

Solvent Mediated Interactions in the Structure of the Nucleosome Core Particle at 1.9 Å Resolution[†]

Curt A. Davey, David F. Sargent, Karolin Luger, Armin W. Maeder and Timothy J. Richmond*

*Institut für Molekularbiologie
und Biophysik
ETH-Hönggerberg, CH-8093
Zurich, Switzerland*

Solvent binding in the nucleosome core particle containing a 147 base pair, defined-sequence DNA is characterized from the X-ray crystal structure at 1.9 Å resolution. A single-base-pair increase in DNA length over that used previously results in substantially improved clarity of the electron density and accuracy for the histone protein and DNA atomic coordinates. The reduced disorder has allowed for the first time extensive modeling of water molecules and ions.

Over 3000 water molecules and 18 ions have been identified. Water molecules acting as hydrogen-bond bridges between protein and DNA are approximately equal in number to the direct hydrogen bonds between these components. Bridging water molecules have a dual role in promoting histone–DNA association not only by providing further stability to direct protein–DNA interactions, but also by enabling formation of many additional interactions between more distantly related elements. Water molecules residing in the minor groove play an important role in facilitating insertion of arginine side-chains. Water structure at the interface of the histones and DNA provides a means of accommodating intrinsic DNA conformational variation, thus limiting the sequence dependency of nucleosome positioning while enhancing mobility.

Monovalent anions are bound near the N termini of histone α -helices that are not occluded by DNA phosphate groups. Their location in proximity to the DNA phosphodiester backbone suggests that they damp the electrostatic interaction between the histone proteins and the DNA. Divalent cations are bound at specific sites in the nucleosome core particle and contribute to histone–histone and histone–DNA interparticle interactions. These interactions may be relevant to nucleosome association in arrays.

© 2002 Elsevier Science Ltd. All rights reserved

Keywords: chromatin; histone; protein–DNA interaction; solvent interaction; X-ray crystal structure

*Corresponding author

Introduction

DNA in eukaryotic organisms is organized in a nucleoprotein complex called chromatin. Genetic

[†] We dedicate this paper to the memory of Max Perutz who was particularly inspirational and supportive to T.J.R. in the early stages of this study.

Present address: K. Luger, Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO 80523-1870, USA.

Abbreviations used: ASA, solvent-accessible surface area; NCP, nucleosome core particle; SHL, superhelix location.

E-mail address of the corresponding author: richmond@mol.biol.ethz.ch

processes, both vital such as transcription and replication, and pathological such as cancer and viral infection, depend on DNA in the context of chromatin. The nucleosome is the fundamental repeating unit of chromatin occurring generally every 157–240 bp. The nucleosome core, the greater part of the nucleosome, comprises an octamer, containing a single histone H3–H4 tetramer and two histone H2A–H2B dimers, and 147 bp of DNA. We have determined the structure of the nucleosome core particle containing recombinant histones and a 147 bp, defined-sequence DNA (NCP147) by X-ray crystallography at 1.9 Å resolution. The significantly improved electron density for this particle compared to previous

reports on the related 146 bp particle results from the increase in DNA length from 146 to 147 bp.^{1,2} This has allowed location of 3130 water molecules, 14 manganese cations and four chloride anions in the refined structure. We also report here the refined structures of the original 146 bp particle (NCP146) and a second particle (NCP146b) containing a different 146 bp DNA sequence at 2.0 Å and 2.6 Å resolution, respectively. Both NCP147 and NCP146 provide significantly more complete and accurate structures of the nucleosome core particle compared to recent reports using 146 bp DNA fragments, endogenous histones or recombinant variants.^{3–5}

The important contribution of solvent molecules to the stability and specificity of protein–DNA interactions has been recognized through a combination of high-resolution structural results and allied thermodynamic studies.^{6–12} Although the X-ray and NMR structures of over 200 DNA-binding proteins have been determined, few are of adequate quality to pinpoint solvent interactions accurately.¹³ Those which are, have on average 12 water molecules located in positions mediating protein–DNA interactions.¹⁴ In contrast, the NCP147 structure reveals ten times this number of water molecules at the protein–DNA interface. In regard to non-specific DNA-binding proteins, water molecules are envisioned to play the role of adapters, allowing for roughly uniform affinity between protein and differing DNA sequences despite DNA-sequence-dependent conformational variability. Although general rules may emerge for non-specific, protein–DNA association from the few relevant structures available, elucidation of important mechanistic detail requires that each individual motif be examined.¹⁵

Even though nucleosomes are ubiquitous on genomic DNA, they display local, preferred-sequence positioning as well as the ability to move diffusively along the DNA.^{16,17} Solvent interactions are likely to contribute substantially to positioning and mobility. Therefore, elucidation of the structural details of all the interactions accounting for histone protein and DNA association is required to understand the molecular basis of eukaryotic genetic processes. We present here for the first time a detailed description of the water molecule and ion structure in the nucleosome core particle.

Results

147 bp versus 146 bp particles

Crystals of NCP146 containing *Xenopus laevis* core histones and a 146 bp palindromic DNA, taken from one-half of a human α -satellite sequence repeat, were used previously for the structure determination of the nucleosome core at 2.8 Å and 2.0 Å resolution.^{1,2} The NCP147, NCP146 and NCP146b structures reported here at

1.9 Å, 2.0 Å and 2.6 Å, respectively (Table 1), show the same features overall, limited only by differences in resolution and influences from DNA disorder. These structures show that the particle comprises DNA arranged in 1.65 turns of a left-handed superhelix surrounding an octamer containing two copies of each of core histone H2A, H2B, H3 and H4 (Figure 1(a)). The two halves of NCP147, NCP146 and NCP146b are related by a pseudo-2-fold symmetry axis (dyad axis). This dyad axis passes through the central base-pair in the sequence of NCP147 and one of the two central base-pairs in NCP146 and NCP146b. NCP146b incorporates a different half-repeat, human α -satellite DNA sequence compared to NCP147 and NCP146.

For NCP147, the electron density and model for the DNA is of comparable quality to the best regions of the protein. The histone proteins are divided between poorly ordered N-terminal tail regions and well-ordered histone-fold domains and histone-fold extensions (Figure 1(b)). Notably for the two 146 bp DNA-containing structures, the electron density for the histone-fold and extension regions is significantly better defined than for most of the DNA. This observation is explained by the overwinding and stretching of the DNA by 1 bp in NCP146 and NCP146b compared to NCP147. This distortion is apparently generated by the base-stacking forces between terminal base-pairs of adjacent particles in the crystals and translocates through the DNA duplex, settling primarily in one 12 bp region on one side of the DNA superhelix—superhelix location -2.0 (SHL -2.0) in NCP146 and SHL -5.0 in NCP146b. However, the disorder in DNA conformation is evidently dispersed at lower occupancy throughout the 146 bp DNA structures, degrading the quality of the electron density overall. We reasoned that the clarity of the DNA structure could be improved by increasing the length of the DNA sequence by a single base-pair to 147 bp. Indeed, NCP147 no longer significantly displays this distortion, as indicated by the mean real-space R -factor¹⁸ values for the DNA nucleotides of 0.076 and 0.135 for NCP147 and NCP146, respectively (Figure 1(c)). This improvement is also apparent from (i) the overall coordinate errors¹⁹ of 0.23 Å and 0.28 Å, and (ii) mean B -factors for the DNA of 61.4 Å² and 82.5 Å² for NCP147 and NCP146, respectively, with both structures refined to virtually identical resolution (Table 1).

The DNA of NCP147 remains roughly twice as mobile or statically disordered in comparison to the protein histone-fold regions based on ratios of mean crystallographic B -factors. The value over all DNA atoms is 61.4 Å² compared to 29.4 Å² for the histone-fold domains and 38.4 Å² for all assigned protein atoms (Figure 1(d)). The DNA mobility base by base is highly non-uniform, being minimal in the region of the histone–DNA contact points where the minor groove faces the histone octamer, and maximal for the bases with phosphate groups

Table 1. Data collection and refinement statistics

	NCP147	NCP146	NCP146b
<i>Data collection</i>			
No. crystals	44	27	5
Resolution (Å)	20–1.94	99–2.0	30–2.6 ^a
Resolution of last shell (Å)	2.01–1.94	2.04–2.0	2.69–2.60
Redundancy	5.1	7.4	3.7
No. unique <i>hkl</i>	151,140	145,317	60,481
Completion % (last shell)	95.6 (88.0)	98.3 (84.5)	89.5 (45.7)
R_{merge} % (last shell)	7.2 (29.1)	6.8 (49.5)	7.2 (15.7)
Unit cell	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
<i>a</i> (Å)	105.95	105.40	105.30
<i>b</i> (Å)	181.17	181.54	175.69
<i>c</i> (Å)	109.49	109.60	109.53
<i>Refinement</i>			
Resolution (Å)	6.0–1.94	6.0–2.0	6.0–2.6
Resolution (last shell Å)	2.01–1.94	2.07–2.00	2.68–2.60
<i>R</i> -factor % (last shell)	20.8 (34.1)	24.0 (31.2)	24.6 (32.9)
Free <i>R</i> -factor % (last shell)	27.5 (37.4)	27.5 (32.2)	30.0 (36.1)
Reflections work/free	141,038/2898	133,711/2716	51,863/1043
<i>No. atoms in model</i>			
Total (mean <i>B</i> -factor Å ²)	15,965 (51.5)	13,023 (58.2)	12,438 (51.6)
Protein (mean <i>B</i> -factor Å ²)	6796 (38.4) ^b	6087 (35.2)	6015 (30.1)
DNA (mean <i>B</i> -factor Å ²)	6021 (61.4)	5980 (82.5)	5980 (74.0)
Solvent (mean <i>B</i> -factor Å ²)	3148 (60.8)	956 (52.1)	443 (43.3)
<i>r.m.s.d. from ideality</i>			
Bond lengths (Å)	0.009	0.010	0.006
Bond angles (deg.)	1.24	1.18	0.96
Dihedral angles (deg.)	19.2	18.9	18.5
Mean coordinate error (Å)	0.23	0.28	0.35

^a X-ray data to the observed diffraction limit of 2.0 Å have not yet been measured.

^b In all, 10.4% of the protein atoms were not included in the model owing to disorder in the histone terminal regions.

on the outside of the particle. The number, as well as accuracy, of solvent molecules modeled in different regions of the structure is correlated with the local *B*-factor magnitude.

