# INSIGHTS INTO MECHANISMS OF NUCLEOSOME REMODELING FROM ANALYSIS OF CRYSTAL STRUCTURES

by

FEI XU

A dissertation submitted to the

Graduate School – New Brunswick

Rutgers, The State University of New Jersey

and

The Graduate School of Biomedical Sciences

University of Medicine and Dentistry of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Biochemistry

Written under the direction of

Professor Wilma K. Olson

And approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October 2007

ABSTRACT OF THE DISSERTATION

INSIGHTS INTO MECHANISMS OF NUCLEOSOME REMODELING FROM ANALYSIS OF CRYSTAL STRUCTURES

BY

FEI XU

Dissertation Director:

Wilma K. Olson

The packaging of DNA into nucleosomes impedes the binding and access of molecules involved in its processing. The yeast SWI/SNF complex is a regulatory factor in yeast, which stimulates the mobilization of chromatin DNA required for transcription. Amino acid point mutations in the genes coding for histones H3 and H4 can partially bypass the requirement of the SWI/SNF complex. The atomic mechanisms, which underlie the observed remodeling, however, are difficult to discern from the crystal structures of nucleosomes bearing these so-called Sin (SWI-INdependent) mutations. The modified nucleosomes are only minimally distorted from the wild-type structure, with the loss of a few direct protein-DNA contacts near the sites of amino acid mutation and no large-scale change in overall structure (Muthurajan *et al.* (2004) *EMBO J.* **23**, 260).

Detailed analysis of the conformation and interactions of the histone proteins and DNA in the SIN-mutant structures reveals additional loss of intermolecular contacts and accompanying variation in the orientation and displacement of individual base-pair steps in the vicinity of histones H2A and H2B. The modified amino acids seemingly transmit information relevant to DNA binding across the nucleosome structure.

Nucleosomal DNA in the wild-type and SIN-mutant structures is represented in a cylindrical frame as the shape of a nucleosome resembles a cylinder with the nucleosomal DNA wrapping along a superhelical pathway around the histone core. This analysis shows an increase in radius at the successive or neighboring dimer steps in the SIN-mutants, compared to the wild-type, either toward the end of nucleosomal DNA or near the histone dimerization interface.

The structural templates of the wild-type and SIN-mutant nucleosomes are threaded on nucleosomal DNA sequences, which measure the relative ease of deforming a DNA sequence on the surface of a histone protein core. Surprisingly, the threading scores can discriminate among different SIN-mutant structures. Highly deformed steps in terms of large deformation scores in the SIN-mutant structures are located near one end and near the central dyad of the nucleosomal DNA. Models, which incorporate the highly deformed steps, are suggestive of ways in which the DNA might loop away from or peel off the nucleosome core particle.

Two nucleosome structures have identical histone cores but different DNA sequences are compared to each other. The difference in the superhelical shapes and the shifted locations of highly deformed steps show how nucleosomal DNA adjusts its conformation according to its sequence specificity.

## Table of Contents

## List of Figures

**List of Tables**

# Chapter 1 Characteristics of Nucleosome Core Particle

In this chapter, the basic characteristics of the nucleosome core particle and important background knowledge of SWI/SNF Independent (SIN)-mutant nucleosome core particles are introduced in the context of DNA folding in a cell.

## 1.1 The organization of compacted DNA[1, 2].

In all eukaryotic cells, DNA is highly compacted and stored in the nucleus. In a human cell, the total length of the extended DNA in the nucleus is approximately 2 meters [1]. However, the diameter of a nucleus is only about 6 μm. How such long and extended DNA is folded and compacted into the extremely small space of a nucleus is one of the most intriguing questions in molecular biology.

As far as is known currently, there are five levels in the hierarchical scheme of DNA compaction [1-3]. The lowest level is the nucleosome core particle, in which about 146 base pair (bp) DNA wrap around a core of eight histone proteins made up of two copies of each of four histones — H2A, H2B, H3 and H4. In addition to the histone core, nucleosomal DNA binds to the linker histone H1 or H5 (a H1 variant) [4], which is responsible for inter-nucleosome interactions. Technically, the nucleosome core particle, together with the linker DNA and histone H1 (or histone H5), is called a nucleosome. The DNA that wraps around the core histone proteins constitutes the beads in the "beads-on-a-string" form of chromatin, and between the "beads" is the so-called linker DNA. Upon the addition of 1 to 2 mM divalent cations, the "beads-on-a-string" form of chromatin fiber condenses in solution into the so-called 30 nm chromatin fiber [2, 3]. The histone tails play an important role in organizing the arrays of nucleosomes in the "beads-

on-a-string" form into the 30 nm fiber [2, 3]. How the " beads-on-a-string" fiber is folded into 30 nm chromation is still a controversial question [2]. The DNA collapses into an even more condensed fiber *in vivo,* called the chromonema fiber with a diameter of 100 nm [3]. The chromonema fiber undergoes further compaction to form the final level of DNA compaction, chromosome with a diameter of 1400nm and length of 2μm (human chromosome 22) [1]. The term, chromatid, used in Figure 1.1, refers to one half of a chromosome. The two sister chromatids that make up a chromosome, come from the two copies of DNA formed upon replication and connected by the centromere. The long-range fiber-fiber interactions that are involved in the formation of the chromonema fiber and chromosomes, are not well understood.

## 1.2 Characteristics of the Nucleosome Core Particle

### 1.2.1 Overall structure

The histone proteins are organized into two H3-H4 pairs and two H2A-H2B pairs (Figure 1.2). At the first step of histone-octamer assembly, two H3-H4 dimers form a tetramer, which binds to the central 60 bp of nucleosomal DNA. Then, the $(H3\text{-}H4)_2$ tetramer and two H2A-H2B dimers, which organize 30 bp on the either end of the DNA, form a histone octamer. The penultimate 10 bp of the nucleosomal DNA are bound by a region of H3 (the N-terminal tail of H3 in red contacting the region about SHL ±6.5 in Figure 1.2 (b) and (c)) that does not form an integral part of the $(H3\text{-}H4)_2$ tetramer [5]

There are tight interactions between the structured regions of the histone proteins and the minor groove of the DNA at 14 independent DNA-binding locations. These interactions make use of two specific DNA-binding motifs on the proteins (see Section 1.2.3) and involve tight hydrogen bonds between the main chain amides of the polypeptide

backbone and the phosphate oxygens of DNA, in combination with electrostatic interactions with cationic amino-acid side chains [5].

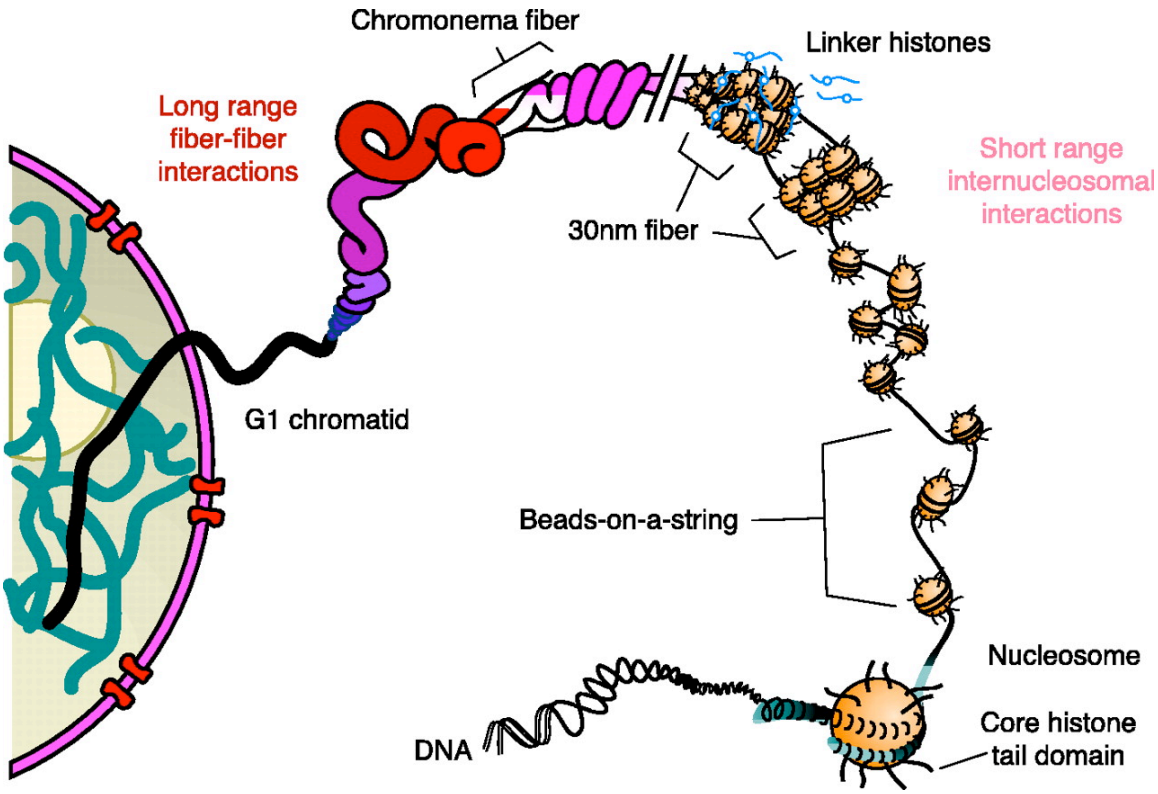



Figure 1.1. Compacted organization of chromatin in the interphase nucleus. There are three levels of chromatin folding. The lowest level is the nucleosome core particles which includes a histone octamer and 146 bp DNA. The second level is the "beads-on-a-string" form, a string of nucleosomes on DNA. The third level is the 30nm fiber, which is formed by the inter-nucleosome interactions with histone protein octamers mediated by linker histones [2]. Details of higher-level orgnization, such as chromonema fiber and chromatids, or chromosome are not yet well understood.

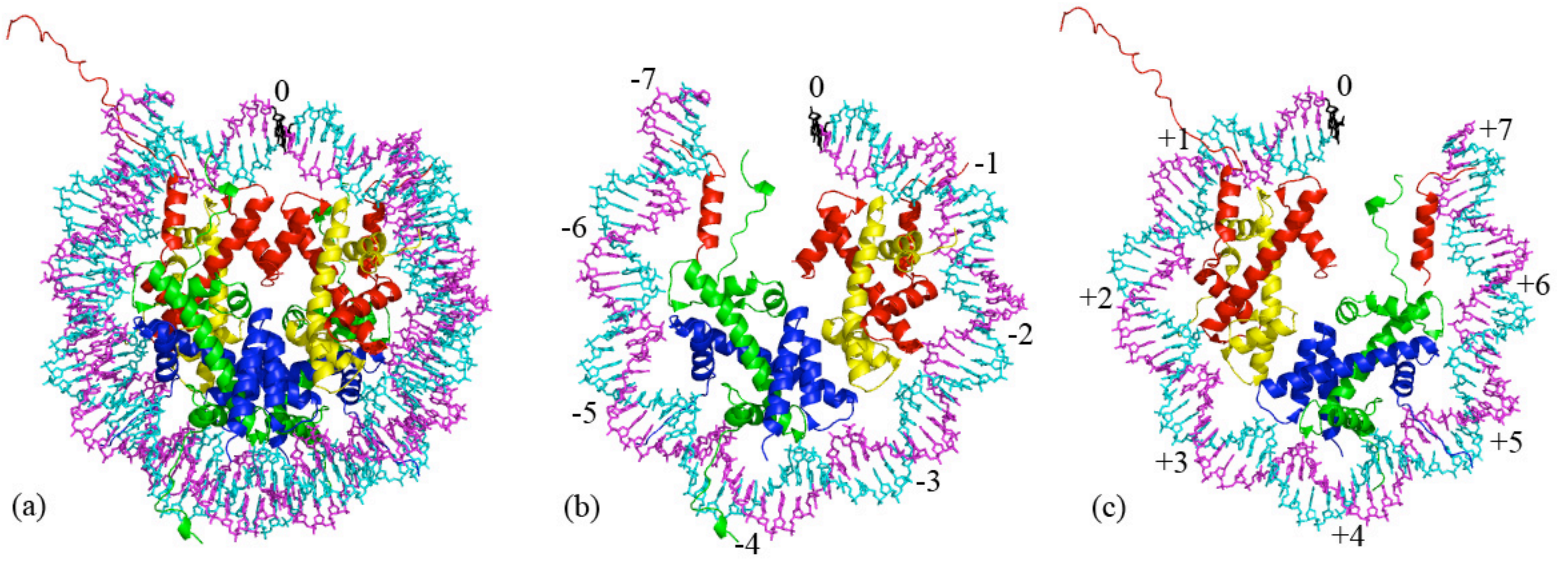



Figure 1.2. Color-coded representation of the 2.8 Å-resolution nucleosome core-particle structure (NDB ID: PD0001) [3]. (a) This view shows the relative locations of the eight proteins in the histone core and the fold of the 146 bp DNA wrapped around it. The histone proteins, H2A, H2B, H3, H4, are shown respectively in red, yellow, green, and blue, and DNA strands I and II in cyan and magenta. The remaining parts of the image illustrate (b) the short half of the DNA with associated proteins, including the dyad axis at base pair 73 (1 to 73 bp), and (c) the long half of the DNA and bound histones, including the dyad axis at base pair 73 (73 to 146 bp). The numbers in black along the DNA in parts (b) and (c) correspond to the superhelical location (SHL). The black base pair at 0 corresponds to the structural dyad at base pair 73, or SHL 0.

### 1.2.2 Nucleosome core-particle DNA

The histone octamer forms a disc-like shape with a flat, left-handed superhelical ramp that wraps about 1.65 superhelical turns of DNA [5] (Figure 1.2). Most of the known nucleosome core particles have pseudo 2-fold symmetry with the central base pair falling on the pseudo 2-fold axis (dyad axis) so that the nucleosome core particle crystallized with 146 base pairs is divided into 73 and 72 bp halves and that with 147 base pairs of DNA into two 73 bp halves (NDB-ID: pd0287) [6]. The dyad axis lies approximately in the plane of the nucleosome disc and bisects the $(H3\text{-}H4)_2$ histone tetramer through its central four-helix bundle structure. Superhelix location numbers (SHL) are defined according to the rotational orientation of the DNA double helix relative to the central base pair. The superhelix location zero, or SHL 0, lies at the central base pair with its major groove facing the protein octamer. For successive helical turns, the location number ranges from −7 to +7 over the whole nucleosomal DNA.

Due to tight interactions with the histone octamer, nucleosomal DNA is said to be highly distorted with a 10.2 bp helical periodicity that deviates from the 10.5 bp turn twisting observed in free DNA [6, 7]. To compensate for one less base pair in the 72-bp half than in the 73-bp half, the DNA in the 73-bp half is undertwisted. The overwound region is located approximately 1.5 to 2 turns away from the dyad axis in the first published nucleosome core particle structure (NDB-ID: pd0001), where the periodicity is 9.4 bp per turn in the 72-bp half versus 10.6 bp per turn for the same region in the 73-bp half [5]. The location of the overwound region varies in different structures [8].

Although the DNA superhelix has an average diameter of 41.8 Å, it is not uniformly bent. Different DNA sequences have different local structural features, and the underlying

histone structure has an irregular shape [6, 9]. The largest radii occur at both SHL ±1.5 and SHL ±4-5. Factors responsible for these distortions in DNA include the α1α2 DNA-binding sites on the H3-H4 pairs at SHL ±1.5, which appears to cause an outward bulge in the DNA, and the adjacent L1L2 and α1α2 DNA-binding sites on the H2A-H2B dimers at SHL ±4 to ±5, which seems to buckle the DNA outwards [5].

### 1.2.3 Histone-DNA interaction

Binding of DNA to the histone surface is primarily mediated by the insertion of arginine side chains into the minor groove of DNA at points where the minor groove faces inward toward the superhelical protein ramp [5]. The binding involves the interaction of arginine with the phosphates on opposite strands separated by approximately five base pairs, or about one half of a helical turn, and suggests that contact with the DNA takes place by a pinching of the minor groove with concomitant opening of the major groove.

The DNA-binding sites on protein can be divided into two types (Figure 1.3). The first, or α1α2-binding motif, is constructed from a pairing of the N-termini of two α-helices, which interact with the phosphate backbone through the dipoles produced in the helices. Such α1α2-binding is found at SHL ±1.5 and ±4.5. The second, or L1L2-binding motif, is formed by two loops from different histone proteins. 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. The L1L2-binding motif is located at SHL ±0.5, ±2.5, ±3.5, and ±5.5.

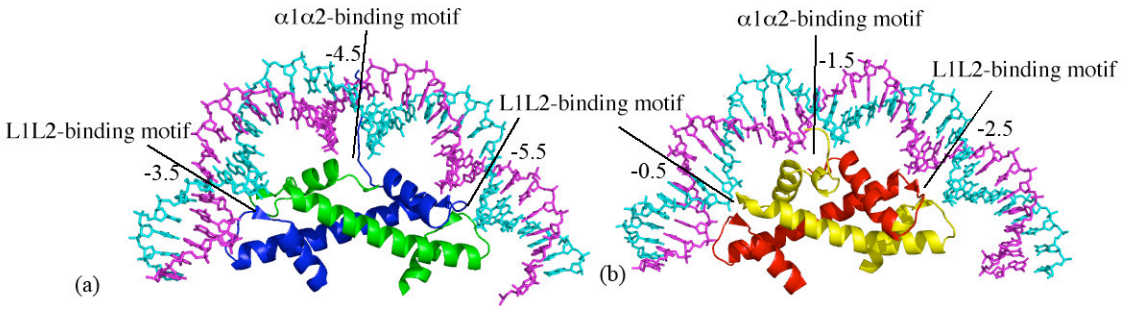



Figure 1.3. Histone-fold pairs. (a) H2A-H2B fold pair. (b) H3-H4 fold pair. The α1α2-binding motif and L1L2-binding motif are shown. The color scheme is the same as in Figure 1.2.

## 1.3 Nuclesome dynamic properties

Although the persistence length of free DNA in solution is 150 bp [10], the same length of DNA (145-147 bp) bends tightly upon associated with the histone protein in the nucleosome core particle.The DNA bending at the sites of protein-DNA interaction in the minor groove, is coupled with significant lateral displacement of the base-pair steps at the same locations [11]. That is to say, the ramp of DNA wrapping around a histone core is sheared at the locations where DNA bends into its minor groove. Certain base-pair steps, such as CA/TG, can accommodate this highly deformed conformation, in terms of low deformation scores [11], which will also be discussed in chapter 6. Furthermore, different base-pair steps show different deformability [12], which may be utilized by nucleosome positioning on genomic DNA.

In contrast to thea static picture of the nucleosome revealed in single crystal structure, nucleosomes display dynamic properties. The DNA bound to the H2A and H2B proteins at the ends of the nucleosome core particle exists in a dynamic equilibrium state of

wrapping and unwrapping, which may facilitate the access of proteins such as remodeling factors or transcription factors to nucleosomal DNA [10]. These motions presumably underlies the exchange of different variants of histones H2A and H2B in nucleosomes [13]. The dynamic properties of nucleosomes play a key role in many biological processes such as DNA remodeling and related initiation of DNA transcription, which will be discussed in the next chapter.


1. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P., *Molecular Biology of the Cell*. 4th ed. 2002, New York: Garland Publishing.
2. Hansen, J.C., *Conformational dynamics of the chromatin fiber in solution: determinants, mehcanisms, and functions*. Ann. Rev. Biophys. Biomol. Struct., 2002. 31: p. 361-92.
3. Horn, P.J., Peterson, C. L., *Chromatin higher order folding: wrapping up transcription*. Science, 2002. 297: p. 1824-1827.
4. Morris, N.R., *A comparison of the structure of chicken erythrocyte and chicken liver chromatin*. Cell, 1976. 9: p. 627-632.
5. Luger, K., Maedar, W., Richmond, R. K., Sargent, D. F., and T. J. Richmond, *Crystal structure of the nucleosome core paricle at 2.8 Å resolution*. Nature, 1997. 389: p. 251-259.
6. Richmond, T.J., adn Davey, A. C., *The structure of DNA in the nucleosome core*. Nature, 2003. 423: p. 145-150.
7. Rhodes, D., and Klug, K., *Sequence-dependent helical periodicity ofo DNA*. Nature, 1981. 292: p. 378-380.
8. Suto, R.K., Edayathumangalam, R. S., White, C. L., Melander, C., Gottesfeld, J. M., Dervan, P. B., and Luger, K., *Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands*. J. Mol. Biol., 2002. 326: p. 317-380.
9. Colasanti, A.V., *Conformatioinal state of double helix DNA*, Ph.D Thesis. 2006, Rutgers, the State University of New Jersey, New Brusnwick.
10. Hargerman, P.J., *Flexibility of DNA*. Annu. Rev. Biophys. Biophys. Chem., 1988. 17: p. 265-86.
11. Tolstorukov, M.Y., Colasanti, A. V., McCandlish, D., Olson, W. K., and Zhurkin, V. B., *A novel 'roll-and-slide' mechanism of DNA folding in chromatin. Implications for nucleosome positioning*. J. Mol. Biol., 2007.
12. Olson, K.W., Gorin, A. A., Lu, X. J., Hock, L. M., and Zhurkin, V. B. , *DNA sequence-dependent deformability deduced from protein-DNA crystal complexes*. Proc. Natl. Acad. Sci. USA, 1998. 96: p. 11163-11168.
13. Luger, K., Hansen, J. C., *Nucleosome and chromatin fiber dynamics*. Current Opinion in Structure Biology, 2005. 15: p. 188-196.

# Chapter 2 SWI/SNF Independent (SIN) mutations of Nucleosomes

## 2.1 SWI/SNF complex

Nucleosome formation occludes the access of the transcription factors and RNA polymerase to the DNA promoter region. Thus, a preliminary step of transcription initiation is nucleosome remodeling (Figure 2.1). Remodeling complexes, such as SWI/SNF in yeast (*S. cerevisiae*), which are recruited by the transcription apparatus, displace the histone octamers and help expose DNA to the transcription factors [1, 2]. The name SWI/SNF is originated from the subunits, which were identified by SWI or SNF mutations in yeast. SWI stand for the yeast mating type SWItching; and SNF refers to Sucrose Non-Fermentation [3].

The expression of about 120 genes, or 2% of the yeast genome, depends on the SWI and SNF genes. The transcription of the yeast HO gene is regulated by the genes SWI1-5. However, the SWI/SNF complex is not crucial for yeast as there are only about 150 copies per yeast cell. Other remodeling complexes, such as RSC, may be essential. There are many copies of RSC in yeast.

*In vitro* experiments have shown that the SWI/SNF complex can remodel nucleosomes without global displacement or large-scale loss of histone octamers [1]. The protein-DNA contacts are altered and the pattern of DNase I cutting of nucleosomal DNA is changed at the target site of the SWI/SNF complex. The mechanism of nuclesome remodeling by the SWI/SNF complex is not understood well yet.

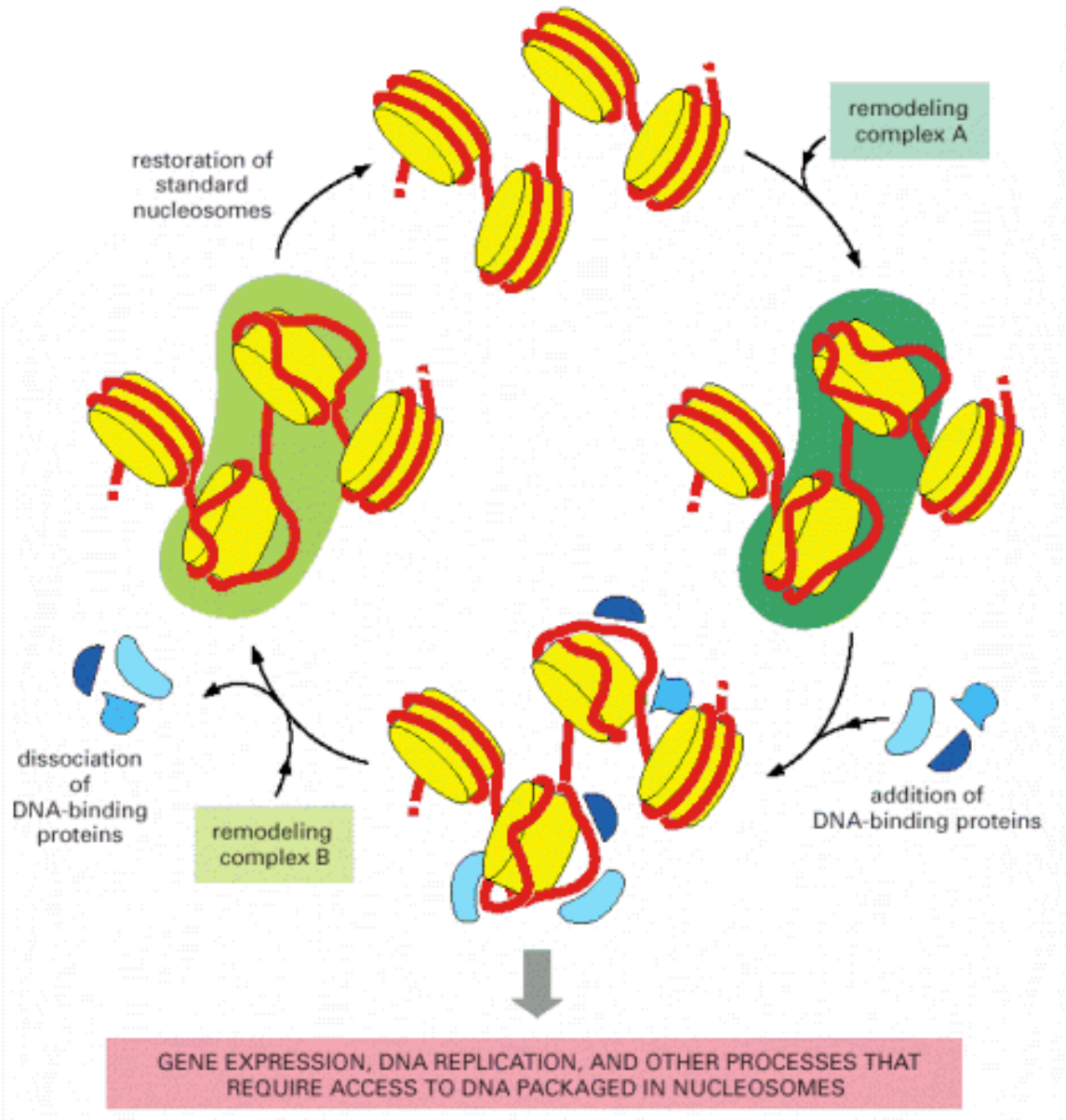



Figure 2.1. A scheme of nucleosome remodeling. The SWI/SNF complex is included in complex A, which is responsible for the disruption of nucleosomes. After the nucleosome remodeling, DNA-binding proteins gain direct access to DNA and biological processes, such as DNA transcription, can be initiated. The remodeling factor B will release DNA-binding proteins from the nucleosomes and restore the nucleosomes to their original state [2].

## 2.2 Identify SIN mutations

Point mutations of the genes that code for histones H3 and H4 in yeast can relieve the requirement of the SWI/SNF complex [1, 4]. These point mutations have been characterized using molecular genetic methods such as HO-LacZ expression profiles. The SWI/SNF complex regulates the transcription of the HO gene. When the LacZ gene is fused to the 5′ end of the HO gene, the expression level of the HO gene is positively related to the activity of the LacZ gene. The amount of β-galactosidase, the expression product of the LacZ gene, is then measured qualitatively in terms of growth conditions on cultures containing only galactose.

The yeast strain *swi1 hht2*Δ has a mutant SWI1 gene and lacks the H3 gene. Insertion of this strain in plasmids carrying the wild type H3 gene, the mutant H3 gene, or no gene as a control [4] reveals the role of point mutations in SIN-mutant nucleosomes. The H3 genes in the plasmids complement the deletion of the H3 gene in the yeast genome, from which the function of the H3 gene could be characterized. The data in Table 2.1 show that the presence of wild type H3 in the plasmid cannot relieve the mutations in the SWI1 gene since the HO-LacZ activity is as low as that of the control. The mutant H3 gene, where arginine116 is altered to histidine (R116H), however, bypasses the requirement of the SWI1 gene since HO-LacZ activity is high. Using similar methods, four other point mutations have been identified: T118I in H3; and V43I, R45H and R45C in H4. Other features of these mutations will be described in detail in section 2.4.

Table 2.1. The low level of LacZ gene transcription caused by mutation of the SWI gene is recovered by point mutations in the histone H3 gene in yeast.

| Strain (plasmid insert) | HO-LacZ activity |
|---|---|
| *swi1 hht2*Δ[a] (none [b]) | <0.1 |
| *swi1 hht2*Δ (*hht2-1*[c]) | 1.0 |
| *swi1 hht2*Δ (HHT2[d]) | <0.1 |

a. The yeast strain, *swi1 hht2*Δ, has the mutant SWI1 gene and the deletion of the HHT2 gene (histone H3 gene). The inserted plasmid carries no gene at all in b, the mutant H3 gene (R116H) in c, and the wild type H3 gene in d. This table is extracted and modified from the work of Kruger, *et al*. [4]

2.3 The X-ray crystal structures of the SIN mutant nucleosomes

Eleven crystal structures of nucleosome core particles reconstituted with histones bearing amino-acid point mutations have been determined [5]. The atomic coordinates of the SIN mutant structures have been deposited in Protein Data Bank [6]. These eleven structures share the same DNA sequence but contain different pairs of mutant histones, *i.e*., point mutations in both copies of H3 or H4. Five of the structures containing the five mutations originally found in yeast genetic screens, include the R116H H3, T118I H3, R45C H4, R45H H4, and V43I H4 mutants (Table 2.2 (a), column 5). Four other structures contain mutants where the four key amino acid residues (arginine 116 in histone H3, threonine 118 in H3, arginine 45 in H4, and valine 43 in H4) are substituted by alanine residues. These mutants include the R116A, T118A mutants of H3 and the R45A and V43A mutants of H4. As the methyl side chain of alanine is small, these alanine-involved mutations were investigated to see whether the SIN mutations are related to the loss of function of specific residues. One of the key H3 amino acids and one of the key H4 amino acid are thus replaced with structurally more disruptive groups in the histones of the crystal structures such as T118H H3 and R45E H4 SIN mutants. Moreover, in the T118H mutant, the imidizole ring of the mutant histidine residue carries a positive charge

compared to the polar noncharged side chain of the threonine residue in the native histone. In the R45E H4 mutant, the side chain of the glutamate residue carries a negative charge compared to the positively charged side chain of arginine in the wild type. These point mutations are clustered near the dyad of the nucleosome at SHL ±0.5 (Figure 2.2 (c), (d)). The general structures of the SIN-mutant nucleosomes are very similar to the wild-type structure except for moderate local effects, such as the loss of a few direct protein-DNA contacts near the sites of SIN mutations [5]. The wild-type nucleosome (structure ID: DNA2) studied in this thesis is unpublished data of K. Luger *et.al.*, which is very similar to the nucleosome core particle (PDB ID: 1AOI) [7] except for the improved resolution ( 2.2 Å of DNA2 vs. 2.8 Å of 1AOI).

