above the inserted guanidino group away from the protein surface.

mGI	Arginine	Type	DNA phosphates	Base position
v 6	R77 1H2A	β	278/19	Ade17
v 5	R33 1H2B	α	268/29	Ade27
v 4	R42 1H2A	β	258/39	Thy37
v 3	R83 1H3	β	246/51	Cyt49
v 2	R63 1H3	α	237/60	Gua58
v 1	R45 1H4	β	228/70	Ade67/Gua68
d 1	R45 2H4	β	217/81	Gua78/Cyt79
d 2	R63 2H3	α	207/90	Cyt88
d 3	R83 2H3	β	198–197/200	Gua97/Gua98
d 4	R42 2H2A	β	186/112	Ade109/Ade110
d 5	R33 2H2B	α	175/122	Thy120
d 6	R77 2H2A	β	165/132	Thy130

visible in the octamer structure. The N-terminal tail of H2B extends between DNA gyres, with 17 residues of 1H2B and 18 of 2H2B visible in the NCP electron-density map. The H2B chain bends abruptly past the CH helix and forms a well ordered accessory helix, CaH, which forms a large portion of the face of the histone core (see Fig. 14). The histone fold of H2B interacts with the histone fold of H4 in the primary interaction binding the (H2A–H2B) dimer to the $(H3–H4)_2$ tetramer.

Histone H3 N-terminal tails extend well beyond the DNA gyres near the entry and exit point of the DNA in the NCP. None of the N-terminal tail is visible in the histone octamer structure, while the tail is visible in electron-density maps of the NCP for 14 residues in 1H3 and for 28 residues in 2H3. An N-terminal accessory helix displays significant interactions in stabilizing the DNA termini. Perhaps the most significant interaction of the histone fold of H3 is the binding of 1H3 to 2H3, which forms the foundation of the nucleosome structure.

The N-terminal tails of H4 extend from a 3_{10} N-terminal accessory helix which carries the tail out from the protein face of the NCP. The two H4 chains display a striking asymmetry in the orientation of the N-terminal tail domains with respect to symmetry neighbors. The tail of 1H4 extends to contact the DNA between ventral mGI 2 and mGI 3 (see Fig. 15), with Arg23 making direct contact with a DNA phosphate. The tail then proceeds outwards to make indifferent contacts with 1H2B of a symmetry neighbor. The N-terminal tail domain of 2H4 is involved in an important packing interaction with a different symmetry neighbor by binding to an acidic region on the dimer surface of the neighboring NCP in the crystal lattice. This region is of particular interest as the interaction between the NCP and its neighbor is mediated by a very well ordered cacodylate ion located near the N-terminus of the 3_{10} accessory helix and the C-terminus of the NH helix of H3.

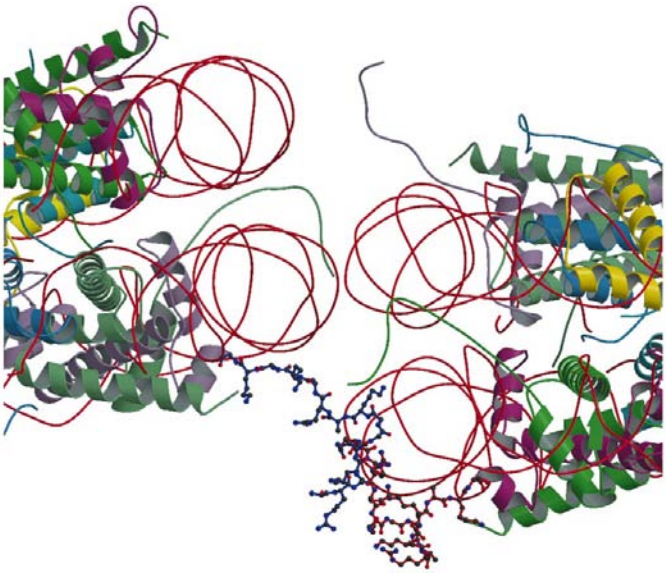

Figure 12
N-terminal tails of H2A are shown as ball-and-stick models to show the interaction of the H2A tails between neighboring NCPs. The C atoms of the N-terminal tail of 2H2A are shown in blue and those of 1H2A are in red. The tail of 2H2A extends across and interacts with the N-terminal tail of 1H2A on the neighboring NCP. At the base of the tail, a short accessory helix (in ribbon rendering) interacts directly with the DNA.

4.6. DNA binding to the octamer

For the purposes of discussion of the structure of the α-satellite DNA palindrome in the NCP, nucleotides were numbered 1–146 5′ to 3′, with nucleotide 1 being the 5′ end of the strand on the side of the NCP here defined as the ventral surface. The 3′ end at nucleotide 146 lies on the dorsal surface of the NCP. The complementary strand is numbered, 5′ to 3′, as nucleotides 147–292, with nucleotide 292 base paired to nucleotide 1. The DNA strand numbered 1–146 is also referred to here as the Watson strand and its complement as the Crick strand.

Binding of DNA to the histone surface is primarily mediated by the insertion of arginine side chains at the minor groove-in position, mGI. The construction of the histone octamer serves to place the mGIs in such a way as to form a left-handed superhelical ramp with the mGIs spaced more or less evenly along the ramp. The positions were determined by locating all arginines inserted into the minor groove and

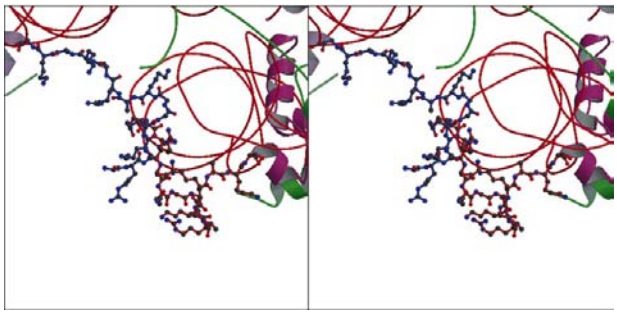

Figure 13
Stereo pair representing the interaction between H2A N-terminal tails of neighboring NCPs. The region shown in this figure is a detail of the region of interaction between N-terminal tails of H2As in the previous figure.

determining which phosphates on either side of the minor groove were bound by putative salt bridges to the guanidino group of the arginine. The base position of individual mGIs was determined by the identity of the base along the Watson strand of the DNA positioned over the guanidino group of the arginine defining the mGI. These results are tabulated in Table 3. Fig. 15 presents the mapping and numbering of the mGIs along the length of the DNA in the NCP. Numbering of mGI positions begins on either side flanking the pseudodyad of the particle at base 73 and proceeds outward to the DNA termini. Temperature factors for the phosphate atoms of the DNA were plotted as shown in Fig. 16. The plot clearly shows the relationship between the temperature factors of DNA P atoms and the position of the phosphate moiety in relation to the surface of the histone core. The lowest temperature factors correspond to minor groove-in, mGI, positions. The relationship of the temperature factor for DNA atoms to position with respect to the histone octamer is shown through color coding of atoms by temperature factor in Fig. 17.

The minor groove-in positions represent binding motifs which are repeated regularly around the superhelical ramp of the octamer core. An arginine residue inserted into the minor groove appears as the central feature of the structure, which includes residues from two protein subunits. Two basic designs are found and are designated here as α and β forms of the mGI.

4.6.1. αmGI DNA-binding motif. The α form is constructed from a pairing of the N-termini of two central histone fold α-helices. This binding motif was termed the paired ends of helices (PEM) motif (Arents *et al.*, 1991), which interacts with the phosphate backbone through the dipoles produced in the helices. The αmGI binding is found at mGI 5 and mGI 2. The binding at mGI 5 is accompanied by the insertion of Arg33 of H2B into the minor groove. This position differs dramatically from all other mGIs in that the inserted arginine occurs on an N-terminal tail at a position not visible in the octamer structure (see Fig. 18). Possible motion of the tail of H2B is restrained by the interaction of Arg32 from the NH helix of H2A hydrogen bonding to Glu35 of H2B. The NH helix of H2B is directed toward the phosphate group of Gua268 at ventral mGI 5 and Gua122 at dorsal mGI 5. The main-chain amides of Ile39 and Ser38 of H2B contact the phosphate through water bridges, with additional contacts from the amides of Ser36. The phosphate also is directly hydrogen bonded to Arg29 from the NH helix of H2A. The NH helix of H2A is not directed toward the DNA phosphates but is accompanied by an N-terminal accessory helix, which interacts with the phosphate of Gua31 at ventral mGI 5 and with Gua177 at dorsal mGI 5. The interactions include direct hydrogen bonding to the main-chain amide N atom of Arg17 and a water bridge from the amide N atom of Ser18 of the

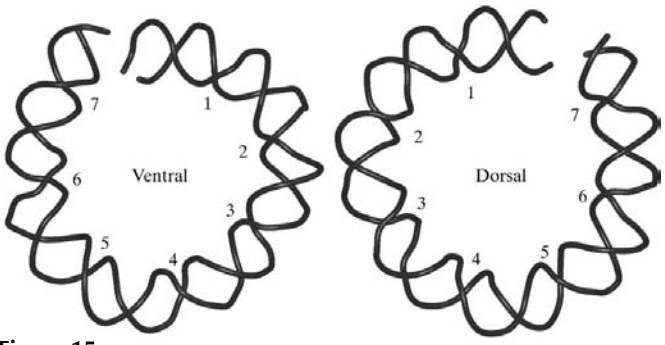

Figure 15
Mapping of minor groove-in (mGI) positions along the superhelical path of DNA in the nucleosome core particle for both ventral and dorsal halves, with the zero position at the pseudo-dyad and numbering outward toward the DNA termini. Positions 1–6 correspond to mGI sites with an arginine inserted into the DNA minor groove.