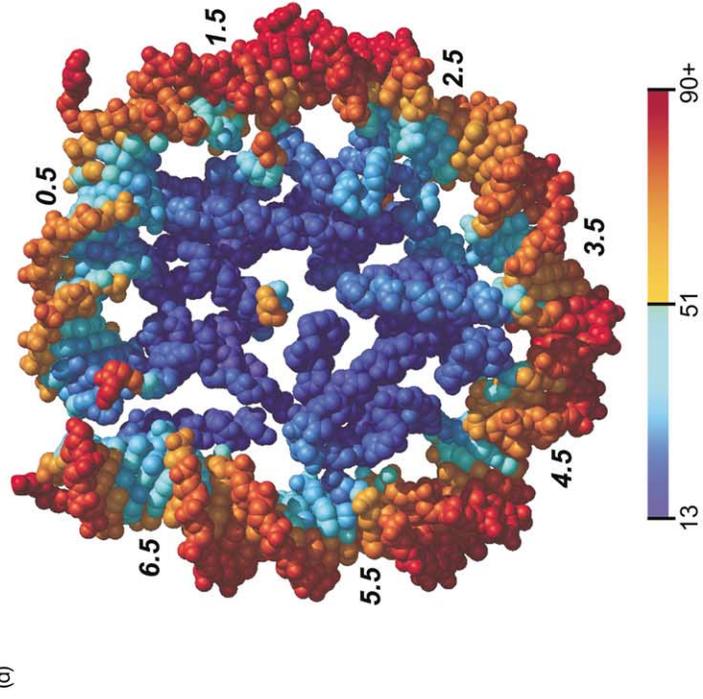
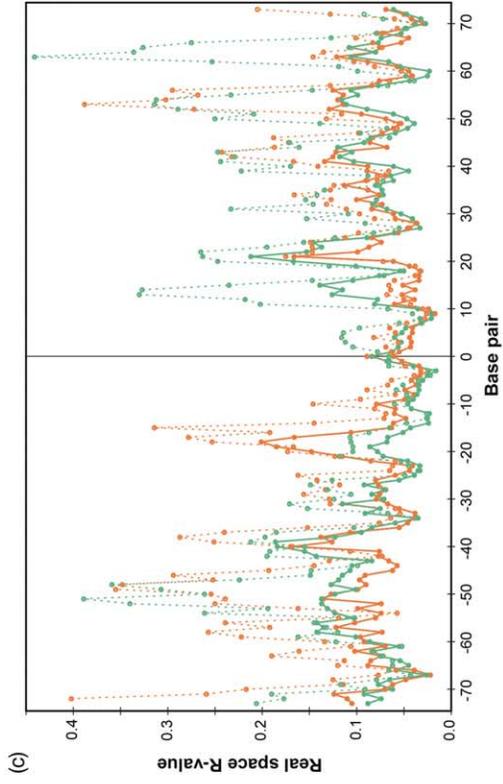
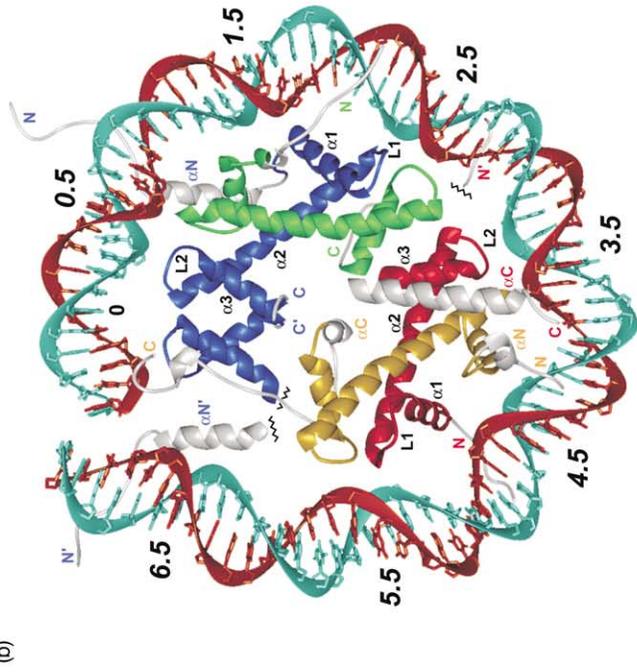
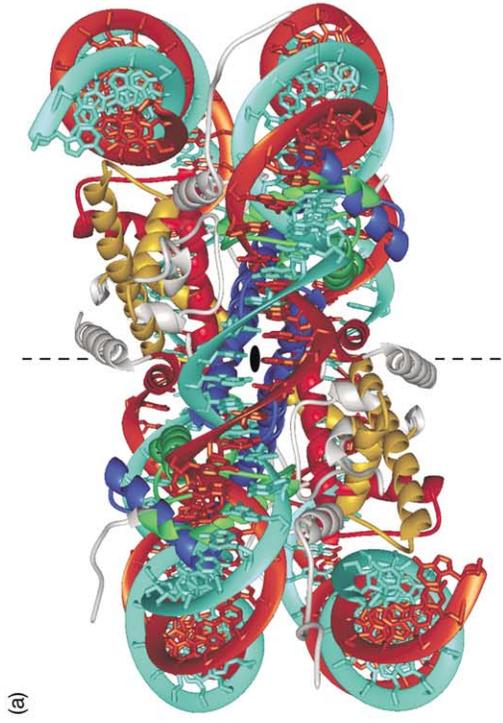
Hydration of histone and DNA

NCP147 has 74,049 Å² of solvent-accessible surface area (ASA), much of which occurs in cavities within the histone octamer or between it and the DNA. This distribution of surface as well as the high resolution of the NCP147 structure has enabled us to model 3130 water molecules, yielding a total of 15,965 atoms for protein, DNA, water molecule and ions (Figure 2(a)). The primary hydration layer of the complete particle contains 2088 water molecules (centers within 3.5 Å of any protein or DNA atom) or about two-thirds of the total. In this layer, 1108 water molecules surround the DNA (Figure 2(b)), and of these, 302 water molecules are found within the minor groove, predominantly where runs of four or five A-T base-pairs occur in the DNA sequence. Previous studies on oligonucleotide structure have described the arrangements seen here in the minor groove as a “spine of hydration”, for which arrays of water molecules are bound regularly to adenine N3 and thymine O2 atoms.²⁰ On the inner face of the DNA superhelix where the minor groove is narrowest,

water molecules are seen to take up the most regular structures, including the five and six-member fused rings that broaden the spine of hydration observed at SHL4.5 (Figure 2(c)). Structural analyses of nucleic acid hydration have revealed that the phosphate groups are the most strongly solvated components of DNA, with the primary O1P and O2P oxygen atoms making hydrogen bonds to water molecules much more frequently than the secondary O3' and O5' oxygen atoms.^{21,22} In the NCP147, 78% of the water molecule hydrogen bonds made to the phosphate groups occur with the O1P and O2P oxygen atoms. Furthermore, solvent hydrogen bonds with the phosphodiester backbone outnumber those with bases by 1.4:1.

Histone DNA-binding sites

The interactions between protein and DNA within the nucleosome core particle can be divided into seven unique DNA-binding sites per half-particle, corresponding to the regions where the minor groove of the double helix faces inward (SHL0.5 to SHL6.5, Figure 1(b)). Six of these sites on the histone octamer cover the central 121 bp of the DNA (SHL0.5 to SHL5.5) and are constructed primarily of paired elements from different histone-fold domains (Figure 3). Binding of the



terminal 13 bp of the DNA superhelix (SHL6.5) is accomplished by the H3- α N' helix (H3 of other half-particle) and the root of the H3 N-terminal tail (SHL-0.7 and SHL6.7) (Figure 4(a)). Additionally, the root of the N-terminal tail of H2B, which is highly basic and poorly ordered, passes through a minor groove channel formed by adjacent DNA gyres (SHL-2.7 and SHL4.7) as it exits the particle (Figure 4(b)).

The reduction of ASA due to the protein-DNA interface in NCP147 is 12,671 Å² or 14.6% of the combined total calculated for the histone octamer at 34,310 Å² and the DNA superhelix at 52,410 Å² separately. Dehydration of DNA phosphate groups owing to hydrogen bonds formed with peptide amide groups, serine and threonine hydroxyl groups, and arginine and lysine side-chain guanidinium and amino groups accounts for 59.4% of the 6713 Å² total lost for the DNA. The van der Waals contacts between histone and the hydrocarbon portions of the phosphodiester backbone account for 34.6% or most of the remainder. Except for an important set of arginine side-chains penetrating the minor groove, the majority of protein-DNA interactions involve the most accessible atoms of the phosphodiester chains.

Direct versus water-bridge histone-DNA interactions

A total of 121 water molecules mediate interactions between histone and DNA throughout NCP147 by making hydrogen bonds simultaneously with both elements (see Supplementary Material). Quantifying the relative importance to the stability of the nucleosome conferred by the water-bridge interactions in comparison to the direct interactions between protein and DNA is not straightforward. Uncertainty over the entropic cost of localized water molecules as well as for the strength of bonds with non-ideal geometry is limiting. Nevertheless, the number of hydrogen bonds can be tabulated in each case. In NCP147, there are 116 hydrogen bonds made directly between protein and DNA and 358 hydrogen bonds made

between bridging water molecules and either protein or DNA (Table 2). In an aqueous system in which non-interacting, directly interacting or water-bridge groups of the protein and DNA have essentially equal access to solvent, as is the case here, the enthalpic contribution to hydrogen bond formation will be, on average, the same for direct versus water-bridge interactions. The additional stability gain afforded by either type of interaction stems from the entropic contribution of solvent liberation.²³ A direct interaction between two groups releases two water molecules, but a water-bridge interaction requires association of three groups *via* a water molecule to free two water molecules. The mean number of hydrogen bonds made to bridging water molecules in NCP147 is 3.0, and as there are 116 direct interactions and 121 bridging water molecules, the contribution to nucleosome core stability from direct and water-bridge interactions is approximately equal. Electrostatic differences due to the larger separation between basic side-chains and phosphate groups for water-bridge versus direct interactions are compensated for by the polarizability of water and by the observation that water-bridge interactions frequently employ two positively charged amino acid side-chains. The "hydrogen bond inventory" analysis, put forward to account semi-quantitatively for the formation of direct hydrogen-bond interactions,²⁴ is evidently applicable to water-bridge interactions if an additional hydrogen bond is required to offset the cost of each bridging water molecule.