Table 2.2. Properties about (a) the crystal structures and (b) the biochemistry experiments of SIN mutant nulceosomes.
(a)

| Mutant protein | Structure ID | Mutated residue in PDB file [a] | Mutation | Mutation property | PDB ID | NDB ID | Resolution | Reference |
|---|---|---|---|---|---|---|---|---|
| | DNA2 (WT) | | | | | | 2.2 | Unpublished data of Luger K. et.al. |
| H3 | R116A | Arg (resdiue No. 516) in chain A<br>Arg (residue No. 716) in chain E | Arg to Ala | Blackout function of the arginine residue | 1P34 | PD0413 | 2.7 | [5] |
| | R116H | Arg (resdiue No. 516) in chain A<br>Arg (residue No. 716) in chain E | Arg to His | Phenotype | 1P3A | PD0414 | 3.0 | [5] |
| | T118A | Thr (residue No. 518) in chain A<br>Thr (residue No. 718) in chain E | Thr to Ala | Blackout function of the threonine residue | 1P3K | PD0419 | 2.9 | [5] |
| | T118H | Thr (residue No. 518) in chain A<br>Thr (residue No. 718) in chain E | Thr to His | Replaced by a residue with different charge | 1P3L | PD0420 | 2.4 | [5] |
| | T118I | Thr (residue No. 518) in chain A<br>Thr (residue No. 718) in chain E | Thr to Ile | Phenotype | 1P3M | PD0421 | 2.4 | [5] |
| H4 | R45A | Arg (residue No. 45) in chain B<br>Arg (residue No. 245) in chain F | Arg to Ala | Blackout function of the arginine residue | 1P3B | PD0415 | 3.0 | [5] |
| | R45C | Arg (residue No. 45) in chain B<br>Arg (residue No. 245) in chain F | Arg to Cys | Phenotype | 1P3F | PD0416 | 2.9 | [5] |
| | R45E | Arg (residue No. 45) in chain B<br>Arg (residue No. 245) in chain F | Arg to Glu | Replaced by a residue with different charge | 1P3G | PD0417 | 2.7 | [5] |
| | R45H | Arg (residue No. 45) in chain B<br>Arg (residue No. 245) in chain F | Arg to His | Phenotype | 1P3I | PD0418 | 2.3 | [5] |
| | V43A | Val (residue No. 43) in chain B<br>Val (residue No. 243) in chain F | Val to Ala | Blackout function of the valine residue | 1P3O | PD0422 | 2.75 | [5] |
| | V43I | Val (residue No. 43) in chain B<br>Val (residue No. 243) in chain F | Val to Ile | Phenotype | 1P3P | PD0423 | 2.7 | [5] |

a. The PDB files containing the atomic coordinates of the nucleosomes are downloaded from Protein Data Bank [6].

(b)

| Mutant protein | Structure ID | Mutation | PDB ID | Resolution Å | Requirement of heat shifting [a] | Temp. °C for 50% shifting [b] | NaCl required for dissociation [c] | MNase accessibility [d] | DNase I cleavage [e] |
|---|---|---|---|---|---|---|---|---|---|
| | DNA2 (WT) | | | 2.2 | Yes (30min @ 37°C) | 43 | 1.1 M | | |
| H3 | R116A | Arg to Ala | 1P34 | 2.7 | Yes | | | | |
| | R116H | Arg to His | 1P3A | 3.0 | Yes (15min @ 37°C) | 39 | 0.85M | Much more | Unpositioned |
| | T118A | Thr to Ala | 1P3K | 2.9 | Yes | | | | |
| | T118H | Thr to His | 1P3L | 2.4 | No | | | | |
| | T118I | Thr to Ile | 1P3M | 2.4 | No | 34 | | More | Unpositioned |
| H4 | R45A | Arg to Ala | 1P3B | 3.0 | No | | | | |
| | R45C | Arg to Cys | 1P3F | 2.9 | No | 33 | 0.89M | | |
| | R45E | Arg to Glu | 1P3G | 2.7 | No | | | | |
| | R45H | Arg to His | 1P3I | 2.3 | No | 33 | | More | |
| | V43A | Val to Ala | 1P3O | 2.75 | Yes | | | | |
| | V43I | Val to Ile | 1P3P | 2.7 | Yes (20min @ 37°C) | 42 | 0.91M | | |

a. The time required for 100% shifting when a nucleosomes is incubated at 37˚C [5]. b. The characteristic temperatures required for 50% heat-shifting when nucleosomes are incubated for 60 minutes [8]. c. The midpoints of the concentrations of salt required for nucleosome dissociation [8]. d. The accessibility of micrococcal nuclease [9, 10]. e. The cleavage pattern of DNase I [9].

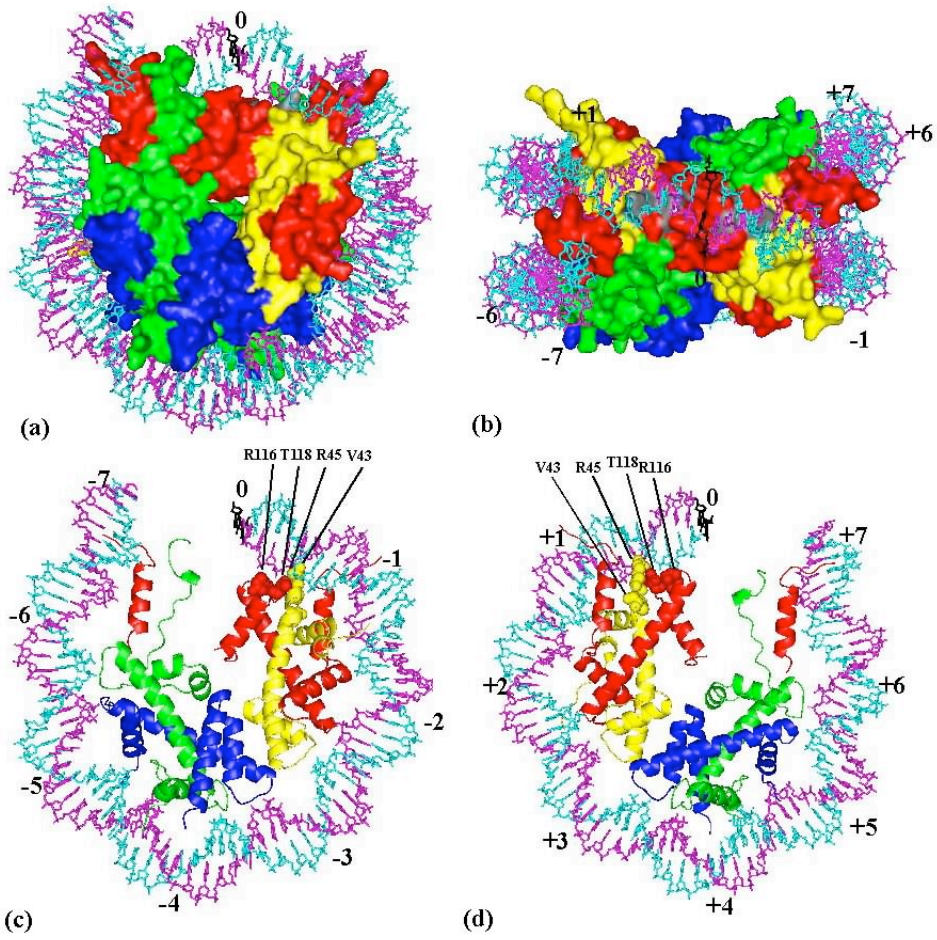



Figure 2.2. Color-coded representation of chemical features in the nucleosome core-particle structure (here called: DNA2) (unpublished data by K. Luger et al.). The top view (a) and the side view (b), with the dyad axis facing forward show the relative locations of the eight proteins in the histone core and the fold of the 146-bp DNA wrapped around it. The histone proteins, H2A, H2B, H3, H4, are shown respectively in red, yellow, green, and blue, and DNA strands I and II in cyan and magenta. The lower pair of images illustrate (c) the short half of the DNA with associated proteins, including the dyad axis at base pair 73 (1 to 73 bp), and (d) the long half of the DNA and bound histones, including the dyad axis at base pair 73 (73 to 146 bp). The numbers in black along the DNA correspond to the superhelical location (SHL). The black base pair at 0 corresponds to the structural dyad at base pair 73 and SHL location 0. The amino acid point mutations in the SIN mutants are mapped in gray on the surface of histone proteins H3, H4 in (a) and (b), and are shown in space-filled representation in (c) and (d).

The local changes of the mutated residues and the DNA at SHL ±0.5, compared to the wild type structure, were emphasized previously [5]. As the long side chain of arginine 45 is inserted into the DNA minor groove at SHL ±0.5, the mutation of this residue to cysteine or alanine introduces an empty space in the DNA minor groove. This empty space, in turn, causes a loss of contacts of histone H4 to DNA atoms in the vicinity of sites of the mutation. When the arginine 45 is changed to a histidine or glutamate residue, the side chains of histidine and glutamate flip away from the nucleosomal DNA, also introducing empty space in DNA minor groove. In contrast, in the T118I and T118H H4 mutants, the mutated isoleucine 118 or histidine 118 residue pushes arginine 45 further into the DNA minor groove, causing a large shift (up to 1.2 Å) of the DNA backbone. Because of this DNA backbone shift, hydrogen bonds between protein and DNA atoms are lost near the sites of mutations. In the R116H and R116A H3 mutants, the hydrogen bonds are also lost or weakened by the mutations of arginine 116 to histidine or alanine. The V43A and V43I H4 mutants undergo more subtle structural changes. In the V43A mutant, the mutation of valine 43 to alanine does not cause any essential local distortion. This thesis presents a detailed analysis of the conformation and interactions of the histone proteins and DNA in SIN-mutant structures, from the atomic level to global DNA remodeling. The atomic contacts of protein to DNA atoms, the base-step parameters, and the superhelical parameters in the SIN-mutant structures are compared to those in the wild type structure. The three-dimensional DNA templates from the mutant structures are used to thread the nucleosomal DNA sequence from the X-ray structures. These analyses provide new insight into the nucleosome-remodeling mechanism at the level of the nucleosome as a whole. This study also provide a method to compare the nucleosome

structures regardless of the length and nucleotide composition of DNA, which will be discussed in Chapter 7.

## 2.4 The SIN-mutant nucleosomes have some characteristics different from the wild-type nucleosomes.

### 2.4.1 The mutations lead to increased nucleosome-sliding rates

The nucleosome-sliding rates have been measured by two groups using heat-shift methods [5, 8]. One method is called high-resolution gel shift assay and repositioning of nucleosome core particle [5, 9]. When histone protein is reconstitute at 4 °C with nucleosomal DNA longer than 146 bp, the histone protein core is usually positioned at specific locations along the DNA sequence. This heterogeneous population of nucleosomes with different positions of the histone core on the same nucleosomal DNA, can be detected by shifted locations on a gel. Surprisingly, reconstituted nucleosomes with the different positioning on DNA, can also be obtained with two strong146-bp positioning sequences, the 5S RNA gene of *Lytechinus variegatus* and the human α-satellite DNA introduced in the crystallizing nucleosomes, at 4 °C. Even more surprisingly, the simple incubation at 37 – 55 °C for a period of the time (20 – 180 min) induces formation of a homogenous population of nucleosomes positioned uniquely at the central position of the 146 bp nucleosomal DNA. The purity of the heat-induced central positioned nucleosomes is high enough for crystallization.

SIN-mutant and wild type nucleosomes, were reconstituted at 4 °C on a 146-bp DNA fragment with three different translational positions offset by 10 bp with respect to each other and subjected to heat-shifting [5]. The central binding position is most stable as it

makes more contacts with protein than the other two positions. Histone proteins bound at the two off-center positions slide to the central position when the nucleosomes are incubated at 37 °C for a period of time. However, even at 4 °C, some SIN mutants (all of the R45 H4, T118H, and T118I H3 mutants) form homogenouse populations at the central positioned nucleosomes. These mutants appear to slide along the DNA fragment to the central position so quickly that it is difficult to obtain samples with the histone protein assembly bound at the two off-center positions. In contrast, all of the R116 H3, T118A H4 and V43 H4 mutants forms heterogeneous population of nucleosomes are formed at 4 °C and shift to the central position upon the incubation at 37 °C for different periods of time. The times required for heat shifting of the wild type nucleosome, and the V43I and R116H, mutants 37 °C are 30 min, 20 min, and 15 min, respectively (Table 2.2 (b), column a). The sliding rates of the R116 H3, T118A H3, and V43 H4 mutants are faster than the wild type but slower than those of the T118H and T118I mutants of H3 and the R45 mutants of H4. That is to say, the energy barrier of these mutants towards repositioning is significantly lowered and the effect is apparently strongest for the T118H, T118I, and R45 mutants.

In the second method [8], a 255-bp DNA fragment from the nucA region of the mouse mammary tumor virus long-terminal repeat (MMTV LTR) was used to reconstitute nucleosomes. The dyad of the majority of bound nucleosomes is located at +70 relative to the MMTV transcriptional start site, *i.e.*, approximately at the center of the DNA fragment. Thus the nucleosomes occupy the central 147 bp and there are two 54 bp protein-free ends. When nucleosomes, which were reconstituted on this DNA fragment and maintained at 0-4 °C, are subjected to thermal incubation for increasing times, they

redistribute to positions +22 and +125, 48 bp upstream and 55 bp downstream of the original position, respectively. The characteristic temperatures sufficient for mobilizing 50% the nucleosomes in 60 min are listed in Table 2.2 (b), column b. If the required temperature is lower, the sliding rate is higher. The order of sliding from high to low corresponds to mutants R45C, R45H in H4 (33 °C) > T118I in H3 (34 °C) > R116H in H4 (39 °C) > V43I in H4 (42 °C) > wild type (44 °C) and is consistent with the findings of the heat-shift experiments. In the heat-shift experiment, the order of the sliding rate from high to low is T118H and T118I in H3, R45A, R45C, R45E, R45H in H4 (no requirement of heating) > R116H in H3 (heating at 37 °C for 15 min) > V43I in H3 (heating at 37 °C for 20 min) > wild type (heating at 37 °C for 30 min).

2.4.2 Mutant nucleosomes dissociate at lower salt concentration [8]

The concentrations of NaCl required for the dissociation of histone proteins and DNA have also been measured. The dissociation is monitored by measuring the loss of quenching intrinsic fluorescence emissions from the hydroxy-phenyl group of tyrosine residues of histone proteins by nearby DNA [10]. The R116H H3, R45C H4, and V43I H4 mutants dissociate at lower NaCl concentrations than the wild type (Table 2.2 (b), column c), suggesting that the interactions between the histone protein and DNA in these mutants are not so strong as those in the wild type. The order of interactions between the histone protein core and DNA from strong to weak is wild type > V43I H4 > R45C H4 >R116H H3, a result consistent with the aftermentioned positioning shifts.

2.4.3 The DNA of some mutant nucleosomes is more accessible to enzymatic cleavage than that in the wild-type nucleosome (Table 2.2 (b), column d) [11]

Micrococcal nuclease (MNase) prefers to cleave protein-free DNA. It normally cleaves chromatin into nucleosome core particles, each containing the histone protein octamer $(H2aH2b\text{-}H3H4)_2$ and 147 bp DNA. If the time of incubation with MNase is prolonged, the 147 bp DNA can also be digested into small fragments and the nucleosome core particle disrupted. Wild-type mononucleosomes, which were reconstituted on 238 bp DNA fragments and digested by MNase, show the obvious accumulation of 147 bp of DNA. The band showing the accumulation of 147 bp of DNA is either completely absent (R116H H3) or a substantially reduced (T118I H3) for mutant histones. That is to say, the DNA bound to these mutants is more accessible to MNase.

In other experiments, nuclei have been isolated from strains that express wild-type or histone H4 mutants and then treated with MNase. Comparison of the MNase digestions shows that the chromatin that contains the R45H H4 mutant is more sensitive to MNase digestion than the chromatin that contains wild-type histone H4.

2.4.4 The positioning of DNA on the histone surface changes in some mutants (Table 2.2 (b), column e) [11]

DNase I digestion shows the precise rotational positioning of the double helix with respect to the histone surface, such that certain sequences are exposed toward solvent and others make contact with the histones. The sites of DNase I cleavage are spaced every 10 or 11 bp. The precise rotational positioning of DNA on the surface of the core histones, as revealed by DNase I cleavage, is substantially lost in the T118I and R116H mutants of histone H3.

## 2.5 Outline of the thesis

The SIN-mutant nucleosomes differ in many ways from the wild-type nucleosomes, with different heat-shifting requirements, NaCl concentration-dependent dissociation of histone proteins from nucleosomal DNA, MNase accessibility, and DNase I cleavage. The structures of the SIN-mutant nucleosome structures, however, show only minimal distortions from the wild-type structure, with a few direct protein-DNA contacts lost near the sites of amino acid mutation and no large-scale change in overall structure [5]. This thesis presents a detailed analysis of the conformation and interactions of the histone proteins and DNA in the SIN-mutant structures, from the level of individual atomic contacts to global DNA remodeling. The base-pair step parameters of the SIN-mutant nucleosome structures are compared to those of the wild-type structure and shown in novel, color-coded graphic representations in Chapter 3. The loss of contacts of histone protein to DNA in the SIN-mutant nucleosome structures, compared to the wild-type, is examined in Chapter 4. The nucleosomal DNA molecules are represented in a cylindrical frame in Chapter 5. The superhelical shapes of the SIN-mutant nucleosomal DNA in these frames are compared with the superhelical fold of the wild-type. Furthermore, elastic scores of nucleosomal DNA sequences, which measure the deformation of the sequences on structural templates from the SIN-mutant and wild-type structures, are computed, using threading methods, in Chapter 6. The highly deformed steps with large deformation scores in the SIN-mutant structures, compared to the wild-type, are identified and used to remodel the wild-type nucleosomal DNA structure. These analyses reveal DNA conformational changes and loss of contacts to the histone proteins *in trans*, sequentially and spatially distant from the mutation sites, in the SIN-mutant nucleosomes

compared to the wild-type. Models, which incorporate these highly deformed steps, are suggestive of ways in which the DNA might loop away or peel off the nucleosome core particle.

Whereas the SIN-mutant and wild-type nucleosomes have comparable DNA sequences but different protein cores, two nucleosome structures have identical histone cores but different DNA sequence. The difference in the superhelical shapes and the shifted locations of highly deformed steps in these two structures show how nucleosomal DNA adjusts its conformation to accommodate different DNA sequence.


1. Lewin, J., *Genes VIII*. 2004: Pearson Prentice Hall.
2. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P., *Molecular Biology of the Cell*. 4th ed. 2002, New York: Garland Publishing.
3. Pazin, M.J., and Kadonaga, J. T., *SWI2/2SNF and related proteins:ATP-driven motorss that disrupt protein-DNA interactions?* Cell, 1997. 88: p. 737-740.
4. Kruger, W., Peterson, C. L., Sil, A, Coburn, C., Arents, G., Moudrianakis, E. N., and Herskowitz, I., *Amino acid substitutions in the structured domains of histones H3 and H4 partially relieve the requirement of the yeast SWI/SNF complex for transcription*. Genes & Development, 1995. 9: p. 2770-2779.
5. Muthuranjan, U.M., Bao, Y., Forsberg, J., Edayathumagalam, R. S., Dyer, P. N., White, C. L., and Luger, K., *Crystal structures of histone Sin mutant nucleosomes reveal altered protein-DNA interactions*. EMBO J., 2004. 23: p. 260-271.
6. Berman H. M., W., J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E., *The Protein Data Bank*. Nucleic Acids Research, 2000. 28: p. 235-242.
7. Luger, K., Maedar, W., Richmond, R. K., Sargent, D. F., and T. J. Richmond, *Crystal structure of the nucleosome core paricle at 2.8 Å resolution*. Nature, 1997. 389: p. 251-259.
8. Flaus, A., Rencurel, C., Ferreira, H., Wiechens, N., and Owen-Hugers, T., *Sin mutations alter inherent nucleosome mobility*. EMBO J., 2004. 23: p. 343-353.
9. Luger, K., Rechsteiner, T. J., and Richmond, T. J., *Preparation of nucleosome core particle from recombinant histones*. Mehtods in Enzymology, 1999. 304: p. 3-18.
10. Luger, K., Rechsteiner, T. J., Flaus, A. J., Waye, M. M., and Richmond, T. J., *Charaterization of nucleosome core particles containing histone proteins made in bacteria*. J. Mol. Biol., 1997. 272: p. 301-311.

11. Kurumizaka, H., Wolffe, A. P., *Sin mutation of histone H3: influence on nucleosome core structure and function*. Molecular and Cellular Biology, 1997. 17(12): p. 6953-6969.

# Chapter 3 Geometrical Parameters of Nucleosomal DNA

## 3.1 Background

### 3.1.1 Geometrical parameters

Six base-pair parameters (three rotations and three translations) uniquely define the relative position and orientation of two base pairs. These parameters — termed Shear, Stretch, Stagger, Buckle, Propeller, and Opening — are illustrated in Figure 3.1. Six rigid body parameters are also required to describe the position and orientation of one base pair relative to another. Two sets of dimer step parameters are commonly used: (1) the set of local base-pair step parameters — Shift, Slide, Rise, Tilt, Roll, and Twist, which describe the stacking geometry of a dinucleotide step from a local perspective; and (2) the set of helical parameters — *X*-displacement, *Y*-displacement, helical twist, helical rise, inclination, tip, which describe the positioning of base pair with respect to the axis of the helix [1, 2].

These parameters, when calculated with respect to the standard reference frame developed by Olson et al. [3], are nearly independent of analytical treatment. Here we compute these parameters, using a rigorous, matrix-based algorithm within 3DNA [2]. Except for a change from the mean base-pair normal used in the calculation of step parameters to a local helical axis in the determination of helical parameters, the mathematics behind the two sets of dimer step parameters are identical and rigorous. Thus, the set of local base-pair step parameters and the set of helical parameter can be deduced from each other [2].

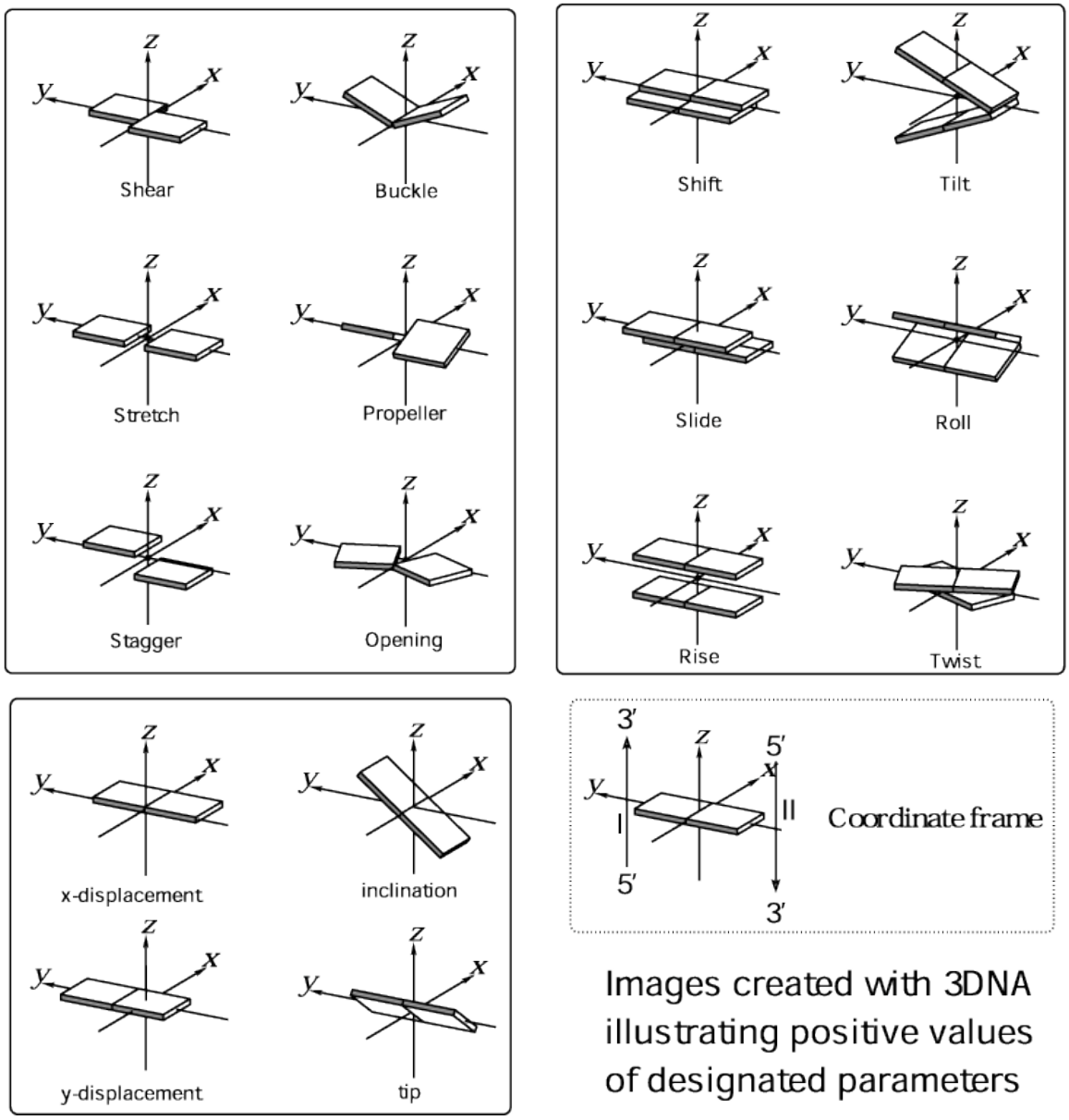



Figure 3.1. Schematic illustration of DNA base-pair parameters. The base-pair reference frame (lower right) is constructed such that the *X*-axis points away from the (shaded) minor groove edge of a base or base pair and the *Y*-axis points toward the sequence strand (I). The relative position and orientation of successive base-pair planes are described with respect to both a dimer reference frame (upper right) and a local helical frame (lower left). Images illustrate positive values of the designated parameters. The relative position and orientation of two complementary bases are described with respect to a local frame (upper left). For illustration purposes, helical twist is the same as Twist and helical rise is the same as Rise.

### 3.1.2 Major- and minor-groove width

The major and minor grooves are the indentations in nucleic acid double helices formed as a consequence of the asymmetry of the base pair. The C1′-N9 (purine) and C1′-N1 (pyrimidine) base-sugar bonds lie, by convention, on the minor-groove side, and the C6/N7 (purine) and C4 (pyrimidine) base atoms and their substituents on the major-groove side (Figure 3.2). Groove width can be defined as the perpendicular distance between phosphate groups on opposite strands (Figure 3.3). The algorithm used in 3DNA [2] to calculate the major- and minor-groove width is that proposed by El Hassan and Calladine [4].

### 3.1.3 Temperature factor (B-factor)

The temperature factor, or B-factor ( $B$ ), is a measurement of the thermal motion of atoms in an X-ray crystal [5]. Mathematically, the B-factor is given as $B = 8\pi^2 \bar{u}^2 / 3$ where $\bar{u}$ is the mean displacement of an atom during the thermal motion. The atomic thermal motion affects the intensity of the X-ray through a factor, $\exp[-B(\sin^2\theta/\lambda^2)]$, where $\theta$ is the reflecting angle of reflection of the X-ray and $\lambda$ is the wavelength.

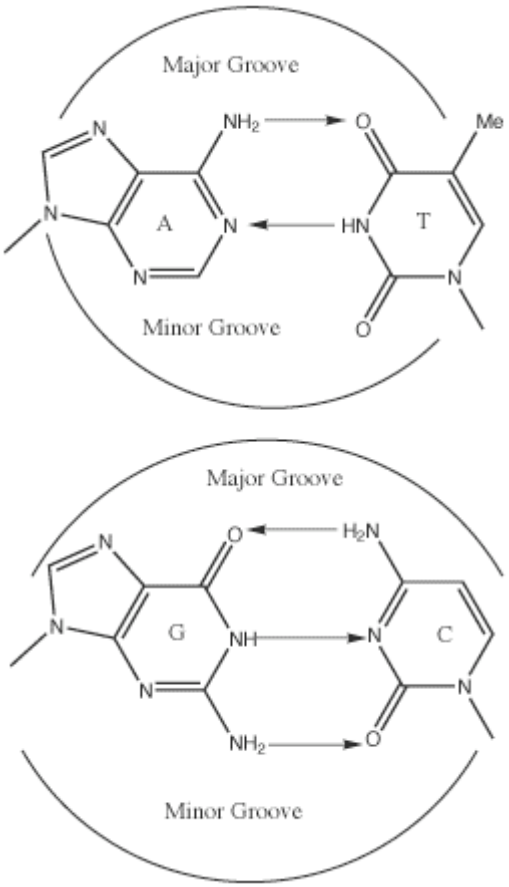



Figure 3.2. Schematic illustration of the minor and major grooves of DNA. The C1′- N9 (purine) and C1′-N1 (pyrimindine) glycosyl bonds by convention lie on the minor-groove edge of the Waston-Crick base pairs. The C6/N7 (purine) and C4 (pyrimidine) base atoms lie on the major-groove edge. The hydrogen bonds between complementary bases are depicted by arrows running from the proton-donor to proton-acceptor atoms.

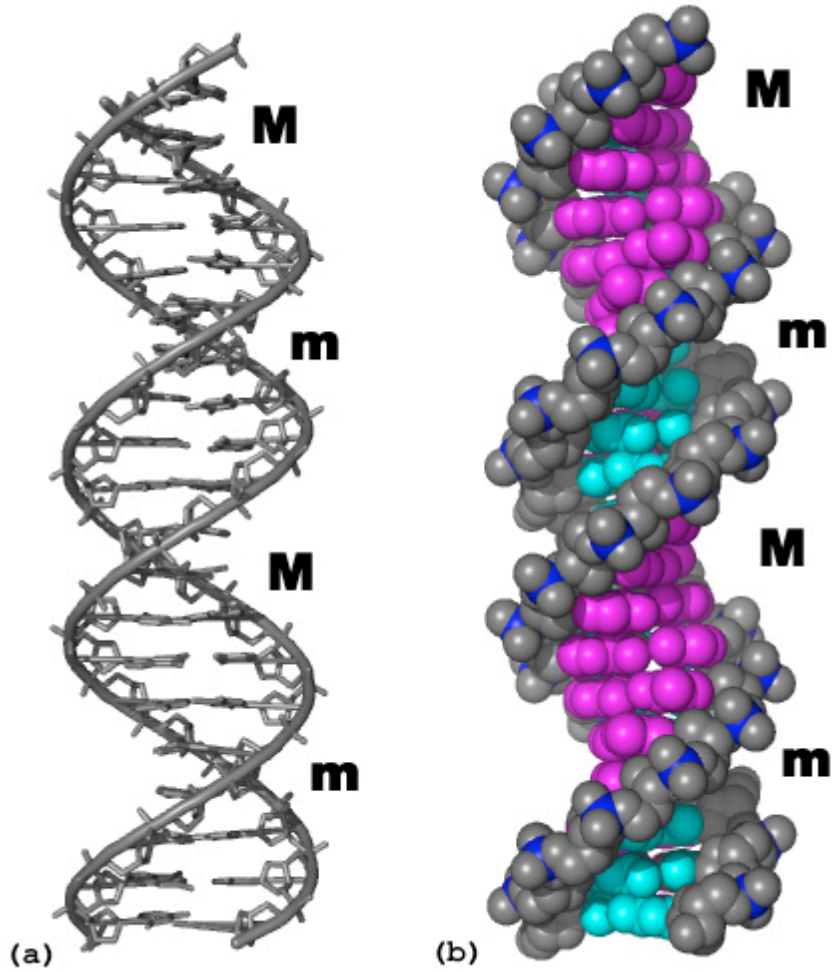



Figure 3.3. Representation of DNA showing the minor and major grooves designates: (a) stick representation of DNA; (b) space-filling representation of DNA. M donates major groove and m means minor groove. The color scheme in (b) highlights base atoms exposed in the major grooves (magenta) and in minor grooves (cyan). Backbone atoms, except for phosphorus, are shown in grey. The phosphorus atom is shown in blue.