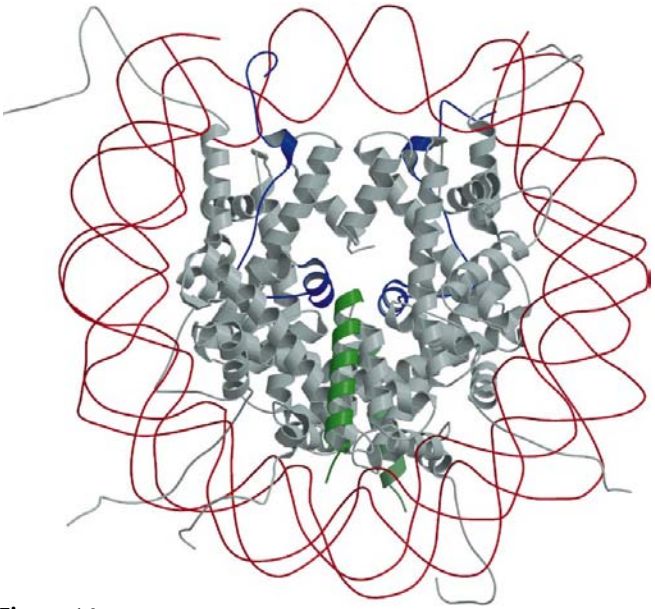

Figure 14
The C-terminal domains of the dimer histone subunits are shown in color, with the remainder of the protein shown in grey. Histone H2A C-terminal domains are presented in blue for clarity. The C-terminal accessory helix of H2B is shown in green.

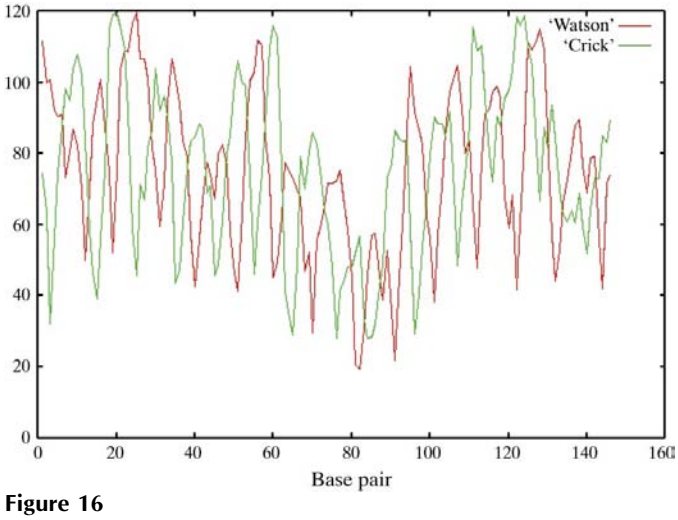

Figure 16
Plot of temperature factors refined for P atoms along the backbone of the DNA in the nucleosome core particle structure. The identity of the strands is as discussed in the text.

H2A N-terminal accessory helix. The N-terminal tail of 1H2B also passes between the DNA gyres and inserts Arg29 into the minor groove between the phosphates of Thy104 and Gua192. Interestingly, the 2H2B tail also passes between the DNA gyres but does not insert into the minor groove between phosphates of Gua249 and Thy48 as might have been predicted from symmetry; instead it is deflected to place the guanidino group between the DNA gyres in proximity to the the phosphates of Thy48 and Gua125.

The αmGI binding sites at 2 are defined by the positions of Arg63 of H3. A generalized view of the αmGI binding motif at this position is shown in Fig. 19. The guanidino group is restrained from deeply inserting into the minor groove by Thr30 of H4. The inserted arginine is positioned on a loop over a hydrophobic pocket. The NH helices of H3 and H4 are both involved in the formation of the binding site. At dorsal mGI 2 the N-terminus of the NH helix of H3 is directed toward the phosphate group of Thy238 at dorsal mGI 2 and Thy91 at ventral mGI 2 and the main-chain amide N atoms of Leu65 and Lys64 are hydrogen bonded to the phosphate. The NH helix of H4 is directed toward the phosphate of Ade61 at dorsal mGI 2 and Thy208 at ventral mGI 2 but is displaced from the DNA backbone and does not make direct contacts with it, although a water bridge is seen from Lys31 of 2H4 to the phosphate at dorsal mGI 2. An unusual feature is the hydrophobic interaction in the major groove between Leu65 of H3 and a thymine. In the octamer structure, Leu65 of H3 is exposed to solvent and reinforced by the adjacent Pro66. The position of this distinctive hydrophobic interaction in the major groove places it in the center of a run of four thymine residues, which is intriguing in light of the importance of polyA tracts in the DNA sequence dependent phasing of nucleosomes.

4.6.2. βmGI DNA-binding motif. The βmGI binding motif is found at mGI 6, mGI 4, mGI 3 and mGI 1. In general terms, the βmGI binding site is constructed from the CL loop of H3 and the NL loop of H4 or the CL loop of H2A and the NL loop of H2B. The loops are joined in a short parallel β-bridge structure which forms a platform over which the phosphate backbones along the minor groove of the DNA are positioned. In most cases, the guanidino group of the inserted arginine is restrained by hydrogen bonding to a threonine on the opposite loop. It may be that this interaction is designed to position the guanidino group in the DNA minor groove. However, the interaction of the guanidino group with a threonine is not invariant. The hydroxyl-carrying or acidic groups bounding the arginine may also serve to position the arginine side chain on the surface of the histone octamer in the absence of DNA. An examination of the binding pattern of the DNA at the mGIs shown in Table 3 revealed that the phosphates directly bound to the arginine were separated by five base pairs or about one half of a helical turn. When one phosphate contacts an arginine at an mGI site, the next phosphate bound will be the one closest to the site on the other strand which is one-half of a helical turn away. This suggests that the binding of the DNA takes place by a pinching of the minor groove with concomitant opening of the major groove. This may provide an alternative explanation for the significance of polyA tracts adjacent to an mGI, as the polyA tract serves to narrow the minor groove in a 5′ to 3′ direction. Thus, DNA binding to the octamer occurs at the mGI by the interaction of the arginine with phosphates on opposite strands and separated by one-half of a helical turn. It would appear that the function of the remainder of the octamer is to place the arginines at appropriate positions. The

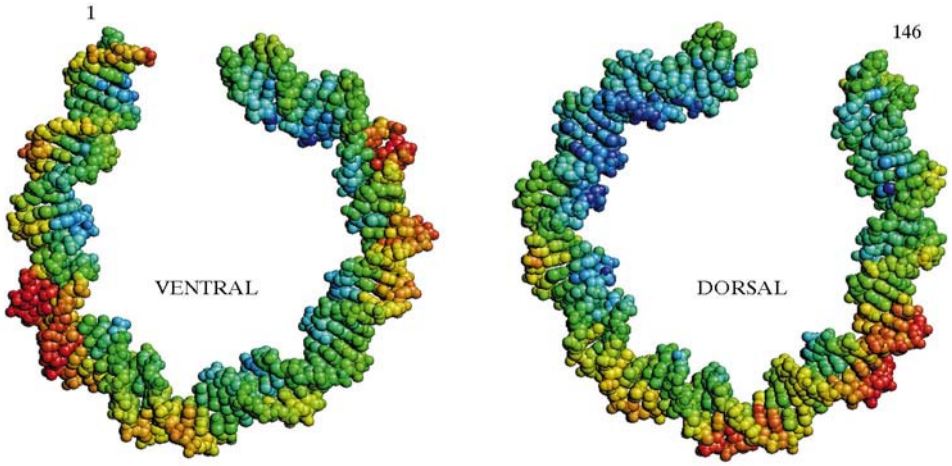

Figure 17
Temperature factors for DNA atoms are represented by color, from blue for lowest temperature factor to red for the highest temperature factor. The DNA is broken into ventral and dorsal halves at the dyad.

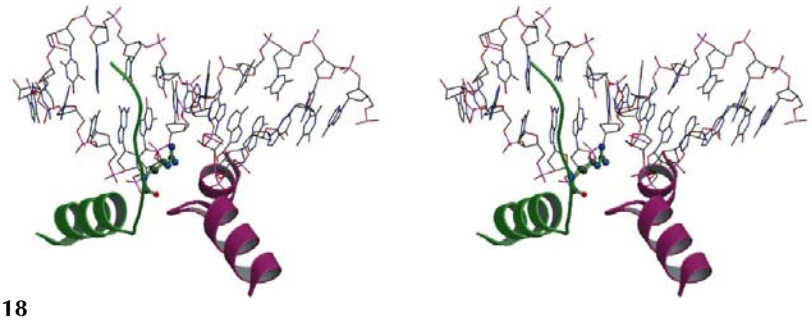

Figure 18
Stereoview of the structure of the αmGI binding site at dorsal mGI 5. The inserted arginine, Arg33 of H2B, is shown in ball-and-stick rendering. The NH helix of H2A is not directed toward a DNA phosphate but is accompanied by an N-terminal accessory helix which interacts directly with a DNA phosphate outside of the minor groove. H2B is shown in green and H2A in lavender.

propensity for the interaction of the arginine and tetrahedral ions such as phosphate or sulfate is shown in the octamer structure, where two sulfate ions are bound at Arg45 of H4 in approximately the same positions as the phosphates at the same arginine at dorsal mGI 1 (see Fig. 20). This is also seen in the structure of the archael histones, HMfA and HMfB (PDB code 1b67), where a sulfate ion is bound in approximately the same position.