The distribution of hydrogen bonds attributed to bridging water molecules is 1.4:1 between the DNA backbone and bases, and 1:2 between the histone main-chain and side-chains. Comparable numbers for direct interactions are 6.7:1 and 1:1.3. Bridging water molecules clearly increase the range over which hydrogen-bonding moieties held apart by tertiary and quaternary structure constraints are able to produce stabilizing interactions between protein and DNA. Since bridging water molecules, on average, make three hydrogen bonds with protein and DNA, they, in addition to the direct hydrogen bonds, will place restraints on

Figure 1. Structure of the 147 bp nucleosome core particle (NCP147) at 1.9 Å resolution. (a) View down the axis of 2-fold pseudo-symmetry (dyad axis, black) with the DNA superhelix axis oriented vertically (broken line). The dyad axis bisects the central base-pair. The 147 bp palindromic DNA sequence shows nearly perfect 2-fold symmetry relating the two 73 bp halves of the DNA superhelix extending from the central base-pair. The DNA strands are cyan and brown. The histone-fold domains of the histone proteins are blue for H3, green for H4, yellow for H2A and red for H2B. The histone-fold extensions and N-terminal tail regions shown are white. (b) View down the DNA superhelix axis showing one half of the structure to illustrate the organization of histone and DNA. Colors are as for (a). The superhelix locations are labeled at the DNA-binding sites of the histone-fold pairs and the H3- α N helix (SHL: 0.5, 1.5, 2.5, 3.5, 4.5, 5.5, 6.5). The central base-pair is indicated (0). The histone-fold substructure for histones H3 and H2B are labeled (α 1, L1, α 2, L2, α 3) as are histone-fold extensions (α N, α C) and segments of the N and C-terminal tails (N, C, N', C'). (c) Real space *R*-value versus base-pair number. Plots are shown separately for both DNA strands (cyan and brown) of both NCP147 (continuous line) and NCP146 (broken line). (d) Temperature factor (crystallographic *B*-factor) values for half of NCP147. The van der Waals surface is shown in the same view as in (b). The atoms are colored by temperature factor as indicated in the scale from 13 to 90+ Å².

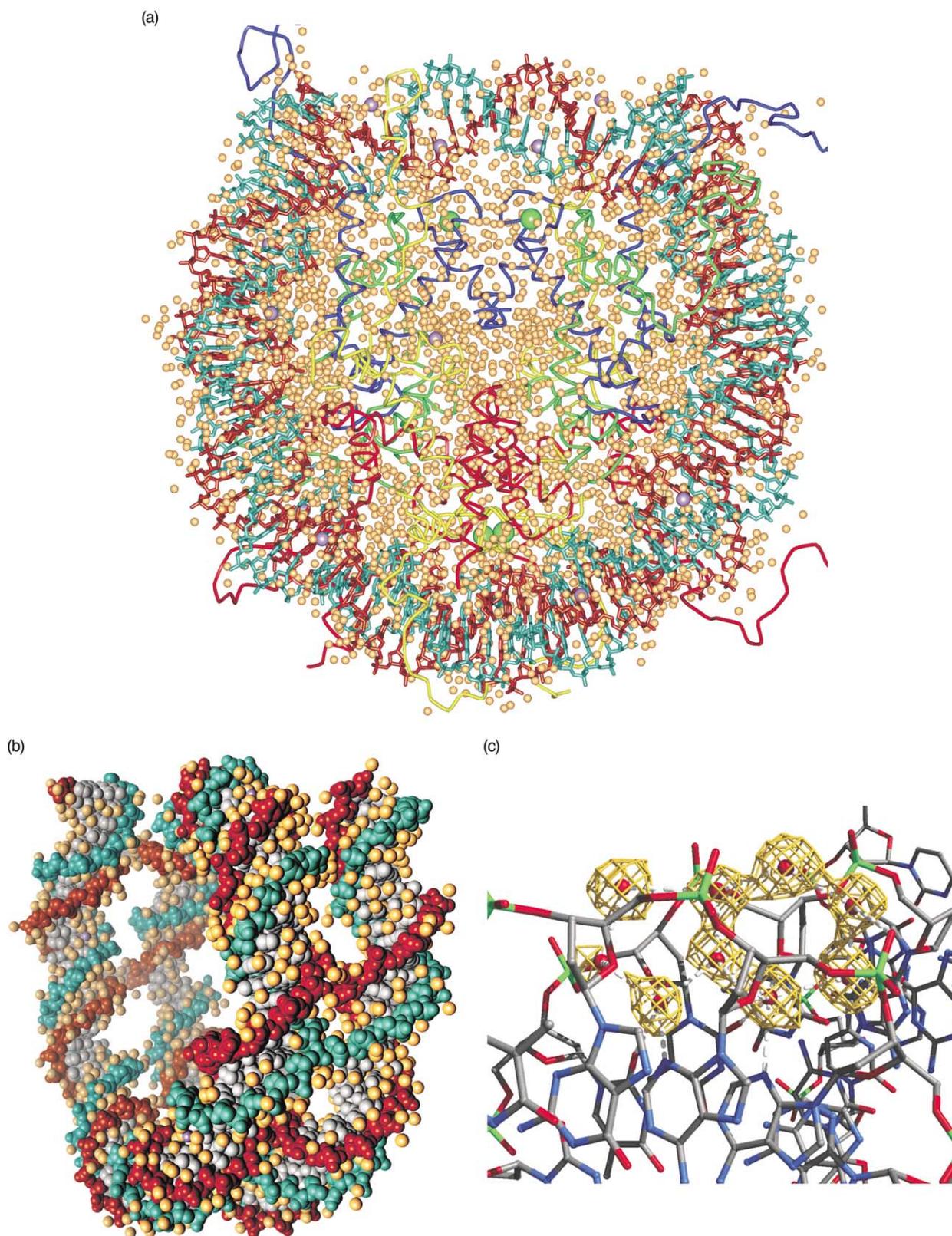


Figure 2. Solvation of NCP147. (a) Water molecules and ions associated with NCP147. The view is as in [Figure 1\(b\)](#). Water molecules (gold) are shown as spheres of half van der Waals radius. Manganese ions (violet) and chloride ions (green) are shown as spheres of van der Waals radius. The path of the histone chains and DNA strands are shown (H3, blue; H4, green; H2A, yellow; H2B, red; DNA strands, cyan and brown). The histone N-terminal tails are not shown in their entirety. (b) Space-filling representation of the DNA and its primary hydration layer. Water molecules (gold) with centers closer than 3.5 Å from any DNA atom center are shown. The DNA superhelix (backbones strands, cyan and brown; bases, silver) is rotated 60° around the dyad axis compared to the view in (a). (c) Minor groove “spine of hydration”. Five and six-member fused rings of water molecules (red) with hydrogen bonds (broken, white) are shown in the DNA minor groove with the simulated-annealing, omit difference electron density ($F_o - F_c$, 1.3 σ contour, yellow) superimposed.

any sequence-dependent conformational variability of DNA bound in the nucleosome.

Bridging water molecules involved in protein–DNA interactions are one of two types, either “assisting” or “facilitating”. Water molecules in the assisting class further link histone main-chain or side-chain groups with DNA phosphate groups that already have one or more direct hydrogen bonds with each other. Water molecules in the facilitating class connect groups that are too distant or inappropriately orientated to permit direct interactions and account for 83% of the bridging water molecules in NCP147, usually involving either basic side-chains or peptide amide groups. In many cases, an arginine or lysine side-chain is fully extended toward a phosphate group and has a water molecule lengthening its reach to a phosphate oxygen atom (e.g. [Figure 3\(a\)](#): H3-K122, H4-R39). Nevertheless, the local environment may override phosphate group interaction, as occurs for those arginine side-chains which have preferred interaction with aspartate side-chains instead of adjacent DNA phosphate groups ([Figure 3\(a\)](#): H3-R116/D123; [Figure 3\(c\)](#): H4-R78/D85).

Side-chain/DNA versus main-chain/DNA interactions

The limited assortment of side-chain types found to make direct hydrogen bonds to DNA is enlarged for water-mediated interactions ([Table 2](#)). Direct interactions involve arginine and lysine residues predominantly (e.g. [Figure 3\(b\)](#), H4-R36, H3-R69; [Figure 3\(c\)](#), H4-K79), followed by serine and threonine (e.g. [Figure 3\(a\)](#), H3-T118; [Figure 3\(c\)](#), H4-T80; [Figure 3\(f\)](#), H2A-T76, H2B-S53) and one occurrence of tyrosine ([Figure 3\(f\)](#), H2B-Y39). Water-mediated interactions additionally include glutamate, glutamine and histidine side-chains (e.g. [Figure 3\(c\)](#), H3-Q85). Water molecules act as adaptors allowing main-chain carbonyl groups to be linked to the DNA phosphodiester chains. The main-chain carbonyl oxygen atoms are brought into proximity to the DNA by the many hydrogen bonds made between neighboring main-chain amide nitrogen atoms and phosphate oxygen atoms (e.g. [Figure 3\(a\)](#), H4-I46). Remarkably, three glutamate carboxyl groups are linked *via* single water molecule adaptors to DNA ([Figure 3\(a\)](#), H3-E50; [Figure 3\(d\)](#), H2A-E41; [Figure 3\(e\)](#), H2B-E32). In all cases, the glutamate side-chain forms additional direct hydrogen bonds with arginine or serine side-chains, which are in turn linked directly or indirectly to phosphate groups. One possible role for these negatively charged carboxyl groups residing close to the negatively charged DNA backbone is the reduction of the overall affinity of the histone octamer for DNA. This adjustment may be important for nucleosome assembly, disassembly and mobility along the DNA.

In contrast to the arginine and lysine side-chains, which are intrinsically flexible and observed in

some cases to have different conformations in the two halves of NCP147, the conformation of the peptide backbone is essentially identical between the two halves. Superposition of the histone-fold domains and the histone H2A and H3- α N helices between copies in the two halves of the NCP147 yields values for the r.m.s. deviation on atomic positions for all protein atoms and for main-chain atoms of 0.76 Å and 0.27 Å, respectively. Amide nitrogen atoms in the conformationally invariant protein backbone are responsible for one-half of all direct hydrogen bonds between histone and DNA backbone ([Table 2](#)). Of 50 such direct interactions, the 18 occurring at the N termini of α -helices are strengthened by the delocalized positive charge of 0.5–0.7 associated with the helix dipole.²⁵ Importantly, of the 27 peptide amide groups involved in water-bridge interactions, 23 are helix-dipole assisted and should be similarly reinforced.