## 3.2 Methods

### 3.2.1 Transfer a numeric table to a color-coded image

A computer program, which can transfer a numerical table to a mosaic picture, has been developed to align the values of selected parameters in different DNA structures and to make visual comparisons of different structure. The mosaic blocks are color-coded

according to the value of the DNA parameter. Rather than viewing a numerical 'spreadsheet', we see a mosaic picture with patterns of color. The similarities and differences in the variation of corresponding parameters in different DNA structures can be detected quickly and clearly by eye.

We divide a given range of values $[t_1, t_2]$ into 21 intervals, which correspond to the 21 colors in the color range discussed later. The length of the central 19 intervals is $(t_1 - t_2)/20$. The length of the first and last intervals is $(t_1 - t_2)/40$, half of the length of the central 19 intervals.

If $x < t_1$ or $x > t_2$, we arbitrarily set $x$ in the first or last interval (interval index, $i = 1$ or 21), that is, $x$ is color-coded as the darkest blue or red. The middle points of the central 19 intervals pinpoint the colors in the color range and the other values within the intervals are approximated by the same colors as that of the middle-point values.

Thus, for any numeric value, $x$, inside or outside the range $[t_1, t_2]$, the $i^{th}$ interval, which $x$ belongs to, can be written as:

$$i = \begin{cases} 1 & x < t_1 \\ \left[11.5 - \dfrac{t_m - x}{v}\right] & t_1 \le x \le t_m \\ \left[\dfrac{x - t_m}{v} + 11.5\right] & t_m \le x \le t_2 \\ 21 & x > t_2 \end{cases}$$

Here $v = (t_2 - t_1)/20$ is the length of interval, $t_m = (t_1 + t_2)/2$ is the value of the middle point over the range of values, and the square parentheses designate the integer, which is closest to the non-integer obtained from the formula above.

The 21 intervals correspond to a series of 10 blue shades for values less than the middle point of the range, $t_m$, 10 red shades for value more than $t_m$, and white for the values around $t_m$.

In python programming [6], a color can be defined as a vector composed of three elements designating red, green and blue. A color vector, $c_i$, in the color range used in my programs is expressed as follows:

$$c_i = \begin{cases} [25.5 * i \quad 25.5 * i \quad 255] & 1 \le i < 11 \\ [255 \quad 255 \quad 255] & i = 11 \\ [255 \quad 25.5 * (21 - i) \quad 25.5 * (21 - i)] & 11 < i \le 21 \end{cases}$$

When $0 \le i < 11$, the color vector, $c_i$, represents a shade of blue. When $i = 11$, the color vector represents white. When $11 < i \le 21$, the color vector $c_i$ represents a shade of red. The value $x$ in the $i^{th}$ interval is shown in a color defined by the color vector $c_i$.

For a range $[-t_1, t_1]$, which is symmetrical around zero, the small values around zero do not show important information. The central 10 colors are sometimes blacked out to emphasize large positive or negative values. The color vector $c_i$ is thus written as below:

$$c_i = \begin{cases} [25.5 * i \quad 25.5 * i \quad 255] & 1 \le i < 6 \\ [0 \quad 0 \quad 0] & 6 \le i \le 16 \\ [255 \quad 25.5 * (21 - i) \quad 25.5 * (21 - i)] & 16 < i \le 21 \end{cases}.$$

3.2.2 A novel three-dimensional representation of nucleosomal DNA

A program, which integrates information about the atoms contacted by protein and the structural parameters of DNA in a three-dimensional molecular graphics representation of a protein-DNA complex structure, has also been developed. The atoms of a DNA molecule are color-coded according to the value of a given structural parameter

and differences are shown by the variation of color in a molecular graphics representation of the structure. The information about the atoms in a DNA molecule contacted by protein is superimposed on the molecular graphics presentation by showing the atoms contacted by protein as enlarged balls and those not contacted by protein as dots. The colors of the balls and sticks (bonds) display the variation of the parameters of a given DNA structure.

I have written a program to create a script for use with the molecular visualization software, Pymol [7]. With the commands in the first part of the script, the backbone atoms of nucleosomal DNA are selected and color-coded according to the values of base-pair step parameters. The color-coding procedure is the same as mentioned in Section 3.2.1. In the second part of the script, DNA atoms contacted by protein atoms are picked up and shown in space-filled representations. At the same time, these enlarged DNA atoms are also colored in shades that reflect the values of the selected parameters.

This method has two advantages: (1) it is convenient to see the variation of the value of a given DNA structure parameter directly in a three-dimensional molecular graphic representation of the structure; (2) examination of the images reveals the parts of the DNA structure and parameters which change in response to protein contact.

### 3.2.3 Root-mean-square displacement

The root-mean-square displacement of corresponding DNA fragments in different nucleosome core particles is computed using a least-squares fitting procedure based on the closed-form solution of absolute orientation using a unit quaternion as first introduced by Horn [8]. Suppose there are two matrices representing two sets of coordinates, the

wild type one ($W$) and the mutant one ($M$). A symmetrical 3×3 matrix, namely a covariance matrix $C$ is obtained as

$$C = \frac{1}{N}\left[W' \cdot M - \frac{1}{N} W' \cdot i \cdot i' \cdot M\right].$$

Here $N$ is the number of atoms in the wild-type or mutant structure, namely, the number of rows in matrix $W$ or $M$. $i$ is an $N \times 1$ matrix composed of only ones. $W'$ and $i'$ are the transposes of matrices $W$ and $i$.

A real $4 \times 4$ symmetric matrix $M$ can be deduced from the elements of matrix $C$ as follow:

$$M = \begin{bmatrix} c_{1,1} + c_{2,2} + c_{3,3} & c_{2,3} - c_{3,2} & c_{3,1} - c_{1,3} & c_{1,2} - c_{2,1} \\ c_{2,3} - c_{3,2} & c_{1,1} - c_{2,2} - c_{3,3} & c_{1,2} + c_{2,1} & c_{3,1} + c_{1,3} \\ c_{3,1} - c_{1,3} & c_{1,2} + c_{2,1} & -c_{1,1} + c_{2,2} - c_{3,3} & c_{2,3} + c_{3,2} \\ c_{1,2} - c_{c,1} & c_{3,1} + c_{1,3} & c_{2,3} + c_{3,2} & -c_{11} - c_{2,2} + c_{3,3} \end{bmatrix}.$$

The eigenvector, which corresponds to the most positive eigenvalue of matrix *M*, is

$$V = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 \end{bmatrix}.$$

The rotation matrix *R* is written as:

$$R = \begin{bmatrix} v_1v_1 + v_2v_2 - v_4v_4 & 2(v_2v_3 - v_1v_4) & 2(v_2v_4 + v_1v_3) \\ 2(v_3v_2 + v_1v_4) & v_1v_1 - v_2v_2 + v_3v_3 - v_4v_4 & 2(v_3v_4 - v_1v_2) \\ 2(v_4v_2 - v_1v_3) & 2(v_4v_3 + v_1v_2) & v_1v_1 - v_2v_2 - v_3v_3 + v_4v_4 \end{bmatrix}.$$

The translational offset ($O$) between the two sets of coordinates is the translational difference between their centroids and written as follow:

$$O = \overline{M} - \overline{W} \cdot R'.$$

Here $\overline{M}$ is a $1 \times 3$ vector composed of the average of each column of matrix $M$. Each element of the $1 \times 3$ vector $\overline{W}$ corresponds to the average of each column of matrix $W$.

The difference matrix $D$, whose elements represent the minimized differences between the corresponding $1 \times 3$ coordinates between the two sets of coordinates, are found to be:

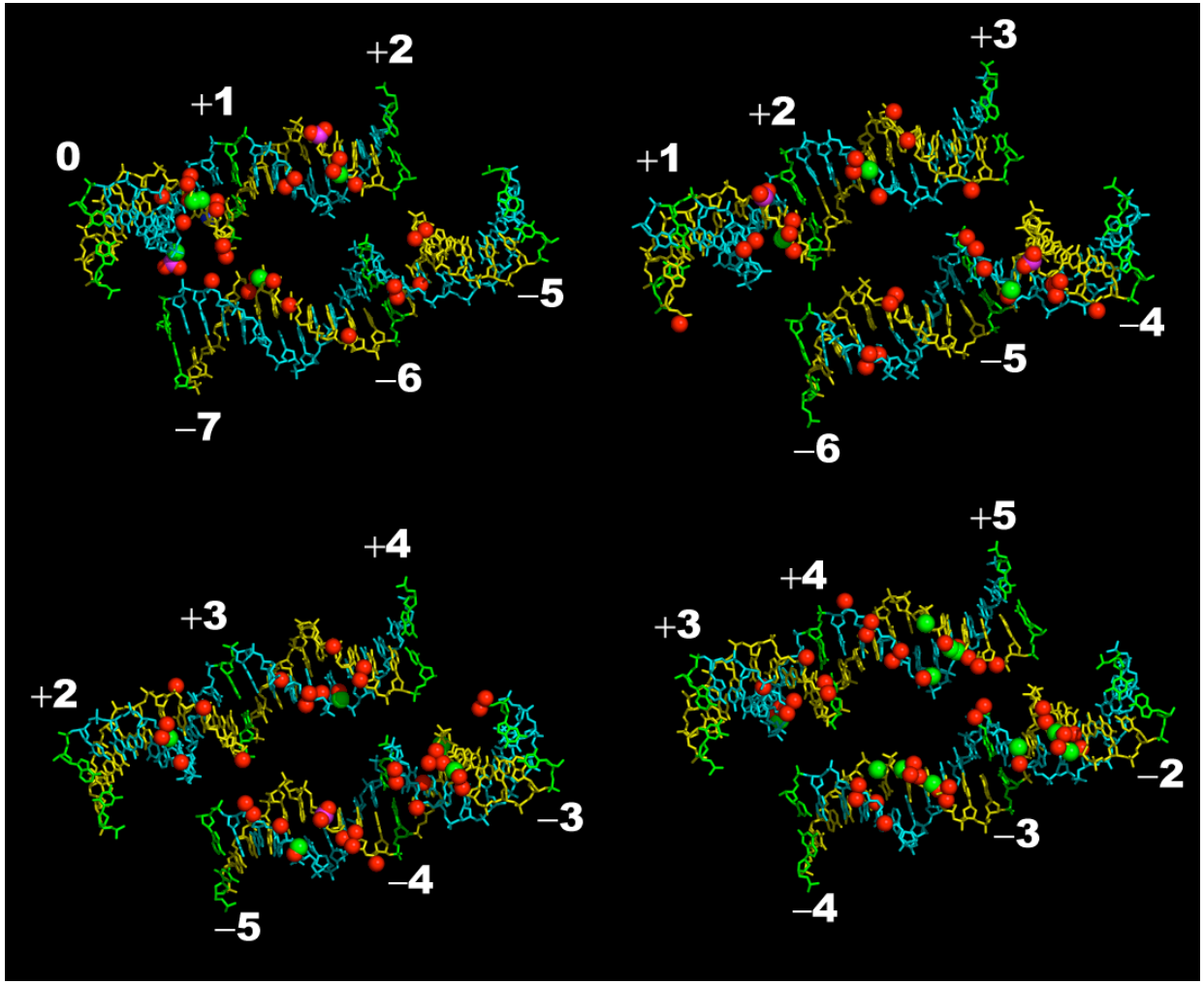
$$D = M - (W \times R' + I \times O)$$

Here $I$ is $N \times 3$ matrix, in which all elements is 1. The value of the root-mean-square deviation (RMSD) is the square root of the average of the square of every element in matrix $D$.

## 3.3 Results

### 3.3.1 DNA super groove in nucleosome core particles

The DNA atoms contacted by histone proteins are clustered at 14 sites in the DNA minor grooves on the inner side of the DNA superhelix. The DNA minor grooves from two juxtaposed stretches of nucloesomal DNA form a “super” minor groove at the site of protein-DNA interaction (Figure 3.4)[9].


+2
+1
0
–5
–6
–7
+3
+2
+1
–4
–5
–6
+4
+3
+2
–3
–4
–5
+5
+4
+3
–2
–3
–4

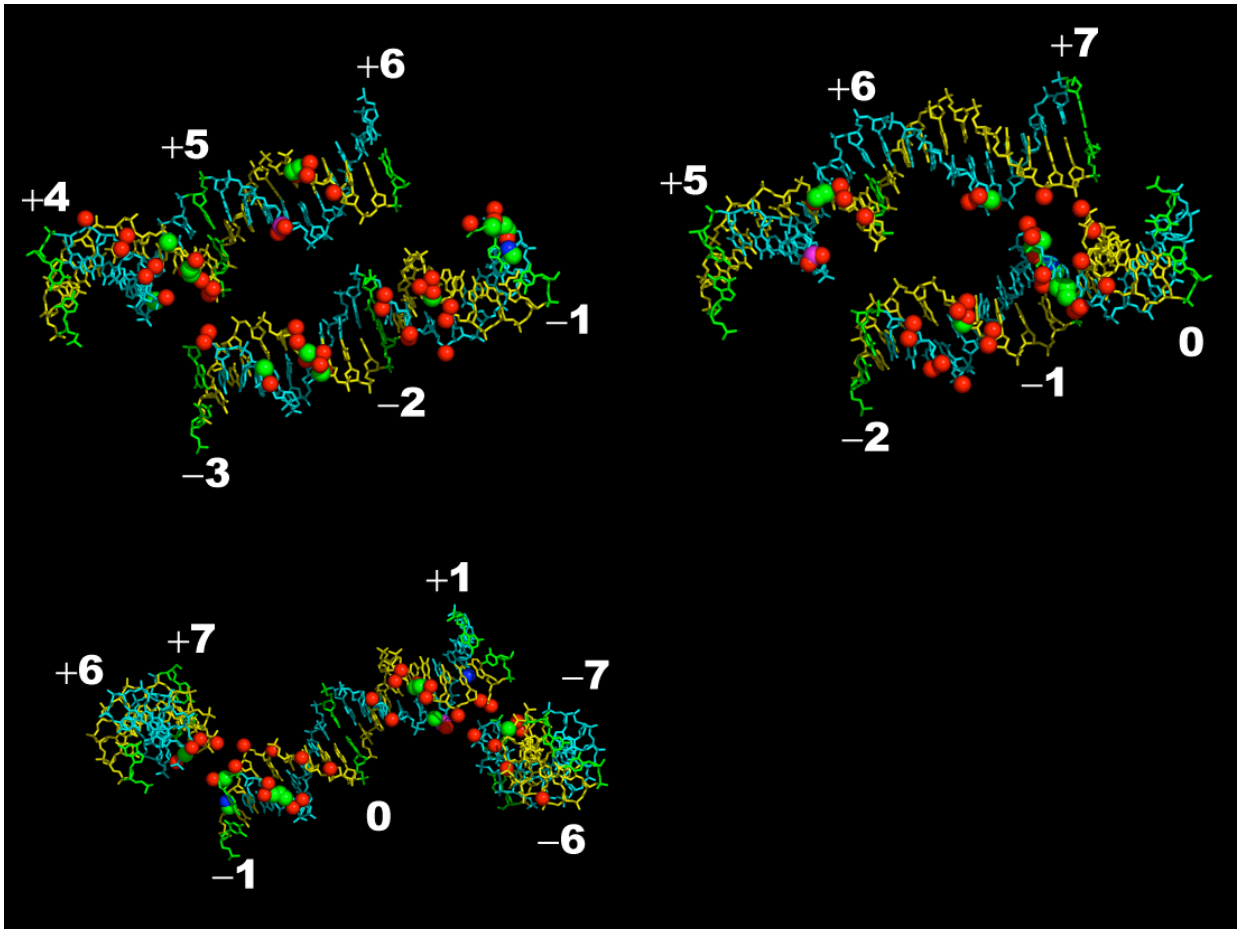



Figure 3.4. The DNA super grooves at the protein-DNA interface in the 2.8 Å resolution nucleosome core particle structure (structure ID: DNA2). Strands I and II are shown in yellow and cyan, respectively. The base pairs in green corresponds to the superhelical locations denoted along the DNA. The DNA atoms contacted by protein are shown in space-filled representations and color-coded according to their elemental identity. carbon atoms are show in green; nitrogen atoms are shown blue; oxygen atoms are shown in red; phosphorus atoms are shown in magenta.

### 3.3.2 Mapping protein-DNA contacts and values of the dimer step parameters along nucleosomal DNA.

The seven protein-DNA interaction sites in the first half of 146 bp nucleosomal DNA (1 to 73 bp) are shown with space-filled representations of the DNA atoms contacted by histone proteins (Figure 3.5) (structure ID: DNA2, unpublished data of Luger, K. *et al.*). At the same time, all the DNA backbone atoms are color-coded according to the values of selected DNA geometrical parameters. As is clear from the figures, the DNA atoms contacted by protein occur where DNA bends into its minor groove and the minor-groove width becomes exceptionally narrow (SHL ±1.5, ±2.5, ±3.5, ±4.5, ±5.5). In Figure 3.5 (d), the space-filled representation of the contacted DNA atoms in blue illustrates how the protein-DNA contacts occur in concert with the narrowing of the minor groove. The narrow minor groove at the protein-DNA contact locations is accompanied by positive Slide, negative Roll, and increased Twist (Figure 3.5 (a), (b), and (c)). This coupling of Slide, Roll and Twist of nucleosomal DNA have been noted by Richmond *et al*. [10] and Tolstorukov *et al*. [11]. However, the present method of transferring a numerical table to a series of color-coded pictures makes a very clear representation of the periodicity and the coupling of the step parameters.

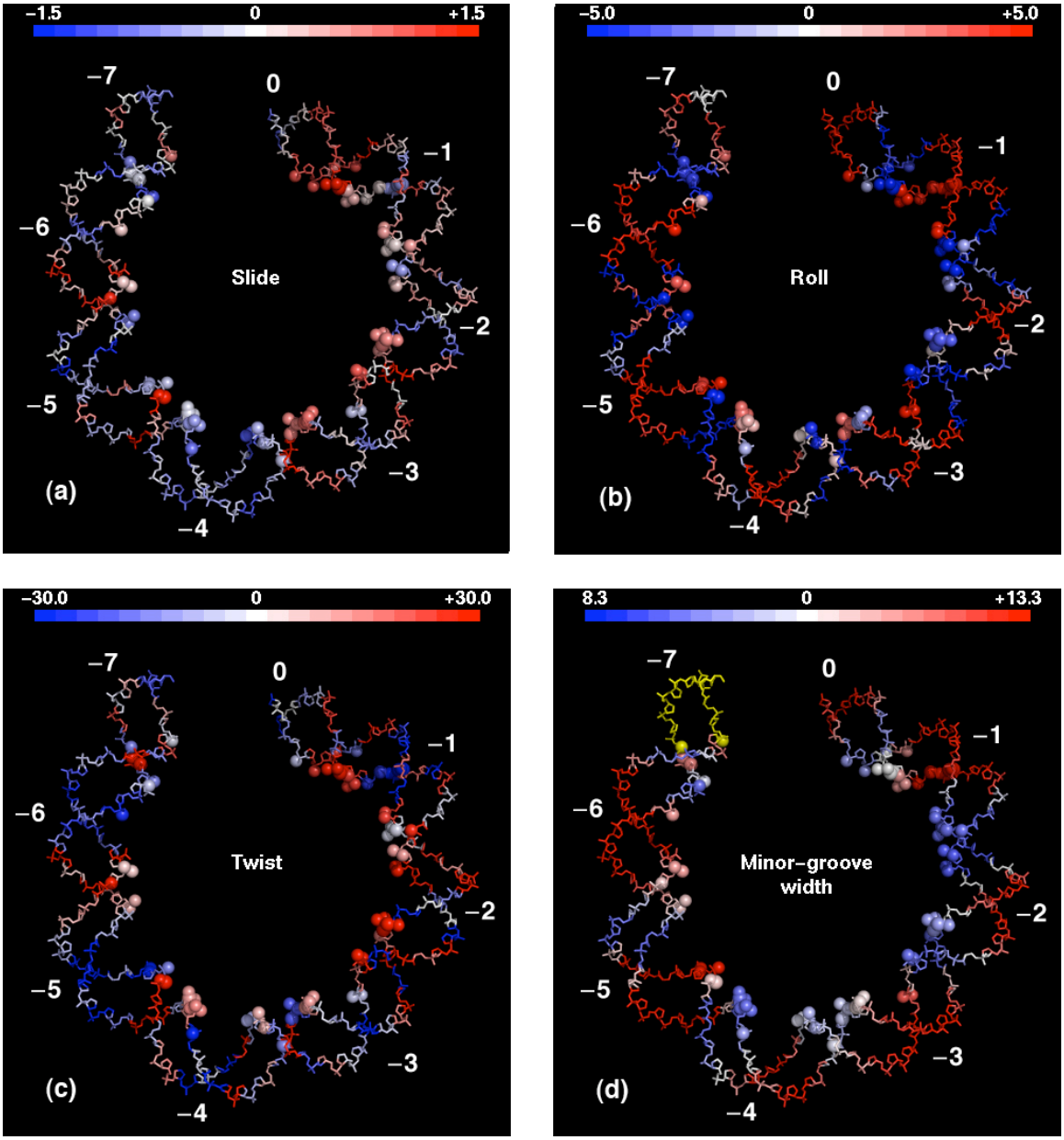



Figure 3.5. The values of DNA step parameters mapped on a graphical representation of base pairs 1 to 73 of nucleosomal DNA (unpublished data of Luger *et al.*, ID: DNA2). The backbone is color-coded according to the values of (a) Slide, (b) Roll, (c) Twist, and (d) minor-groove width. The atoms represented by enlarged balls are contacted by protein within 3.4 Å. The respective color-coded value ranges are as follows: −1.5 Å to +1.5 Å (Slide), −5.0˚ to +5.0˚ (Roll), −30.0˚ to +30.0˚ (Twist), +8.3 Å to +13.3 Å (minor-groove width). The color scheme is shown on the top of each part of the figure, with values less than or equal to the designated minimum in dark blue, middle values in white, and values more greater than or equal to the designated maximum in dark red. The DNA fragment is shown in yellow in (d) where the minor-groove width cannot be measured.

3.3.3 The variation of base-step parameters along nucleosomal DNA

For DNA in SIN mutants and wild type nucleosomal structures, the values of base-step parameters are mapped on each nucleotide and aligned with respect to the common DNA sequence (Figure 3.6). In addition to minor-groove width, the three local step parameters — Slide, Roll, and Twist — have a periodicity of about 10 bp (Figure 3.6). Nucleosomal DNA bends into its minor groove and major groove, alternatively, about every 10 bp. At the locations where DNA bends into the major groove (SHL ±5, ±4, ±3, ±2, ±1), Slide is negative, Roll is positive, Twist is reduced, and the minor-groove width is enlarged. When nucleosomal DNA bends into its minor groove (SHL ±5.5, ±4.5, ±3.5, ±2.5, ±1.5, ±0.5), and Slide is positive, Roll is negative, Twist is increased, and the minor-groove width is reduced. In addition, the DNA atoms contacted by protein are clustered at the sites along the minor grooves into which DNA bends.

The step parameters in the local helical frames of SIN-mutant and the wild-type nucleosomal DNA are also studied. The values of the *X*-displacement, *Y*-displacement, Inclination and helical-Twist are mapped on each nucleotide along the same set of DNA molecules (Figure 3.7). The local helical parameters show a 10 bp periodicity and are coupled with each other. The values of *X*-displacement, *Y*-displacement, Inclination, and Tip in a helical frame are closely related to Slide, Shift, Roll, and Tilt in a local frame, respectively. Like the values of Shift and Tilt, *Y*-displacement and Tip do not have obvious periodicity (data not shown), while the periodicity of *X*-displacement and Inclination is even stronger than that of Slide and Roll (Figure 3.7). The periodicity of *X*-displacement in nucleosomal DNA structure is even stronger than that of Slide. These trends are a direct result of the algorithms used to define parameters [2]. At the locations

where DNA bends into the minor groove (SHL ±1.5, ±2.5, ±3.5, ±4.5), the local helical step parameters show the coupling of increased helical Twist with increased *X*-displacement and reduced Inclination.

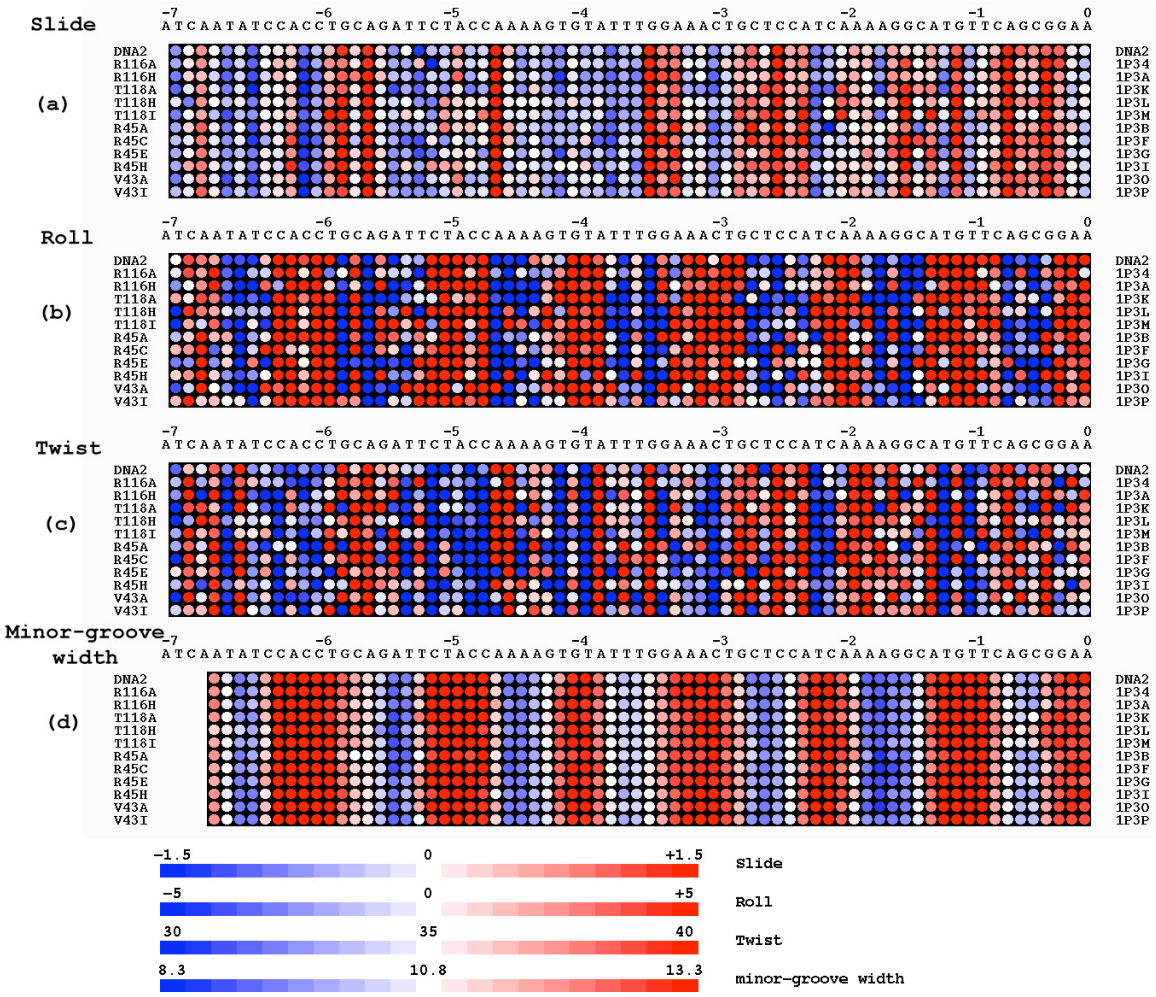



Figure 3.6. The local step parameters and minor-groove width of SIN-mutants and the wild-type nucleosome (DNA2) are mapped on each base-pair step along the bound DNA, from base pair 1 to 73. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The values of (a) Slide, (b) Roll, (c) Twist, and (d) minor-groove width are shown in shades of blue (negative or small values), white (zero or medium values), and red (positive or large values). The ranges of Slide, Roll, Twist and minor-groove width are –1.5 Å to +1.5 Å, –5° to +5°, 30° to 40°, 8.3 Å to 13.3 Å, respectively.

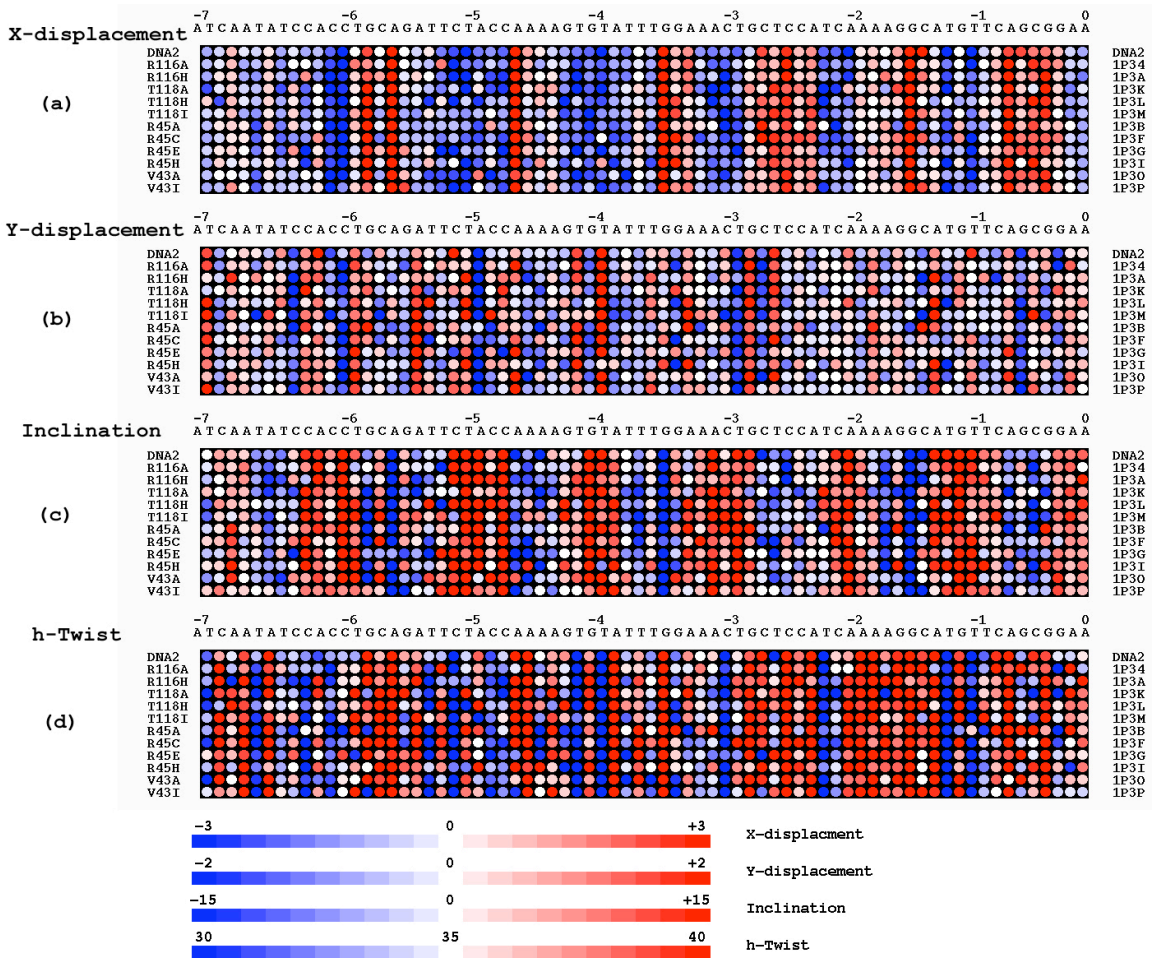



Figure 3.7. The local helical parameters of SIN-mutants and the wild-type nucleosomal DNA (DNA2) are mapped on each base-pair step along the first half of the nucleosomal DNA from base pair 1 to 73. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The values of (a) *X*-displacement, (b) *Y*-displacement, (c) Inclination, and (d) helical-Twist are shown in shades of blue (negative or small values), white (zero or medium values), and red (positive or large values). The ranges of *X*-displacement, *Y*-displacement, Inclination, and helical-Twist are –3 Å to +3 Å, –2 ˚ to +2 ˚, -15˚ to +15˚, and –30˚ to 40˚, respectively.