The lowest temperature factors seen on the DNA backbone in Fig. 16 occur adjacent to the nucleosomal DNA dyad center at a point at which the DNA minor groove contacts dorsal mGI 1. Examination of the structural details of the binding site reveal the guanidino group of Arg45 on the N-terminal loop of the histone fold of 2H4 inserting into the minor groove and approximately equidistant from the phosphates of Gua217 and Gua81. Fig. 20 presents a comparison of the environment of Arg45 of H4 in both the NCP and the octamer structures. Interestingly, the phosphate positions seen in the NCP are nearly the same as the positions of sulfate ions bound in the octamer structure. The binding site is further stabilized through arginines and lysines from 2H4 (Arg35, Lys31, Lys44, Arg39), Arg116 from 2H3 and Lys115 from 2H2A contributing to ionic stabilization of the DNA backbone. In addition, in the NCP Arg36 of H4 hydrogen bonds directly with the phosphate of Ade207. The guanidino group of Arg45 is restrained in both the octamer and NCP structures by hydrogen bonding with Thr118 of 2H3. In the NCP structure, this interaction constrains the insertion of the guanidino group into the minor groove. In the octamer, the same interaction of Arg45 with Thr118 appears to restrain the motion of the guanidino group, which extends above the protein surface much like a grappling hook to initiate binding of DNA to the octamer.

At dorsal mGI 3, the interaction of the N-terminus of the mH helix of H3 with DNA phosphates is carried through by direct hydrogen bonding of the main-chain amide N atom of Ser86 to the phosphate and by water bridges between the phosphate O atom and the amide N atom of Ser87 as well as the hydroxyl group of Ser87 (see Fig. 21). At ventral mGI 3, the side chain of Gln85 of H3 is also water bridged to the

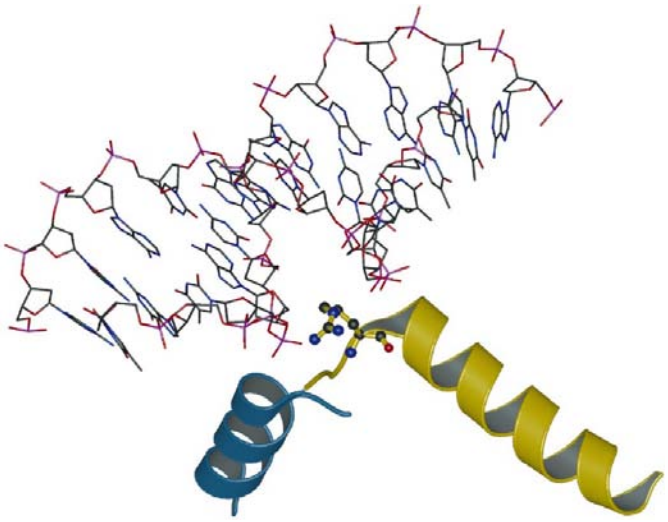

Figure 19
mGI motif found at mGI 2 is formed from the NH helices of H3 (yellow) and H4 (blue). The inserted arginine, Arg63 of H3, is shown in ball-and-stick rendering. The N-termini of the paired helices are directed at phosphate groups on the DNA backbone.

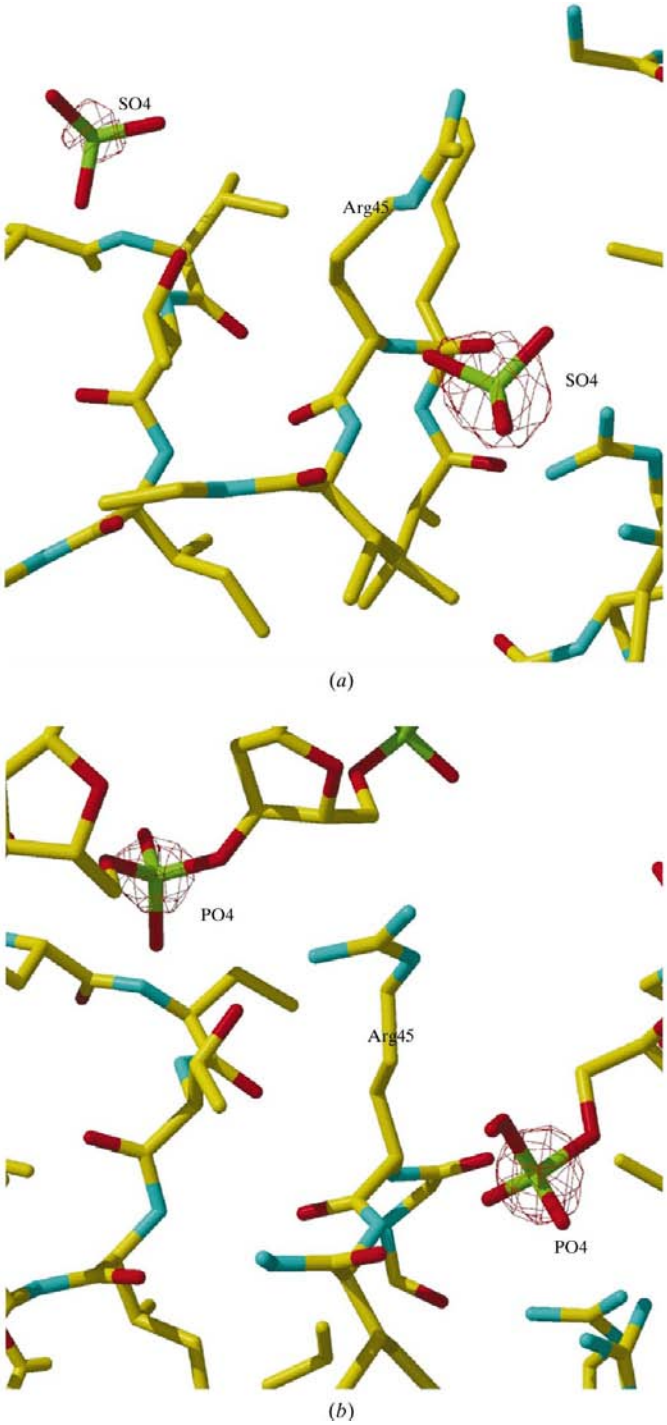

Figure 20
A comparison of the environment of Arg45 of histone H4 in the octamer structure and at the dorsal mGI 1 of the nucleosome core particle. In the octamer structure, in (*a*) sulfate ions are bound in approximately the same positions as the DNA phosphates as shown in (*b*). Electron density is shown at five times the σ level of the $2F_o - F_c$ map.

equivalent phosphate. At dorsal mGI 3, Gln85 reaches across the minor groove to form a water bridge with the phosphate of Ade102, while the side chain is restrained by interactions with the C-terminus of the CH helix of 2H4. These interactions are formed through hydrogen bonding to the amide N atom of Ala83 and by a water bridge to the amide N atom of Met84 of 2H4. The phosphate group bound at the N-terminus of the mH helix of H3 at ventral mGI 3 is Cyt50 and at dorsal mGI 3 is Cyt196, which are identical positions by palindrome symmetry of the DNA. The inserted arginine, Arg83 of H3, does not make any apparent bonds at either ventral mGI 3 or dorsal mGI 3 and does not interact with Thr80 of H4 on the opposite side of the β-bridge. Thr80 of H4 is hydrogen bonded to a DNA phosphate through the amide N atom and the hydroxyl group. A hydrogen bond to the same phosphate is also formed by the adjacent amide N atom of Lys79 of H4. This interaction binds the phosphate of Cyt247 at ventral mGI 3 and Cyt101 at dorsal mGI 3, which are also identical by palindrome symmetry of the DNA sequence. Stabilizing interactions on both sides of the minor groove are provided by hydrogen bonding of side chains of Arg72 of H3 and Lys79 of H4 to DNA phosphates. At ventral mGI 3, Arg78 of H4 can be seen to have two conformations, one of which involves hydrogen bonding to the phosphate group of Ade248 and a water bridge to the side chain of Asn84 of 1H2B. The second conformation shows the guanidino group of Arg78 directly hydrogen bonded to Asp85 of 1H4. This conformation is the only one seen at dorsal mGI 3.

A chloride ion bound in a pocket formed by main-chain amide N atoms stabilizes the C-terminus of the β-bridge structure formed between the NL loop of H2A and the CL loop of H2B at mGI 4 (see Fig. 22). The inserted arginine, Arg42 of 1H2A, makes a hydrogen-bond contact with the O4′ of Ade257 at ventral mGI 4, while at dorsal mGI 4 the guanidino group of Arg42 is bound to a water. Thr88 of H2B does not interact with the guanidino group of the inserted arginine. The main-chain amides of Ser87 and Thr88 of H2B hydrogen bond with the phosphate of Gua40 at ventral mGI 4 and of Gua185 at dorsal mGI 4. The hydroxyl group of Ser87 is also positioned to make hydrogen-bond contact with a phosphate O atom but only clearly does so at dorsal mGI 4. Further stabilizing interactions are provided through direct hydrogen bonding of Arg35 of H2A and Arg86 of H2B to phosphates on opposite sides of the minor groove.

The βmGI binding site at 6 differs somewhat from the other βmGIs. The inserted arginine, Arg77, is placed on the CL loop of H2A and the threonine which might be expected to hydrogen bond to it, Thr52 of H2B, points inward toward the histone core. The inserted arginine is flanked by Thr72 of H2A on one side and Asp51 of H2B on the other. Main-chain amide N atoms of Thr76 and Arg77 of H2A hydrogen bond directly to a DNA phosphate, Cyt132 at dorsal mGI 6 and Ade19 at ventral mGI 6. Arg77 of H2A is inserted deeply into the minor groove and hydrogen bonds to the O4′ on the ribose moiety of Gua131 (see Fig. 23).

An interesting feature of the DNA in the NCP structure is the path taken by the double helix between neighboring particles. As seen in Fig. 24, the 5′ ends are abutted in a quite unnatural manner while the hydrophobic stacking of base rings is easily maintained moving from one NCP to another.