At each histone DNA-binding site, the same protein chemical species are found involved in both direct and indirect interactions with DNA phosphate groups. The well-defined main-chain conformation of the histone proteins could dictate the particular set of interactions used, with the DNA conforming completely to the histone DNA-binding template. Conversely, the histone-bound DNA might display sequence-dependent conformational variation such that the interactions made would depend on DNA sequence. In this respect, the region from SHL4.5 through SHL5.5 was compared in NCP147 with NCP146b as it contains the largest number of DNA sequence differences between the two structures ([Figure 5](#)). In a total of 27 interactions occurring for NCP147 and 17 for NCP146b, 11 changes take place between three possible interaction states: (1) bound directly, (2) bound indirectly, and (3) unbound, clearly demonstrating DNA sequence effects. Such alternative interactions may help accommodate binding of different sequences in the nucleosome generally, and in addition facilitate DNA mobility within the nucleosome core.

Side-chains penetrating the minor groove

Although the majority of interactions between histone and DNA engage the DNA phosphate groups, important interactions also occur for side-chains penetrating into the DNA minor groove. Together, the 20 side-chains that bind in the minor groove account for a loss of 2962 Å² ASA or nearly one-fourth of the total reduction in ASA for NCP147. These side-chains typically reside in the narrowest regions of the minor groove where it faces directly inward, sandwiched between flanking deoxyribose moieties. Most of the ASA reduction applicable to the DNA minor groove (1522 Å²) results from the desolvation of the deoxyribose hydrocarbon atoms (1179 Å²), with lesser amounts for the O4' atoms (137 Å²) and minor groove base substituents (206 Å²). For comparison,

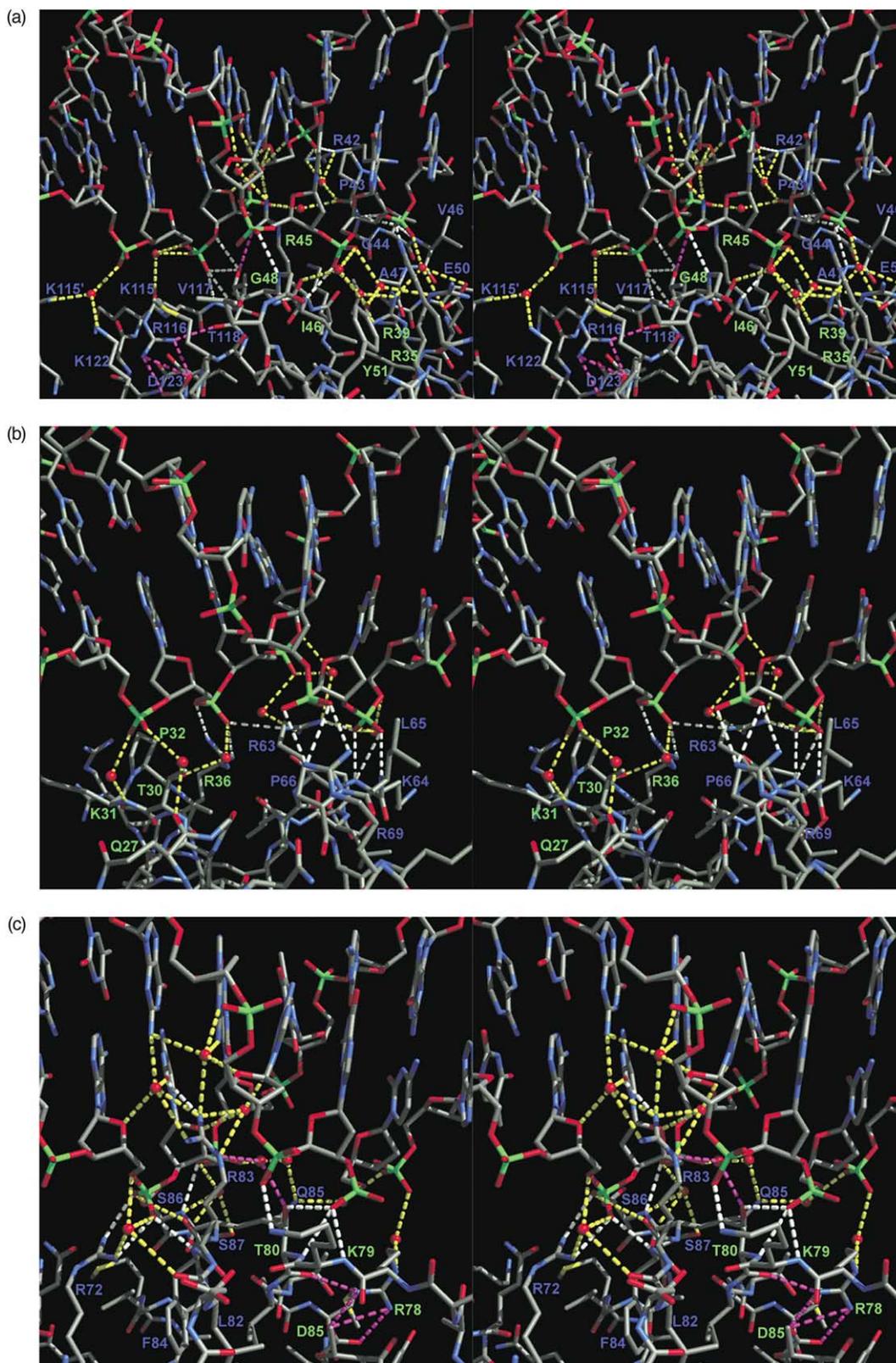


Figure 3 (legend opposite)

the reduction in major groove ASA overall is only 61 \AA^2 , a consequence of the interactions of H3-L65 and H4-P32 with thymine methyl groups (Figure 3(b)).

An arginine side-chain is located in the minor groove of all histone DNA-binding sites with one exception. Additionally, a proline, arginine, tyrosine, and histidine side-chain from the DNA

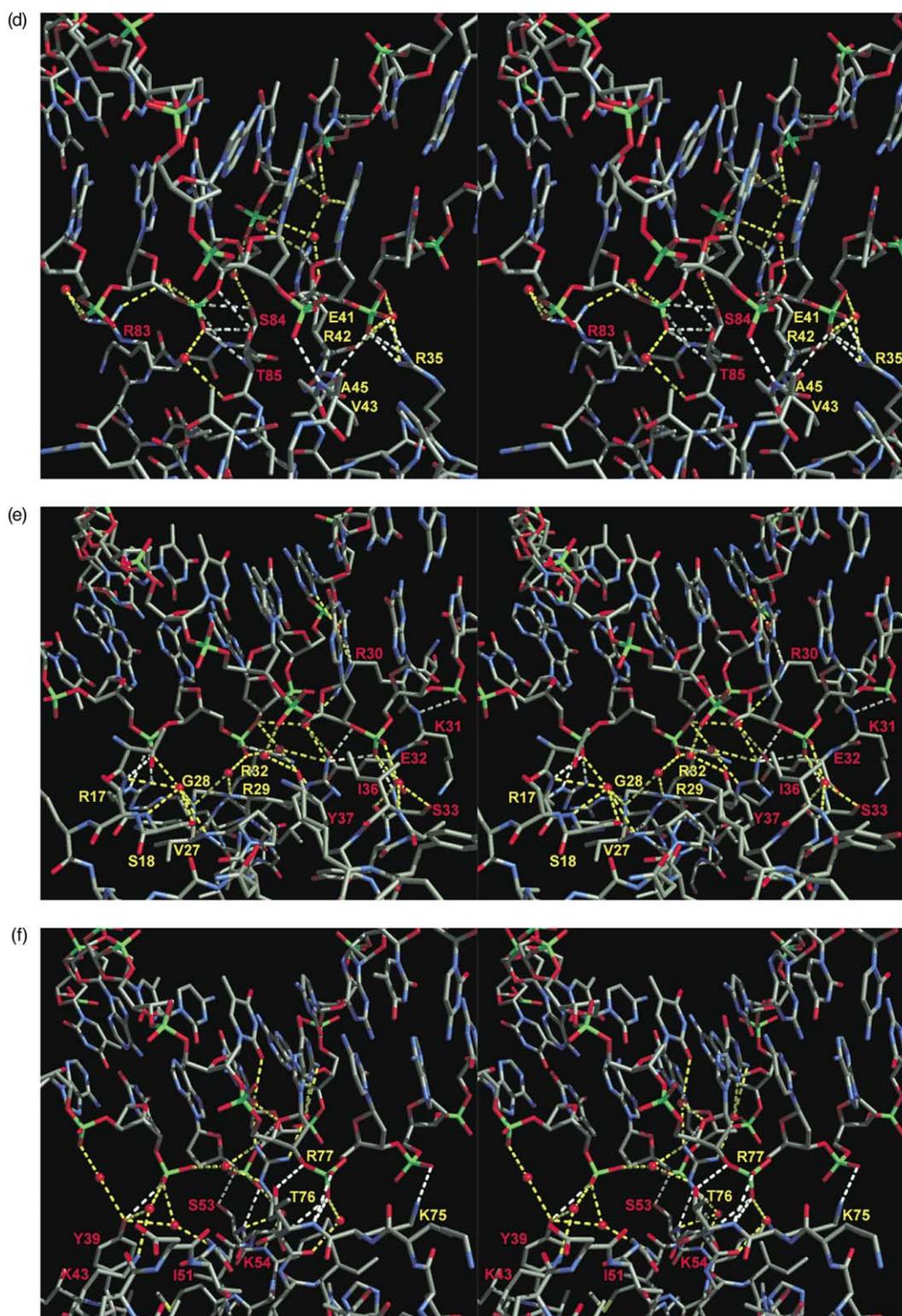


Figure 3. Histone-fold DNA-binding sites (stereographs). The DNA-binding sites are shown for one-half of the H3–H4 tetramer at (a) SHL0.5, (b) SHL1.5, (c) SHL2.5, and for one H2A–H2B dimer at (d) SHL3.5, (e) SHL4.5, (f) SHL5.5. Amino acid residues interacting with the DNA are labeled (H3, blue; H4, green; H2A, yellow; H2B, red) and shown with connecting direct (white) and water-bridge (yellow) hydrogen bonds, and surrounding water molecules (red spheres). Further hydrogen bonds (magenta) highlight specific arginine residue interactions.

gyre-binding region of the H3 N-terminal tail make contacts in the channel formed by the two flanking minor grooves at SHL–0.7 and SHL6.7 (Figure 4(a)). Multiple arginine and lysine side-

chains from the gyre-binding region of the H2B N-terminal tail also pass through a minor groove channel at SHL–2.7 and SHL4.7, but this highly basic region of H2B is poorly ordered and