### 3.3.4 The stretching of the short half of nucleosomal DNA

The DNA of the wild-type and SIN-mutant nucleosome structures is divided into a 72 bp short half and a 73 bp long half. The 72 bp short half of nucleosomal DNA shows

localized stretching in the vicinity of SHL −2 with larger values of Twist compared to those in the longer half near SHL +2 (Figure 3.8). This phenomena was first observed in the first X-ray structure of the nucleosome core particle [12], but it is clearly shown in the color-coded mapping of Twist values in the DNA of the wild-type and SIN-mutant nucleosomal structurues.

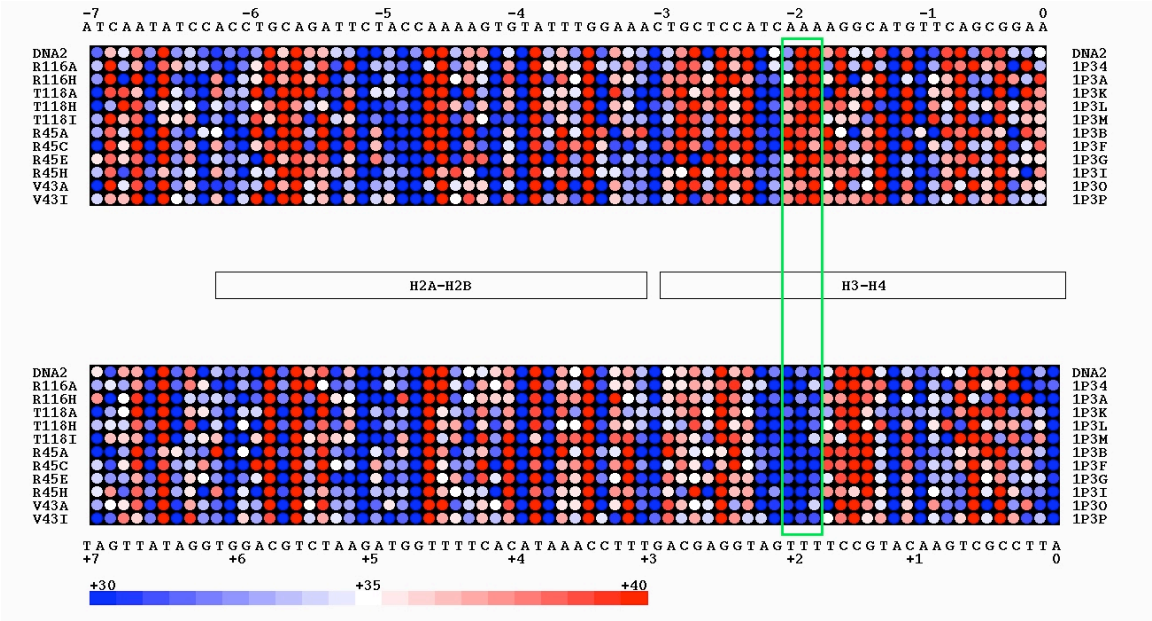



Figure 3.8. The values of Twist of the wild-type (ID: DNA2) and SIN mutant nucleosome structures are mapped on each base-pair step along the bound DNA. Base-pair steps 1 to 72 are shown from left to right in the upper part of the figure. To emphasize the pseudo-symmetry of the nucleosome structures, base-pair steps 73 to 145 are arranged from right to left in the lower part of the figure. The green rectangle shows the over-twisted region around SHL −2 compared with SHL +2. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The color scheme is shown with minimal values of 30˚ or less in dark blue, middle values near 35 ˚ in white, and maximal values of 40˚ or more in dark red.

### 3.3.5 Difference in step parameters of SIN mutant nucleosomes compared to the wild type (DNA2).

The loss of contacts in the mutant structures, which will be mentioned in Chapter 4, is accompanied by notable local DNA conformational changes compared to the wild type structure. For example, the Slide of dimer step 19 (AT/AT) near SHL –5.5 decreases from 0.06 Å in the wild type to an average value of –0.43 Å in the mutants (Figure 3.9), in concert with the loss of contacts of protein to the DNA phosphate atoms at the same location.

Dimer step 25 (CC/GG) near SHL –4.8 shows increased Tilt and decreased Twist in the mutant nucleosomes (4.2˚ Tilt in the wild type vs. an average value of 13.0˚ in the mutants and 32.3˚ Twist in the wild type vs. 26.5˚ Twist averaged over all the mutants) (Figure 3.9). In the adjacent dimer step 26 (CA/TG), Roll also changes from –18.3˚ in the wild type to –6.1˚ on average in the mutants and Shift changes from –0.45 Å in the wild type to –1.11 Å in the mutant. The phosphate atoms in base pair 25 (C/G) of DNA strand II concomitantly lose many contacts to histone proteins (Figure 4.3).

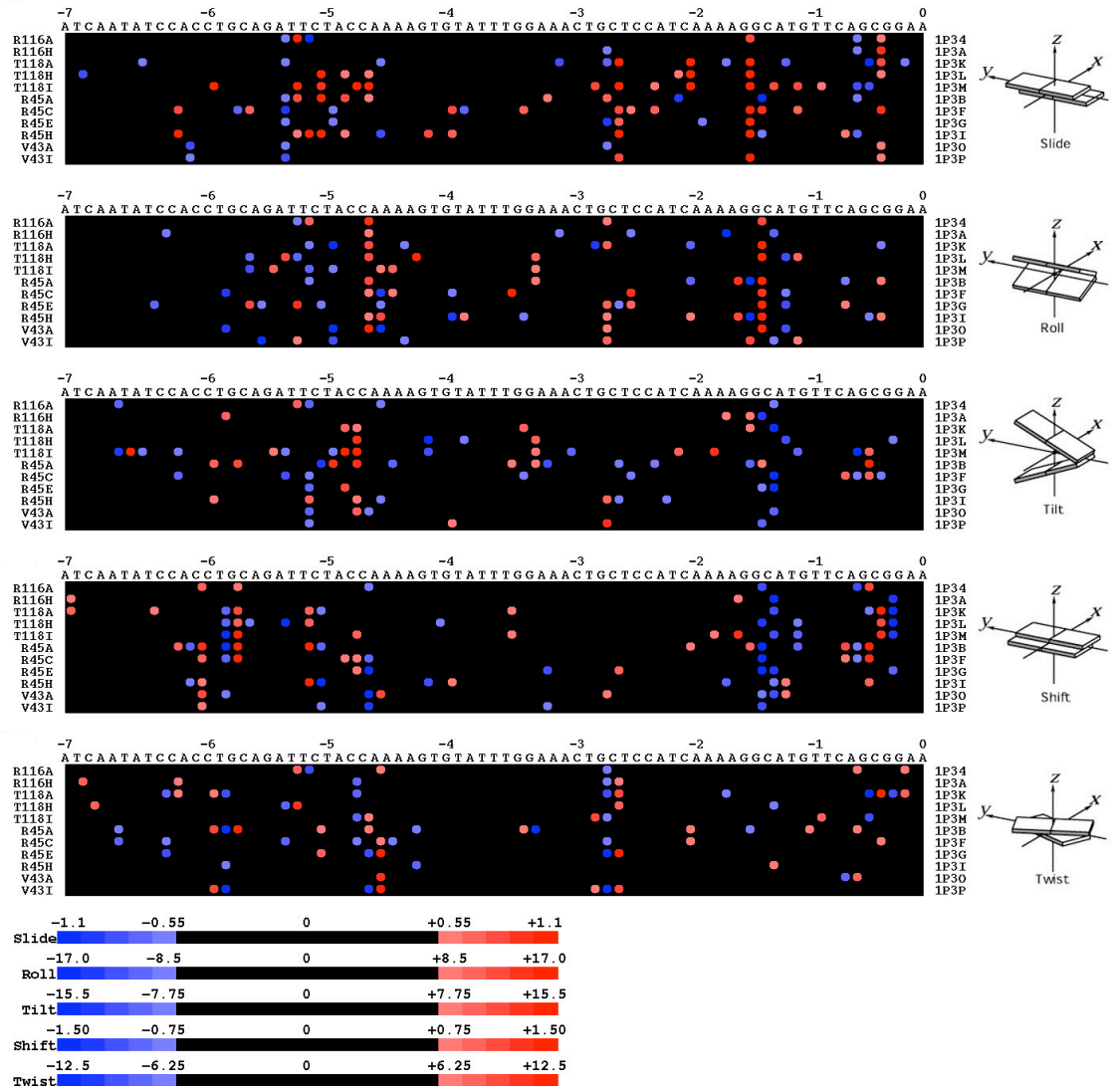



Figure 3.9. The difference in step parameters of SIN mutant compared to the wild type (DNA2), are mapped on each base-pair step along the nucleosomal DNA from base pair 1 to 73. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The difference in values of the mutant vs. the values of wild type is shown in shades of blue (negative) and red (positive). Differences in Slide, Roll, Tilt, Shift, and Twist between –0.5 Å to 0.5 Å, –7.4˚ to 7.4˚, –6.6˚ to 6.6˚, –0.66 Å to 0.66 Å and –5.5˚ to 5.5˚, respectively, are shown in black.

The contact losses near SHL ±5.5 and –4.5 and the nearby DNA conformational changes

(at SHL –5.5 and –4.5) suggest how the protein induces changes in local DNA stretch.

Other differences in the mutant nucleosomes occur at dimer step 46 (GC/GC) between SHL –3 to –2.5, where both Slide and Twist are decreased (Slide from 1.06 Å in the wild type to 0.78 Å in the mutants, Twist from 46˚ in the wild type to 37.2˚ in the mutants) (Figure 3.9). Roll is also increased from –11.7˚ in the wild type to –3.0˚ in the mutants at dimer step 46. Dimer step 47 has increased Slide 0.76 Å and decreased Twist (37.9˚). The value of Slide at dimer step 47 is near zero in the wild type. Sugar atoms of DNA located three nucleotides away in strand I at SHL –3, lose many contacts to protein (Figure 4.4). It is notable that the contact loss and DNA conformational changes near SHL ± 3 occur near the dimerization interface of H3H4 and H2AH2B. The dimerization of H3H4 and H2AH2B may thus be disrupted during nucleosome remodeling [13].

Dimer steps 58 (GG/CC) and 59 (GC/GC) at SHL –1.5 also undergo local conformational distortions (Figure 3.9) accompanied by the loss contacts from phosphate atoms at base pair 60 and from sugar atoms at base pair 59 (Figures 4.3 and 4.4). This result is consistent with the values of root-mean-square displacements reported below. Interestingly, the values of DNA conformational parameters in the mutants are quite similar to those of the wild type in the longer half of nucleosomal DNA (Figure 3.10).

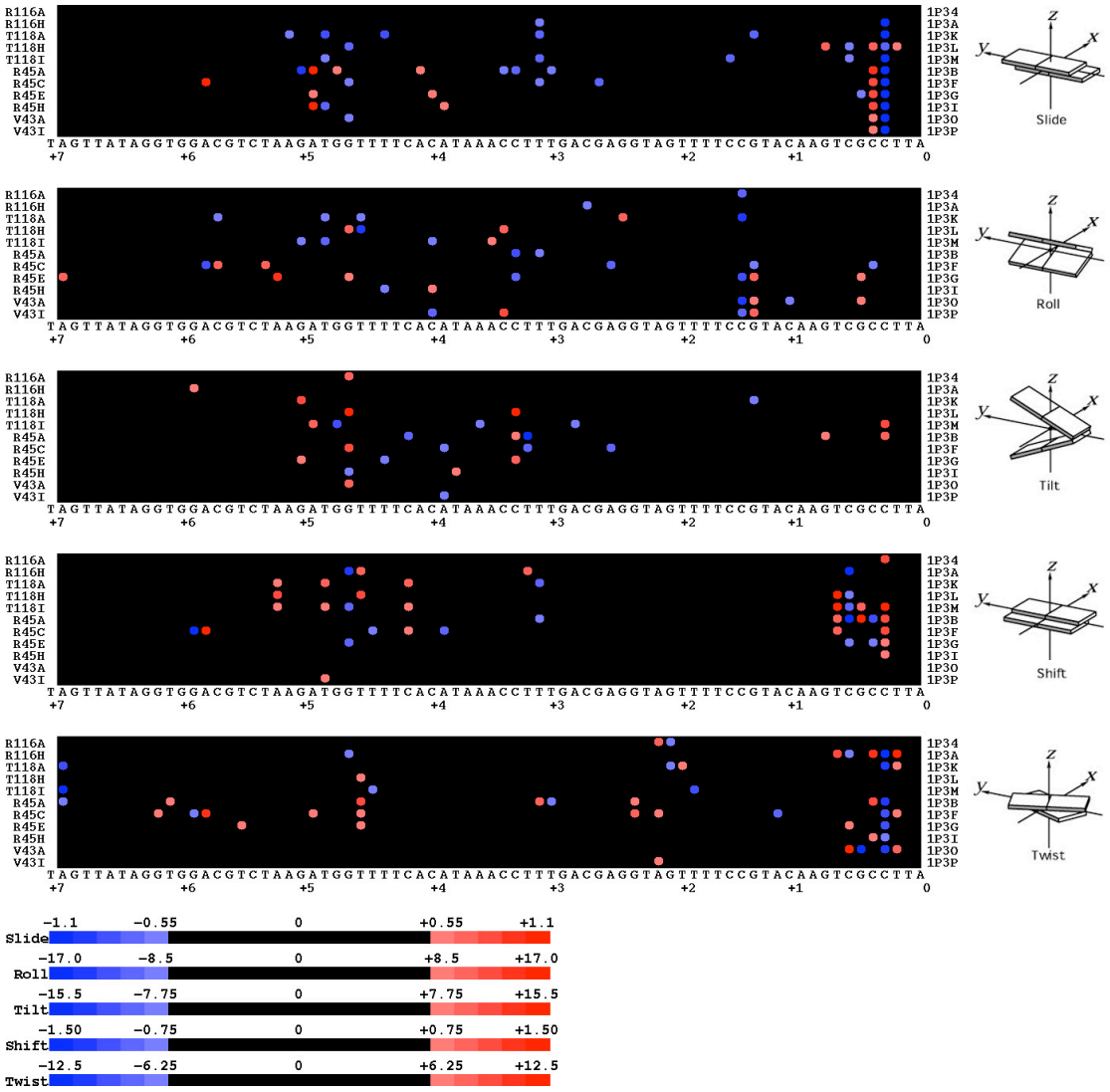



Figure 3.10. The difference in step parameters of SIN mutants compared to the wild type (DNA2), are mapped on each base-pair step along the nucleosomal DNA from base pair 73 to 146. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The difference in values of the mutant vs. the wild type is shown in shades of blue (negative) and red (positive). Differences in Slide, Roll, Tilt, Shift, and Twist between –0.5 Å to 0.5 Å, –7.4˚ to 7.4˚, –6.6˚ to 6.6˚, –0.66 Å to 0.66 Å, and –5.5˚ to 5.5˚, respectively, are shown in black.

3.3.6 Root-mean-square displacement (RMSD) of DNA and protein atoms in SIN-mutant nucleosomes compared with the wild type

The 146 bp DNA is divided into 14 fragments of about 10 bp according to superhelical location. The RMSD values of the DNA base atoms in each DNA fragment in the SIN mutation nucleosomes are compared with the values in the corresponding DNA fragment of the wild-type nucleosome (Figure 3.11 (a)). The base atoms are selected since they are most likely to reflect changes in base-pair step parameters. The RMSD values of all comparable protein atoms in the two copies of histones H3, H4, H2A and H2B, excluding the atoms in the mutated residues, are also calculated (Figure 3.11 (b)).

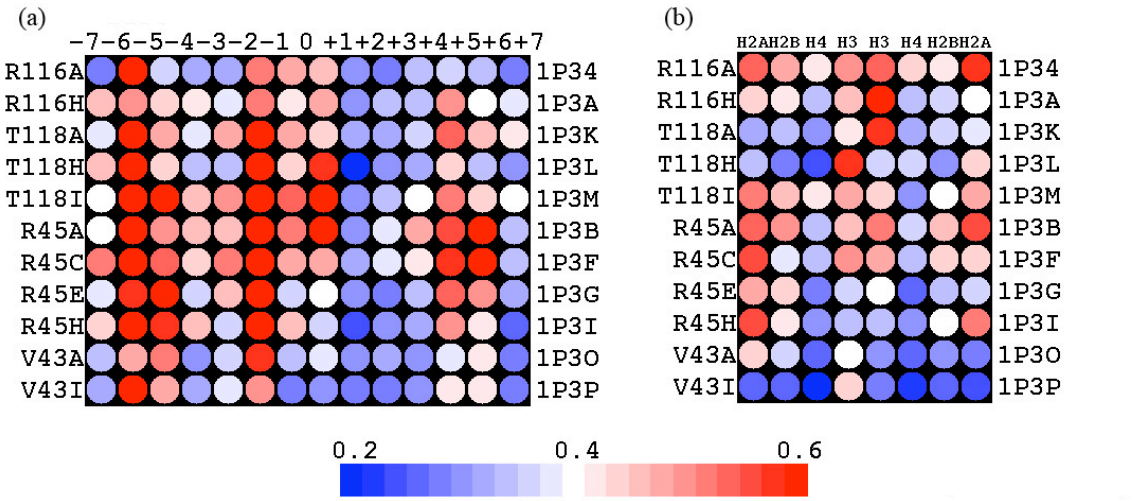



Figure 3.11. (a) RMSD values of DNA base atoms, compared with those in the wild type, are mapped on each superhelical location along the nucleosomal DNA. (b) RMSD values of protein atoms are mapped on each histone subunit. The structure identifiers are shown at the left edge of the diagrams. Superhelical location numbers and histone subunit names are shown at the top of the diagrams in (a) and (b), respectively.

The RMSD values of DNA base atoms are asymmetrical due to the asymmetry of the nucloesomal DNA with respect to the dyad axis. Large RMSD values occur predominantly at SHL −6 to −5, which is consistent with the loss of contacts between

histone and DNA and the DNA conformational changes found at those locations. Correspondingly, RMSD values of protein atoms in histone H2A, which bindsat DNA SHL –6 to –5, are relatively large, suggesting that histone H2A may affect DNA conformation through local protein-DNA interactions.

The RMSD values of all DNA atoms are also examined (Figure 3.12). The RMSD values are large at DNA superhelical positions from SHL −1 to +1, where the conformational change of DNA may be caused by the amino-acid point mutations (located at SHL ±0.5). The RMSD values of protein atoms in histones H3 and H4 bearing the amino-acid point mutations, are also large in this region (data not shown). Notably, the T118I H3, R45A H4, R45H H4 SIN mutants, with the exceptionally large RMSD values of DNA from SHL −6 to −5, have high sliding rates in the heat-shifting experiments discussed in Chapter 2. The templates from these three SIN mutants also have high threading scores compared to the other SIN mutants and the wild type (discussed in Chapter 6). The superhelical radius of base pair 22 at SHL −5 also decreases by 2.34 Å in the R45H H4 SIN mutant compared to the wild type [14].

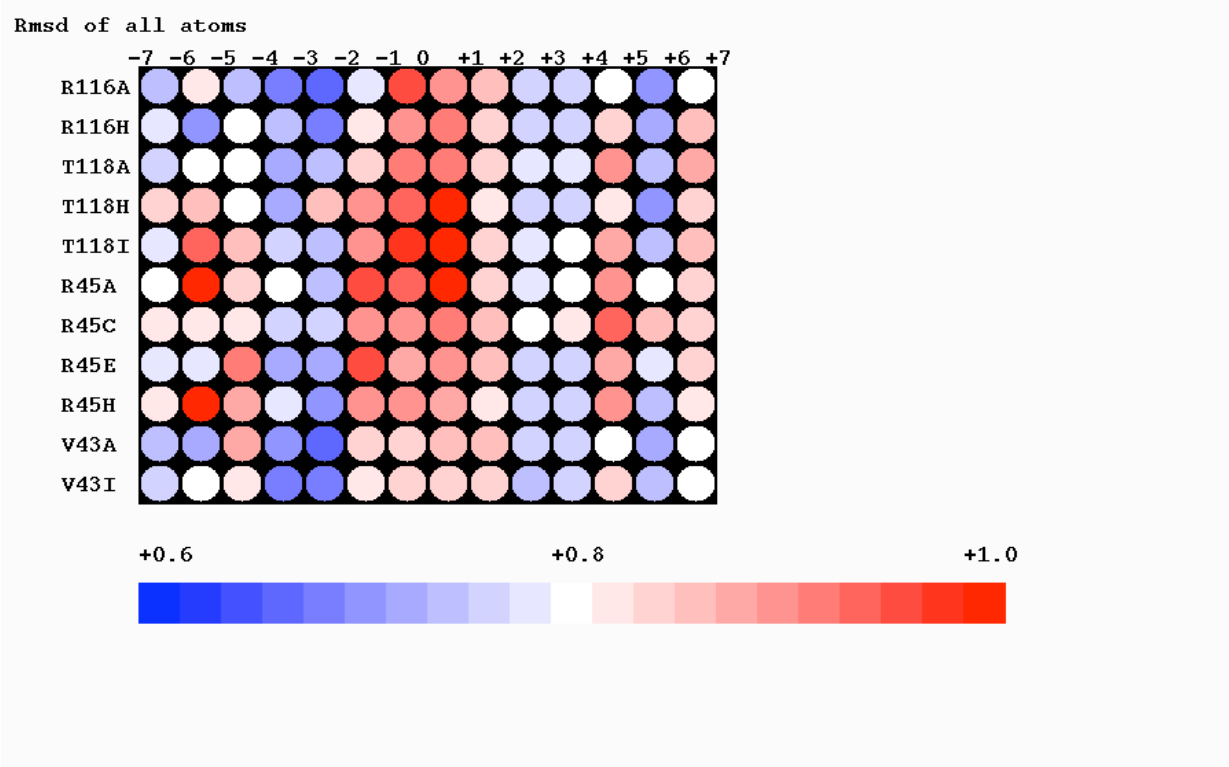



Figure 3.12. RMSD values of all DNA atoms in each superhelical location, compared with that in the wild type, are mapped on each superhelical location along the NCP DNA. The structure IDs are shown at the edge of the diagram.

### 3.3.7 B-factors of DNA atoms and histone protein atoms in the wild-type and SIN-mutant nucleosomes.

The mean values of the B-factors of DNA and protein atoms at different sites on the nucleosomes are compared in Figure 3.13. In general, the B-factors of the protein atoms are much smaller than those of the DNA atoms (41 Å on average for protein atoms vs. 88 Å on average for DNA atoms). The B-factors of the protein atoms also spread over a smaller range than those of the DNA atoms (the standard deviation is 7 Å for protein atoms vs. 20 Å for DNA atoms). The conformational changes in DNA seen at SHL –6 to –5 in the SIN-mutant structures are accompanied by high values of the B-factor in that region (Figure 3.13 (a)). On the other hand, histone H2A with low B-factors in the

vicinity of DNA at SHL –6 to –5 also shows large RMSD values in the SIN-mutant structures compared with the wild type structure (Figure 3.13 (b)). This suggests that the observed DNA conformational changes at SHL –6 to –5 may at that site not be related entirely to the uncertainty (or intrinsic deformability) of the crystal structures, but at least partially to the protein-DNA interactions with histone H2A.

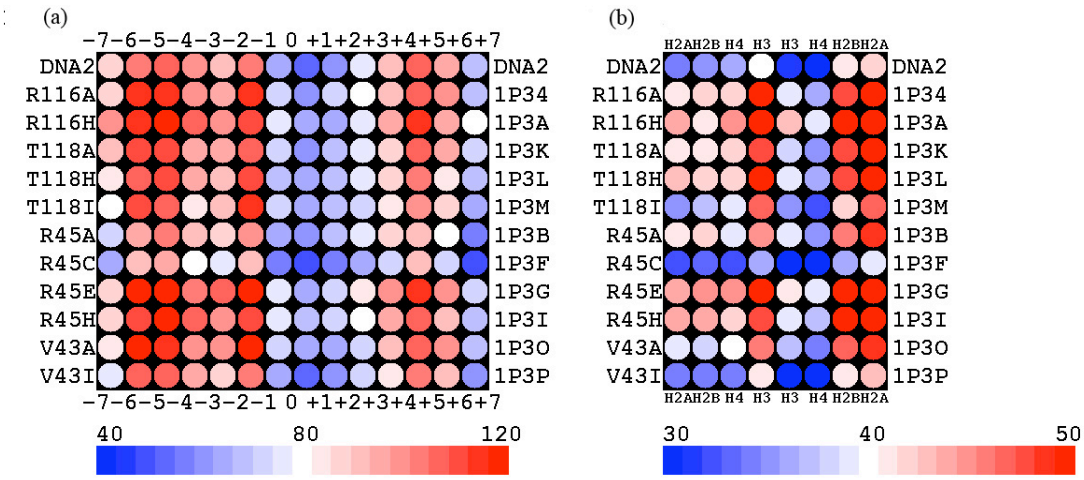



Figure 3.13. (a) The average B-factors of DNA base atoms are mapped on each superhelical location along nucleosomal DNA. (b) The average B-factors of protein atoms are mapped on each histone subunit. The structure identifiers are shown at the left edges of the diagrams. Superhelical location numbers and histone subunit names are shown at the top and bottom of the diagram in (a) and (b), respectively.

## 3.4 Summary and discussion

Two visualization methods are introduced in this chapter to show the structural characteristics of nucleosomal DNA clearly. At SHL ±1.5, ±2.5, ±3.5, ±4.5, DNA bends into the minor groove with accompanying occurrences of positive Slide, negative Roll, increased Twist, and small minor-groove width. And at the same time, the protein-DNA contacts cluster at SHL ±1.5, ±2.5, ±3.5, ±4.5. At SHL ±1, ±2, ±3, ±4, ±5, DNA bends

into the major groove with accompanying occurrence of negative Slide, positive Roll, and decreased Twist. The local helical parameters show the similar trends. That is, at the locations where DNA bends into minor groove (SHL ±1.5, ±2.5, ±3.5, ±4.5), there is a coupling of increased helical Twist with increased *X*-displacement and reduced Inclination. For nucleosomal DNA, there is a 10-bp periodicity in the values of Slide, Roll and Twist in the local frame and in *X*-displacement, Inclination, and helical Twist in the helical frame. However, Shift and Tilt, and closely related values of *Y*-displacement and Tip in a helical frame, do not show any obvious periodicity.

In order to match the 73-bp half of nucleosomal DNA, the short 72 bp half is locally stretched with the over-twisted region found at SHL −2. The differences between the 72 bp half and the 73 half, which is described in Section 5.1and in Chapter 5 is related to the asymmetry of the crystal structures.

The values of step parameters of the SIN-mutant structures are compared to the wild-type. DNA far from the sites of SIN mutations undergoes appreciable conformational change in the solved crystal structures. These changes occur where the DNA is bound by histone subunits H2A, H2B. The site-specific mutation occurs in histone H3 or H4. The DNA conformational change toward the end of DNA is accompanied by the loss of contacts in the same region. The changes in contacts will be discussed in Chapter 4.

As is well know, the H2AH2B dimer stacks with histone subunits H3, H4 in the vicinity of SHL 0.5 (Figure 2.2 (b)). That is to say, the SIN mutations in the vicinity of DNA at SHL 0.5 may transmit information to DNA at SHL –5.5 to - 4.5 through allosteric interaction between histones H3H4 and histones H2AH2B. This suggests how SIN mutations may help the ends of DNA peel off histone proteins more easily.

Another change in the SIN-mutant structures occurs at SHL ±3, locations near the dimerization interface of histones H3H4 and H2AH2B. Many contacts of protein to DNA are lost at SHL ±3, as will be mentioned in Chapter 4, and obvious DNA conformational changes occur at SHL –3. This suggests that the SIN mutations may also disrupt the dimerization of histones H3H4 and H2aH2b so as to affect the mobilization of nucleosomes. The later has also been proposed by Flaus, *et al.* [13] on the basis of observations that histones are removed or exchanged during the nucleosome remodeling [15, 16].

The root-mean-square displacement of DNA atoms in the SIN-mutant structures compared to the wild type is small compared to the resolutions of X-ray crystal structures. The resolution of the SIN-mutant structures ranges between 2.4 and 3.0 Å [5], but the extreme displacements noted here are greater than the coordinate errors, which vary from 0.34 to 0.37 Å. Hence, the B-factors of histone proteins and nucleosomal DNA are studied. In the SIN-mutant structures, histone-H2A contacts to DNA in the vicinity of SHL −5.5 show high RMSD values and low B-factors. This suggests that the difference in histone H2A may cause the DNA conformational change at SHL −6 to −5.

1. Dickerson, R.E., *Definitions and nomenclature of nucleic acid structure parameters*. Nucleic Acids Research, 1989. 17: p. 1797-1803.
2. Lu, X.J., and Olson, W. K., *3DNA: a software package for the analysis rebuilding and visualization of three-dimensional nucleic acid structures*. Nucleic Acids Research, 2003. 31: p. 5108-5121.
3. Olson, K.W., Bansal, M., Burley, S. K., Deckerson, R. E., Gerstein, M., Harvey, S. C., Heinemann, U., Lu, X. J., Neidle, S., Shakked, Z., Sklenar, H., Suzuki, M., Tung, C.-S., Wethof, E., Wolgerger, C., and H. M. Berman, *A standard reference frame for the description of nucleic acid base-pair geometry*. J. Mol. Biol., 2001. 313: p. 229-237.

4. El Hassan, M.A., Calldine, C. R. , *Structure and conformation of helical nucleic acids: analysis program (SCHNAaP)*. J. Mol. Biol., 1995. 273: p. 648-668.
5. Blundell, T.L., and Johnson, L. N., *Protein crystallography*. 1976, London: Academic Press Inc. 565.
6. Rossum, G.V., *Python reference manual*, ed. F.L.J. Drake. 2006.
7. Delano, W.L. *The pymol molecular graphic system*. 2002 Available from: http://www.pymol.org.
8. Horn, B.K.P., *Closed-form solution of absolute orientation using unit quaternions*. J.Opt. Soc. Am. A, 1987. 4: p. 629-642.
9. Suto, R.K., Edayathumangalam, R. S., White, C. L., Melander, C., Gottesfeld, J. M., Dervan, P. B., and Luger, K., *Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands*. J. Mol. Biol., 2002. 326: p. 317-380.
10. Richmond, T.J., adn Davey, A. C., *The structure of DNA in the nucleosome core*. Nature, 2003. 423: p. 145-150.
11. Tolstorukov, M.Y., Colasanti, A. V., McCandlish, D., Olson, W. K., and Zhurkin, V. B., *A novel 'roll-and-slide' mechanism of DNA folding in chromatin. Imaplication for nucleosome positioning*. 2006.
12. Luger, K., Maedar, W., Richmond, R. K., Sargent, D. F., and T. J. Richmond, *Crystal structure of the nucleosome core paricle at 2.8 Å resolution*. Nature, 1997. 389: p. 251-259.
13. Flaus, A., Owen-Hugers, T., *Mechanisms for ATP-dependent chromatin remodelling: farewell to the tuna-can octamer?* Current Opinion in Genetics & Development, 2004. 14(2): p. 165-173.
14. Colasanti, A.V., *Conformatioinal state of double helix DNA*, in *Genetics and Molecular Biology*. 2006, Rutgers, the state university of New Jersey: New Brusnwick.
15. Brono, M., Flaus, A., Stockdale, C., Rencurel, C., Ferreira, H., *Histone H2A/H2B dimer exchange driven by ATP-dependent chromatin remodeling activities*. Mol Cell 2003. 12: p. 1599-1606.
16. Mizuguchi, G., Shen, X., Landry, J., Wu, WH., Sen, S., Wu, C., *ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex*. Science, 2004. 303: p. 343-348.