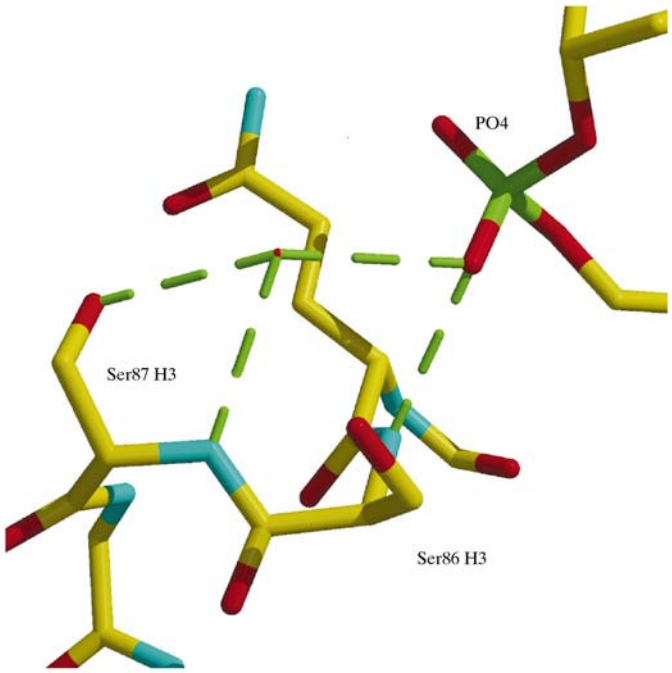

Figure 21
Direct binding at the N-terminus of the mH helix of H3 to DNA at mGI 3 is made through induced dipole of the α-helix and by direct hydrogen bonding of main-chain amides as well as a water bridge to DNA phosphate.

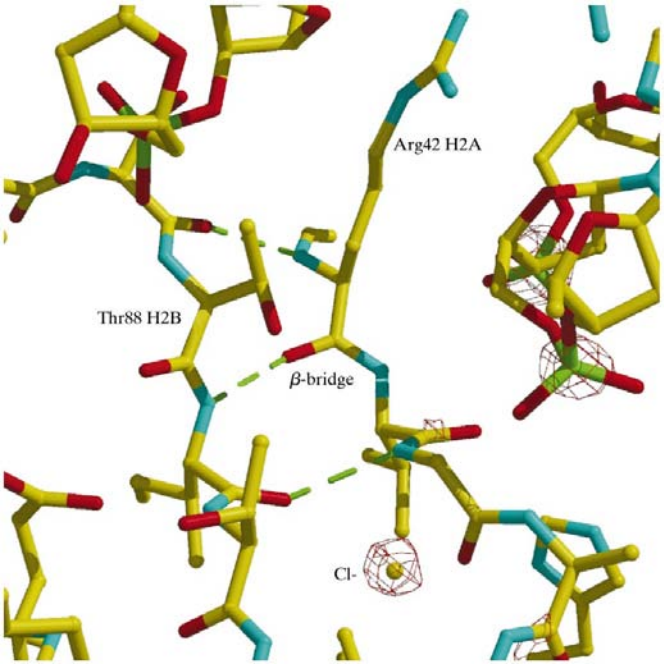

Figure 22
The structure of the β-bridge at mGI 4 is stabilized by the binding of a chloride ion in a pocket formed by main-chain amide N atoms. The guanidino group of Arg42 of H2A is inserted into the DNA minor groove. Electron density is shown at five times the σ level of the $2F_o - F_c$ map.

The DNA used in this study was not treated to remove the 5′ phosphates (Fig. 25).

4.7. Folding of the DNA into the nucleosome core particle

The role of arginines in the binding and folding of DNA into the NCP has long been known. Using NCPs prepared by micrococcal nuclease digestion of chicken erythrocyte chromatin, it was determined (Ichimura *et al.*, 1982) that 14 arginine residues were involved in binding to DNA and were released for chemical modification only above 0.6 *M* NaCl. Although the precise identity of the 14 arginine residues was not established, it is likely that they were the 14 arginines seen to insert into the minor groove of the DNA in the NCP structure, 12 of which have been used to define the mGI positions. The guanidino group of arginine binds strongly through salt bridges to tetrahedral-type anions such as the phosphate group on the DNA backbone. In the octamer structure, sulfate ions are bound to the exposed side chains of several of the arginines at latent mGI sites in the approximate positions which would be occupied by phosphate groups when DNA is bound as seen in Fig. 20. The arginine side chains at the latent mGI sites extend above the protein surface to provide anchor points for attachment of DNA to the histone octamer. It has long been known that the central approximate 100 base pairs of DNA in the nucleosome core particle are more resistant to micrococcal nuclease than the termini (van Holde *et al.*, 1980). The strength of the binding is further suggested by hydroxyl radical footprinting of nucleosome cores at high salt and high temperature (Bashkin *et al.*, 1993), which demonstrated that translational and rotational positioning is maintained even at 0.8 *M* NaCl or at 348 K. Mechanisms for displacement of DNA from the histone core *in vitro* involve melting of the DNA duplex or disruption of the binding by very high salt concentrations or by cationic molecules such as the protamines which possess polyarginine segments.

The necessity of mutation of the native α-satellite sequence to provide a synthetic *Eco*RI site may have contributed to the failure of the palindrome to phase with perfect symmetry by altering the helical parameters of the native sequence at the dyad. Energy minimizations of the central 22 base pairs using the *DUPLEX* program (Hingerty *et al.*, 1989) clearly demonstrated the tendency of the *Eco*RI recognition sequence to bend the fragment away from the histone surface with the bend centered at the dyad (unpublished results). We have previously postulated (Uberbacher *et al.*, 1988) that the histones bind DNA cooperatively and that upon initial binding the histones 'communicate' to their symmetry-related counterparts which properties of the DNA to select. From an examination of the temperature factors of the phosphates along both strands of DNA in Fig. 16 it is clear that the phosphodiester backbone is more stably positioned on the side of the dyad possessing the longer DNA strand. The lowest temperature factors are those for the phosphates bound at dorsal mGI 1. The DNA must bind first on one side of the histone core in a conformation which provides the lowest free energy of binding and then be accommodated on the other side by binding in a slightly less favorable mode. However, since the histone binding sites are not strictly uniform in nature, the binding is not uniformly misdirected and the base positions at ventral mGI 5 and 6 are as expected by palin-

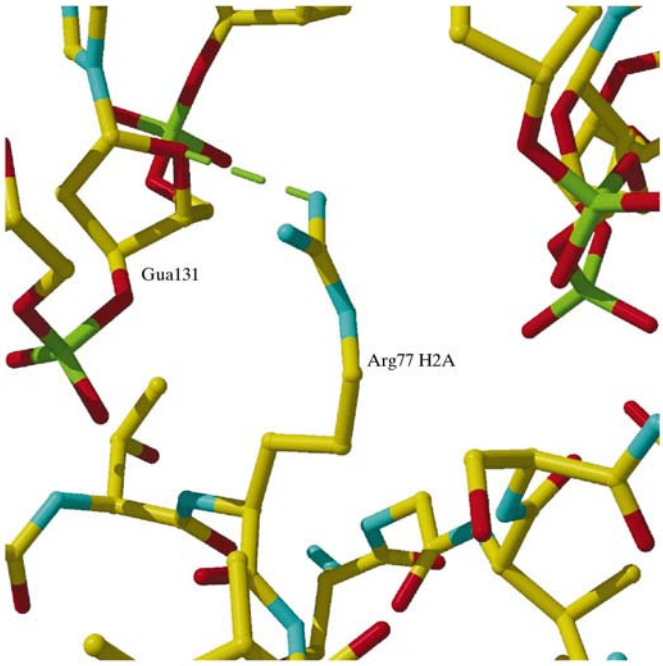

Figure 23
The structure of dorsal mGI 6 has the inserted arginine, Arg77 of H2A, hydrogen bonding to the O4′ atom of Gua131.

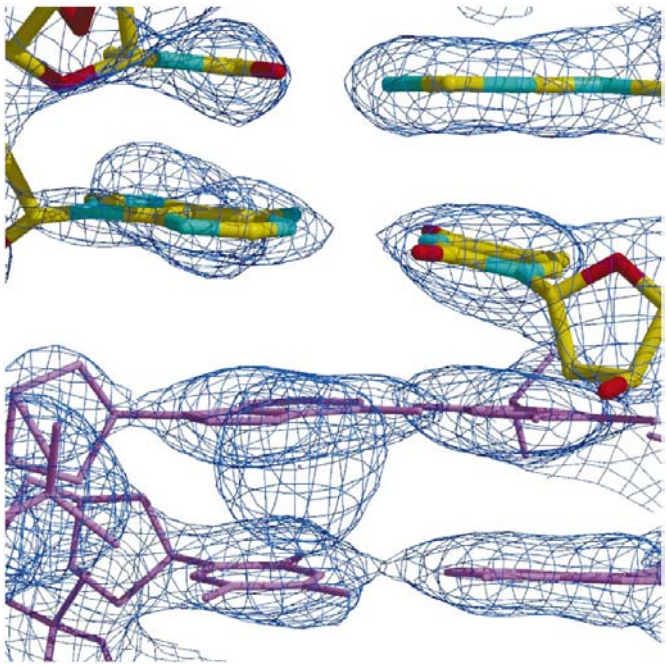

Figure 24
Base stacking is continuous between neighboring NCPs in the crystal lattice as the DNA exits one NCP and enters the next. Electron density is diplayed at one times the σ level of the $2F_o - F_c$ map.

drome sequence symmetry. The DNA between ventral mGI 2 and 3 is underwound compared with its symmetry-related counterpart (see Fig. 26), as is the DNA between ventral mGI 4 and 5. This is in partial disagreement with the results of Luger *et al.* (1997), who found only a difference between the DNA segments between mGI 2 and 3. Interestingly, these two regions are flanked by αmGIs on one side (mGI 2 and mGI 5). The mixture of αmGI and βmGI motifs may account for some of the plasticity in DNA binding seen here while still providing for precise phasing.