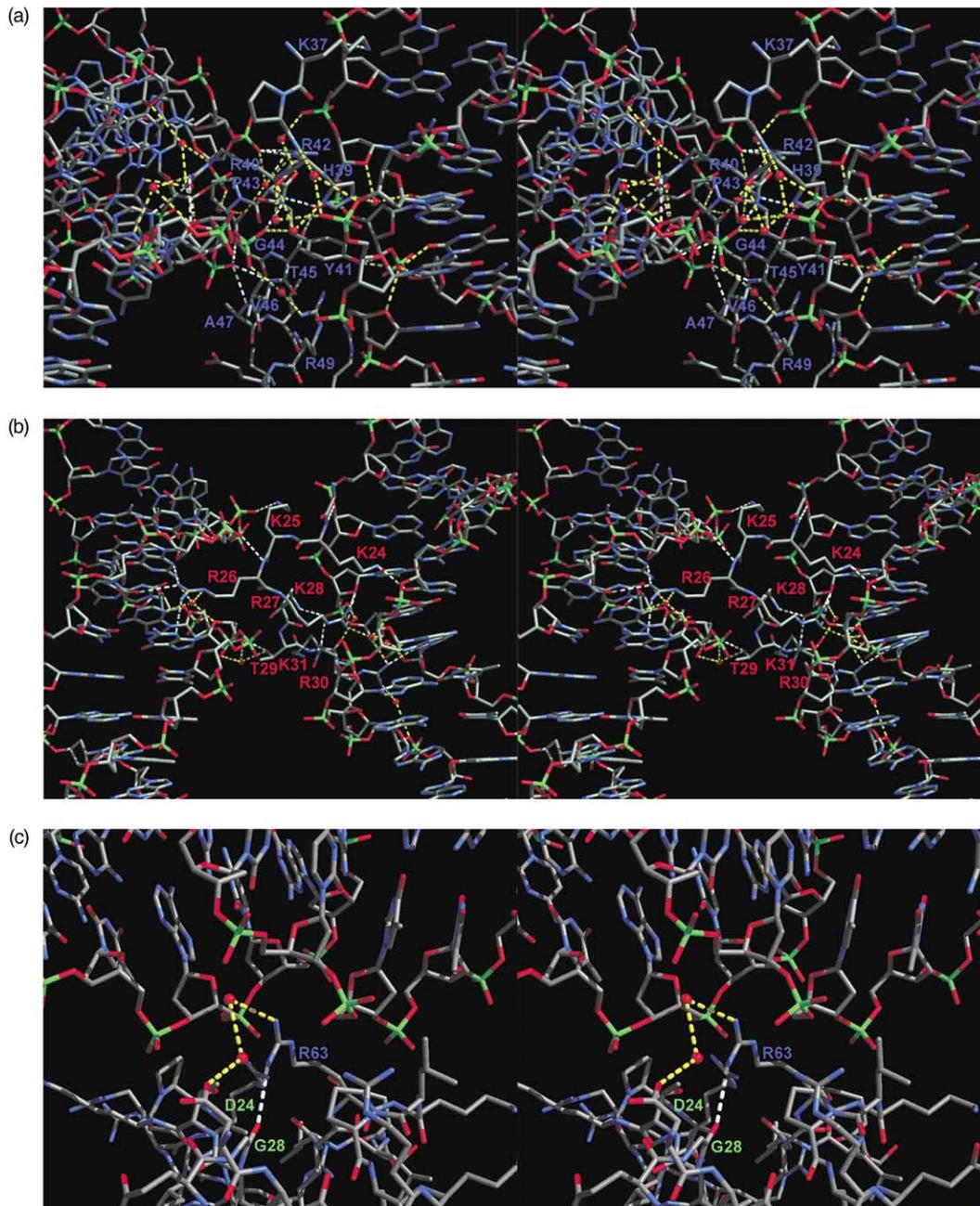


Figure 4. DNA-binding by the H3- α N helix, the H3 and H2B N-terminal tails, and influence of the H4 N-terminal tail (stereographs). Colors are as for Figure 3. (a) The H3- α N helix, N-terminal to the H3 histone-fold, binds SHL6.5. The base of the H3 N-terminal tail (amino acid residues 39–46) passes through the adjacent minor groove channel comprising DNA SHL6.7 and SHL-0.7. (b) The base of the H2B N-terminal tail (amino acid residues 24–31) interacts with the minor groove channel comprising DNA SHL4.7 and SHL-2.7. (c) The H4 N-terminal tail directs H3-R63 side-chain orientation at SHL-1.5. Hydrogen bonds that link H3-R63 with H4-D24 *via* two water molecules are shown (yellow).

specific associations are not observed (Figure 4(b)).

Well-ordered side-chains residing in the minor groove form two different types of interaction with base substituents depending on the degree of extension into the groove. Most interact with the base edges only *via* water molecules in the primary hydration layer of the bases. In several cases, in addition to forming multiple water bridges, the terminal nitrogen atoms of the

arginine guanidinium group engage in hydrogen bonds with acceptor cytosine O2 and adenine and guanine N3 atoms (Figure 3(c), H3-R83; Figure 3(e), H2B-R30; Figure 4(a), H3-R40). Depending on the orientation of the guanidinium group, further van der Waals contacts or hydrogen bonds with the deoxyribose O4' atoms, or water bridges to phosphodiester O3' atoms are made (Figure 3(d), H2A-R42; Figure 3(e), H2B-R30; Figure 3(f), H2A-R77; Figure 4(a),

Table 2. Number of hydrogen bonds mediating histone protein–DNA interaction^a

SHL of protein site	Direct						Indirect (H ₂ O mediated)								
	Phosphodiester chains				Minor groove		Phosphodiester chains				Minor groove				
	Side-chain		Peptide		Side-chain	Total	Side-chain		Peptide		DNA	Side-chain		DNA	
	Arg, Lys	Ser, Thr	NH + ^b	NH	Arg	Tyr	Arg, Lys	Other ^c	NH	CO	DNA	Arg	Tyr	Total ^d	
–0.5	1	1	3	5	2	12	8 ^e	4		7	15	5	12	51 (17)	
0.5	2	1	3	5	3	14	10 ^e	4		5	11	6	11	47 (15)	
–1.5	3		2			5	3	0	1	1	5			10 (4)	
1.5	6		2			8	3	1	1	1	8			14 (6)	
–2.5	3	2	1	3	1	10	1	4	2	2	6	5	7	27 (9)	
2.5	3	1	1	3	1	9	2	3	2	2	7	5	9	30 (9)	
–3.5	3	1	1	3	1	9	6	0		3	8	2	5	24 (9)	
3.5	3	2	1	3	1	10	3	2		2	8	2	4	21 (8)	
–4.5	5		1	2	1	9	3	3	6	2	7	1	3	25 (8)	
4.5	5		1	2	1	9	5	4	7	1	8	1	2	28 (7)	
–5.5	0	1	1	2	1	5	2	3	3	3	9	3	7	30 (11)	
5.5	1	3 ^f	1	2	1	8	3	3	3	1	6	2	4	22 (8)	
–6.5	1	1		1		4	0		1	2	4	1 ^g	1	5	14 (5)
6.5	1	1		1		4	1		1	0	3	2 ^g	2	6	15 (5)
Totals	37	14	18	32	13	116	50	31	27	32	105	35	3	75	358 (121)

^a N-terminal tail regions (H3, 1–36; H4, 1–24; H2A, 1–15; H2B, 1–29) and C-terminal tail of H2A (121–128) are not included.

^b Helix-dipole assisted: peptide amide group at N terminus of an α -helix.

^c Gln/Glu/Ser/Thr/Tyr.

^d Number of water molecules in parentheses.

^e One hydrogen bond to His.

^f One hydrogen bond to Tyr.

^g Hydrogen bonds to His.

H3-R40). In addition to the arginine residue, three types of amino acid residues from the gyre-binding region of the H3 N-terminal tail reside in the minor groove (Figure 4(a)). Amino acid P43 makes hydrophobic contacts with the DNA backbone and van der Waals contacts with the primary hydration layer of the minor groove. Amino acid residues H39 and Y41 interact with bases *via* individual bridging water molecules, with Y41 forming an additional hydrogen bond with a deoxyribose O4' atom. Interestingly, the side-chains of H39, Y41, and R49, all of which interact with the DNA at SHL6.7, form a cation– π –cation T-stack interaction.²⁶

When examined collectively over all 12 histone-fold DNA-binding sites, the arginine side-chain most closely associated with the minor groove displays variable contacts with the surrounding DNA and protein. These side-chains have elevated mobility compared to other histone components interacting with the DNA, and the members of symmetrically related pairs often have significantly different conformations. In the cases of SHL \pm 0.5, \pm 1.5, \pm 2.5 and \pm 3.5, a threonine hydroxyl group is in near proximity to the arginine side-chain such that at SHL \pm 0.5 a direct hydrogen bond is made between it and the arginine guanidinium group (Figure 3(a), H4-R45/H3-T118). However, at SHL + 2.5 and SHL \pm 3.5, the link is made *via* a water molecule (Figure 3(c), H3-R83/H4-T80; Figure 3(d), H2A-R42/H2B-T85).

Influence of the H4 N-terminal tail at SHL–1.5

At SHL \pm 1.5, the symmetrically located amino acid residues H3-R63 interact with the phosphodiester backbones on opposite sides of the minor groove, but their guanidinium groups have orthogonal orientations and opposite directions along the DNA (Figures 3(b) and 4(c)). This difference is a consequence of dissimilar paths for the two H4 N-terminal tails due to crystal packing. At SHL–1.5, the guanidinium group does not form hydrogen-bonding interactions with the DNA, but instead makes van der Waals contacts with both DNA backbones, a hydrogen bond with the peptide carbonyl oxygen atom of H4-G28, and a salt link with the carboxyl group of H4-D24 6 Å away connected indirectly *via* hydrogen bonds with two intervening water molecules (Figure 4(c)). These interactions with the H4 tail orient the arginine guanidinium group parallel with the walls of the minor groove. The H4 tail in this case associates through side-chains H4-R17 and H4-R19 with two phosphate groups on one of the DNA strands adjacent to the H3-R63. At SHL + 1.5, the H3-R63 side-chain is positioned precisely between two phosphate groups on opposing DNA strands, making hydrogen bonds with both and forming three additional water-bridge interactions with phosphate groups (Figure 3(b)). As a consequence of the interparticle association mediated by the H4 N terminus in the SHL + 1.5 half of NCP147, the

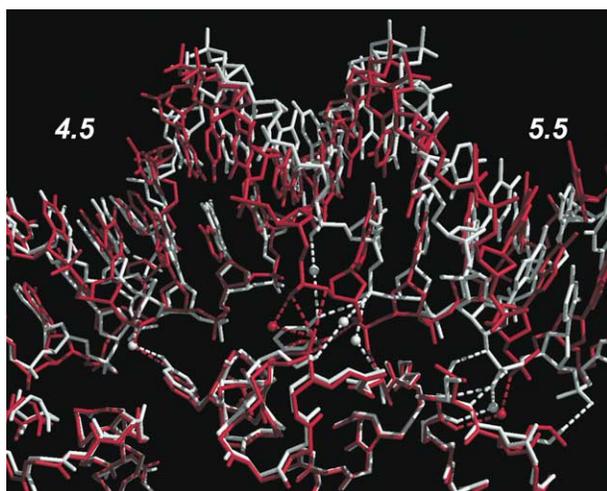


Figure 5. Variability of histone–DNA interactions between NCP147 and NCP146b. The superposition of NCP147 (white) with NCP146b (red) is shown for the region covering SHL4.5 to SHL5.5. The alignment was made by least squares minimization of distances between corresponding atoms of the histone-fold domains only, resulting in an r.m.s. deviation of 0.80 Å. Direct and water-mediated hydrogen bonds (dashed) between histone and DNA phosphate groups that are substantially different for the two structures are shown. For clarity, only the histone main-chain and side-chain atoms relevant to the differences observed are shown.