# Chapter 4 The Contacts Loss of Histone Proteins with DNA in SIN-Mutants Nucleosome Structures

The contacts of protein with DNA are thought to induce the conformational changes seen in high resolution structure. Here we define that the contacts of protein with DNA in terms of the distances between a protein atom and a DNA atom falling within distances from 3.0 Å to 4.0 Å and examine the number and type of contacts. The number of the contacts of the histone protein to DNA atoms in nucleosome structures reflects how well histone protein interacts with nucleosomal DNA. Comparison of the differences in the number of the contacts in the SIN mutants to the wild type may help understand nucleosome remodeling mechanisms.

## 4.1 Method

The distances between atoms in a structure are calculated using an algorithm, which assigns atoms to a 5Å×5Å×5Å cubic grids and scans only atom pairs within each grid or in neighboring grids. The contacts of protein to DNA atoms within distances of 3.0 Å to 4.0 Å, at increments of 0.2 Å, in the wild-type and SIN-mutant structures were kindly provided by Dr.Yun Li.

## 4.2 Results

### 4.2.1 The total number of contacts of histone protein to nucleosomal DNA

The number of contacts between histone and phosphate atoms is about three times (186 contacts vs. 59 contacts at 3.4 Å or less) that between histone and sugar atoms in the wild-type structure (Table 4.1, row a). The contacts to DNA bases (data not shown) are limited (there are only four protein-DNA contacts of within a cutoff limit of 3.4 Å in the

wild type), a finding which is consistent with the fact that there are no base-specific interactions between the histone proteins and nucleosomal DNA. The relative numbers of contacts in different parts of the wild-type structure is as follows: phosphate-H3H4 > phosphate-H2AH2B> sugar-H3H4 > sugar-H2AH2B (Table 4.1, row a). There are fewer contacts in the SIN mutants than in the wild type (30 contacts or less on average). Although the point mutations occur in either histone H3 or H4, contacts between histone H2AH2B and DNA decrease considerably (Table 4.1, column b). In all of the H3 mutants as well as in the R45E and V43A H4 mutants, histones H2AH2B lose a large number of contacts to the DNA phosphate atoms. Although, the R45C and R45H H4 mutants do not lose so many contacts as the other mutants, they do show relatively large decreases in the number of contacts with histones H3 and H4 (Table 4.1, column c). These two mutants also stand out at the functional level in that they can relieve the requirement of the SWI/SNF complex in yeast in *vivo* [1]. Moreover, nucleosomal arrays reconstituted with the R45C H4 mutant completely lose the ability to undergo magnesium-dependent condensation [2].

4.2.3 Trends in contact losses between histone and DNA atoms as a function of the contact cutoff

The number of protein-DNA contacts in the wild-type nucleosome (DNA2) obtained with different values of the contact cutoff (3.0 Å and 4.0 Å) are summarized in Table 4.2. As expected, the total number of contacts increases as the value of the cutoff is increased. The contacts of histone protein to DNA phosphate atoms are always dominant in numbers (Table 4.2, column a plus column b). Although the cutoff ranges from 3.0 Å to

4.0 Å, the number of contacts between histone H3H4 and DNA phosphate atoms exceed those from other part of the structure.

Table 4.1. The number of contacts between atoms of the different histone subunits and different parts of DNA.

| Mutant protein | Structure ID | H2AH2B[b] | H3H4[c] | Phosphate H2AH2B[d] | Sugar H2aH2B[e] | Phosphate H3H4[f] | Sugar H3H4[g] |
|---|---|---|---|---|---|---|---|
| | DNA2 (WT)[a] | 111 | 138 | 86 | 25 | 100 | 34 |
| H3 | R116A | **−26** | −12 | **−21** | −5 | −2 | −9 |
| | R116H | **−26** | −14 | **−17** | −9 | −7 | −7 |
| | T118A | **−21** | −13 | **−14** | −7 | **−13** | 0 |
| | T118H | **−18** | **−26** | **−11** | −7 | **−22** | −3 |
| | T118I | **−32** | −10 | **−29** | −3 | **−12** | 3 |
| H4 | R45A | −3 | 3 | 1 | −4 | 11 | −8 |
| | R45C | −6 | **−16** | −2 | −4 | −6 | **−10** |
| | R45E | **−29** | **−15** | **−19** | **−10** | −1 | **−13** |
| | R45H | −2 | **−17** | −2 | −2 | −6 | −9 |
| | V43A | **−22** | **15** | **−18** | −4 | −3 | **−12** |
| | V43I | 12 | 0 | **−11** | −1 | 1 | −2 |

a. Numerical values in the highlighted row denote the number of contacts between protein and DNA backbone atoms within 3.4 Å or less in the wild-type structure. Numbers in the other rows represent the difference in the number of contacts of the proteins in the mutants compared to the wild type. Contacts are subdivided according to phosphate atoms and sugar atoms contacted by H2A and H2B (H2AH2B), H3 and H4 (H3H4). Values highlighted in red point to the largest loss of contacts (absolute values greater than 15 in the H2AH2B or H3H4 columns, column b and c and values greater than 10 on phosphate or sugar columns, column d to g).

Table 4.2. The number of the lost protein-DNA contacts (a) averaged over all the SIN-mutant nucleosomes, (b) in the R45H H4 SIN mutant (PDB ID: 1P3I), and (c) in the T118I H3 SIN mutant (PDB ID: 1P3M) with respect to the wild type one (DNA2) as a function of the cutoff limit between 3.0 Å to 4.0 Å.

(a)

| Cutoff (Å) | Total | Phosphate H2AH2B[a] | Sugar H2AH2B[b] | Phosphate H3H4[c] | Sugar H3H4[d] |
|---|---|---|---|---|---|
| 3.0 | -13.5 | -3.4 | -0.6 | -9.5 | 0.5 |
| 3.2 | -9.2 | -6.4 | -4.0 | 1.4 | 0.1 |
| 3.4 | -29.7 | -13.4 | -5.0 | -5.6 | -5.5 |
| 3.6 | -71.4 | -21.3 | -9.9 | -21.8 | -16.7 |
| 3.8 | -122.9 | -35.7 | -22.2 | -40.9 | -21.8 |
| 4.0 | -142.1 | -42.1 | -29.5 | -40.1 | -24.2 |

Numbers in the rows represent the average difference in the number of contacts of DNA with proteins in all SIN mutants compared to the wild type. The number of lost contacts is subdivided into four columns (a – d), corresponding to contacts of phosphate and to sugar atoms to H2A and H2B (H2AH2B) in columns a and b, H3 and H4 (H3H4) in columns c and d.

(b)

| Cutoff (Å) | Total | Phosphate H2AH2B[a] | Sugar H2AH2B[b] | Phosphate H3H4[c] | Sugar H3H4[d] |
|---|---|---|---|---|---|
| 3.0 | -14 | -4 | 0 | -11 | 1 |
| 3.2 | 1 | -1 | -1 | 1 | 1 |
| 3.4 | -19 | -3 | -2 | -6 | -8 |
| 3.6 | -59 | -14 | -8 | -17 | -20 |
| 3.8 | -94 | -27 | -13 | -25 | -29 |
| 4.0 | -122 | -30 | -17 | -27 | -42 |

Numbers in the rows represent the differences in the numbers of contacts of the protein in the R45H H4 SIN mutant (PDB ID: 1P3I) compared to the wild type. See footnote above.

(c)

| Cutoff (Å) | Total | Phosphate H2AH2B[a] | Sugar H2AH2B[b] | Phosphate H3H4[c] | Sugar H3H4[d] |
|---|---|---|---|---|---|
| 3.0 | -26 | -11 | -2 | -14 | 1 |
| 3.2 | -21 | -16 | -4 | -8 | 7 |
| 3.4 | -41 | -29 | -3 | -12 | 4 |
| 3.6 | -80 | -42 | -2 | -29 | -5 |
| 3.8 | -158 | -58 | -24 | -64 | -9 |
| 4.0 | -202 | -76 | -35 | -75 | -8 |

Numbers in the rows represent the differences in the numbers of contacts of the protein in the T118I H4 SIN mutant (PDB ID: 1P3M) compared to the wild type. See footnote above.

The average numbers of lost contacts of SIN-mutant nucleosome structures compared to the wild type (DNA2) are also reported for different cutoff distances (Table 4.2 (a)). At shorter cutoff limits (3.2 Å and 3.4 Å), the DNA phosphate atoms lose more contacts to histones H2A and H2B (H2AH2B) than to histones H3 and H4 (H3H4). When the cutoff is 3.6 Å or larger, the number of lost contacts from DNA phosphate atoms to histones H2AH2B is similar to the number of contacts lost from histones H3H4. Although the contacts between DNA phosphate atoms and histones H3H4 constitute the majority of close contacts, histones H2AH2B lose a greater percentage of contacts compared with histones H3H4.

The number of lost contacts in the R45H H4 SIN mutant (PDB ID: 1P3I) for different cutoff distance is reported in Table 4.2 (b). Histones H3H4 in the R45H H4 SIN mutant lose eleven close contacts (within 3.0 Å) to phosphate atoms. These differences are compensated by contacts at larger distances. The number of the lost contacts between histones H3H4 and DNA phosphate atoms increases for cutoffs of 3.2 Å and 3.4 Å. The DNA phosphate atoms lose more close contacts to histones H3H4 than to histones H2AH2B (compare Table 4.2 (b) columns a and c), but lose comparable numbers of contacts to H2AH2B and histones H3H4 at greater distances (3.6, 3.8 and 4.0 Å). The number of lost contacts between histone H3H4 and DNA sugar atoms is more than doubled of the number contacts of histone H2AH2B lost from DNA sugar atoms (compare Table 4.2 (b) columns b and d). Histones H2AH2B lose more close contacts (within 3.0, 3.2, and 3.4 Å) from DNA phosphate and sugar atoms than histones H3H4 does. At the larger cutoff of distance (within 3.6 to 4.0 Å), the number of lost H2AH2B-DNA contact is quite comparable to number lost between histone H3H4 and DNA.

Compared to the R45H H4 SIN mutant, the T118I H3 SIN mutant (PDB ID: 1P3M) loses more contacts overall (compare Table 4.2 (b) and (c)). Except for contacts within 3.0 Å or 3.8 Å, the DNA phosphate atoms lose more contacts from histone H2A2B than from histone H3H4 (compare Table 4.2 (c), columns a and c). DNA sugar atoms lose fewer close contacts (within 3.0, or 3.2, or 3.4 Å) from both histone H2AH2B and histone H3H4 (compare Table 4.2 (c), columns b and d). DNA sugar atoms lose considerably more contacts at larger cutoff limits (3.8 and 4.0 Å) from histone H2AH2B than from histone H3H4. In summary, histones H2AH2B, in the R45H H4 SIN mutant, tend to lose more contacts than histones H2AH2B.

### 4.2.3 Mapping the numbers of contacts on every base pair of nucleosomal DNA

The contacts (within 3.4 Å) of protein to DNA phosphate atoms are mapped and aligned on each nucleotide of DNA in the wild-type and SIN-mutant structures (Figure 4.1). Although, the nucleotides, which contact the protein in SIN-mutant structures, are similar to those of the wild type, the number of contacts is smaller in the SIN mutants. The phosphate atoms contacted by protein in strand I are located around SHL −6.4, −5.3, −4.3, −3.3, −2.3, −1.3 and −0.5 in the first half of DNA, and around SHL +6.8, +5.7, +4.7, +3.7, +2.7, +1.7 and +1 to +0.7 in the second half. Similarly, the contacts of protein to phosphate atoms in strand II occur at SHL −6.7, −5.7, −4.7, −3.7, −2.7, −1.7, −1.1 to −0.8 and SHL +0.5, +1.3, +2.3, +3.3, +4.3, +5.3 and +6.3. That is, the sites of direct contact occur on either side of the sites of bending noted in Chapter 3. For example, at SHL +3.5, where DNA bends into its minor groove, the protein-DNA contacts involve phosphate atoms at SHL +3.7 on strand I and SHL +3.3 on strand II. From the viewpoint

of protein, the side chains of basic amino acids, such as arginine and lysine, insert into the DNA minor groove at 14 locations along DNA [3].

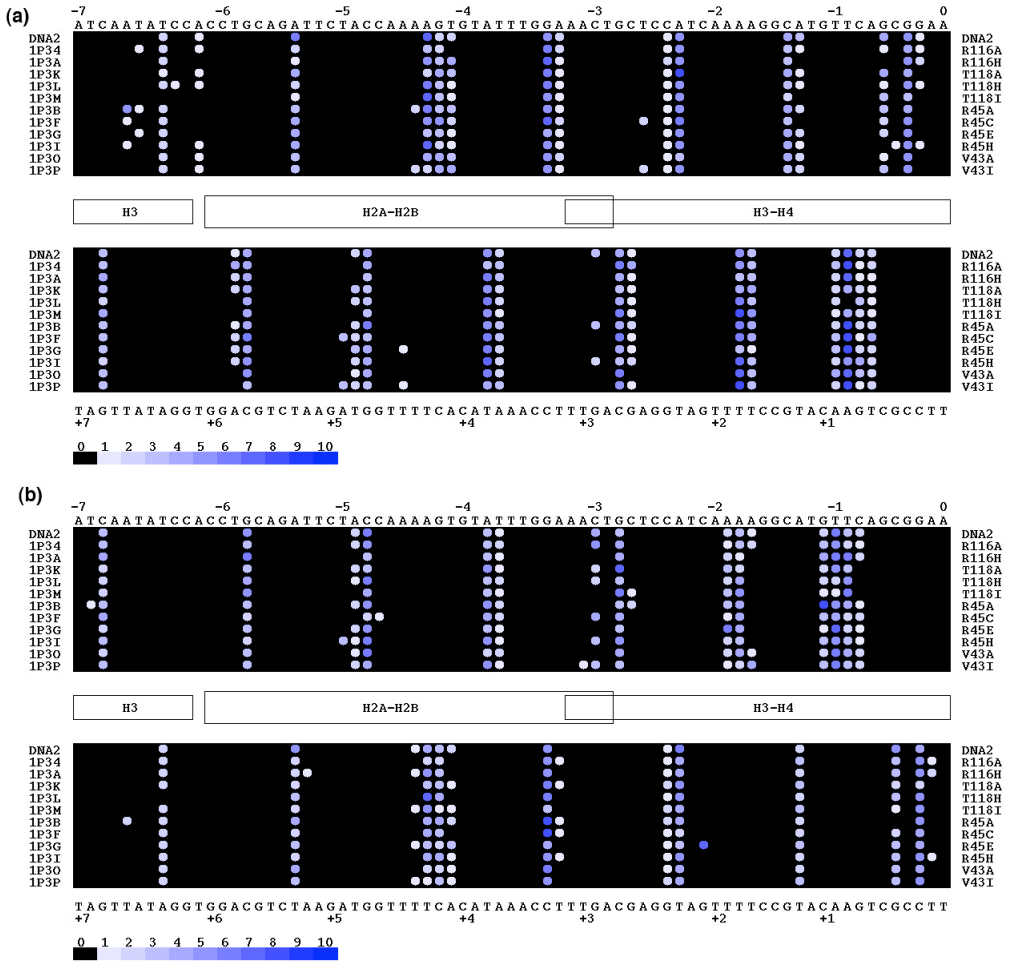



Figure 4.1. The numbers of close contacts to DNA sugar atoms in (a) DNA strand I, (b) DNA strand II are mapped on each nucleotide along nucleosomal DNA. The PDB IDs, except for the wild type (DNA2), are shown at the left edge of the diagram. The number of contacts from 1 to 10 is shown in different shades of blue.

Compared to the contacts of protein with DNA phosphate atoms, the number of contacts of protein to sugar atoms is less and the pattern of contacts is more irregular (compare (Figure 4.1 to Figure 4.2). Sugar contacts occur at SHL −6.5, −3.5, +2.3, −1.5, +0.5, +2, +3.7, +4.7, +5.7 on strand I and SHL −5.7, −4.7, −3.7, −2.8, −1.9, −1, +0.5, +2.3, +3.5,

+6.5 on strand II. That is, contacts toward the two ends of DNA, in the different strands: at SHL −6.5, and +5.7, histone protein contacts DNA sugar atoms primarily on strand I and at SHL −5.7 and −4.7, primarily on strand II. At the central positions of DNA — SHL −3.5, −2.3, −1.5, +0.5, +2, +3.7, +4.7 on strand I and SHL −3.7, −2.8, −1.9, −1, +0.5, +2.3, +3.5, +6.5 on strand II — histone protein contacts of DNA sugar atoms from both strands on either side of the same minor grooves.

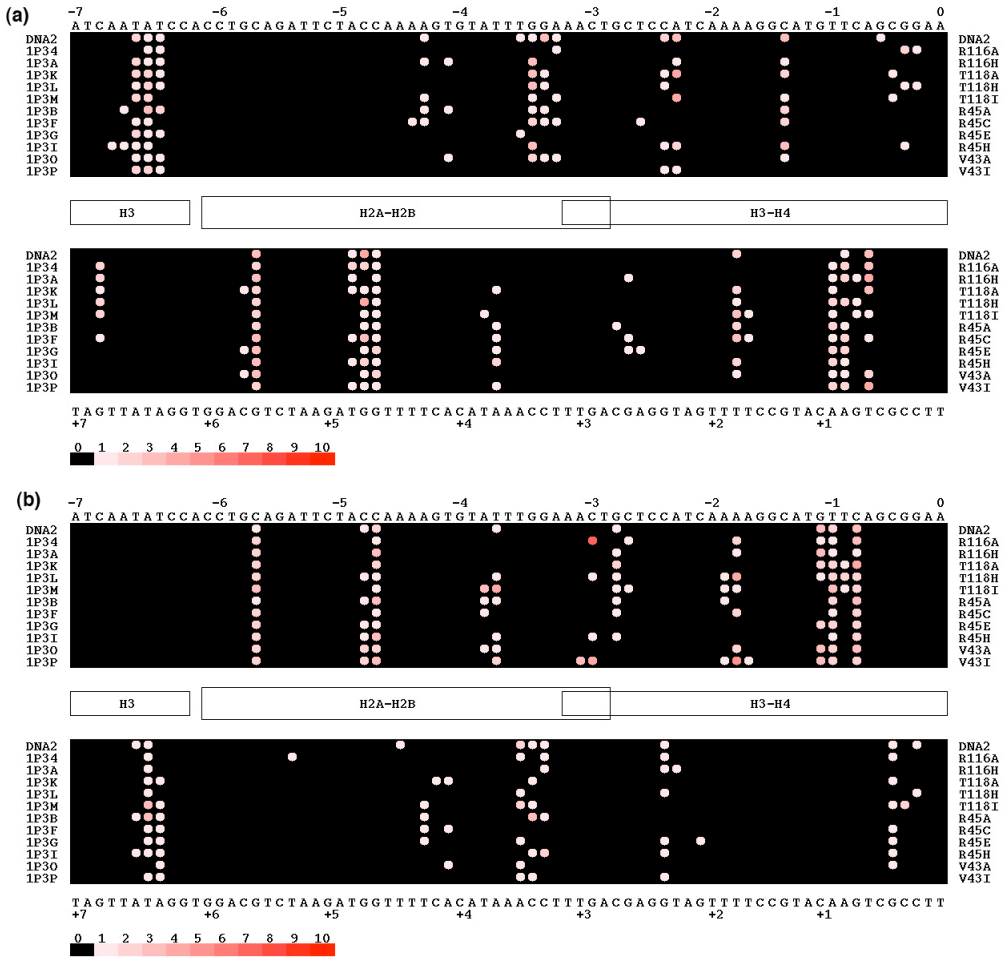



Figure 4.2. The numbers of close contacts to DNA sugar atoms in (a) DNA strand I, (b) DNA strand II are mapped on each nucleotide along nucleosomal DNA. The PDB IDs, except for the wild type (DNA2), are shown at the left edge of the diagram. The number of contacts from 1 to 10 is shown in different shades of blue.

4.2.4 Mapping the redistribution of contacts between histone protein and DNA-phosphate atoms at different distances in SIN-mutant structures compared to the wild type (DNA2).

The redistribution of contacts between protein and DNA phosphate atoms at cutoff limint of 3.0, 3.2, and 3.4 Å is reported in Figure 4.3, 4.4, and 4.5, respectively. The close contacts (of 3.0 Å or less) lost in the SIN mutants compared to the wild type occur at SHL +3 on strand I and SHL −3 on strand II, near the dimerization interface of histones H3H4 and H2AH2B (Figure 4.3). The loss of contacts at SHL −1, +2.3 and +3.3 on strand I at the 3.0 Å cutoff is compensated by an increase in the number of contacts at longer distances (3.2 Å). The blue dots designating the loss of contacts at the 3.0 Å cutoff in Figure 4.3 disappear in Figure 4.4, where the loss of contacts at the 3.2 Å cutoff is plotted. The difference in the loss of close contacts found for 3.2 Å and 3.0 Å occurs at SHL −4.3, −1.3 in DNA strand I, and SHL −5, +4.3 in DNA strand II. More contacts are lost from histones to DNA phosphate atoms for the cutoffs between 3.2 and 3.4 Å at SHL −5.5 and −0.5 on strand I and SHL −5.8, +0.5, and +5.3 on strand II (compare Figure 4.4 to Figure 4.5).

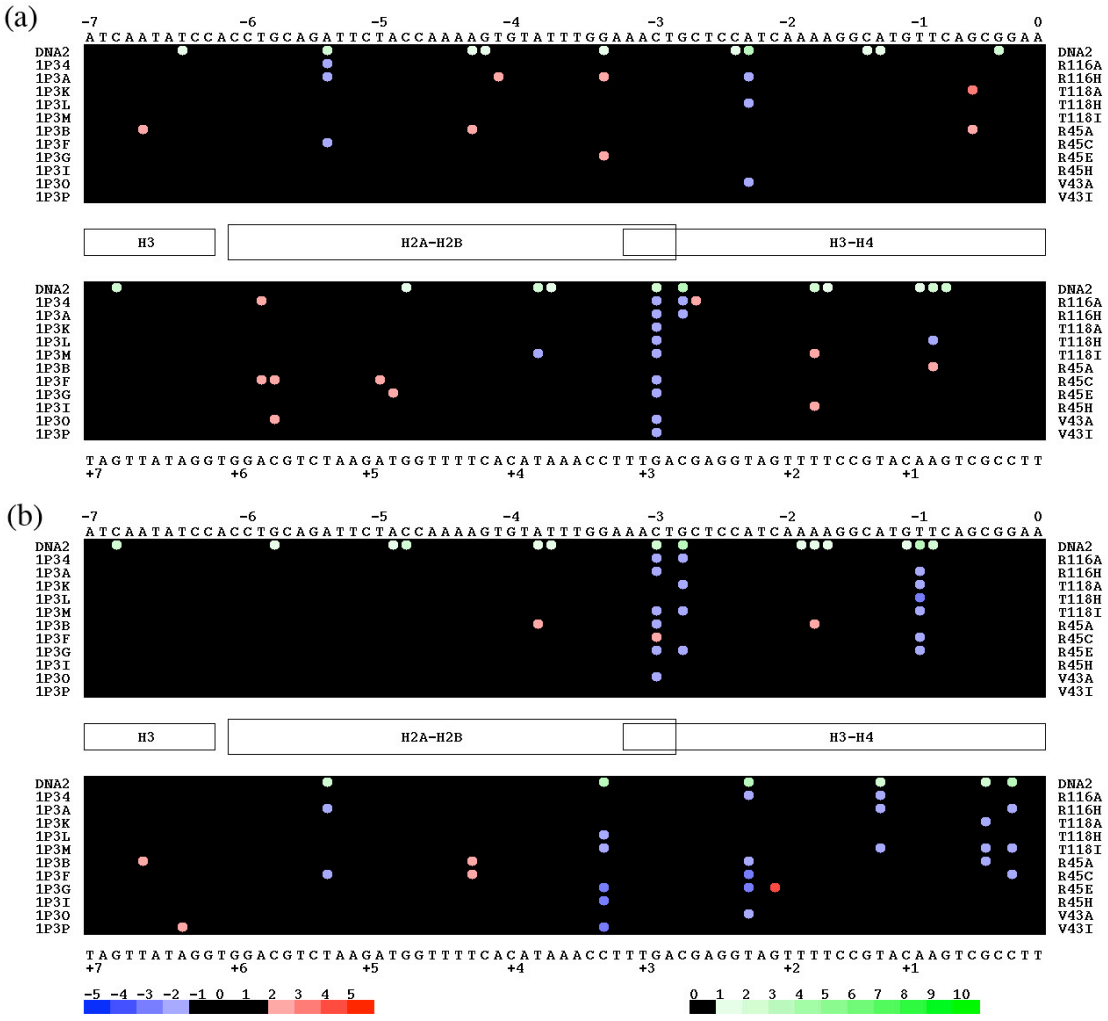



Figure 4.3. Difference in the numbers of close contacts within 3.0 Å to DNA phosphate atoms in (a) DNA strand I and (b) DNA strand II compared with the wild type. Data are mapped on each nucleotide along the nucleosomal DNA. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The mutation locations are shown at the right edge of the diagram. The number of contacts in the wild type is shown in different shades of green, and the difference in contacts to the mutants vs. the wild type in shades of blue (negative) and red (positive). Differences of –1, 0, +1 are shown in black.

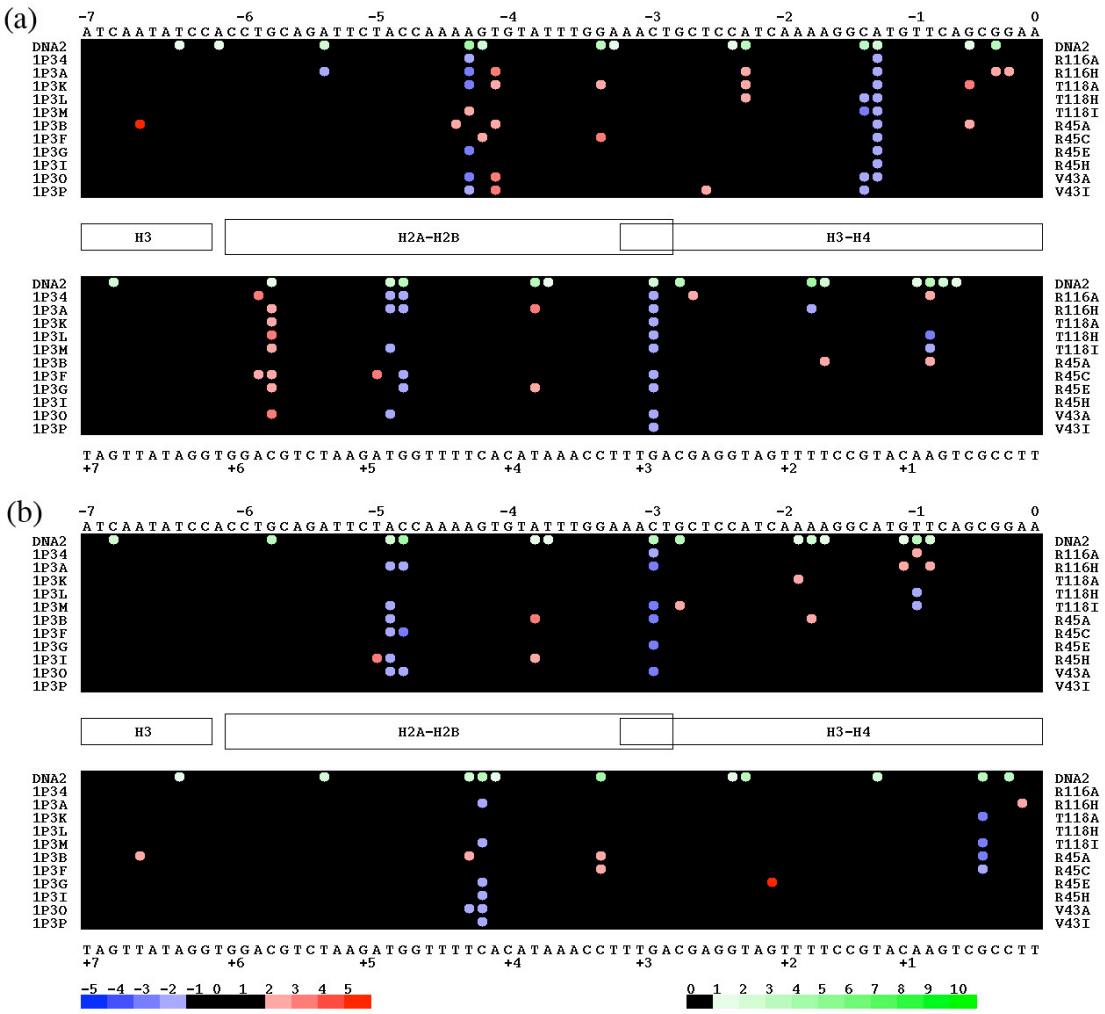



Figure 4.4. Difference in the numbers of close contacts within 3.2 Å to DNA phosphate atoms in (a) DNA strand I and (b) DNA strand II compared with the wild type. Data are mapped on each nucleotide along the nucleosomal DNA. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The mutation locations are shown at the right edge of the diagram. The number of contacts in the wild type is shown in different shades of green, and the difference in contacts to the mutants vs. the wild type in shades of blue (negative) and red (positive). Differences of –1, 0, +1 are shown in black.