A further very important consequence of the asymmetry in the DNA is that the non-equivalence of the two sides of the NCP as it packs into the crystal lattice provides an explanation to the question of how the NCP packed. The original intention of crystallizing NCPs containing a DNA palindrome (Harp *et al.*, 1996) was to produce a twofold symmetric particle for which the dorsal–ventral asymmetry did not exist. Because of the shape of the NCP, it can, like a coin, land in one of two orientations, resulting in twofold disorder in the crystal if the two orientations are not equivalent. The single orientation was intriguing in that it meant that some other factor was involved which allowed the NCPs to pack in only one orientation. If we refer to the short side of the DNA as the 'left' side and the long side of the DNA as the 'right' side, an analysis of crystal packing shows that the DNA exiting the NCP is abutted to DNA exiting a symmetry neighbor. The left side of the DNA of the nucleosome core particle abuts the right side of the DNA of a symmetry neighbor and the right side of the DNA abuts the left side of the DNA of another symmetry neighbor in the same plane. This can be clearly seen in Fig. 27. If the NCP is shown in a dorsal view, the two symmetry neighbors sharing the same continuous path of DNA will present their ventral side. Thus, between dyad centers of two such symmetry-related NCPs there are 145 base pairs of DNA, excluding the base pairs at either dyad. This fact means that the NCP does not have two equally likely choices of packing with either the ventral or dorsal surface down, but can only pack in one orientation as determined by the orientation of the NCP to which it joins DNA termini. Unfortunately, the question of how the packing of NCPs in the crystal relates to the packaging of nucleosomes in higher order structure is unclear. Certainly, the incorrect bond polarity at the interface of the DNA termini between symmetry neighbors is an artifact. Additionally, the stacking of NCPs in the direction of the superhelical axis places the symmetry neighbors with alternating polarity of the dyad axis; that is, in a head–tail–head fashion which would not be expected from any model of the condensed (30 nm) fiber with the linkers disposed toward the center.

Folding of DNA onto the superhelical ramp provided by the histone octamer core results in overwinding of the duplex and an increase in the helical twist angle between neighboring base pairs. The helical repeat of DNA in solution has been measured as 10.5 base pairs per turn of the double helix. The helical repeat of the DNA bound to the NCP reflects the overwinding resulting from bending and is measured as approximately 9.9 base pairs per turn.

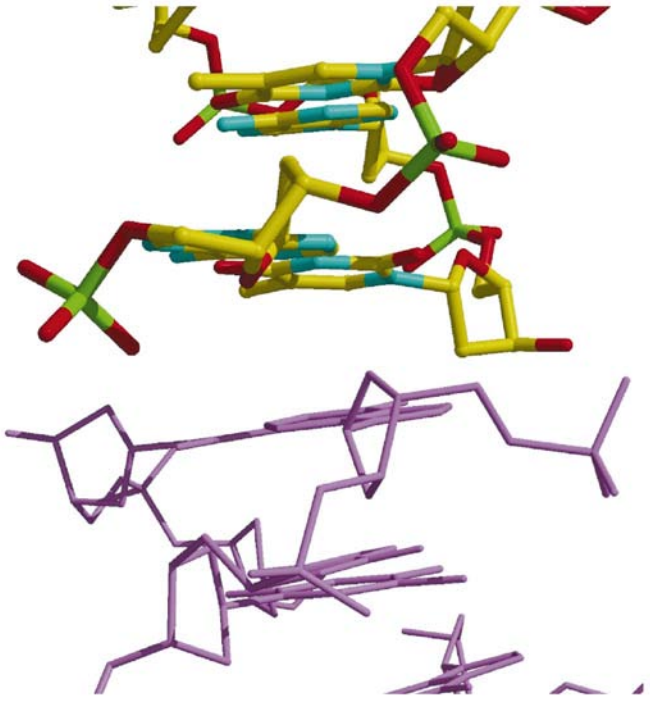

Figure 25
DNA from a neighboring NCP abuts the DNA termini 5′ to 5′ requiring a shift in the position of the terminal phosphate groups.

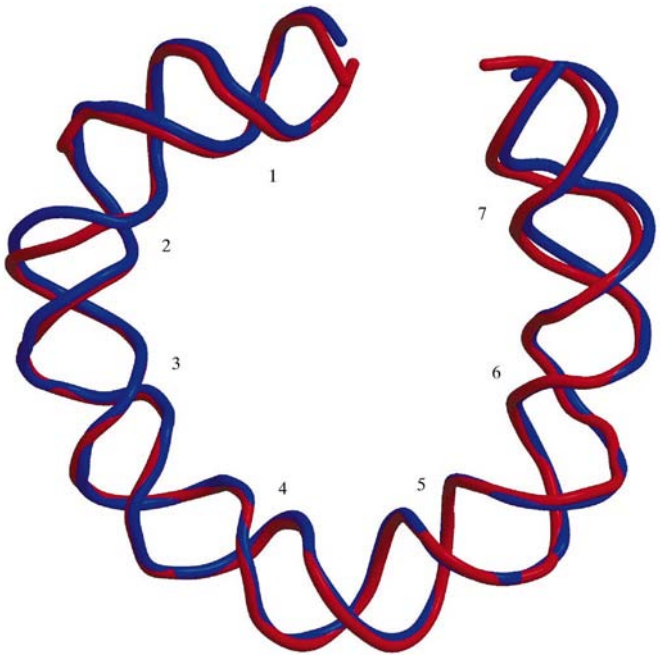

Figure 26
Overlapping of the phosphodiester backbone trace of the two halves of the DNA demonstrates binding asymmetry. The minor groove-in sites are numbered as in the previous figure. The 73 base-pair dorsal half is shown in blue and the 72 base-pair ventral half in red. Asymmetry in the path of the DNA backbone are revealed between positions 2 and 3 and, to a lesser degree, between positions 5 and 6.

4.8. Influence of the solution environment on the crystal structure

NCP crystals taken directly from the mother liquor do not diffract well and the unit-cell parameters are highly variable. The primary source of variability appears to be stacking along the long axis of the crystal, the *c* axis. For this reason, the crystals have been stabilized by adding an organic cosolvent such as MPD. With increasing MPD concentration, the length of the *c* axis is reduced and the diffraction limit extended to an optimum level of MPD beyond which diffraction begins to suffer. It was once thought that MPD and similar organic cosolvents served to lower the activity of the bulk water in the crystal by organizing water structure and thus removing water available for hydration of protein structures. However, MPD acts conversely to increase the sphere of hydration of proteins (Timasheff, 1993; Pittz & Timasheff, 1978) and it now appears that the dielectric constant is the operative parameter (Flock *et al.*, 1996; Arscott *et al.*, 1995). Of particular concern to the structure of the NCP is the effect of MPD on the structure of DNA. It is not clear what (if any) impact MPD may have on bending owing to phased A-tracts (Sprous *et al.*, 1995; Dickerson *et al.*, 1996). The effect of an organic cosolvent on the mobility of core particles in gel electrophoresis (Pennings *et al.*, 1992) may explain the response of the *c* unit-cell parameter to varying levels of MPD. A decrease in the dielectric constant increases the charge-repulsion effect of phosphates and decreases bending at the DNA termini. This allows the establishment of appropriate base stacking between termini of neighboring particles and decreases the variability of the orientation of the dyad axis with respect to the *c* axis.

4.9. Ions and hydration

A total of 15 Mn^{2+} ions were identified in the structure. Most were bound directly to the N7 atom of purine bases at an average distance of 2.45 $\pm$ 0.19 Å. Ten Mn^{2+} ions were bound to the N7 of guanine and two were bound to the N7 atom of adenines at the 5′ termini. MN461 appears to be bound octahedrally through water bridges in the minor groove of the DNA, with one of the water bridges clearly visible. Of the identified waters, HOH256, may be a similarly coordinated Mn^{2+}. Two Mn^{2+} ions, MN466 and 473, appear to be shared between the N7s of adjacent guanine bases. Potassium ions were similarly bound to the N7 atoms of purines, with adenines and guanines equally likely. The average distance of the potassium ions from an N7 atom was 3.36 $\pm$ 0.19 Å for the seven potassiums.

Chloride ions were identified as being bound in two positions which are equivalent by dyad symmetry. The chloride is tightly bound in a pocket formed by main-chain amide N atoms of Ala45, Gly46, Ala47 of H2A and by the amide N atom of Ser91. The amide N atom of Ala45 of H2A is in direct contact with a DNA phosphate. The proximity of the negatively charged phosphate is problematic but it may be supposed that the chloride bound to the octamer prior to DNA binding in the reconstitution reaction. A sulfate ion in the low-temperature octamer structure binds strongly in the same location, providing an ion bridge stabilizing the two histone chains in the same manner that the chloride ion in the NCP stabilizes the C-terminus of the β-bridge as discussed above.

The single most intense peak in the $2mF_o - DF_c$ electron-density map was identified as a cacodylate ion. The identity of the ion was confirmed using omit maps and by replacing the cacodylate with a Mn^{2+} ion. In a $2mF_o - DF_c$ map calculated with the cacodylate replaced by Mn^{2+}, the electron density displayed distinct lobes in which the cacodylate was fitted (see Fig. 28). The cacodylate is placed within a pocket formed by side chains of Gln76 and Asp77 of 2H3 and the main-chain carbonyl group of Leu22 of 2H4 on the NCP and by side chains of Glu64 of 1H2A and His49 of 2H2B as well as the main-chain carbonyl group of Val48 of 1H2B on the symmetry neighbor.

4.9.1. Structural water. A total of 349 waters were identified in the structure. 42, or 10%, are hydrogen bonded to main-chain amide N atoms. Of those, a total of 24 are hydrogen bonded only to amide N atoms, four are involved in water bridges between adjacent main-chain amide N atoms on the same chain and one makes a water bridge between chains 2H3 and 2H4, and 13 are involved in forming water bridges between main-chain amide N atoms and DNA phosphates. There are 43 identified waters which are hydrogen bonded only to DNA phosphates, while 24 are bound to DNA bases. Many of the remaining waters are involved in forming water networks with the histone core of the NCP.