H4-D24 side-chain is too far away (12 Å) to affect the conformation of H3-R63.

Histone ion-binding sites

Specific ion-binding sites in protein and DNA structures can be identified using electron density maps showing locations of anomalous scattering centers in combination with information on composition and metal ion coordination geometry. This approach has allowed identification of 18 bound ions in NCP147. Five are bound to protein alone and are described here, and 13 are associated with the DNA (C.A.D. & T.J.R. unpublished results). The highest peak found in the anomalous difference Fourier map overall stands 17.7 times above the r.m.s. variation in electron density. It corresponds to an exceptionally well-ordered manganese ion bound between particles in the most extensive site of interparticle protein–protein contact in the crystals (Figure 6(a)).² The Mn^{2+} is coordinated with nearly perfect octahedral geometry by a carboxylic oxygen atom of H3-D77 from one particle and the peptide carbonyl oxygen atom of H2B-V45 from the other, as well as four water molecules (Figure 6(b)). These water molecules are bound by seven protein acceptor groups coming from an aspartate, a glutamate and a glutamine side-chain, and three main-chain carbonyl groups distributed between both particles. The cation binding site is centered on the C terminus of the H2B- α 1 helix, which adds electrostatic stabil-

ization from the helix dipole. Nevertheless, the other copy of each half-binding site has no metal ion bound, and therefore metal binding requires interparticle interactions.

This region of interparticle protein–protein interaction additionally includes the neighboring basic amino acid residues H4-K20 and H4-R23 of the H4 N-terminal tail which make direct hydrogen bonds with amino acid residue H2B-Q44 and acidic amino acid residues H2A-E56 and H2B-E110 from the histone-fold and α C helix of the H2A–H2B dimer (Figure 6(c)). The arrangement of these three histone dimer amino acid residues suggests that they are pre-disposed for binding the arginine side-chain. This interaction site and that centered on the Mn^{2+} must be tightly linked, as H4-R23 and H4-L22 in the ion site are adjacent in the H4 tail sequence.

In addition to the manganese ions seen in NCP147, four chloride ions are found in two symmetrically related pairs of sites in the histone-fold domains (Figure 7). These occur at the N termini of the H3 and H2B- α 3 helices and are absent at the analogous H4 and H2A locations. Of the α -helices in the histone-fold domains, only the α 3 helices do not engage in dipole-assisted hydrogen bonds from peptide amide to DNA phosphate groups. Cl and Br anions have been observed in other crystal structures bound to the N termini of α -helices where they are stabilized by the positive charge associated with the helix dipole.^{25,27} At the H3- α 3 sites, the Cl ion is ligated in an irregular hexacoordinate geometry to the peptide amide group of H3-K122 and five water molecules (Figure 7(a)). The H3 side-chains M120, P121 and K122 form a “van der Waals cup” within which the anion-polyhydrate resides. In contrast, the Cl ion bound to the H2B sites has the peptide amide groups of both H2B-S88 and H2A-G46 in its coordination sphere. The close packing of the N termini from the H2B- α 3 helix and H2A- α 2 helix enables Cl^- to bind both simultaneously (Figure 7(b)). With the N terminus of the H2A- α 2 helix also engaged in direct interactions with the DNA, the Cl^- is situated strikingly close, only 4.6 Å, to a phosphate oxygen atom. However, one of the two tetrahedrally coordinated water molecules supports this configuration by making a hydrogen bond to this phosphate oxygen atom.

The observation that chloride anions are bound to the H3 and H2B- α 3 helices raises the question of why similar species are not found at the analogous positions of H4 and H2A. In the case of the H4- α 3 helix, the end of the helix is blocked by the H3-Q85 side-chain accepting a hydrogen bond from the H4-A83 peptide amide group. This means of capping is not possible for the α 3 helices of the other three histones because the amino acid residues residing at the homologous position to H3-Q85 are either serine or glycine. For the H2A- α 3 helix, the ion binding site is apparently occluded by a bridging water molecule between

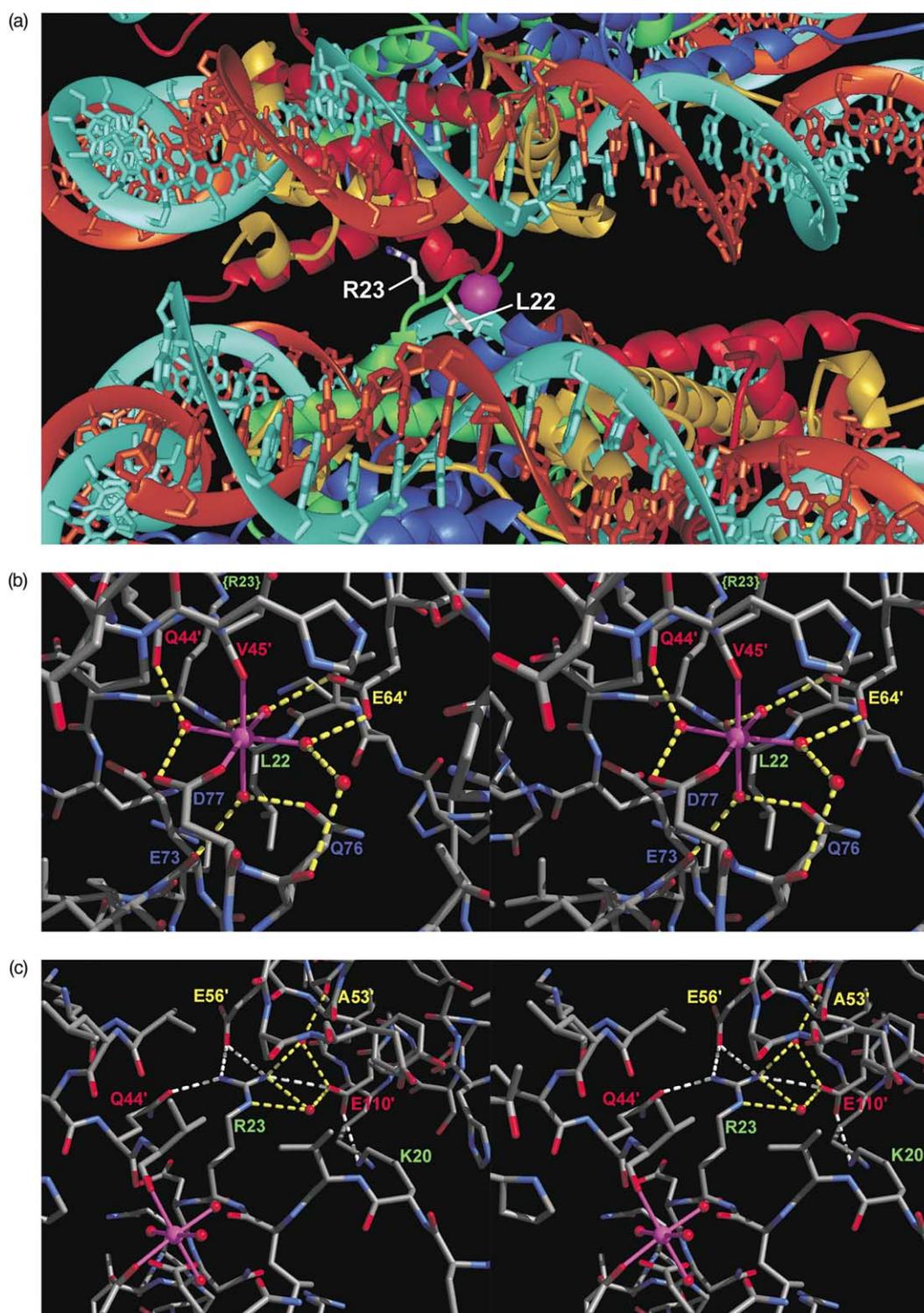


Figure 6. Specific histone-histone interparticle interactions in the NCP147 crystals. (a) Two adjacent side-chains in the histone H4 N-terminal tail, H4-R23 and H4-L22, and a Mn^{2+} ion (magenta) participate in extensive interactions with a highly acidic region of the H2A-H2B dimer of a neighboring nucleosome core. (b) (stereograph) A Mn^{2+} ion (magenta) connects H3-D77 in the H3-H4 tetramer of one particle with H2B-V45 in the H2A-H2B dimer of an adjacent particle. Both of these amino acid residues enter the coordination sphere of the divalent cation. The mean length over all six Mn^{2+} -oxygen bonds (magenta) is 2.25 Å. Hydrogen bonds made directly between histone moieties (white) or *via* water molecules (yellow) are shown. (c) (stereograph) The guanidinium group of H4-R23, in a site adjacent to the Mn^{2+} ion of (b), makes four direct hydrogen bonds to three surrounding side-chains of the neighboring H2A-H2B dimer.