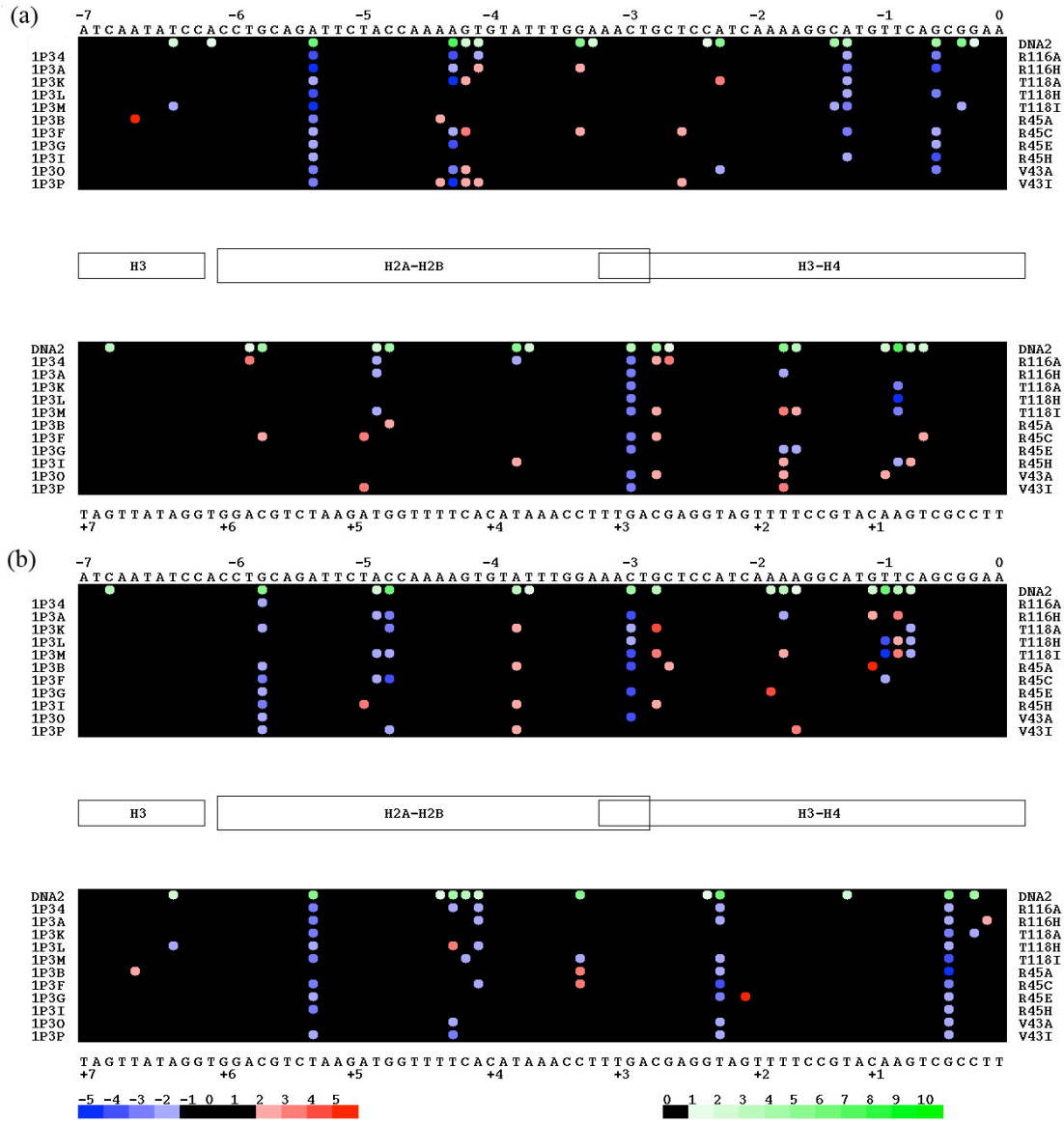



Figure 4.5. Difference in the numbers of close contacts within 3.4 Å to DNA phosphate atoms in (a) DNA strand I and (b) DNA strand II compared with the wild type. Data are mapped on each nucleotide along the nucleosomal DNA. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The mutation locations are shown at the right edge of the diagram. The number of contacts in the wild type is shown in different shades of green, and the difference in contacts to the mutants vs. the wild type in shades of blue (negative) and red (positive). Differences of –1, 0, +1 are shown in black.

Loss of contacts between protein and DNA phosphate atoms at the larger cutoff (3.6, 3.8 and 4.0 Å) are reported in Figure A4.1 to A4.3 in Appendix. The number of lost contacts

from the same locations increases with increased cutoff. As the cutoff increases, more locations along the nucleosomal DNA lose contacts, especially at SHL −6.5, −3.3, −2.3, +0.7, +1.7, +6.8 on strand II and SHL −6.8, −3.8, −1.7, +1.2, +5.3, +6.4 on DNA strand II.

4.2.5 Mapping the redistribution of contacts between histone protein and DNA sugar atoms in SIN mutants compared to the wild type.

The number of lost contacts of histone protein to DNA sugar atoms at different cutoff limits is reported in Figure 4.6 and 4.7, and Appendix Figure A4.3 to A4.7). Only one base pair, located at SHL −4.7 on DNA strand II, loses contacts at a distance less than 3.0 Å (Appendix Figure A4.4). In addition to the loss of contacts at SHL −4.7 on strand II, an other nucleotide, located at SHL +4.7, from strand I loses contacts within 3.2 Å (Figure 4.6 (b)). DNA sugar atoms at SHL −3.3, −2.3, −1.3, +0.6, and +1.9 on DNA stand I and SHL −1.1, +3.5 on DNA strand II lose more contacts within 3.4 Å (Figure 4.7).

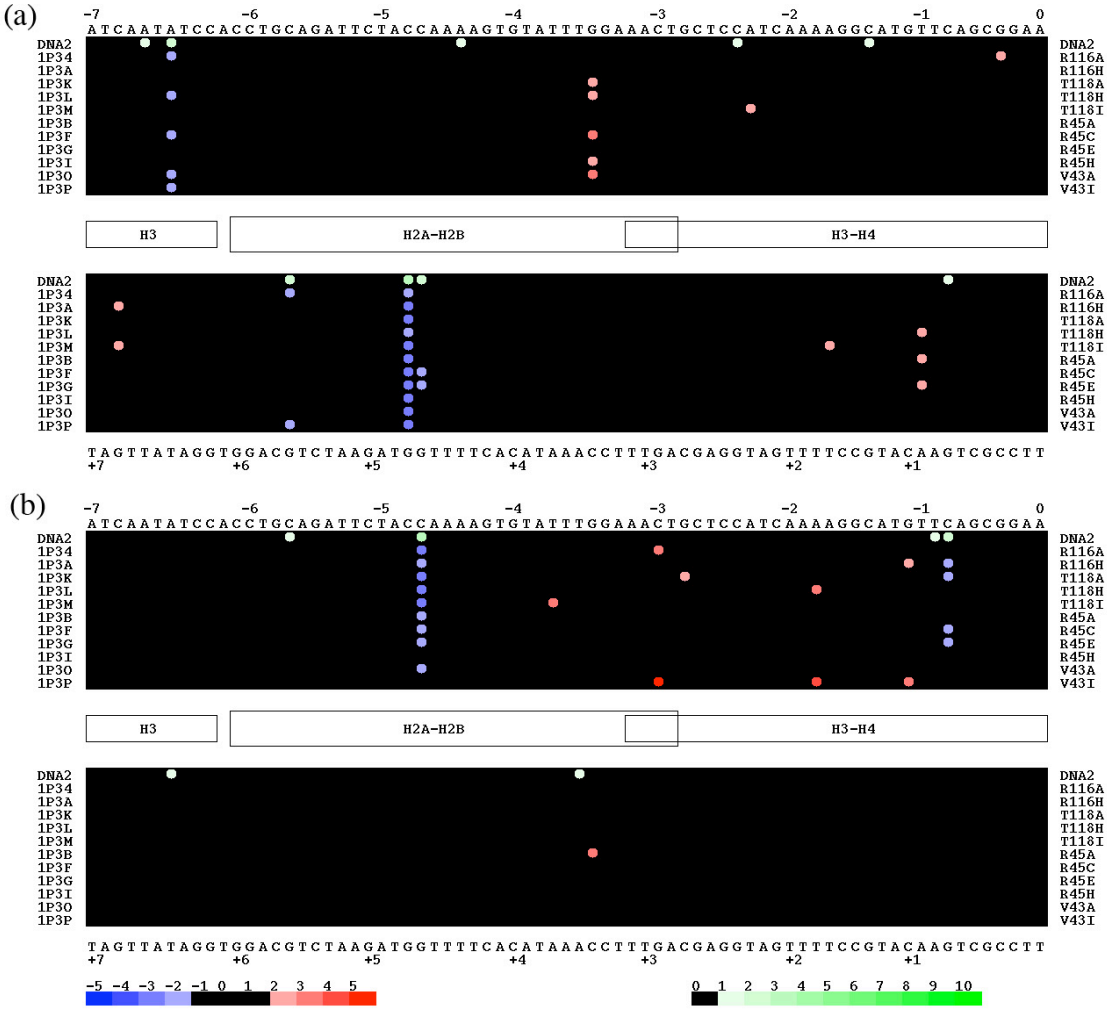



Figure 4.6. The difference in the numbers of close contacts within 3.2 Å to DNA sugar atoms in (a) DNA strand I, (b) DNA strand II compared with those of the wild type. Data are mapped on each nucleotide along the nucleosomal DNA. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The mutation locations are shown at the right edge of the diagram. The number of contacts in the wild type is shown in different shades of green, and the difference in contacts to the mutants vs. the wild type in shades of blue (negative) and red (positive). Differences of –1, 0, +1 are shown in black.

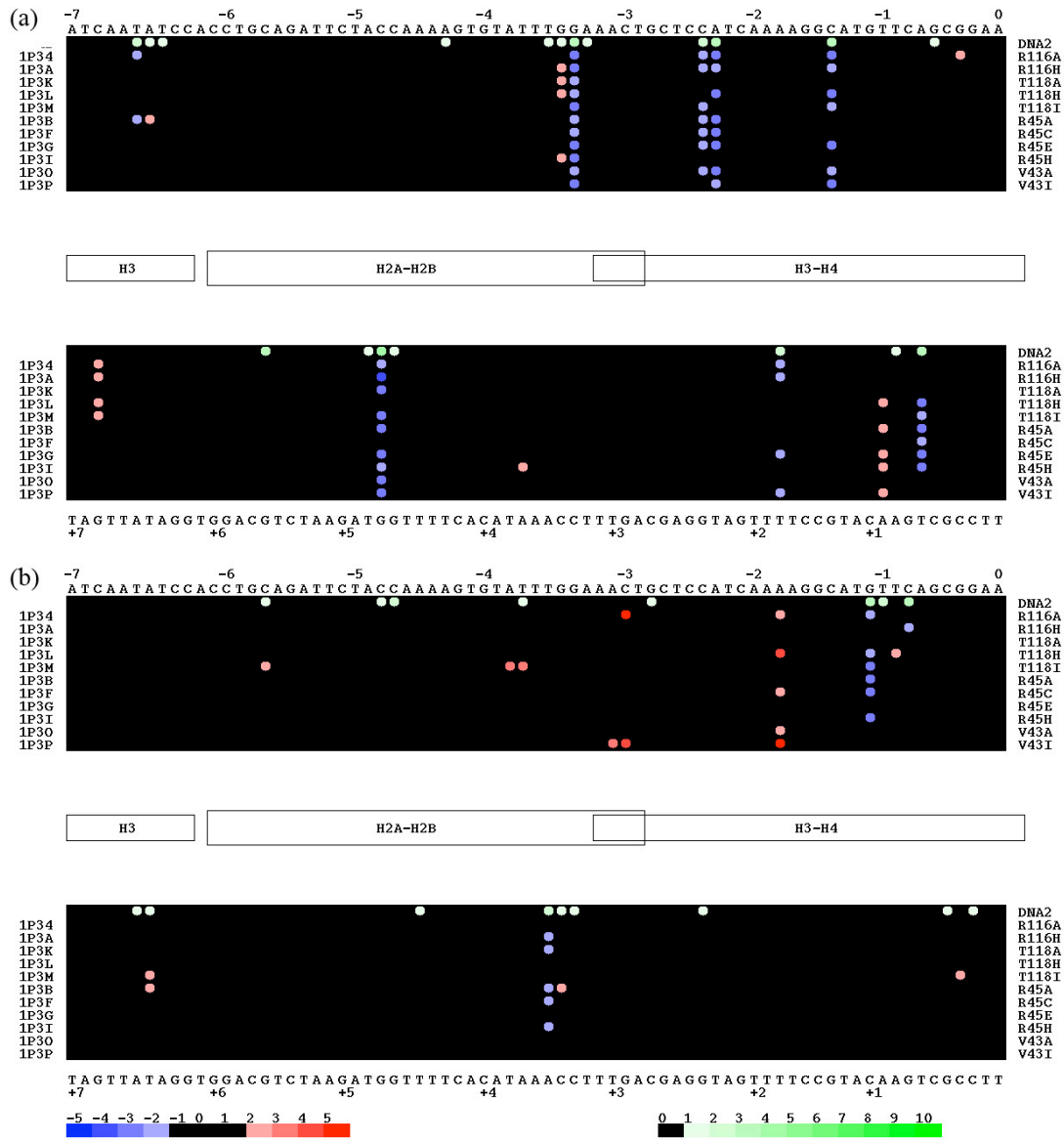



Figure 4.7. The difference in the numbers of close contacts within 3.4 Å to DNA sugar atoms in (a) DNA strand I, (b) DNA strand II compared with those of the wild type. Data are mapped on each nucleotide along the nucleosomal DNA. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The mutation locations are shown at the right edge of the diagram. The number of contacts in the wild type is shown in different shades of green, and the difference in contacts to the mutants vs. the wild type in shades of blue (negative) and red (positive). Differences of –1, 0, +1 are shown in black.

When the cutoff of contacts increases from 3.6 Å to 4.0 Å, other contacts are lost from the aforementioned locations, such as at SHL +4.7 from DNA strand I (Appendix Figure

A4.5 to A4.7). At the same time, more nucleotides lose contacts at other sites along nucleosomal DNA (marked by the increasing numbers of columns of blues dots in Appendix Figure A4.5 to A4.7. At the 4.0 Å cutoff, the loss of contact from histone protein to DNA sugar atoms is distributed throughout the nucleosomal DNA with losses occurring at the same locations where contacts to DNA phosphate atoms are lost (compare Appendix Figures A4.3 to A4.7).

### 4.2.6 Mapping the redistribution of contacts of arginine and lysine residues in histone protein in SIN mutants compared to the wild type

The loss of contacts of arginine and lysine to DNA phosphate atoms is plotted in Figure 4.8. The major contacts between arginine and lysine and DNA phosphate atoms are conserved in the SIN mutants compared to the wild type except at SHL and −4.3 and −0.5 on DNA strand I, and SHL −5.9 and +0.5 on DNA strand II. The losses at ±0.5 are close to the site of point mutations. The losses at more distant sites reflect a *trans*-effect of the SIN mutations to protein-DNA interactions toward one end of the nucleosomal DNA.

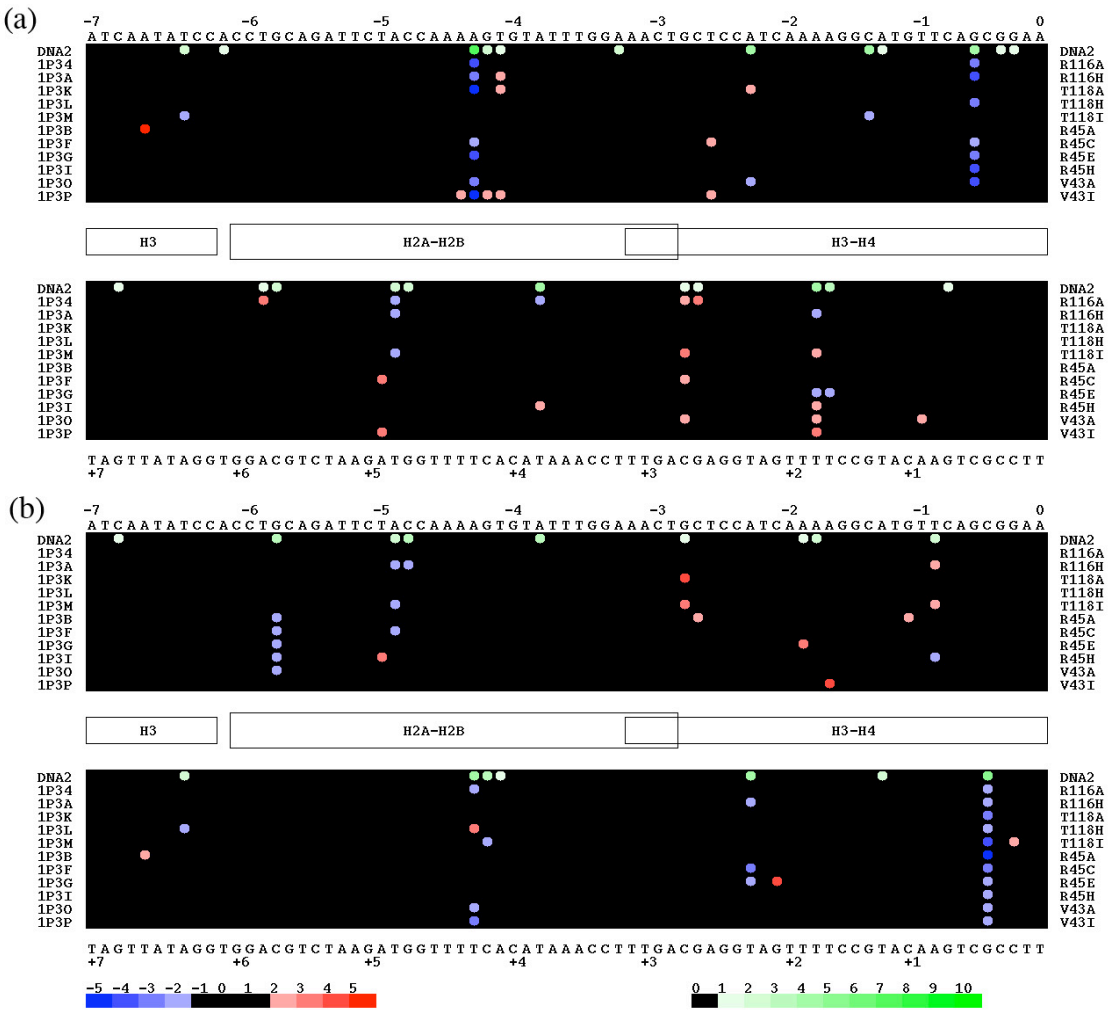



Figure 4.8. The difference in the numbers of close contacts within 3.4 Å of arginine and lysine residues to DNA phosphate atoms in (a) DNA strand I, (b) DNA strand II compared with those of the wild type. Data are mapped on each nucleotide along the nucleosomal DNA. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The mutation locations are shown at the right edge of the diagram. The number of contacts in the wild type is shown in different shades of green, and the difference in contacts to the mutants vs. the wild type in shades of blue (negative) and red (positive). Differences of –1, 0, +1 are shown in black.

Comparison of the number of lost contacts to DNA phosphate atoms from all protein atoms (Figure 4.8) with the number of lost from arginine and lysine residues and at the same cutoff limit (Figure 4.8) reveals the sites where other than arginine and lysine interact with DNA. The contacts of serine atoms to DNA phosphate atoms are lost from

nucleotide 19 at SHL −5.5 on DNA strand I and nucleotide 128 at SHL +5.5 on strand II. Contacts in the wild type found between a threonine residue and the DNA phosphate atoms of nucleotide 61 at SHL −1.3 in DNA strand I are lost in the SIN mutants. Other contacts of threonine residues to DNA phosphate atoms found in the wild type at nucleotide 103 at SHL +3 in DNA strand I and nucleotide 249 at SHL −3 in strand II disappear in the SIN mutants. Contacts of phenylalanine residues with the phosphate at nucleotide 197 at SHL +2.3 on strand II, are also lost in the SIN mutants. The loss of contacts of polar residues, such as serine and threonine, to DNA phosphate atoms occurs in the region near the histone dimerization interface at SHL ±3 and toward the two ends of DNA at SHL ±5.5.

In contrast to the more conserved contacts of arginine and lysine to DNA phosphate atoms, most of the contacts of DNA sugar atoms lost in the SIN mutants involves arginine and lysine residues (compare Figure 4.7 and Figure 4.9).

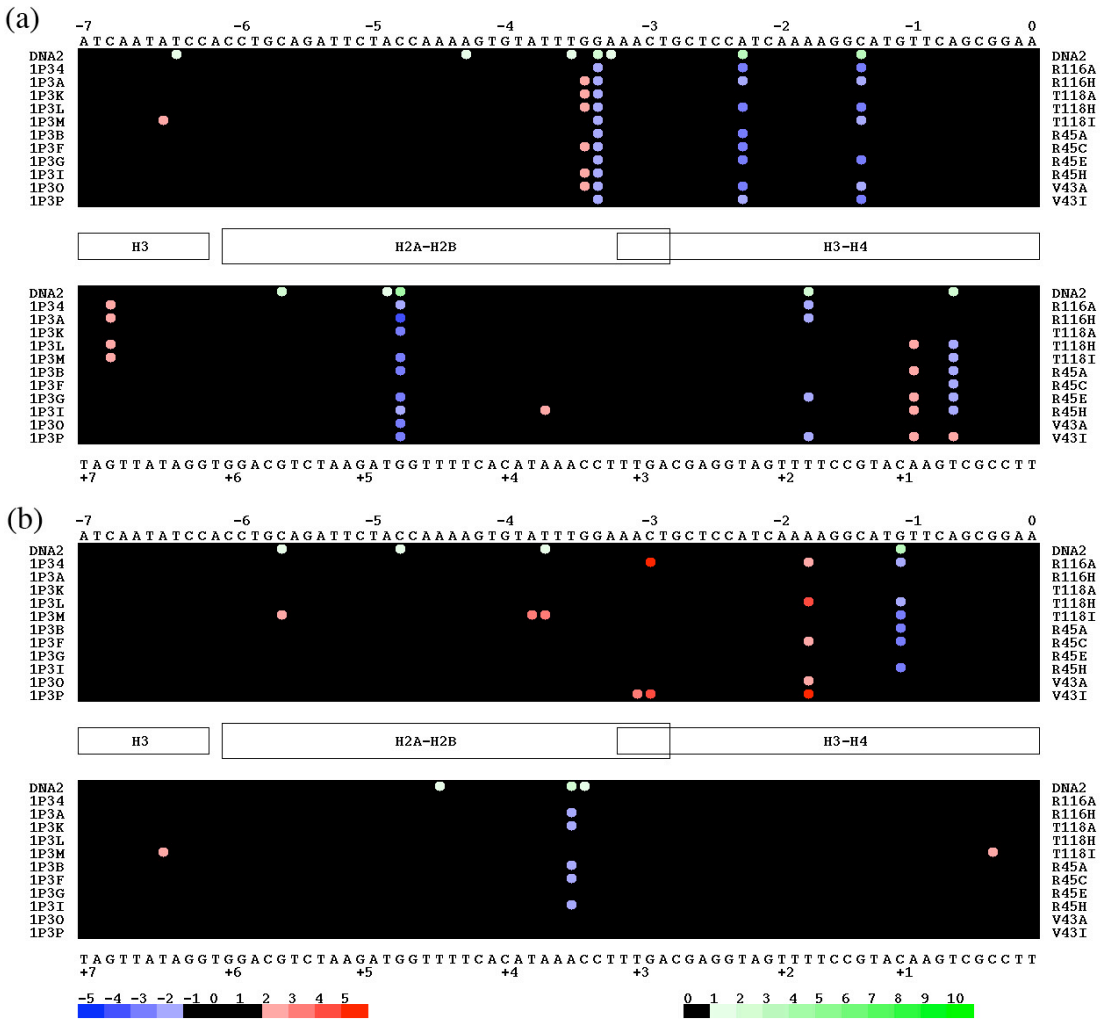



Figure 4.9. The difference in the numbers of close contacts within 3.4 Å of arginine and lysine residues to DNA sugar atoms in (a) DNA strand I, (b) DNA stand II compared with those of the wild type. Data are mapped on each nucleotide along the nucleosomal DNA. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The mutation locations are shown at the right edge of the diagram. The number of contacts in the wild type is shown in different shades of green, and the difference is contacts to the mutants vs. the wild type in shades of blue (negative) and red (positive). Differences of –1, 0, +1 are shown in black.

## 4.3 Summary and discussion

In both the wild type and SIN mutant nucleosomes, contacts of histone protein to DNA phosphate atoms play an important role. The relative numbers of contacts is as follows:

phosphate-H3H4 > phosphate-H2AH2B> sugar-H3H4 > sugar-H2AH2B. These trends persist as the cutoff limit increases from 3.0 Å to 4.0 Å, except for the greater number of contacts at longer distance.

The number of lost contacts between histone subunits and different parts of DNA is compared at the different cutoffs ranging from 3.0 Å to 4.0 Å at increments of 0.2 Å. Although the number of contacts of histones H2AH2B to DNA phosphate atoms is much less than that between histones H3H4 and DNA phosphate atoms, the number of lost histones H2AH2B contacts is larger than or comparable to that lost between histones H3H4 and DNA phosphate atoms. A considerable number of contacts (~ 9) between histone H3H4 and DNA phosphate atoms is lost at the 3.0 Å cutoff. As the cutoff increases from 3.2 Å to 4.0 Å, the number of lost contacts of histones H2AH2B to DNA phosphate atoms is greater than or similar to that lost from histones H3H4. Case studing of the R45H H4 and T118I H3 SIN mutants shows these trends.

Compared to the number of contacts lost to DNA phosphate atoms, the number lost to DNA sugar atoms is less; for example, only two nucleotides, which are located at SHL −4.7 on strand I and SHL +4.7 on strand II, are lost within the 3.2 Å cutoff. With the increase of cutoff distance, the number of lost contacts and the locations of lost contacts to DNA sugar atoms also increase.

The lost contacts from the DNA phosphate atoms are plotted in terms of the different cutoffs (3.0 Å to 4.0 Å). The closest contacts, within 3.0 Å are lost from locations near the dimerization interface of histones H3H4 and H2AH2B (SHL +3 on strand I and SHL −3 on strand II). The regions toward the two ends of DNA tend to lost\e contacts at larger cutoffs. The contacts lost between 3.0 Å and 3.2 Å occurs in nucleotides at SHL −4.3,

−1.3 on strand I, and SHL −5 and +4.3 on strand II. The contacts lost between 3.2 and 3.4 Å occurs in nucleotides at SHL −5.5 and −0.5 on strand I, and SHL −5.8, +0.5, +5.3 on strand II. With increase of the cutoff distance from 3.4 to 4.0 Å, more contacts are lost at more locations along the nucleosomal DNA.

Further study shows that the contacts to DNA of arginine and lysine residues in the histone proteins of the SIN mutants do not change much compared to the wild type. Rather than arginine and lysine, serine residues are involved in the contact losses at the two ends of DNA (at SHL ±5.5). At the dimerization interface (SHL ±3), contacts of threonine residues from DNA phosphate atoms are lost.

To summarize, the major difference in the SIN-mutant structures from the wild type is the amino-acid point mutation at histone protein H3 or H4 near SHL ±0.5. However, the mutations not only cause a loss of contacts to DNA locally near SHL ±0.5 but also a loss overall the DNA. The loss of contacts of serine residues near SHL −5.5 and +5.5 suggest that the ends of DNA can be peeled off the histone octamer more easily in the mutant. In addition to the contacts lost toward the ends of DNA, the DNA (at SHL ± 3) near the dimerization interface of H3H4 and H2AH2B also undergoes a loss of contacts to threonine residues in the SIN-mutant structures. The dimerization of H3H4 and H2AH2B may thus be disrupted during nucleosome remodeling [4]. This suggests that the nucleosome remodeling may be related to the destabilization of the interactions between histone subunits.

1. Kruger, W., Peterson, C. L., Sil, A, Coburn, C., Arents, G., Moudrianakis, E. N., and Herskowitz, I., Amino acid substitutions in the structured domains of histones H3 and H4 partially relieve the requirement of the yeast SWI/SNF complex for transcription. *Genes & Development*, 1995. 9: p. 2770-2779.

2. Horn, P.J., Crowley, K. A., Carruthers, L. M., Hansen, J. C., and Peterson, C. L., The SIN domain of the histone octamer is essential for intramolecular folding of nucleosomal arrays *Nature Structure Biology*, 2002. 9(3): p. 167-171.
3. Harp, J.M., Hanson, B. L., Timm, D. E. and Bunick, G. J., Asymmetries in the nucleosome core particle at 2.5 Å resolution *Acta Crystallog. Sec. D*, 2000. 56: p. 1513-1534.
4. Flaus, A., Owen-Hugers, T., Mechanisms for ATP-dependent chromatin remodelling: farewell to the tuna-can octamer? *Current Opinion in Genetics & Development*, 2004. 14(2): p. 165-173.

# Chapter 5 A Cylindrical Model of a Nucleosome

The shape of a nucleosome resembles a cylinder with the nucleosomal DNA wrapping along a superhelical pathway around the central histone core. It is natural and convenient to represent nucleosomal DNA in a cylindrical frame. In this chapter, SIN mutant and wild-type nucleosomal DNA molecules are translated to their geometric centers and rotated until their shortest principal axis points along the long axis of a cylinder and their longest principal axis points toward the origin of the dyad base pair (base pair 73). Instead of transforming the Cartesian coordinates of all nucleosomal DNA atoms into a cylindrical frame, we reduce DNA to the base-pair level and only transform the centers of base pairs into a cylindrical frame.

## 5.1 Cylindrical model

### 5.1.1 Translate nucleosomal DNA to its geometric center.

$M$ is an $n \times 3$ matrix, $n$ is the number of atoms of a nucleosomal DNA structure. The three columns of the matrix $M$ represent the $x$, $y$, $z$-coordinates of the atoms.

$$M = \begin{pmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & z_n \end{pmatrix}$$

The geometric center $O$ is the average of the above coordinates

$$O = \begin{pmatrix} \langle x \rangle & \langle y \rangle & \langle z \rangle \end{pmatrix},$$

where $\langle x \rangle = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\langle y \rangle = \frac{1}{n}\sum_{i=1}^{n} y_i$, $\langle z \rangle = \frac{1}{n}\sum_{i=1}^{n} z_i$.

The coordinates of a structure are translated to the geometric center using the equation below:

$$M_C = M - i \cdot O$$

Here $i$ is an $n \times 1$ column vector composed of only ones.

5.1.2 Align a nucleosomal DNA along its principal axis.

Following previous studies of nucleosomal DNA [1], a 15 bp fragment is cut from each end of the 146 bp DNA in the SIN mutant and the wild-type nucleosomal structures. The principal axes are determined from the coordinates of the shortened structures, which are contained in an $m \times 3$ matrix, $M_{CT}$, when $m$ is the number of atoms in the truncated structure.

The covariance matrix of the three-dimensional data set contained in matrix $M_{CT}$ is

$$C = \frac{1}{m-1}\left[ M_{CT}^T \cdot M_{CT} - \frac{1}{m} M_{CT}^T \cdot i \cdot i^T M_{CT} \right].$$

Here $i$ is an $m \times 1$ column vector consisting of only ones.

$M_C^T$ and $i^T$ are the transpose of matrix $M_C$ and vector $i$, respectively.

The eigenvalues and eigenvectors of the covariance matrix are given by:

$$\Lambda = \begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 \end{pmatrix},$$

$$\Gamma = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}.$$

Here we arrange the eigenvalues such that $\lambda_1 \le \lambda_2 \le \lambda_3$. The three eigenvalues $\lambda_1$, $\lambda_2$, $\lambda_3$ measures of how the coordinates vary along the three dimensions corresponding to the eigenvectors $\begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix}$, $\begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix}$, $\begin{pmatrix} c_1 & c_2 & c_3 \end{pmatrix}$, respectively.

The nucleosome core particle has a cylinder-like shape with a radius greater than its height. The two circular surfaces of a cylinder resemble the top and bottom surfaces of the histone core and the nucleosomal DNA wraps around the sidewall of the cylinder.

We, thus, align the axis of least variance ( $\lambda_1$ ) of nucleosomal DNA along the vertical axis (*Z*-axis) of the cylinder. The coordinates of the atoms along this axis have the least variance.