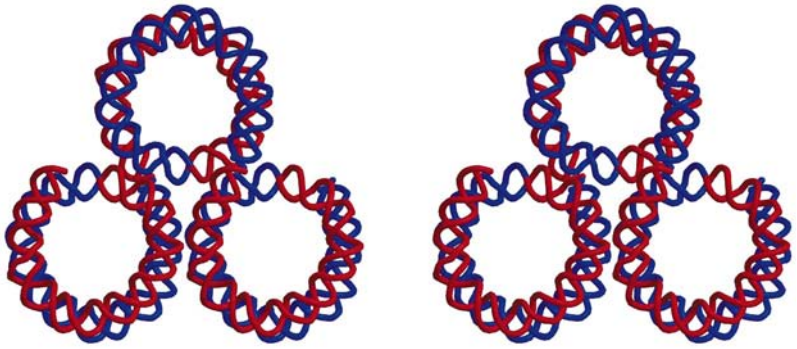

Figure 27
Stereo packing diagram showing the continuous path of DNA from the NCP to symmetry neighbors. DNA from three NCPs are shown with the 72 base-pair ventral half colored red and the 73 base-pair dorsal half in blue. This shows that the NCPs pack in only one orientation depending on the orientation of NCP already situated in the crystal.

4.10. Microgravity crystallization

Microgravity experiments were undertaken under the auspices of Dr Daniel Carter using the DCAM double-diffusion device (Carter, Wright *et al.*, 1999) originally in order to correct the tendency of NCP crystals to grow more quickly along one lattice axis, resulting in hollow crystals.

NCP crystals diffract weakly, probably because of the large thermal vibrations, particularly in the DNA and the disordered tail regions of the histones. This necessitates large crystals. Additionally, from previous experiments with double-diffusion crystallization, it was known that large crystals could be grown using the gradients developed over time in the crystallization chamber. Microgravity experiments were completely successful in the production of large fully formed crystals. Microgravity also appears to have resulted in significant improvement in crystal quality, as seen in Fig. 2. The reason for increased diffraction quality for crystals grown in a microgravity environment is still somewhat controversial, but may be a consequence of the lack of convection in the microgravity environment as well as the lack of sedimentation of crystals as they grow. During crystal growth, a depletion zone develops around the crystal in which a lowered concentration of nucleosome core particles is found. The entirely diffusion-controlled crystal-growth process reduces the rate of addition to the crystal lattice of larger aggregates with correspondingly lower diffusion coefficients. It may also inhibit addition of impurities by slowing the rate of growth of the crystal, thereby allowing more precise formation of the energy-minimized structure of the crystal lattice (Carter, Lim *et al.*, 1999). The actual value of microgravity crystallization for NCP crystals has not been completely established owing to complicating factors affecting crystallization both on the Space Shuttle and on board the Mir space station. These included unavoidable factors such as launch delays, vibrations and temperature fluctuations. Completion of the International Space Station will allow for more thorough and unbiased experiments on the value of the microgravity environment for crystallization of biological macromolecules.

4.11. Stacking interactions between histone faces

A striking feature of the surface of the histone face of the NCP is the presence of a number of acidic residues on the dimer face (see Fig. 29). These residues include Glu56, Glu61, Glu64, Asp90, Glu91 and Glu92 of H2A and Glu110 of H2B on the ventral face of the NCP. The N-terminal tail of 2H4 from one NCP lies along the acidic patch, forming numerous hydrogen bonds and water bridges between the basic residues on the H4 tail from the dorsal surface of the neighboring NCP and the acidic residues on the opposing dimer surface. The H4 tail is well visualized to Lys8. It is significant to note that there is no evidence of acetylation of lysines on the H4 tail although the histones are native chicken erthrocyte octamers. The cacodylate ion lies in a pocket formed on one side by Asp77, Gln76 and Glu73 on the NH helix and the NL loop of 2H3 as well as the main-chain carbonyl of Leu22 of 2H4 (see Figs. 30

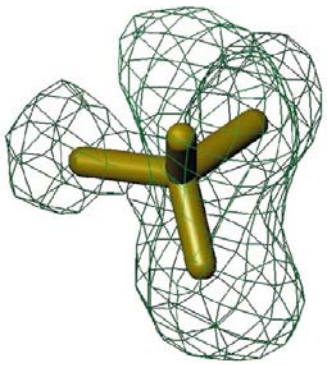

Figure 28
The electron density displayed at three times the σ level is shown for the $2mF_o - DF_c$ electron-density map caculated with the cacodylate ion replaced with an Mn^{2+} ion at the position of the As atom. A model of the cacodylate has been placed within the resulting density to demonstrate the correspondence with the structure of the cacodylate ion.

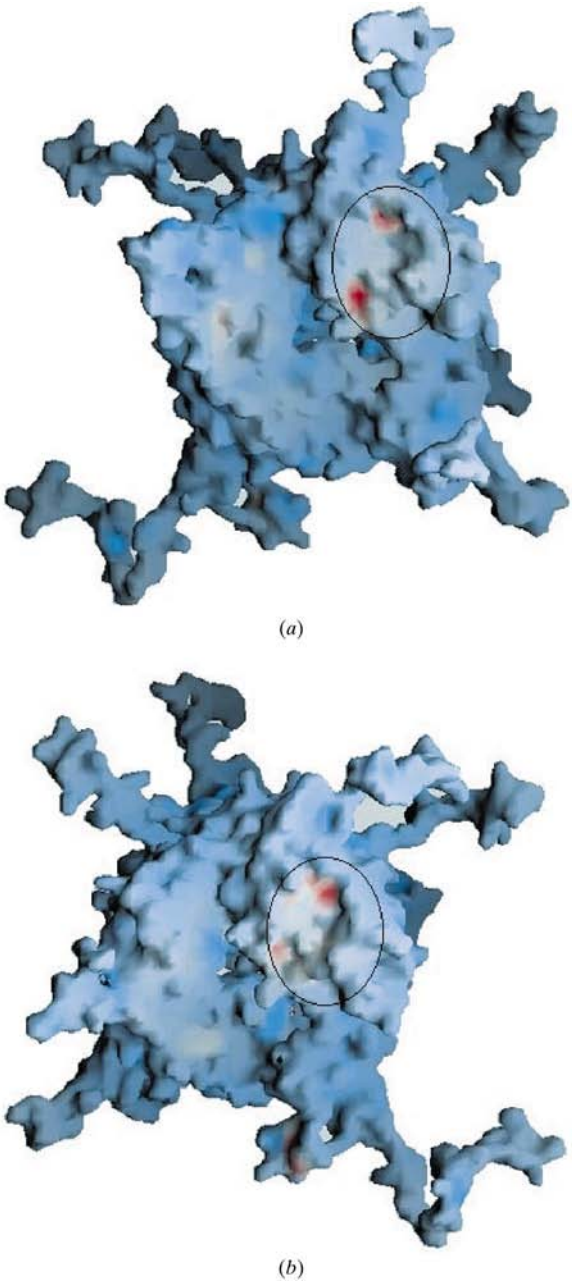

Figure 29
Electrostatic surface potential map of the histone octamer core in both ventral (*a*) and dorsal (*b*) views. An acidic patch on the dimer face in both views is the most striking feature of the map. The acidic patch is indicated by an ellipse in both views.

and 31). The opposing NCP forms the other side of the pocket, containing the cacodylate ion with the main-chain carbonyl groups of Gln47 and Val48 and the side chain of His49 of 1H2B. This interaction appears to be responsible for the asymmetry seen in the dimer histones in Fig. 5. The asymmetry is related to the interaction between neighboring NCPs along the acidic patch on the dimer surface (see Fig. 32) and is a consequence of the fact that only the ventral dimer face of one NCP and the dorsal tetramer face of the adjacent NCP are involved. This stacking interaction may be physiologically relevant; however, it seems unlikely owing to the alternating orientation of the NCPs in the stack. This arrangement does not correspond to current models of higher order chromatin structure.

5. Conclusions and future directions

The general features of the binding of DNA to the left-handed superhelical ramp provided by the histone octamer were essentially correctly identified by Arents *et al.* (1991) from the structure of the histone octamer and modeling of the DNA. However, the central role of the arginine residues inserting into the minor groove at the mGI sites are here seen in detail along with such features as the hydrophobic interaction of H3 Leu65 in the major groove. The structure is, of course, similar to nucleosome structures seen at lower resolution. However, as a result of the concatenation of crystal-growth techniques

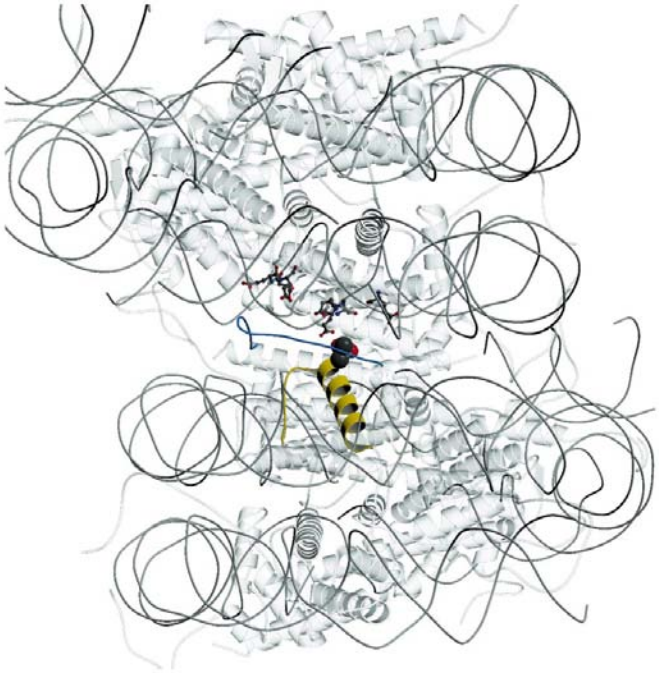

Figure 30
Stacking interactions between histone faces of neighboring NCPs in the crystal lattice. A central feature of the stacking interactions is a strongly bound cacodylate ion. The N-terminal tail of H4, shown in blue, and the NH helix and NL loop of H3 form a pocket containing the cacodylate ion on one particle. The neighboring particle interacts with the cacodylate through the NH helix and NL loop of H2B. The N-terminal tail of H4 is also involved with significant interaction with a patch of acidic residues, rendered as ball-and-stick, on the dimer face of the neighbor NCP. Note that the orientation of the dyad alternates between the two NCPs. Features of the interaction are identified and shown in detail in the next figure.