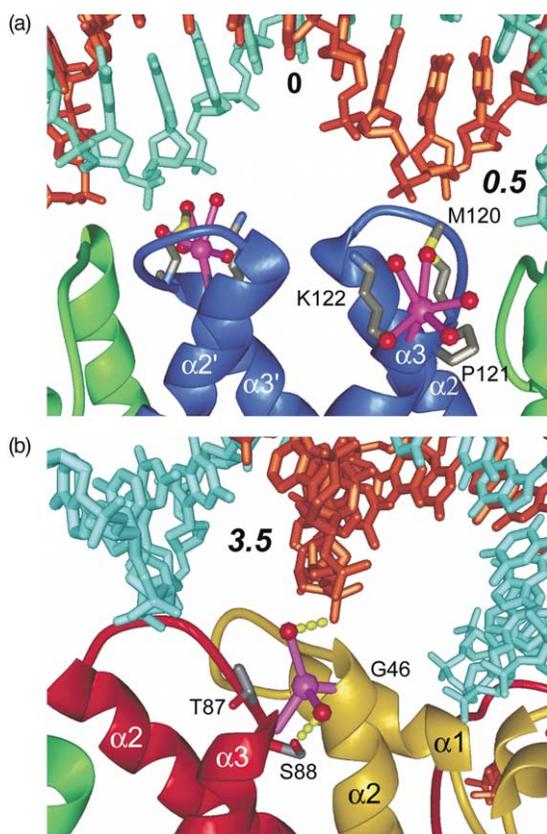


Figure 7. Chloride ion-binding sites in NCP147. Cl⁻ ion and coordinate bonds (magenta) and water molecules (red) of hydration are shown along with amino acid side-chains making direct contact with the ion-hydrates. (a) Dyad-axis related chloride-binding sites of the H3-α3 helices near the center of the DNA (0). Cl⁻ ion is coordinated, with irregular geometry, to the peptide amide group of H3-K122 and five water molecules. (b) Chloride-binding site of the H2B-α3 helix (dyad-axis related site not shown). Cl⁻ ion is coordinated with tetrahedral geometry to the peptide amide groups of H2B-S88 and H2A-G46 and two water molecules. Hydrogen bonds (yellow) from histone and DNA to coordination water molecules are shown. The mean peak height in the anomalous difference map corresponding to the dyad-axis related chloride ions is 3.7σ for (a) and 3.1σ for (b). The mean distances between Cl⁻ ions and coordinated peptide nitrogen and water oxygen atoms are 3.25 Å and 3.53 Å for (a) and 3.28 Å and 3.40 Å for (b), respectively.

the peptide amide group of H2A-R81 at the α3 helix N terminus and the peptide carbonyl group of H2A-Q104 in the H2A C-terminal histone-fold extension.

Discussion

The nucleosome core particle containing a DNA molecule of 147 bp yields crystals in which the DNA is far better ordered than in previous 146 bp particles. It permits for the first time extensive modeling of the solvent component. As is typical

for protein–DNA complexes,¹⁴ somewhat less than one-seventh of the ASA of the isolated histone octamer and DNA components is lost in the complete complex. A disproportionate part of this reduction is the consequence of the insertion of a relatively small number of amino acid side-chains into the DNA minor groove. Most are arginine, which is ideally suited for facilitating the narrowing of the minor groove, simultaneously forming interactions with many of the, mostly hydrogen-accepting, polar components, and making extensive hydrophobic contacts with the deoxyribose moieties. The majority of buried surface area, however, is a consequence of desolvation of phosphate groups through hydrogen-bonding; the predominant mode of direct histone–DNA association. Importantly, the detailed visualization of structured water at the protein–DNA interface shows that water molecules not only contribute significantly to the stability of DNA-binding, but take part in adapting the histone surface to conformational variation in the DNA. This latter observation suggests that bridging water molecules may play a principal role in facilitating nucleosome mobility by providing an interaction pathway for changing phosphate group position.

Compared to most DNA-binding proteins, the general lack of a pronounced DNA sequence-dependency for the histone octamer is striking. The relative difference in DNA-binding energies spans only about 7 kcal/mol (29.23 kJ), with unusual repetitive telomeric sequences²⁸ having the lowest affinity for the histone octamer and synthetic random sequences selected for high affinity²⁹ with the highest. However, this range of binding energy is probably an underestimate, as it does not take into account the poor binding properties of sequences that include significant stretches of poly(dG–dC) and poly(dA–dT) shown to be preferentially excluded from nucleosomes.^{30,31} Furthermore, it has been argued that the conditions of assembly used might suppress the range of values markedly.³² Nevertheless, the current estimate for the range of binding energies for “all possible sequences” equates to a mean difference of approximately 0.5 kcal/mol (2.09 kJ) per histone–DNA binding site; the equivalent of a weak hydrogen bond per site. Consequently, the basis of nucleosome positioning and stability would then rely on all the protein–DNA interaction sites within the nucleosome, and as little as a single bridging water molecule would make a significant contribution. For any particular DNA stretch of 147 bp, some sites could dominate over others, as is likely for the H3–H4 tetramer region with 5S DNA.³³

Sequence-specificity in protein–DNA association has two possible structural origins relying on either “direct readout” of the DNA primary structure through chemically specific bonds to the nucleotide bases or “indirect readout” in which sequence-dependent conformational features of the DNA are recognized by sterically

complementary protein contacts.^{6,34} Water enters this scheme in the first case as an extension of the chemistry of the base, and in the second as a conformational adapter. For the nucleosome core particle, the few side-chains in a position to make significant contacts with the DNA bases are situated in the minor groove, where the potential discriminator for direct readout is the adenine C2 *versus* the guanine N2 group. However, the minor groove is not obviously employed for recognition by either of the two DNA sequences examined here. Several direct interactions are observed between arginine and both pyrimidine O2 and purine N3 atoms, but these sites are not specific for base-pair recognition. Although bridging water molecules can confer sequence specificity in protein–DNA association,^{11,35} the arginine side-chains in the nucleosome core minor groove appear capable of accommodating numerous hydration configurations. Therefore, the role of these side-chains must be largely non-specific in nature, promoting general stabilization of the complex by facilitating the collapse of the minor groove, which in NCP147 has a mean width of $3.0(\pm 0.55)$ Å at the sites of the associated arginine side-chains. The binding of the H4 N-terminal tail to the minor-groove-associated arginine residue at SHL–1.5, which reduces its interaction with the DNA backbones, suggests a mechanism for increasing the mobility of the histone octamer along DNA. Interestingly, a critical role for the H4 tail has been found for chromatin remodeling by the factor ISWI.³⁶

In contrast to base interactions, the abundant contacts between histone protein and phosphodiester backbone could, in principle, result in significant sequence bias for DNA-binding, even though only one-fifth of all phosphate groups are contacted directly or indirectly by histone. The geometric layout of the direct interactions could impose strong restraints on the local conformation of DNA at each binding site in order to obtain the appropriate alignment of interacting groups. At least for the two adjacent DNA-binding sites on the H2A–H2B dimer (SHL4.5 and SHL5.5), however, variability in phosphate position and orientation can be accommodated through substitution of direct by water-mediated interactions and *vice versa*. These alternative sets of interactions would likely dampen sequence effects. As a result, the sequence-specific contribution to the binding affinity is likely to originate from the sequence-dependent conformational energetics of DNA bound in a tight superhelix and from the more restraining histone binding sites (C.A.D. & T.J.R., unpublished results). Conversely, DNA sequences selected on the basis of their conformational adaptability to maximize nucleosome stability would approach the maximum number of direct interactions possible *via* replacement of water-mediated interactions with direct interactions. Although the enthalpic contribution to binding would remain approximately the same in this case, an entropic

gain due to the liberation of solvent molecules would occur. This scenario is consistent with the results of recent equilibrium thermodynamic measurements showing entropic stabilization for two selected, defined-sequence nucleosome core particles (the human alpha satellite sequence used for NCP146 and the *Lytechinus variegatus* 5 S RNA sequence) over a population of particles containing mixed-sequence DNA.³⁷

The extensive hydrogen-bonding network centered on the Mn cation and H4-R23 at an interparticle crystal contact suggests that this interface or possibly a variant of it may be important for the structure of the chromatin fiber.^{1,38} The low abundance of Mn²⁺ *in vivo*^{39,40} relative to Mg²⁺ implies that Mg²⁺ could be the preferred metal ion to mediate such interparticle association. The coordination chemistry of either metal ion is likely to be highly similar, however, since both ions favor octahedral coordination geometry, and the harder Mg cation would be compatible with the liganding oxygen atoms of protein and water molecules observed. Structural conservation for the interchange of Mn²⁺ and Mg²⁺ in proteins is well documented and exemplified by DNA and RNA polymerase structures.^{41,42} The recent structure of a yeast nucleosome shows a different interparticle association from that seen here, but which is also dependent on metal cation.⁴ Comparison of the structures reveals that amino acid substitutions in yeast would weaken the interparticle site seen for the *X. laevis* octamer, and account for the creation of the yeast interparticle sites relative to *X. laevis*. More generally, however, these observations from crystal contacts may not be directly applicable to the nucleosome higher-order structure, but purely suggestive of the type of interactions that can occur.

Materials and Methods

Crystal preparation

Nucleosome core particles were prepared from recombinant *X. laevis* histones and 146 and 147 bp palindromic DNA fragments derived from human α -satellite DNA, as described previously.^{1,43} The 146 bp and 147 bp sequences contain, respectively an *EcoRI* and *Hinfi* restriction endonuclease site at their centers. Crystals were grown by vapor diffusion over the course of one to four weeks in droplets containing 4 mg/ml of core particle, 70–85 mM MnCl₂, 50–60 mM KCl, and 20 mM potassium cacodylate (pH 6.0), equilibrated against 40–46 mM MnCl₂, 30–40 mM KCl, and 10–20 mM potassium cacodylate (pH 6.0).¹ Crystals were transferred to 37 mM MnCl₂, 40 mM KCl, 20 mM potassium cacodylate (pH 6.0), 24% (v/v) 2-methyl-2,4-pentanediol, and 2% (w/v) trehalose for stabilization and improvement of diffraction quality. The 5' to 3' sequences of the DNA molecules are listed with the central base-pair seen in the respective crystal structure indicated in parentheses.

NCP147: ATCAATATCCACCTGCAGATACTACCAA
AAGTGTATTTGGAACTGCTCCATCAAAAGGCATG

TTCAGCTGGA(A)TCCAGCTGAACATGCCTTTTGAT
GGAGCAGTTTCCAAATACACTTTTGGTAGTATCTG
CAGGTGGATATTGAT.

NCP146: ATCAATATCCACCTGCAGATTCTACCAA
AAGTGTAATTTGGAACTGCTCCATCAAAAGGCATG
TTCAGCTGA(A)TTCAGCTGAACATGCCTTTTGATGG
AGCAGTTTCCAAATACACTTTTGGTAGAATCTGCAG
GTGGATATTGAT.

NCP146b: ATCTCCAAATATCCCTTGCGGATCGTA
GAAAAGTGTGTCAAACCTGCGCTATCAAAGGGAA
ACTTCAAAGTGA(A)TTCAGTTGAAGTTTCCCTTTGAT
AGCGCAGTTTGACACACTTTTTCTACGATCCGCAA
GGGATATTTGGAGAT.