As a nucleosome is not an ideal circular cylinder with a fixed radius in the *XY*-plane, the two largest eigenvalues are not exactly the same. It is hard to decide which axis (*X* or *Y*-axis) of the structure to assign to the largest eigenvalue without changing the phase of the original *X* and *Y* coordinates. Here we use Twist, a local base-pair parameter describing the rotational relationship between successive steps along the normal direction of the middle frame, as an indicator of the correct phase between the *X* and *Y*-coordinates. In other words, the values of Twist, which are obtained with the 3DNA software package [2], for the original and re-oriented structures must be same. For all the SIN-mutant and wild-type structures, the rotation matrix $R_1$, the re-ordered $\Gamma$ corresponding to the eigenvalues in the order $\lambda_3$, $\lambda_2$, $\lambda_1$, is written as follow:

$$R_1 = \begin{pmatrix} c_1 & b_1 & a_1 \\ c_2 & b_2 & a_2 \\ c_3 & b_3 & a_3 \end{pmatrix}.$$

After rotation with matrix $R_1$, the structure needs to be rotated around the *Z*-axis by an angle $\varphi$ so that the vector, $\begin{pmatrix} d_x & d_y \end{pmatrix}$ (the origin of the local base-pair frame) along the dyad (base pair 73) of the nucleosomal DNA is located in the *X, Z*-plane. Given, the origin vector of base pair 73, the $\varphi$ angle is given by:

$$\varphi = \begin{cases} -\tan^{-1}(|d_x/d_y|), & d_x > 0, d_y > 0 \\ -\pi + \tan^{-1}(|d_x/d_y|), & d_x < 0, d_y > 0 \\ -\pi - \tan^{-1}(|d_x/d_y|), & d_x < 0, d_y < 0 \\ -2\pi + \tan^{-1}(|d_x/d_y|), & d_x > 0, d_y < 0 \end{cases}$$

The resulting rotation matrix, $R_2$, is as follows:

$$R_2 = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The matrix, $M_{New}$, containing the transformed Cartesian coordinates of the nucleosomal DNA with its geometric center located at (0, 0, 0), the *Z*-axis aligned along the principal axis with least variance, and the origin of the dyad base pair located in the *X, Z*-plane, can be obtained as:

$$M_{New} = (M - i \cdot O) \cdot R_1 \cdot R_2 .$$

Here the matrix $M$ contains the original coordinates, $O$ is the geometric center of the original coordinates ($1 \times 3$ vector), and $I$ is an $n \times 1$ vector composed of ones, and $n$ is the total number of atoms of the complete nucleosome structure (without 15 bp shortened from each end).

5.1.3 Characterizing the base pairs in the cylinder frame

Each DNA base pair can be approximated as a slab and described in a local middle frame of the two pared bases. Successive base pairs (base-pair steps or dimer steps) can also be described in terms of a local middle frame of the base-pair step. The origin, short axis, long axis, and normal vector of dimer steps in nucleosomal DNA are computed with 3DNA software package [2]

Starting with the origin of the i$^{th}$ dimer step as $\left(o_{i,x}, \quad o_{i,y} \quad o_{i,z}\right)$, the normalized radial vector is defined as

$$\vec{r_i} = \frac{1}{\sqrt{{o_{i,x}}^2 + {o_{i,y}}^2}}\left(o_{i,x}, \quad o_{i,y}\right),$$

and the length of the radius as

$$\left|r_i\right| = \sqrt{{o_{i,x}}^2 + {o_{i,y}}^2} .$$

The phase angle, $\theta_i$, of the dimer origin is given by the ratio of its *X* and *Y* component, $O_x$ and $O_y$:

$$\theta_i = \begin{cases} \tan^{-1}(\left|o_{i,x}/o_{i,y}\right|), & o_{i,x} > 0,\ o_{i,y} > 0 \\ \pi - \tan^{-1}(\left|o_{i,x}/o_{i,y}\right|), & o_{i,x} < 0,\ o_{i,y} > 0 \\ \pi + \tan^{-1}(\left|o_{i,x}/o_{i,y}\right|), & o_{i,x} < 0,\ o_{i,y} < 0 \\ 2\pi - \tan^{-1}(\left|o_{i,x}/o_{i,y}\right|), & o_{i,x} > 0,\ o_{i,y} < 0 \end{cases}$$

The vectors along the short axis, the long axis, and the normal vector of the i$^{th}$ dimer are written respectively as $\overrightarrow{v_{i,x}}$, $\overrightarrow{v_{i,y}}$, and $\overrightarrow{v_{i,z}}$. The projections of Shift, Slide, and Rise (the base-pair step parameters describing the translational relationship between neighboring base pairs) on the radial direction of the cylindrical model are

$$r_{Shift,i} = (\vec{r_i} \cdot \overrightarrow{v_{i,x}})Shift ,$$

$$r_{Slide,i} = (\vec{r_i} \cdot \overrightarrow{v_{i,y}})Slide ,$$

$$r_{Rise,i} = (\vec{r_i} \cdot \overrightarrow{v_{i,z}})Rise .$$

Shift, Slide, and Rise are the base-pair step parameters, describing the translational relationship between two successive base pairs. Similarly, the lengths of the projections of Shift, Slide, and Rise on the *Z*-axis of the cylindrical model are:

$z_{Shift,i} = (\vec{n} \cdot \overrightarrow{v_{i,x}})Shift$ ,

$z_{Slide,i} = (\vec{n} \cdot \overrightarrow{v_{i,y}})Slide$ ,

$z_{Rise,i} = (\vec{n} \cdot \overrightarrow{v_{i,z}})Rise$ ,

where $\vec{n} = (0, \quad 0, \quad 1)$ is a unit vector along the *Z*-axis of the cylindrical model.

## 5.2 Results

### 5.2.1 The spatial relationship of superhelical turns of nucleosomal DNA

For every dimer step of nucleosomal DNA from the best-resolved structure (PDB ID: 1KX5, NDB ID: PD0287), the *Z*-coordinate of the origin of the local middle frame is plotted against the cylindrical phase angle (Figure 5.1(a)). The superhelical pathway of nucleosomal DNA is thus projected onto the sidewall of the cylindrical model. As we define the phase angle of the dyad base pair (base pair 74) as 0 °, the left and right edges of the figure cut the sidewall of the cylinder model along the dyad position and flattened plot in Figure 5.1(a) shows the spread of DNA on the sidewall surface. The superhelical ramp of nucleosomal DNA wrapping around the histone core is clear from the plot of the *Z*-coordinates against the phase angle of each base-pair. The spatial relationship of the number of DNA helical turns away from the dyad is also shown clearly. The two strands in the plot (Figure 5.1 (a)) represent the first (base-pair steps 1 to 73) and second (base-pair steps 74 to 146) halves of the DNA. Near the dyad, the *Z*-coordinates are approximately zero. The Z-coordinates increase or decrease gradually, but not smoothly with notable kinks in the region toward the two ends of the DNA. The locations at positive and negative integral number of turns from the dyad, where DNA bends toward the major groove in the two halves of the molecule are located at alternate sites on the

nucleosome surface, *i.e.*, SHL +2 lies between SHL −5 and −6 in terms of its angular location. Correspondingly, the half-integral locations where DNA bends into its minor groove alternate on the two halves of the nucleosome surface, contributing to the DNA supergroove identified in complexes of polyamide drugs with nucleosomes [3] (Figure 3.4 in Chapter 3).

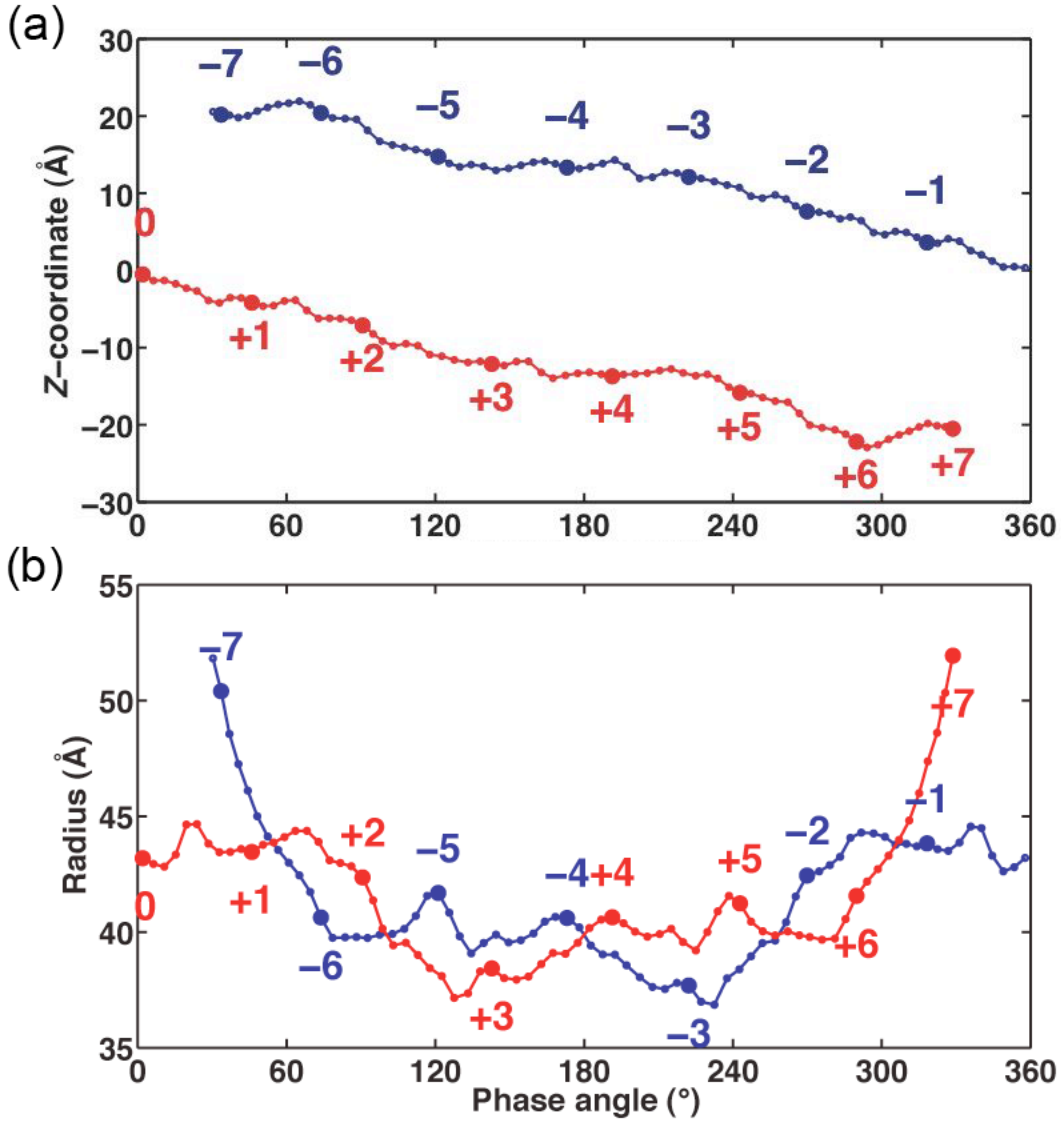



Figure 5.1. (a) *Z*-coordinate and (b) superhelical radius vs. phase angle of each base-pair step of nucleosomal DNA from the best-resolved nucleosome structure (PDB ID: 1KX5). The curves of the first (base-pair step 1 to 73) and second (base-pair step 74 to 146) halves are shown in red and blue, respectively. The origins of the base-pair steps with intergral superhelical numbers are highlighted with larger markers and labeled with the

corresponding superhelical numbers. The superhelical number of each dimer step is assigned as that of the first base pair in the dimer step. The superhelical number is obtained from [1].

The radii, defined by the distance of the dimer step origin to the *Z*-axis, are also plotted against the phase angle (Figure 5.1 (b)). The two ends of the DNA (SHL ±7) are further from the *Z*-axis than other locations. The DNA near the H3H4-H2AH2B dimerization interface (SHL ±3) has the smallest radius. The two halves of nucleosomal DNA are not aligned, like a perfect superelical pathway, along a smooth and perpendicular sidewall of a cylinder. That is, the difference in radii of base-pair steps with the same phase angle is up to 3.7 Å. The two curves representing the radii of base pairs in the two halves of nucleosomal DNA wind together and alternately shift up and down. Crossovers of the two curves, where base pairs in the two halves of DNA have the same radii, occur near SHL ±1 (or ∓6.5), ±4, and ±5.5 (or ∓2.5), (phase angles of approximately 60 °, 100 °, 180 °, 250 °, 300 °).

The *X*- and *Y*-coordinates of the dimer-step origins are plotted in Figure 5.2, which shows the two halves of nucleosomal DNA from a top view looking down the cylindrical axis of the superhelical strand. The symmetry of the two halves of the DNA is evident in the right part of Figure 5.2, where we rotate the second half (SHL 0 to +7) around the *X*-axis by 180°. The two halves overlap well with each other (root-mean-square deviation (RMSD) is 0.38 Å.

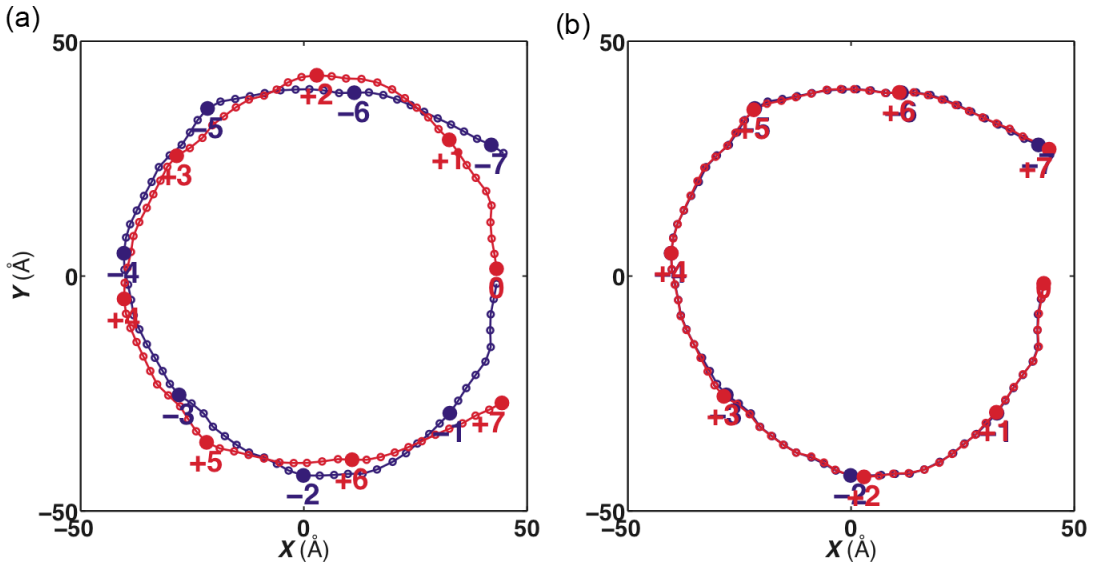



Figure 5.2. X-coordinate vs. Y-coordinate of the origin of each base-pair step of the first (base pair 1 to 73) and second (base pair 75 to 147) half superimposed together (a) without any rotation and (b) with the second half rotated around X axis by 180°. Nucleosomal DNA is from the best-resolved nucleosome structure (PDB ID: 1KX5). The first and second halves are shown in red and blue, respectively. The origins of the base pairs with integral superhelical numbers are highlighted with larger markers and labeled with the corresponding superhelical numbers. See the legend to Figure 5.1 for assignment of superhelical numbers to dimer steps.

The projected pathway is not a perfect circular shape as is clear from the difference of radii reported in Figure 5.1 (b) and Figure 5.2 (a). However, the two projected halves are relatively round from SHL −3 to +3 (the average of radius is 42.0 Å and the standard deviation is 2.43 Å). The two DNA ends (SHL ±6 to ±7) deviate most from the circle (the average of radius is 45.3 Å and the standard deviation is 3.20 Å). Sharp local turns occur at SHL ±4, ±5, ±6, between which the DNA is relatively straight.

The origins of each base-pair local frame along the nucleosomal DNA (NDB ID: PD0287) are projected on *X, Z*-plane, forming a curve with the shape of the letter ‘w’ (Figure 5.3), left part). The first half of the DNA (dimer step 1 to 73) and the second half of the DNA (dimer 74 to 146) have positive and negative *Z*-coordinates, respectively.

The projected pathways of the two halves are quite symmetrical (RMSD 0.58 Å) except for the region from SHL ±2 to ±1 (RMSD 0.85 Å) and ±5.5 to ±6.5 (RMSD 0.81 Å), when the second half of DNA is rotated around the *X*-axis by 180° (Figure 5.3 (b)).

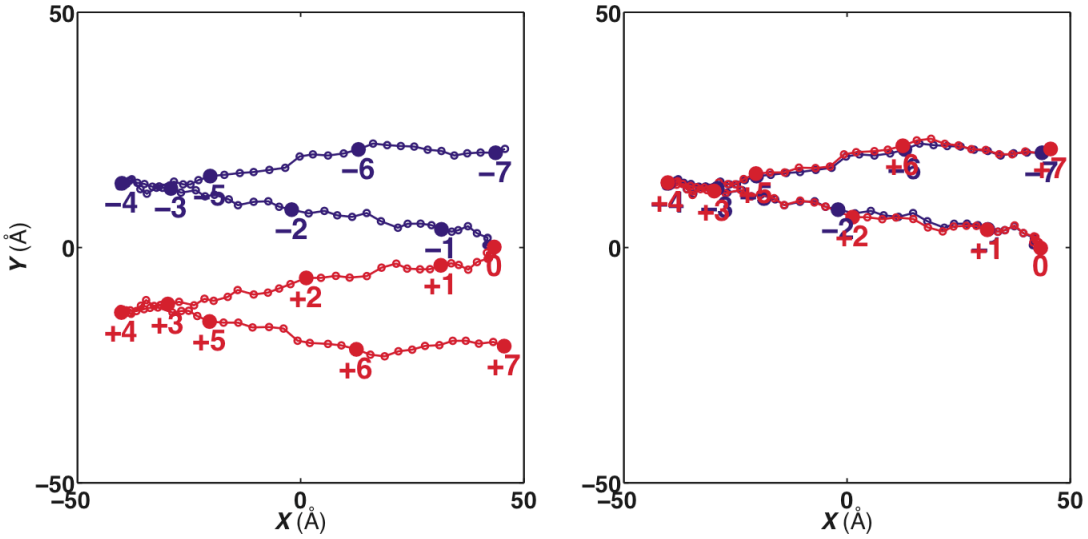



Figure 5.3. X-coordinate vs. Z-coordinate of the origin of each base-pair step of the first (base pair 1 to 73) and second (base pair 75 to 147) half superimposed together (a) without any rotation and (b) with the second half rotated around *X*-axis of 180˚. Nucleosomal DNA is from the best-resolved nucleosome structure (PDB ID: 1KX5). The first and second half are shown in red and blue, respectively. The origins of the base pairs with integral superhelical numbers are highlighted with larger markers and labeled with the corresponding superhelical numbers. See the legend to Figure 5.1 for the assignment of superhelical numbers to dimer steps.

5.2.2 Contribution of Shift, Slide and Rise of each dimer step along the *Z*-axis of nucleosomal DNA.

Nucleosomal DNA winds around the histone core, not along a smooth ramp but rather in a zig-zag way (Figure 5.1 (a)). The contribution of Shift, Slide and Rise of each dimer step along the nucleosomal DNA in the direction of the global *Z*-axis measure how the local step parameters can change superhelical pitch. These contributions are plotted against superhelical position in Figure 5.4. The contribution of Shift at each step is

generally small *i.e.*, no more than 1.5 Å for all the base-pair steps except for the dimer step 45 TG/CA. Contributions of Shift cancel each other so that the accumulation curve along the superhelical axis is close to zero (Figure 5.5). Most of the base-pair steps with relatively large Shift contributions (between 1.0 to 1.5 Å or −1.0 to −1.5 Å) occur at the central locations, SHL −3 to +3.

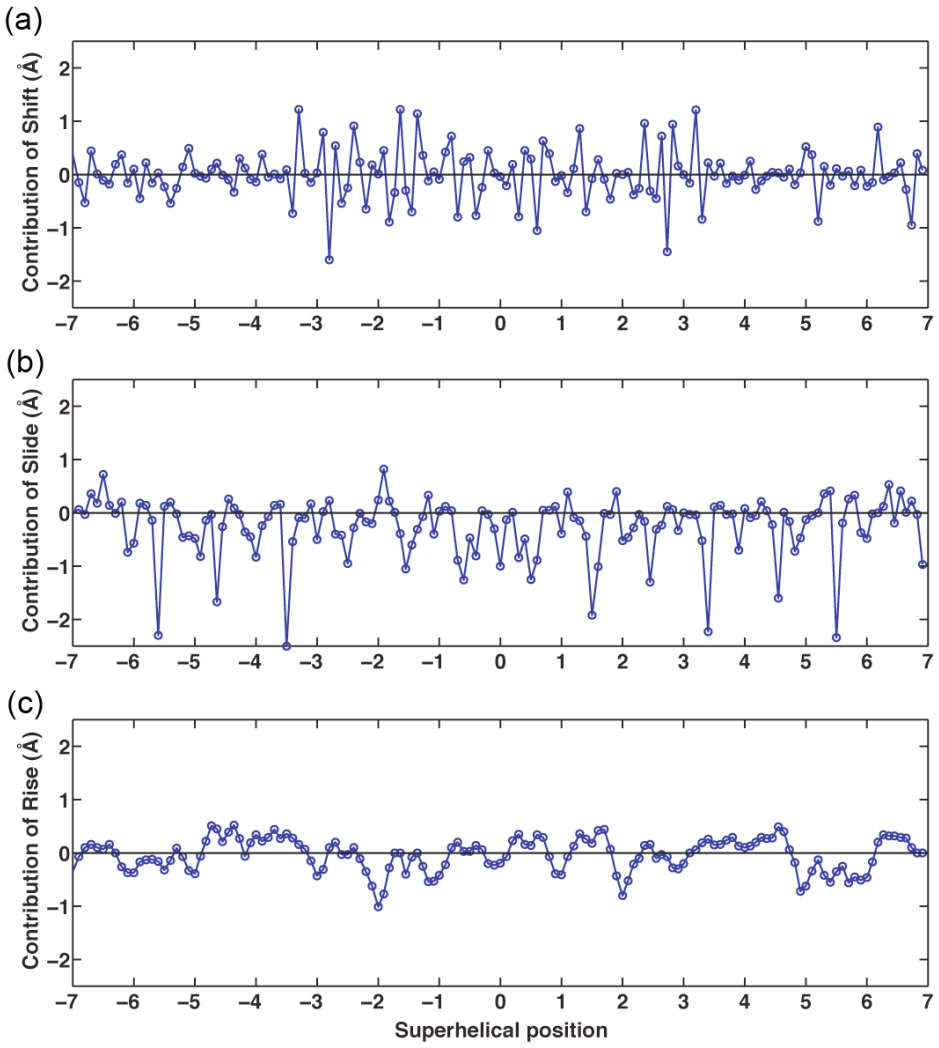



Figure 5.4. Contributions of (a) Shift, (b) Slide, and (c) Rise in the direction of the Z-axis vs. superhelical position of each base-pair step along DNA from the best-resolved structure (PDB ID: 1KX5).

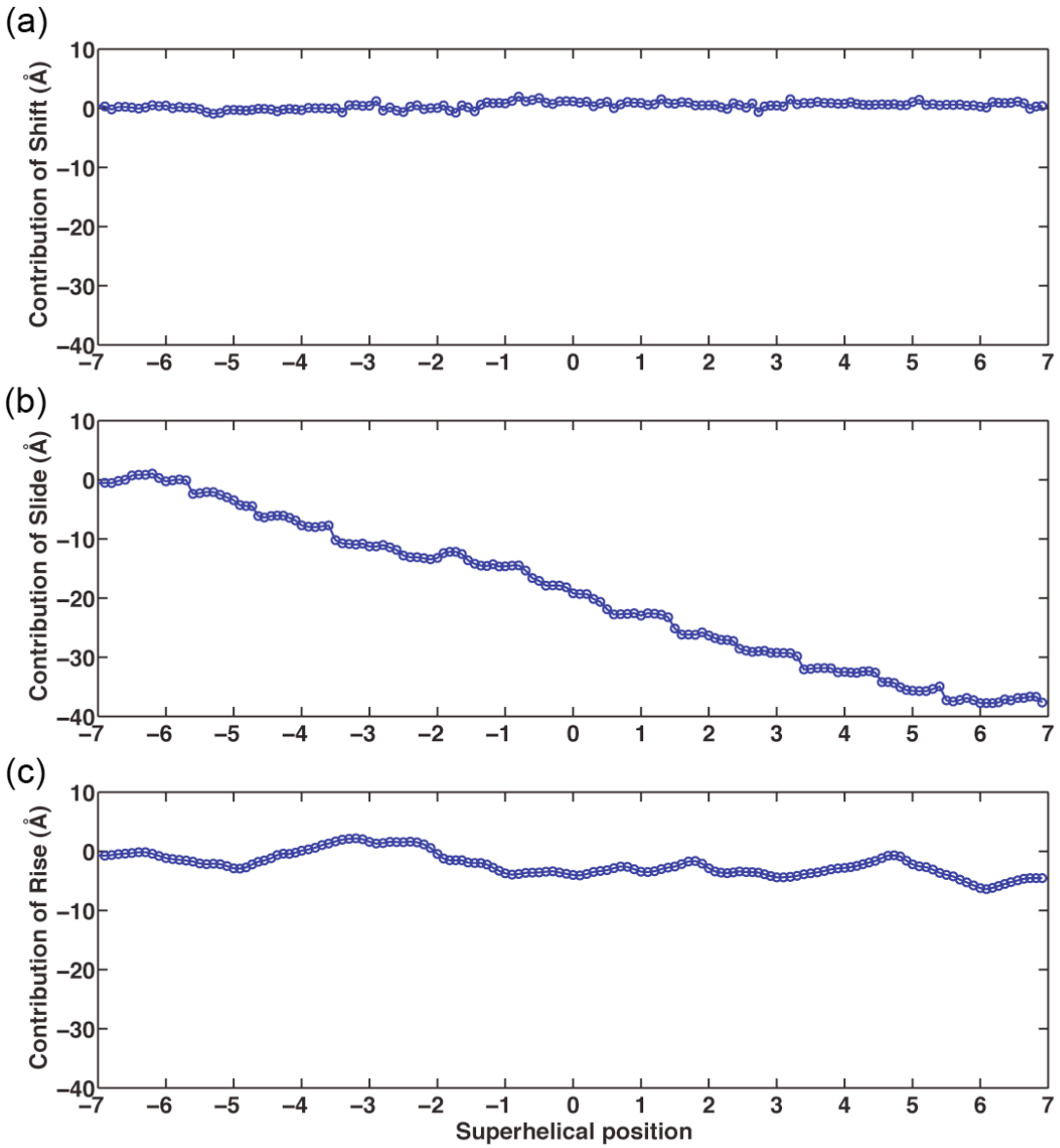



Figure 5.5. The accumulation of the contributions of (a) Shift, (b) Slide, and (c) Rise in the direction of the Z-axis vs. superhelical position of each base-pair step along DNA from the best-resolved structure (PDB ID: 1KX5).

In contrast to Shift, Slide introduces large changes in superhelical pitch and its contribution to *Z*-coordinates is cumulative. The contribution of Slide is more than 1.5 Å and up to 2.45 Å at SHL ±5.5, –4.5, ± 3.5, +1.6, where five of the six base-pair steps are CA/TG or TG/CA steps with large positive Slide, negative Roll, and increased Twist

(Table 5.1 (b)). The contribution of Slide is the determining factor of the superhelical pitch [4] as its accumulation curve decreases from 0 to −35 Å, approximately (Figure 5.5 (b)). The contribution of Rise is even smaller at local base-pair steps (average of 0.30 Å with standard deviation of 0.34 Å over the 146 steps). Similar to Shift, there are no base-pair steps with contributions of Rise, more than 1.5 Å or less than −1.5 Å, in this nucleosome structure. The accumulation of the contribution of Rise is also close to zero (Figure 5.5 (c)).

Table 5.1. Large contributions (larger than 1.5 Å) of (a) Shift and (b) Slide in the direction of the Z-axis and the values of step parameters of specific base-pair steps from the best-resolved nucleosome structure (PDB ID: 1KX5).

(a)

| Step | SHL | Contribution of Shift (Å) | Step parameters | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Shift (Å) | Slide (Å) | Rise (Å) | Tilt (°) | Roll (°) | Twist (°) |
| 45 TG/CA | −2.8 | −1.60 | −1.67 | 0.82 | 3.25 | −5.86 | 3.08 | 37.42 |

(b)

| Step | SHL | Contribution of Slide (Å) | Step parameters | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Shift (Å) | Slide (Å) | Rise (Å) | Tilt (°) | Roll (°) | Twist (°) |
| 16 CA/TG | −5.5 | −2.30 | 0.07 | 2.47 | 3.20 | −2.79 | −11.13 | 46.94 |
| 26 CA/TG | −4.5 | −1.67 | 0.29 | 2.46 | 3.00 | −2.61 | −8.46 | 43.36 |
| 38 TG/CA | −3.5 | −2.50 | −0.43 | 2.58 | 3.25 | −2.30 | −18.43 | 50.04 |
| 89 CC/GG | 1.6 | −1.92 | 0.48 | 1.94 | 3.66 | 1.07 | −8.31 | 45.87 |
| 109 CA/TG | 3.5 | −2.23 | 0.80 | 2.32 | 3.30 | 2.19 | −15.62 | 47.15 |
| 121 TG/CA | 4.6 | −1.60 | −0.04 | 2.29 | 3.25 | −2.75 | −10.72 | 41.67 |
| 131 TG/CA | 5.6 | −2.34 | −0.30 | 2.54 | 3.18 | 3.10 | −16.23 | 46.41 |

5.2.3 The Contributions of Shift, Slide, and Rise in the direction of the radial vectors. The projections of Shift, Slide, and Rise on the direction of the radial vectors are plotted against the superhelical position of each base-pair step from the best-resolved structure (PDB ID: 1KX5) in Figure 5.6. The contribution of Shift is always less than1.5 Å. The largest contributions of Shift, between either 1.0 and 1.5 Å or −1.0 and −1.5 Å, occurs at the base-pair steps located at SHL −3 to +3.

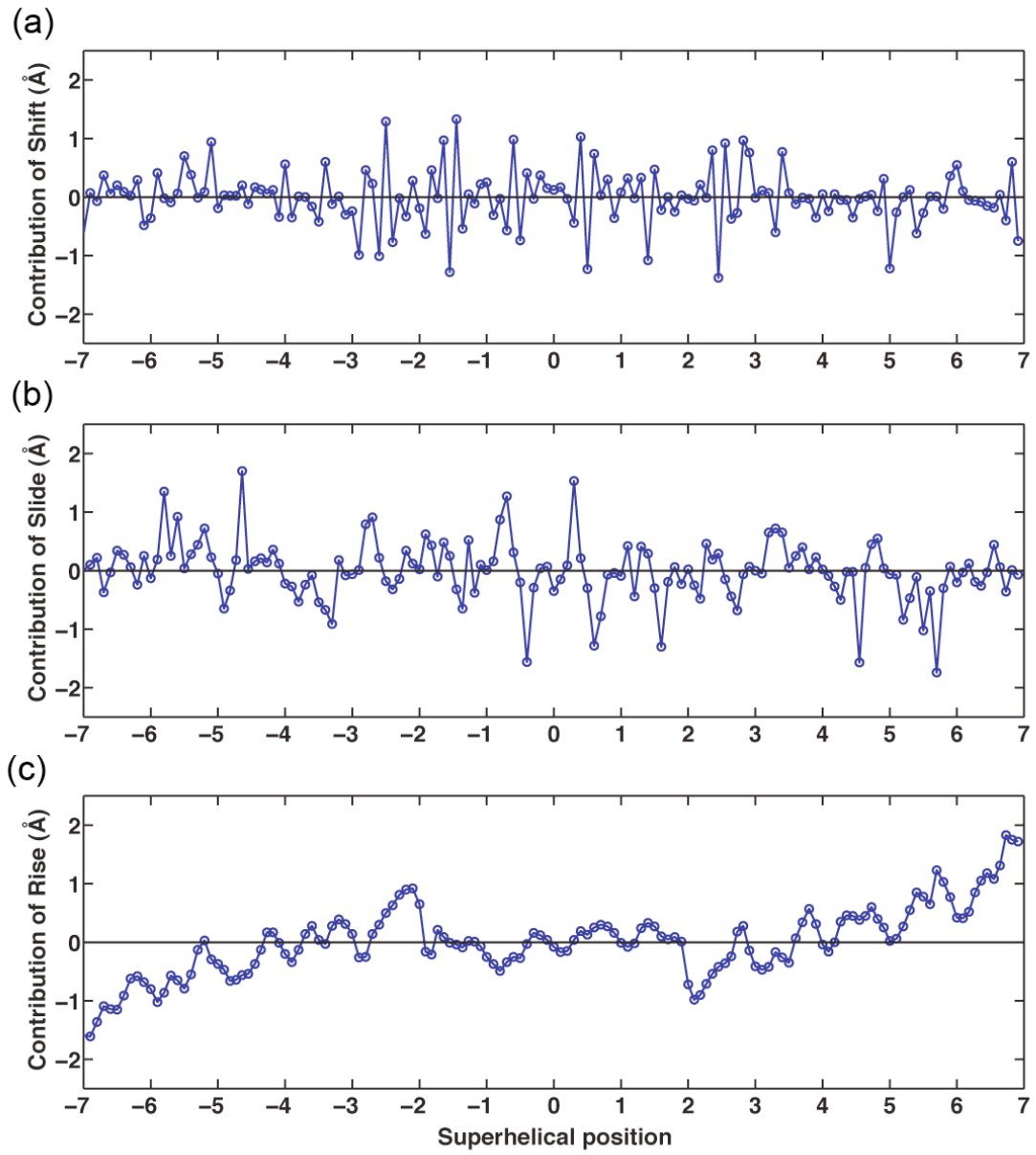



Figure 5.6. Contributions of (a) Shift, (b) Slide, and (c) Rise in the direction of the radial vector vs. superhelical position of each base-pair step along DNA from the best-resolved structure (PDB ID: 1KX5).