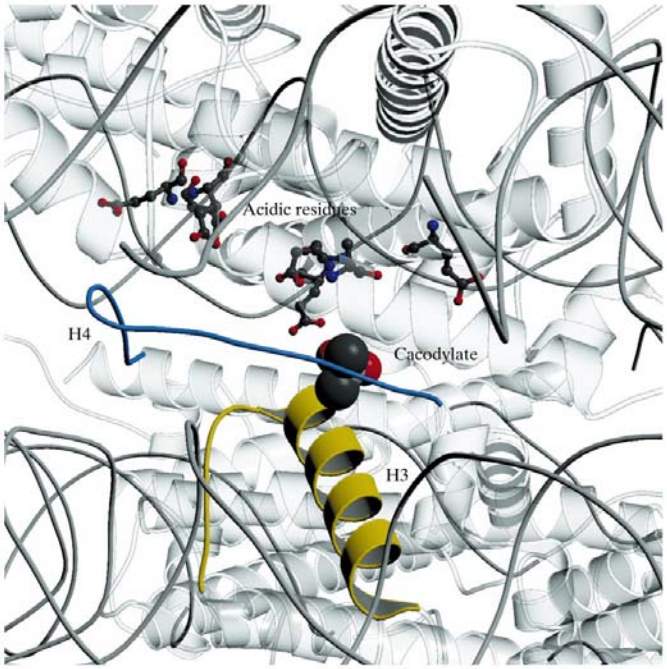

Figure 31
Enlarged view of the previous figure showing details of the stacking interactions between two neighboring NCPs involving the dorsal $(H3–H4)_2$ tetramer face of one NCP and the ventral (H2A–H2B) dimer face of the adjacent NCP. The acidic residues forming the surface acidic patch on the dimer face are rendered as ball-and-stick. The H4 tail is in blue and the region of H3 involved in the interaction is in yellow. The cacodylate is shown in CPK rendering.

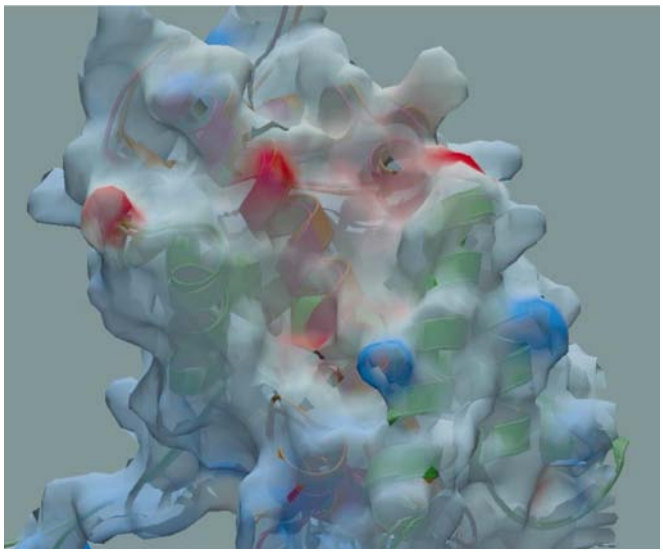

Figure 32
In this figure, a *GRASP* electrostatic potential surface has been been laid over a model of the dimers which have been superimposed by least-squares fitting of the dorsal dimer onto the ventral dimer to demonstrate that the symmetric remodeling of the histones is related to the asymmetry in the stacking interaction involved the acidic region on the dimer face.

and crystal-engineering technology, the structure here represents a significant improvement in the resolution of the structure and the determination of the atomic position for both the protein and DNA components of the NCP.

The well established importance of polyA-tract DNA, distributed approximately in phase with the helical repeat of DNA, to the ability of a DNA fragment to bind to the histone octamer is here given a novel interpretation. Apart from the controversy concerning curvature in polyA-tract DNA (Sprous *et al.*, 1995, 1999; Haran *et al.*, 1994; Dickerson, 1998; Olson & Zhurkin, 2000), it is generally agreed that polyA tracts serve to narrow the minor groove. In nucleosomal DNA sequences, AA dinucleotide steps flanking the minor groove-in position bring DNA phosphates closer together and provide an improved substrate for insertion of an arginine side chain at the mGI binding site. The path of the phosphate backbone will also more closely approximate the separation of the histone fold loops, forming the foundation of a βmGI binding motif. Thus, the width of the minor groove at the point at which DNA contacts the mGI binding motif may be the important factor determining rotational phasing of nucleosomes.

Additionally, the packing of the NCP in the crystal lattice is unlikely to reflect real packing of nucleosomes connected by linker DNA. The strong interaction of the protein faces of the NCP in the crystal structure involves a distinctive acidic surface on the dimer which binds to the tetramer surface of a neighboring NCP, particularly through the N-terminal tail of H4. This interaction is certainly suggestive of strong binding between nucleosomes in the condensed or 30 nm fiber even though the head-to-tail orientations of the two particles in the lattice is not appropriate to any current model of higher order structures. The extremely tight binding of a cacodylate ion within the dimer–tetramer interface connecting two NCPs provides a putative mechanism explaining the mutagenic properties of dimethyl arsenate. In addition to its use in crystallization buffers, cacodylate has been used as an herbicide, is the major metabolite of inorganic arsenic compounds and is known to cause chromosomal aberrations (Kashiwada *et al.*, 1998).

To understand better the functions of different covalent modifications that occur to the histones of the NCP, atomic positions for the histone tails must be extended and more precisely defined. Although these regions are sites of inherent disorder within the molecule, crystal-packing contacts have provided some clues about possible chromatin interactions. Sites of biological significance for acetylation, ADP-ribosylation, phosphorylation, methylation and ubiquitination all occur at positions on the histone tails. Improved structural characterization of histone tails using native variant histones and with those engineered to contain specific modifications may provide additional insight into these regulatory processes. The role of these modified regions in interactions within and between NCPs should also prove amenable to structural characterization.

Recent studies assessing the potential for cryogenic helium in macromolecular cryocrystallography (Hanson *et al.*, 1999, 2000) bode well for further elucidating the structure of the histone tails. An open-flow helium cryostat capable of holding a crystal at a temperature below 20 K at a third-generation synchrotron source has been shown to increase diffraction resolution, lower *B* factors and improve electron density for a number of crystal systems including the NCP. Although first experiments with NCP crystals showed some evidence of structure modulation owing to the extreme temperature, we are confident of acquiring greater detail about histone-tail atomic positions, modifications and interactions as well as providing more details of the structure of DNA in the nucleosome.

With the solution of the NCP structure, opportunities arise to use this structure as the anchor for structural studies of other chromatin assemblages. Modified nucleosomes occur at the centromere, with histone H3 replaced by H3-like proteins. These modified nucleosomes serve as sites of attachment for fibrils during sister chromatid disjunction at metaphase. We have produced nucleosome core particles containing the H3-like protein D6H3 from *Caenorhabtidis elegans*. This protein contains three cysteine residues rather than the single residue in H3. A structural model (Harp *et al.*, in preparation) shows that one of the cysteines is structurally homologous to that in H3 and that one of the cysteines occurs at the C-terminus of the mH helix of D6 in a position to form a disulfide bond across this hinge region of the tetramer. This is in contrast to H3, for which cysteines are not within disulfide-bond forming distance. A third cysteine residue is located on the N-terminal tail domain of D6H3.

Other studies need to be directed towards increasing the database of DNA sequences known to atomic resolution. One simple set of experiments includes changing the length of the DNA sequence to look at the effects of 145 and 147 bp DNA fragments on the NCP structure. We have also engineered a palindrome from the other half of the full-length α-satellite DNA sequence used in this structure. Our identification of correctly phasing α-satellite DNA sequences (Uberbacher *et al.*, 1988) and the development of the specific palindromic DNA for crystallization made high-resolution NCP structures attainable (Harp *et al.*, 1996). However, NCPs with non-palindromic DNA that diffract to lower resolution have been produced and could reflect more accurately the DNA sequences found in nature. A 3.1 Å NCP structure using native α-satellite human X chromosome DNA will be described in a separate publication.

Binding of the NCP with other DNA-interacting proteins offers an opportunity to provide atomic resolution models of DNA-modulating complexes. One of those complexes is that formed between the NCP and HMG 14/17 (Mardian *et al.*, 1980; Uberbacher *et al.*, 1982; Postnikov & Bustin, 1999; Trieschmann *et al.*, 1998). The regions of highest *B* factor on the DNA correlate to binding sites for HMG 14/17. Structures of complexes will help to establish the role of HMG 14/17 and other DNA-interacting proteins in gene regulation and chromatin assemblage.

The structure of the NCP leaves many important questions unanswered and it should be emphasized that the nucleosome core particle is not a physiological unit but an artifact of

nuclease digestion. The structure of the nucleosome has been 25 years in the making (see Kornberg & Lorch, 1999) and is not yet complete. The absence of linker DNA resulted in an asymmetry in the binding of the C-terminal domains of H2A and may have obscured the physiological role of the N-terminal tails of H3 which emerge between the gyres of DNA near the termini of the DNA in this structure. Future crystallographic studies must focus on higher order structures of chromatin including the linker DNA and the linker histones before these structural issues can be resolved.