Crystallographic data collection

Data collection statistics are summarized in Table 1. The crystals of all constructs belong to space group $P2_12_12_1$ with one particle per asymmetric unit. Crystals of the NCP147 and NCP146 constructs are isomorphous. The NCP146b crystals have a significantly shorter b -axis. Prior to data collection, crystals were transferred to 4 °C, flash-cooled in liquid propane at $-120(\pm 2)$ °C, and transferred into a N_2 gas stream at $-170(\pm 5)$ °C. X-ray intensities were recorded at the E.S.R.F. Grenoble using a 30 cm MAR image plate scanner at a wavelength of 0.93 Å for NCP147 (ID14-4) and 0.85 Å for NCP146 and NCP146b (ID9). Prior to data collection, NCP147 crystals were screened for diffraction quality and oriented at the Swiss–Norwegian beamline (E.S.R.F.). X-ray data were combined and processed using DENZO/SCALEPACK.⁴⁴

Model refinement

The 2.8 Å resolution model of NCP146 was used to provide starting phases for NCP147 and NCP146, and as input for molecular replacement for NCP146b.¹ Models were refined using the programs O and CNS.^{45,46} Simulated annealing was used to reduce model bias in the starting phases and provide omit maps for accurate model building. Solvent molecules were located initially as peaks ($>3\sigma$) in $F_o - F_c$ electron density maps with a suitable hydrogen-bonding environment, and confirmed or removed based on omit map density in the final rounds of refinement. Ion binding sites were located based on significant peaks ($>3\sigma$) in anomalous difference electron density maps or residual peaks ($>3\sigma$) in $F_o - F_c$ maps subsequent to full refinement with a water molecule at each of these positions. In addition, water molecules refined at ion-binding sites attained B -factors which were unreasonably low compared to the surrounding associated protein or DNA atoms. All ion-binding sites were confirmed by a final anomalous difference map, calculated using phases from the fully refined model. Graphic Figures were prepared with WebLabViewer 4.0 (Accelrys) and MidasPlus.⁴⁷

Protein Data Bank accession code

The coordinates of NCP146, NCP146b and NCP147 have been deposited in the Protein Data Bank with accession codes 1kx3, 1kx4 and 1kx5, respectively, and will be released six months after the date of publication.

Acknowledgments

We are grateful to the staffs of the ID14-4, ID9 and SNBL beamlines at the E.S.R.F. for their assistance during data collection, and to Dr I. Berger for comments on the manuscript. We appreciate the support of the Swiss National Fund.

References

- Luger, K., Maeder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, **389**, 251–260.
- Luger, K., Maeder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (2000). The atomic structure of the nucleosome core particle. In *Biomolecular Stereodynamics*, vol. 11, Adenine Press, New York, pp. 190–200.
- Harp, J. M., Hanson, B. L., Timm, D. E. & Bunick, G. J. (2000). Asymmetries in the nucleosome core particle at 2.5 Å resolution. *Acta Crystallog. sect. D*, **56**, 1513–1534.
- White, C. L., Suto, R. K. & Luger, K. (2001). Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *EMBO J.* **20**, 5207–5218.
- Suto, R. K., Clarkson, M. J., Tremethick, D. J. & Luger, K. (2000). Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nature Struct. Biol.* **7**, 1121–1124.
- Otwinowski, Z., Schevitz, R. W., Zhang, R. G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q. *et al.* (1988). Crystal structure of trp repressor/operator complex at atomic resolution. *Nature*, **335**, 321–329.
- Billeter, M., Qian, Y. Q., Otting, G., Mueller, M., Gehring, W. & Wuethrich, K. (1993). Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain–DNA complex. *J. Mol. Biol.* **234**, 1084–1093.
- Keller, W., Koenig, P. & Richmond, T. J. (1995). Crystal structure of a bZIP/DNA complex at 2.2 Å: determinants of DNA specific recognition. *J. Mol. Biol.* **254**, 657–667.
- Garner, M. & Rau, D. (1995). Water release associated with specific binding of gal repressor. *EMBO J.* **14**, 1257–1263.
- Lundback, T. & Hard, T. (1996). Sequence-specific DNA-binding dominated by dehydration. *Proc. Natl Acad. Sci. USA*, **93**, 4754–4759.
- Schwabe, J. W. (1997). The role of water in protein–DNA interactions. *Curr. Opin. Struct. Biol.* **7**, 126–134.
- Woda, J., Schneider, B., Patel, K., Mistry, K. & Berman, H. M. (1998). An analysis of the relationship between hydration and protein–DNA interactions. *Biophys. J.* **75**, 2170–2177.
- Luscombe, N. M., Laskowski, R. A. & Thornton, J. M. (2001). Amino acid–base interactions: a three-dimensional analysis of protein–DNA interactions at an atomic level. *Nucl. Acids Res.* **29**, 2860–2874.
- Jones, S., van Heyningen, P., Berman, H. M. & Thornton, J. M. (1999). Protein–DNA interactions: a structural analysis. *J. Mol. Biol.* **287**, 877–896.
- Murphy, F. V. & Churchill, M. E. (2000). Non-sequence-specific DNA recognition: a structural perspective. *Struct. Fold. Des.* **8**, R83–R89.

16. Meersseman, G., Pennings, S. & Bradbury, E. M. (1992). Mobile nucleosomes—a general behavior. *EMBO J.* **11**, 2951–2959.
17. Flaus, A. & Richmond, T. J. (1998). Positioning and stability of nucleosomes on MMTV 3′LTR sequences. *J. Mol. Biol.* **275**, 427–441.
18. Brändén, C.-I. & Jones, T. A. (1990). Between objectivity and subjectivity. *Nature*, **343**, 687–689.
19. Luzzati, V. (1953). Resolution d'une structure cristalline lorsque les positions d'une partie des atomes sont connues: traitement statistique. *Acta Crystallog.* **6**, 142–152.
20. Kopka, M. L., Fratini, A. V., Drew, H. R. & Dickerson, R. E. (1983). Ordered water structure around a B-DNA dodecamer. A quantitative study. *J. Mol. Biol.* **163**, 129–146.
21. Schneider, B., Patel, K. & Berman, H. M. (1998). Hydration of the phosphate group in double-helical DNA. *Biophys. J.* **75**, 2422–2434.
22. Egli, M., Tereshko, V., Teplova, M., Minasov, G., Joachimiak, A., Sanishvili, R. *et al.* (1998). X-ray crystallographic analysis of the hydration of A- and B-form DNA at atomic resolution. *Biopolymers*, **48**, 234–252.
23. Dunitz, J. D. (1994). The entropic cost of bound water in crystals and biomolecules. *Science*, **264**, 670.
24. Fersht, A. (1999). *Structure and Mechanism in Protein Science. A Guide to Enzyme Catalysis and Protein Folding*, Freeman, New York.
25. Hol, W. G. (1985). The role of the alpha-helix dipole in protein function and structure. *Prog. Biophys. Mol. Biol.* **45**, 149–195.
26. Gallivan, J. P. & Dougherty, D. A. (1999). Cation–π interactions in structural biology. *Proc. Natl Acad. Sci. USA*, **96**, 9459–9464.
27. Fiedler, T. J., Davey, C. A. & Fenna, R. E. (2000). X-ray crystal structure and characterization of halide-binding sites of human myeloperoxidase at 1.8 Å resolution. *J. Biol. Chem.* **275**, 11964–11971.
28. Filesi, I., Cacchione, S., De Santis, P., Rossetti, L. & Savino, M. (2000). The main role of the sequence-dependent DNA elasticity in determining the free energy of nucleosome formation on telomeric DNAs. *Biophys. Chem.* **83**, 223–237.
29. Thåström, A., Lowary, P. T., Widlund, H. R., Cao, H., Kubista, M. & Widom, J. (1999). Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J. Mol. Biol.* **288**, 213–229.
30. Simpson, R. T. & Kunzler, P. (1979). Chromatin and core particles formed from the inner histones and synthetic polydeoxyribonucleotides of defined sequence. *Nucl. Acids Res.* **6**, 1387–1415.
31. Rhodes, D. (1979). Nucleosome cores reconstituted from poly (dA-dT) and the octamer of histones. *Nucl. Acids Res.* **6**, 1805–1816.
32. Drew, H. R. (1991). Can one measure the free energy of binding of the histone octamer to different DNA sequences by salt-dependent reconstitution? *J. Mol. Biol.* **219**, 391–392.
33. Dong, F. & van Holde, K. (1991). Nucleosome positioning is determined by the (H3–H4)₂ tetramer. *Proc. Natl Acad. Sci. USA*, **88**, 10596–10600.
34. Jen-Jacobson, L. (1997). Protein–DNA recognition complexes: conservation of structure and binding energy in the transition state. *Biopolymers*, **44**, 153–180.
35. Janin, J. (1999). Wet and dry interfaces: the role of solvent in protein–protein and protein–DNA recognition. *Struct. Fold. Des.* **7**, R277–R279.
36. Clapier, C. R., Laengst, G., Corona, D. F. V., Becker, P. B. & Nightingale, K. P. (2001). Critical role for the histone H4 N terminus in nucleosome remodeling by ISWI. *Mol. Cell. Biol.* **21**, 875–883.
37. Gottesfeld, J. M. & Luger, K. (2001). Energetics and affinity of the histone octamer for defined DNA sequences. *Biochemistry*, **40**, 10927–10933.
38. Luger, K. & Richmond, T. J. (1998). The histone tails of the nucleosome. *Curr. Opin. Genet. Dev.* **8**, 140–146.
39. Collins, K. D. (1997). Charge density-dependent strength of hydration and biological structure. *Biophys. J.* **72**, 65–76.
40. Christianson, D. W. (1997). Structural chemistry and biology of manganese metalloenzymes. *Prog. Biophys. Mol. Biol.* **67**, 217–252.
41. Doublet, S., Tabor, S., Long, A. M., Richardson, C. C. & Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature*, **391**, 251–258.
42. Cramer, P., Bushnell, D. A., Fu, J., Gnatt, A. L., Maier-Davis, B., Thompson, N. E. *et al.* (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. *Science*, **288**, 640–649.
43. Luger, K., Rechsteiner, T. J. & Richmond, T. J. (1999). Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol.* **304**, 3–19.
44. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
45. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110–119.
46. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography & NMR System: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D*, **54**, 905–921.
47. Ferrin, T. E., Huang, C. C., Jarvis, L. E. & Langridge, R. (1988). The MIDAS display system. *J. Mol. Graph.* **6**, 13–27.

Edited by J. O. Thomas

(Received 11 February 2002; accepted 12 April 2002)



<http://www.academicpress.com/jmb>

Supplementary Material comprising of one Table is available on IDEAL.