The authors wish to acknowledge the important contributions of many individuals who have contributed to this project. Edward Uberbacher has made many significant contributions to the development of the DNA palindrome used in the construction of the nucleosome core particle. Arthur Roberson, Elise Palmer, Andreas Geweiss, Matthew Davis, Melissa York and Steve Plichta have made important contributions in the laboratory. The authors also wish to acknowledge Daniel Carter and present and former colleagues at NASA's Marshall Space Flight Center and New Century Pharmaceuticals for making the microgravity environment available for crystallization of the nucleosome core particles. Research reported here was performed as part of dissertation research at the University of Tennesseee by JMH. Diffraction data for this study were collected at Brookhaven National Laboratory in the Biology Department single-crystal diffraction facility at beamline X12-C in the National Synchrotron Light Source. This facility is supported by the United States Department of Energy Offices of Health and Environmental Research and of Basic Energy Sciences under prime contract DE-AC02-98CH10886, by the National Science Foundation and by National Institutes of Health Grant 1P41 RR12408-01A1. Research sponsored by grants from NIH (GM-29818), NASA (NAG8-1568), the Office of Biological and Environmental Research, US Department of Energy and the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the US Department of Energy under Contract No. DE-AC05-00OR22725. The submitted manuscript has been authored by a contractor of the US Government under contract No. DE-AC05-00OR22725. Accordingly, the US Government retains a non-exclusive royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for US Government purposes.

References

Adams, P. D., Pannu, N. S., Read, R. J. & Brunger, A. T. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 5018–5023.

Akker, F. van den & Hol, W. (1999). *Acta Cryst.* D**55**, 206–218.

Arents, G., Burlingame, R., Wang, B.-C., Love, W. E. & Moudrianakis, E. N. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 10148–10152.

Arents, G. & Moudrianakis, E. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 11170–11174.

Arscott, P. G., Ma, C., Wenner, J. R. & Bloomfield, V. A. (1995). *Biopolymers*, **36**, 345–364.

Bashkin, J., Hayes, J. J., Tullius, T. D. & Wolffe, A. P. (1993). *Biochemistry*, **32**, 1895–1898.

Baxevanis, A. D., Arents, G., Moudrianakis, E. N. & Landsman, D. (1995). *Nucleic Acids Res.* **23**, 2685.

Bentley, G. A., Lewit-Bentley, A., Finch, J. T., Podjarny, A. J. & Roth, M. (1984). *J. Mol. Biol.* **176,** 55–75.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–474.

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., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Carter, D. C., Lim, K., Ho, J. X., Wright, B. S., Twigg, P. D., Miller, T. Y., Chapman, J., Keeling, K., Rubel, J., Vekilov, P. G., Thomas, B. R., Rosenberger, F. & Chernov, A. A. (1999). *J. Cryst. Growth*, **196**, 623–637.

Carter, D. C., Wright, B., Miller, T., Chapman, J., Twigg, P., Keeling, K., Moody, K., White, M., Click, M., Rubel, J. A., Ho, J. X., Adcock-Downey, L., Bunick, G. & Harp, J. (1999). *J. Cryst. Growth*, **196**, 602–609.

Chambon, P. (1978). *Cold Spring Harbor Symp. Quant. Biol.* **42**, 1209–1234.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Dickerson, R. E. (1998). *Nucleic Acids Res.* **26**, 1906–1926.

Dickerson, R. E., Goodsell, D. & Kopka, M. L. (1996). *J. Mol. Biol.* **256**, 108–125.

Finch, J. T., Brown, R. S., Rhodes, D., Richmond, T., Rushton, B., Lutter, L. C. & Klug, A. (1981). *J. Mol. Biol.* **145**, 757.

Flock, S., Labarbe, R. & Houssier, C. (1996). *Biophys. J.* **70**, 1456–1465.

Furey, W. & Swaminathan, S. (1997). *Methods Enzymol.* **277**, 590–620.

Hanson, B. L., Harp, J. M., Kirschbaum, K., Parrish, D. A., Timm, D. E., Howard, A., Pinkerton, A. A. & Bunick, G. J. (2000). Submitted.

Hanson, B. L., Harp, J. M., Parrish, D. A., Bunick, C. G., Kirschbaum, K., Pinkerton, A. A. & Bunick, G. J. (1999). *J. Appl. Cryst.* **32**, 814–820.

Haran, T. E., Kahn, J. D. & Crothers, D. M. (1994). *J. Mol. Biol.* **244**, 135–143.

Harp, J. M., Hanson, B. L., Timm, D. E. & Bunick, G. J. (1999). *Acta Cryst.* D**55**, 1329–1344.

Harp, J. M., Palmer, E. L., York, M., Geweiss, A., Davis, M. & Bunick, G. J. (1995). *Electrophoresis*, **16**, 1861–1864.

Harp, J. M., Timm, D. E. & Bunick, G. J. (1998). *Acta Cryst.* D**54**, 622–628.

Harp, J. M., Uberbacher, E. C., Roberson, A. & Bunick, G. J. (1996). *Acta Cryst.* D**52**, 283–288.

Hauser, L. (1985). PhD thesis. University of California, Irvine, USA.

Hingerty, B. E., Figueroa, S., Hayden, T. L. & Broyde, S. (1989). *Biopolymers*, **28**, 1195–1222.

Holde, K. E. van, Allen, J. R., Tatchell, K., Weischet, W. O. & Lohr, D. (1980). *Biophys. J.* **32**, 271–282.

Holt, C. von, Brandt, W. F., Greyling, G. G., Lindsey, G. G., Retief, J. D., Rodrigues, J. A., Schwager, S. & Sewell, B. (1989). *Methods Enzymol.* **170**, 431–523.

Ichimura, S., Mita, K. & Zama, M. (1982). *Biochemistry*, **21**, 5329–5334.

Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.

Kashiwada, E., Kuroda, K. & Endo, G. (1998). *Mutation Res.* **413**, 33–38.

Kornberg, R. (1974). *Science*, **184**, 868–871.

Kornberg, R. D. & Lorch, Y. (1999). *Cell*, **98**, 285–294.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

Luger, K., Mader, A., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997). *Nature (London)*, **389**, 251–260.

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

McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.

Mardian, J. K. W., Paton, A. E., Bunick, G. J. & Olins, D. E. (1980). *Science*, **209**, 1523–1536.
Merritt, E. A. & Murphy, M. E. (1994). *Acta Cryst.* D**50**, 869–873.
Nicholls, A., Sharp, K. & Honig, B. (1991). *Proteins Struct. Funct. Genet.* **11**, 281–296.
O'Halloran, T. V., Lippard, S. J., Richmond, T. J. & Klug, A. (1987). *J. Mol. Biol.* **194**, 705–712.
Olins, A. L. & Olins, D. E. (1973). *J. Cell Biol.* **59**, 252a.
Olins, A. L. & Olins, D. E. (1974). *Science*, **183**, 330–332.
Olson, W. K. & Zhurkin, V. B. (2000). *Curr. Opin. Struct. Biol.* **10**, 286–297.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Palmer, E., Geweiss, A., Harp, J. M., York, M. H. & Bunick, G. J. (1996). *Anal. Biochem.* **231**, 109–114.
Pennings, S., Meerssemans, G. & Bradbury, E. M. (1992). *Nucleic Acids Res.* **20**, 6667–6672.
Pittz, E. P. & Timasheff, S. N. (1978). *Biochemistry*, **17**, 615–623.
Postnikov, Y. V. & Bustin, M. (1999). *Methods Enzymol.* **304**, 133–155.
Ramakrishnan, V. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 11328–11330.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–146.
Rhodes, D., Brown, R. S. & Klug, A. (1989). *Methods Enzymol.* **170**, 420–428.
Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D. & Klug, A. (1984). *Nature (London)*, **311**, 532–537.
Roussel, A. & Cambillau, C. (1991). *Silicon Graphics Partners Directory*, p. 86. Silicon Graphics, Mountain View, CA, USA.
Sayle, R. & Milner-White, E. J. (1995). *Trends Biochem. Sci.* **20**, 374–376.
Sprous, D., Young, M. A. & Beveridge, D. L. (1999). *J. Mol. Biol.* **285**, 1623–1632.
Sprous, D., Zacharias, W., Wood, Z. A. & Harvey, S. C. (1995). *Nucleic Acids Res.* **23**, 1816–1821.
Studitsky, V. M. (1997). *Science*, **278**, 1960–1964.
Timasheff, S. N. (1993). *Annu. Rev. Biophys. Biomol. Struct.* **22**, 67–97.
Trieschmann, L., Martin, B. & Bustin, M. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 5468–5473.
Uberbacher, E. C. & Bunick, G. J., (1986). *Purification of Nucleoprotein Particles by Elution Preparative Gel Electrophoresis.* Available from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161, USA.
Uberbacher, E. C. & Bunick, G. J. (1989). *J. Biomol. Struct. Dyn.* **7**, 1–18.
Uberbacher, E. C., Mardian, J. K. W., Rossi, R. M., Olins, D. E. & Bunick, G. J. (1982). *Proc. Natl Acad. Sci. USA*, **79**, 5258–5262.
Uberbacher, E. C., Wilkinson-Singley, E., Harp, J. M. & Bunick, G. J. (1988). *DNA Bending and Curvature*, edited by W. K. Olson, M. H. Sarma, R. H. Sarma & M. Sundaralingam, Vol. 3, *Structure and Expression*, pp. 139–158. Schenectady, NY, USA: Adenine Press.
Wang, B.-C. (1985). *Methods Enzymol.* **115**, 90–112.
Widom, J. (1998). *Annu. Rev. Biophys. Biomol. Struct.* **27**, 285–327.
Woodcock, C. F. L. (1973). *J. Cell Biol.* **59**, 368a.
Yang, T. P., Hansen, S. K., Oishi, K. K., Ryder, O. A. & Hamkalo, B. A. (1982). *Proc. Natl Acad. Sci. USA*, **79**, 6593–6597.