Some CA/TG steps have large contributions (more than 1.5 Å or less than −1.5 Å) of Slide along both the *Z*-axis and in the direction of radial vectors. These CA/TG steps are located at SHL −4.5, +4.6 and +5.8, where the large positive Slide is coupled with

negative Roll and increased Twist (compare Table 5.1 (b) to Table 5.2 (a)). Other steps, such as dimer steps 70 TG/CA and 77 CA/TG, have large contribution from Slide only along the radial vectors.

Table 5.2. Large contributions (larger than 1.5 A) of (a) Slide and (b) Rise in the direction of the radial vectors and the values of step parameters of specific base-pair steps from the best-resolved nucleosome structure (PDB ID: 1KX5).

(a)

| Step | SHL | Contribution of Slide (Å) | Step parameters Shift (Å) | Slide (Å) | Rise (Å) | Tilt (°) | Roll (°) | Twist (°) |
|---|---|---|---|---|---|---|---|---|
| 26 CA/TG | –4.5 | 1.70 | 0.29 | 2.46 | 3.00 | –2.61 | –8.46 | 43.36 |
| 70 TG/CA | –0.3 | –1.56 | 0.88 | 1.76 | 3.40 | 3.79 | –2.57 | 43.52 |
| 77 CA/TG | 0.4 | 1.53 | –0.91 | 1.74 | 3.45 | –2.54 | –2.05 | 43.56 |
| 121 TG/CA | 4.6 | –1.57 | –0.04 | 2.29 | 3.25 | –2.75 | –10.72 | 41.67 |
| 133 CA/TG | 5.8 | –1.74 | –0.06 | 1.87 | 3.74 | –0.93 | –0.58 | 42.03 |

(b)

| Step | SHL | Contribution of Rise (Å) | Step parameters Shift (Å) | Slide (Å) | Rise (Å) | Tilt (°) | Roll (°) | Twist (°) |
|---|---|---|---|---|---|---|---|---|
| 2 TC/GA | –6.9 | –1.59 | 0.78 | 0.02 | 3.46 | –0.63 | 5.22 | 37.91 |
| 3 CA/TG | –6.8 | –1.61 | –0.17 | 0.13 | 3.35 | 1.71 | 6.59 | 30.96 |
| 144 TG/CA | 6.8 | 1.83 | 1.07 | 0.47 | 3.45 | 6.45 | 4.01 | 33.40 |
| 145 GA/TC | 6.9 | 1.75 | –0.82 | –0.03 | 3.16 | –5.55 | 5.34 | 33.87 |
| 146 AT/AT | 7.0 | 1.72 | 0.89 | –0.97 | 3.26 | 2.00 | 2.20 | 33.88 |

The contribution of Rise is near zero at the region near the dyad (SHL −2 to +2), but increases gradually as the base pairs approach toward the two ends (Figure 5.6 (c)). The base-pair steps with the largest contribution of Rise (larger than 1.5 Å or less than −1.5 Å) occur at the very ends of the nucleosomal DNA. Here, we liken Rise in the normal direction of a terminal dimer-step middle frame to the velocity of an object in a kind of

circular movement. When the projection of the velocity in the radial direction increases as the object moves, the moving trajectory of the object becomes a curve with increasing radius instead of a perfect circle. In terms of the curvature of nucleosomal DNA, the increasing contribution of Rise suggests that the two ends of nucleosomal DNA have a flipped-out shape compared to the central region (SHL −2 to +2), where the DNA fragment is more like a circle.

5.2.4 *Z*-coordinate and relative contributions of Shift, Slide, and Rise along the *Z*-axis of the R45H H4 SIN-mutant nucleosome compared to the wild-type structure

4.2.4.1 *Z*-coordinates of the R45H H4 SIN mutant

The *Z*-coordinate of each base-pair step of the R45H H4 SIN mutant (PDB ID: 1P3I) is compared to that of the wild-type structure. As there is no dimer step with a large difference in *Z*-coordinates (larger than 0.8 Å or less than −0.8 Å) compared to the wild-type DNA, no dimer steps with large differences in *Z*-coordinates are listed in Table 5.3. In contrast, the large differences in the contributions of Shift and Slide along the *Z*-axis do occur at some base-pair steps (Table 5.3(a) and (b)). These differences, which are discussed in the next two subsections, must be compensated by other structure features for the Z_coordinates to remain relatively constant.

5.2.4.2 Shift contribution of the R45H H4 SIN mutant

Compared to the wild type, three dimer steps, at SHL −4.2, −1.7, and −1.3 in the R45H H4 SIN mutant have relatively large differences in the Shift contribution to pitch (Table 5.3(a)). Notably, the large contribution of Shift in dimer step 60 CA/TG is coupled with relatively large change in Twist (46.83° in the R45H mutant vs. 38.95° in the wild type). And at the same time, the nomal vector of dimer step 60 CA/TG seems largely parallel to

the global *X*-axis since the normal vector is (0.83, 0.56, 0.02). That is to say, since the normal axis is approximately perpendicular to the superhelical axis, the increase of Twist of dimer step 60 CA/TG in the R45H mutant will affect the superhelical pitch, thereby canceling the Shift contribution to the superhelical pitch.

Table 5.3. Values of the step parameters of dimer steps of the R45H H4 mutant nucleosome structure (PDB ID: 1P3I) , with large differences (more than 0.8 Å or less than −0.8 Å) in the contribution of (a) Shift and (b) Slide to superhelical pitch compared to the corresponding steps of the wild-type structure (DNA2).

(a)

| Step | SHL | Structure ID | Shift (Å) | Slide (Å) | Rise (Å) | Tilt (°) | Roll (°) | Twist (°) | Contribution of Shift (Å) |
|---|---|---|---|---|---|---|---|---|---|
| 31 GT/AC | −4.2 | R45H | −1.50 | −0.05 | 3.48 | 0.39 | 1.42 | 29.86 | 1.17 |
| | | DNA2 | −0.36 | −0.94 | 3.04 | 3.09 | −1.31 | 26.33 | 0.26 |
| 56 AA/TT | −1.7 | R45H | −0.92 | −0.12 | 3.49 | −7.74 | −9.34 | 36.45 | −0.73 |
| | | DNA2 | 0.17 | −0.55 | 3.41 | −3.88 | −5.23 | 37.00 | 0.13 |
| 60 CA/TG | −1.3 | R45H | −0.89 | 0.93 | 3.89 | 16.83 | 11.60 | 46.83 | 0.88 |
| | | DNA2 | 0.03 | 0.65 | 3.52 | 8.96 | 10.90 | 38.95 | −0.03 |

(b)

| Step | SHL | Structure ID | Shift (Å) | Slide (Å) | Rise (Å) | Tilt (°) | Roll (°) | Twist (°) | Contribution of Slide (Å) |
|---|---|---|---|---|---|---|---|---|---|
| 21 TC/GA | −5.1 | R45H | 1.27 | 0.75 | 1.87 | 20.06 | 9.73 | 23.17 | 0.65 |
| | | DNA2 | −0.22 | −0.42 | 2.73 | 9.36 | 9.83 | 26.15 | −0.36 |
| 22 CT/AG | −5.0 | R45H | −2.60 | 0.66 | 3.33 | −8.41 | 19.72 | 23.90 | 0.61 |
| | | DNA2 | −0.28 | −0.42 | 3.62 | −1.93 | 16.02 | 29.94 | −0.41 |
| 27 AA/TT | −4.5 | R45H | −0.39 | −0.42 | 3.47 | −8.41 | 9.18 | 37.01 | 0.42 |
| | | DNA2 | 0.32 | 0.50 | 3.48 | 0.44 | −4.64 | 41.90 | −0.49 |
| 47 CT/AG | −2.7 | R45H | −1.56 | 1.10 | 3.27 | −9.44 | −4.48 | 34.26 | −0.96 |
| | | DNA2 | −0.91 | 0.01 | 3.13 | 0.00 | −3.44 | 28.22 | −0.01 |
| 58 GG/CC | −1.5 | R45H | 0.22 | 1.90 | 3.73 | −17.14 | −57.2 | 39.03 | −1.77 |
| | | DNA2 | −0.22 | 0.32 | 3.28 | −9.06 | −31.2 | 34.35 | −0.28 |
| 77 AG/CT | 0.4 | R45H | 0.83 | 1.73 | 3.63 | −0.43 | −6.57 | 43.47 | −1.61 |
| | | DNA2 | 1.00 | 0.80 | 3.21 | −0.34 | −2.57 | 36.08 | −0.74 |
| 124 AG/CT | 4.9 | R45H | 0.16 | 1.66 | 3.32 | −3.78 | 11.94 | 33.35 | 1.56 |
| | | DNA2 | 0.24 | 0.04 | 3.39 | −1.39 | 12.79 | 29.71 | 0.04 |

5.2.4.3 Slide contribution of the R45H H4 SIN mutant

Compared to the wild type, the relatively large difference (larger than 0.8 Å or less than −0.8 Å) in the Slide contribution to the superhelical pitch in the R45H H4 mutant occurs primarily at the two ends of the nucleosomal DNA and near the dyad.

There are four steps with large changes in the contribution of Slide at the two ends of the nucleosomal DNA (Table 5.3 (b)). Three of the dimer steps (21 TC/GA, 22 CT/AG, 27 AA/TT) are clustered at one end of the nucleosomal DNA so that the effect of Slide asymmetric. In addition to the large difference in the Slide contribution in the direction of the *Z*-axis, dimer step 21 TC/GA in the R45H H4 mutant is a highly deformed step in terms of its high deformation score (Table 6.6 in Chapter 6). The values of Rise and Tilt (1.87 Å and 20.06 º) of dimer step 21 TC/GA are two major factors causing the high deformation scores when nucleosomal DNA sequences are threaded on the R45H template. Dimer step 124 AG/CA with a large difference (1.52 Å) in the Slide contribution is located at the other end of the nucleosomal DNA. Noticeably, the neighboring step, dimer step 121 GG/CC, is highly deformed with decreased Twist (21.52°) and Rise (2.07°) (Table 6.6 in Chapter 6).

Three other dimer steps with decreased contributionss of Slide to superhelical pitch occur at central positions. The values of Slide at two of these dimer steps (47 CT/AG and 58 GG/CC), found respectively at SHL −2.7 and −1.5, increase dramatically compared to those of the wild type structure (1.10 Å, and 1.90 Å in the SIN mutant vs. 0.01 Å and 0.32 Å in the wild type). Dimer step 58 GG/CC is also a highly deformed step with a large deformation score, primarily caused by Tilt (−17.14°) and Roll (−57.21°) (Table 6.6 in Chapter 6). Dimer step 77 CG/CG lies near the site of amino-acid mutations. As this

dimer steps is located near SHL +0.5 in the vicinity of the point mutations, the changes in Slide may be related to the R45H SIN mutations.

The Slide contributions to the superhelical pitch at some dimer steps are compensated by the variation of Twist and Tilt. For example, the short axis of dimer step 27 AA/TT has a large component along *X*-axis since the short axis vector is (–0.85, 0.50, 0.16). At the same time, the Tilt of dimer step 27 decreases from 0.44° in the wild type to –8.41° in the R45H mutant (Table 5.3 (b)). That is to say, since the short axis is roughly perpendicular to the superhelical axis, the variation of Tilt at dimer step 27 in the R45H mutant affects the superhelical pitch. Like the decrease in Tilt, the decrease of Twist at dimer step 27 (37.01° in the R45H mutant vs. 41.90°) (Table 5.3 (b)) appear to contribute to the superhelical pitch because the normal vector of the dimer step (–0.50, –0.87, 0.02) has a large component along the *Y*-axis. The variation of Tilt and Twist at dimer step 27 in the R45H mutant may cancel the Slide contribution to the superhelical pitch and thus account for the observed constancy of the superhelical pitch of the R45H SIN mutant compared to the wild type. Like Tilt and Twist in dimer step 27, Tilt and Twist in dimer steps 22 CT/AG, 47 CT/AG, and 58 GG/CC may cancel the Slide contribution to superhelical pitch. In dimer step 77 AG/CT, the Twist difference in R45H mutant compared to the wild type may compensate for Slide contribution as the normal vector is approximately perpendicular to the gobal *Z*-axis.

There is no significant change (larger than 0.8 Å or less than –0.8 Å) in the contribution of Rise to the superhleical pith in the R45H H4 SIN mutant compared to the wild type (data not shown).

5.2.5 The trends of variation in *Z*-coordinate and relative contribution of Shift, Slid e and Rise along the *Z*-axis in all SIN-mutant nucleosomes compared to the wild-type structure

The variation of the *Z*-coordinate and the contributions of Shift and Slide along the *Z*-axis of each base-pair step in all the SIN mutants compared to the wild-type are summarized in Figure 5.7 and Appendix Figure A5.1. As the difference in the contribution of Rise along the *Z*-axis in the SIN mutants compared to the wild-type is trivial, the data are not shown.

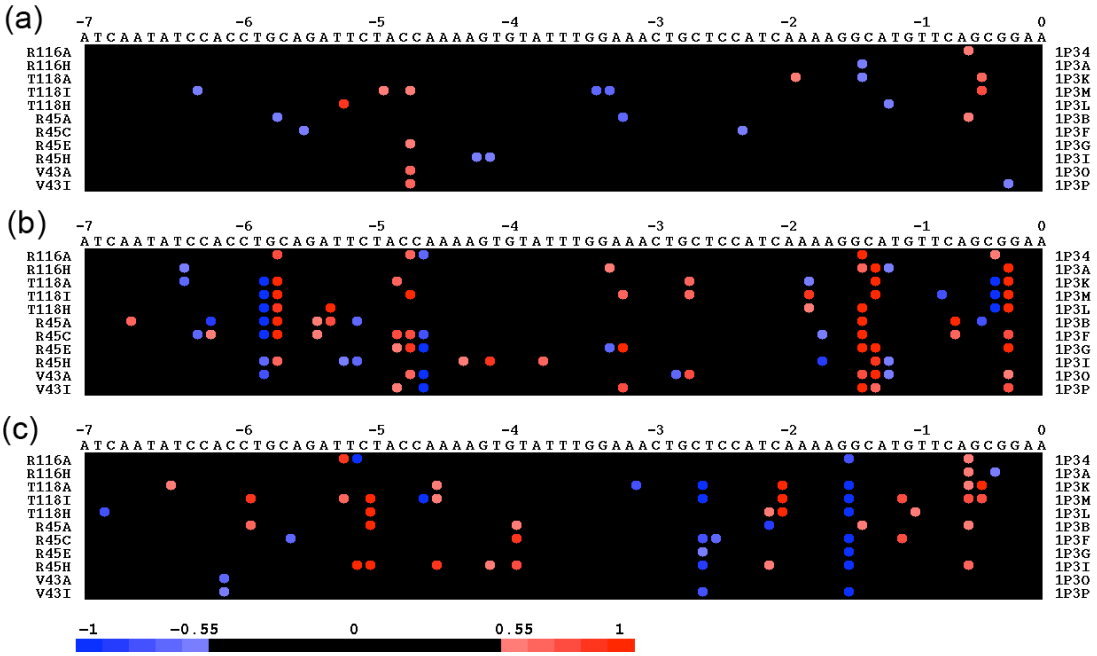



Figure 5.7. The difference in (a) Z-coordinates, (b) the contribution of Shift in the radius direction and (c) the contribution of Slide in the Z-axis of SIN mutant compared to the wild type (DNA2), are mapped on each base-pair step along the first half of the nucleosomal DNA from base-pair step 1 to 73. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The differences in values of the mutants vs. the values of wild type are shown in shades of blue (negative) and red (positive). Differences between -0.55 Å to 0.55 Å, are shown in black.

Like the trends of the R45H H4 SIN mutant, the difference in *Z*-coordinates of all the SIN-mutant structures compared to the wild type is not significant (Figure 5.7(a) and Appendix Figure A5.1(a)). However, in the SIN-mutant structures, there are some common base-pair steps with the large difference in Shift and Slide contributions to the superhelical pitch (see columns of dots in Figure 5.7 (b) and (c)).

The base-pair steps with large difference in the Shift contribution occur at one end of the nucleosomal DNA and near the dyad position (Figure 5.7 (b)). The contribution of Shift at dimer step 14 TG/CA in the SIN mutants is more negative (0.74 Å in the wild type vs. –0.16 Å averaged over all the SIN mutants). In contrast, at the neighboring step 15 GC/GC, contribution of Shift in the SIN mutants is more positive (–0.64 Å in the wild type vs. 0.06 Å averaged over all the SIN mutants). Near the dyad, three dimer steps 59 GC/GC, 60 CA/TG, and 70 GG/CC have increased contribution of Shift (the increase is up to 0.96 Å).

Compared to the first half of DNA (base-pair steps1 to 73), the plot of the difference in Shift contribution of the second half (base-pair steps 74 to 145) appears to be much quieter (compare Figure 5.7 (b) to Appendix Figure A5.1 (b)). That is, the differences between the second halves in the SIN mutants and the wild type are much smaller. Except for one dimer step near SHL +0.5, in the vicinity of the sites of SIN mutations, other base-pair steps with large differences in the contribution of Shift are scattered at the region near the DNA end (SHL –6 to –4).

The large differences of the Slide contribution to superhelical pitch in all the SIN mutants occur at the base-pair steps at central positions, from SHL –3 to +3 (Figure 5.7(c)). Dimer step 47 CT/AG has decreased contribution of Slide and, at the same time. The

contribution of Slide is much more negative in the SIN mutants (–0.28 Å in the wild type vs. an average of –1.33Å over all the SIN mutants) in the SIN mutant at dimer step 58 GG/CC. The step is adjacent to dimer steps 59 GC/GC and 60 CA/TG with increased contribution of Shift to superhelical pitch.

In the second half of DNA (base pair 74 to 146), except for the large differences at dimer steps 75 CC/GG and 76 CG/CG, the differences in the contributions of Slide to the super helical pitch in the SIN mutants compared to the wild type are small compared to the wild type (Appendix Figure A5.1).

5.2.6 Superhelical radius, and relative contribution of Shift, Slide and Rise along the direction of radial vectors of the R45H H4 SIN-mutant nucleosome compared to the wild-type structure

One end of the nuclesomal DNA in the R45H H4 SIN mutant has increased radii at the neighboring dimer steps from SHL –7 to –6 and decreased radii at SHL –5 (Figure 5.8 (a)). The increase of radius at dimer step 7 AT/AT, 9 CC/GG, 11 AC/GT, and 12 CC/GG (SHL –7 to –6) is up to 0.71 Å. Surprisingly, the decrease of radius at dimer step 21 TC/GA and 22 CT/AG at SHL –5.0 is 1.43 Å and 1.02 Å, respectively (Table 5.4 (a)). In addition to the difference of radii at the one end of the nucleosomal DNA, dimer step 58 GG/CC and 59 GC/GC at SHL –1.5 in the R45H H4 SIN mutant undergoes the increase of radius of 0.83 and 0.67 Å, respectively (Table 5.4 (a)). Furthermore, dimer step 58 GG/CC is also a highly deformed step with the large ‘energy’ contribution from Tilt (–9.06 °) and Roll (–31.16 °) (Table 6.6 in Chapter 6). The increase of the radii in the R45H H4 SIN mutant suggests the possible ways of nucleosome remodeling, *i. e*. one end

of the nucleosomal DNA peels away from the histone core and DNA near the dyad bulges out.

Table 5.4. (a) Values of the step parameters of dimer steps of the R45H H4 mutant nucleosome structure (PDB ID: 1P3I) , with large differences in radii (more than 0.8 Å or less than –0.8 Å) compared to the corresponding steps of the wild-type structure (DNA2). (b) The step parameters and the comparative contribution of Shift to superhelical pitch and (c) the step parameters and the comparative contribution of Slide to superhelical pitch at steps, where the difference in the contribution is greater than0.8 Å or less than -0.8 Å.

(a)

| Step | SHL | Structure ID | Shift (Å) | Slide (Å) | Rise (Å) | Tilt (°) | Roll (°) | Twist (°) | Radius (Å) |
|---|---|---|---|---|---|---|---|---|---|
| 21 TC/GA | −5.1 | R45H | 1.27 | 0.75 | 1.87 | 20.06 | 9.73 | 23.17 | 39.45 |
| | | DNA2 | −0.22 | −0.42 | 2.73 | 9.36 | 9.83 | 26.15 | 40.88 |
| 22 CT/AG | −5.0 | R45H | −2.6 | 0.66 | 3.33 | −8.41 | 19.72 | 23.9 | 40.29 |
| | | DNA2 | −0.28 | −0.42 | 3.62 | −1.93 | 16.02 | 29.94 | 41.31 |
| 58 GG/CC | −1.5 | R45H | 0.22 | 1.9 | 3.73 | −17.14 | −57.21 | 39.03 | 45.07 |
| | | DNA2 | −0.22 | 0.32 | 3.28 | −9.06 | −31.15 | 34.35 | 44.24 |

(b)

| Step | SHL | Structure ID | Shift (Å) | Slide (Å) | Rise (Å) | Tilt (°) | Roll (°) | Twist (°) | Contribution of Shift (Å) |
|---|---|---|---|---|---|---|---|---|---|
| 12 CC/GG | −6 | R45H | 1.34 | −0.61 | 3.61 | 0.11 | 17.08 | 30.24 | −1.30 |
| | | DNA2 | 0.40 | −0.49 | 3.49 | 1.59 | 13.61 | 30.83 | −0.39 |
| 21 TC/GA | −5.1 | R45H | 1.27 | 0.75 | 1.87 | 20.06 | 9.73 | 23.17 | −1.15 |
| | | DNA2 | −0.22 | −0.42 | 2.73 | 9.36 | 9.83 | 26.15 | 0.19 |
| 22 CT/AG | −5.0 | R45H | −2.60 | 0.66 | 3.33 | −8.41 | 19.72 | 23.90 | 2.59 |
| | | DNA2 | −0.28 | −0.42 | 3.62 | −1.93 | 16.02 | 29.94 | 0.28 |
| 33 GT/AC | −4.0 | R45H | −0.28 | 0.25 | 2.62 | −1.87 | −2.85 | 24.02 | 0.27 |
| | | DNA2 | −1.14 | −0.55 | 2.85 | −1.85 | 12.57 | 25.95 | 1.09 |
| 68 GC/GC | −0.5 | R45H | 0.23 | 0.95 | 3.83 | 3.13 | −16.01 | 35.84 | 0.21 |
| | | DNA2 | −0.75 | 0.82 | 3.38 | −3.16 | −5.64 | 39.07 | −0.70 |

(c)

| Step | SHL | Structure ID | Shift (Å) | Slide (Å) | Rise (Å) | Tilt (°) | Roll (°) | Twist (°) | Contribution of Slide (Å) |
|---|---|---|---|---|---|---|---|---|---|
| 10 CA/TG | −6.2 | R45H | −1.27 | 1.34 | 3.64 | −7.44 | 6.14 | 37.13 | −1.20 |
| | | DNA2 | −1.07 | 0.28 | 3.40 | 0.23 | 6.80 | 30.35 | −0.24 |
| 76 CC/GG | 0.3 | R45H | 0.24 | 0.79 | 3.07 | 5.86 | 1.33 | 36.13 | 0.70 |
| | | DNA2 | −0.63 | 1.90 | 3.51 | −2.63 | −1.80 | 43.41 | 1.69 |

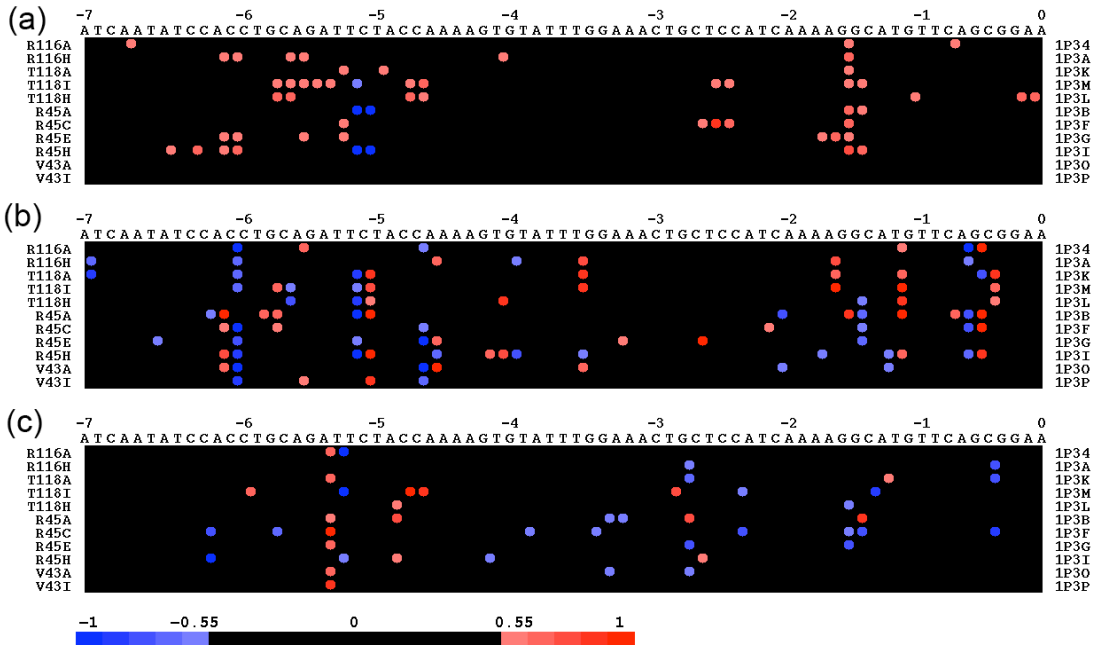



Figure 5.8. The difference in (a) radius, (b) the contribution of Shift in the radial direction and (c) the contribution of Slide in the radial direction of SIN mutants compared to the wild type (DNA2), are mapped on each base-pair step along the first half of the nucleosomeal DNA from base-pair step 1 to 73. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The difference in values of the mutants vs. the values of wild type are shown in shades of blue (negative) and red (positive). Differences between -0.55 Å to 0.55 Å, are shown in black.

Dimer steps with large difference in the Shift contribution to the superhelical pitch in the R45H H4 SIN mutant occur in two regions, toward the one end of the nucleosomal DNA and near the site of SIN mutations (Table 5.4 (b)). Four dimer steps 12 CC/GG, 21 TC/GA, 22 CT/AT, and 33 GT/AC with large variation of Shift are located from SHL –6 to –4. The decreased contribution of Shift may induce decrease of the radius at dimer steps 21 TC/GA. In contrasst, despite of the increased contribution of Shift, dimer step 22 CT/AT has the decreased radii (Table 5.4(a)). Dimer step 68 GC/GC at SHL –0.5, in the vicinity of the site of SIN mutations, has increased contribution of Shift.

Like the Shift contribution, the difference in the Slide contribution in the direction of the radial vectors also occurs in the same two regions mentioned above (Table 5.4 (c)). Dimer step 10 CA/TG with a more negative Slide contribution is located at SHL −6.2. Dimer step 76 CC/GG with a decreased Slide contribution is near the site of SIN mutations at SHL −0.5.

5.2.7 The trend of variation in superhelical radius, and relative contribution of Shift, Slide and Rise along the direction of radial vectors of all the SIN-mutant nucleosomes compared to the wild-type structure

The variation of radii of all the SIN mutants compared to the wild-type is shown in Figure 5.8 (a) and Appendix Figure A5.2 (a). Strikingly, the dots designating the difference in radii in the SIN mutants compared to the wild type cluster in rows rather than in the columns (Figure 5.8 (a)). That is, the R116H H3, T118H H3, T118I H3, R45E H4, and R45H H4 SIN mutants differ from the other SIN mutants in showing increased radii over several neighboring or successive dimer steps at one end of the nucleosomal DNA. For example, the five successive dimer steps from step 15 to 19 (from SHL −5.8 to −5.4) in the T118I H3 SIN mutant have large radii than the wild-type. Even more interestingly, except for the R116H H3 mutant, all other mutants mentioned (T118H, T118I, R45E, and R45H mutants) above have higher sliding rates in heat-shifting experiments [5].

However, some mutants (the R45A and R45C H4 mutants) with high sliding rates in heat-shifting experiments do not show significant increase of radii at the one end of nucleosomal DNA. Rather than an increase of radii toward the end of the nucleosome, the R45A H4 SIN mutant has decreased radii at SHL −5.1 and −5.2 and increased radii at

SHL −1.5 and −1.4. The R45C H4 SIN mutant also has a high sliding rate in the heat-shifting experiment. In addition to the increased radius at dimer step 20 TT/AA at SHL −5.3, the radius increases at base-pair steps near SHL −3.5 (dimer steps 47 CT/AG, 48 TC/GA, and 49 CC/GG) and SHL −1.5 (dimer steps 58 GG/CC and 59 GC/GC). The R116A H3, and the V43A and V43I H4 SIN mutants, which have slow sliding rates in the heat-shifting experiments, do not show any significant increase of radii toward the one end of nucleosomal DNA.

In addition to the increased radii found in the SIN mutants toward the one end of the nucleosomal DNA, dimer step 58 GG/CC at SHL −1.5 has increased radii in all the SIN mutants compared to the wild type (Figure 5.8 (a)).

The dimer steps with large difference in the contributions of Shift and Slide in the direction of the radial vectors are concentrated in the same two regions where the increase of radii occurs (Figure 5.8 (a) and (b)). That is, dimer steps 12 CC/GG, 21 TC/GA, 22 CT/AG (at SHL −6, −5.2, −5.1) at one end of the nucleosomal DNA and 62 TG/CA (at SHL −1.1) near the dyad have large differences in the Shift contribution; and dimer step 19 AT/AT at SHL −5.3 has an increased contribution of Slide.

Interestingly, the second half of the SIN-mutant nucleosomal DNAs is similar to the wild type in terms of superhelical radius and Shift and Slide contribution in the radial direction (Appendix Figure A5.2). Only dimer step 76 CC/GG at SHL +0.3 has a decreased contribution of Slide.

## 5.3 Summary and discussion

Nuclosomal DNA is represented in a cylindrical frame as the shape of a nucleosome resembles a cylinder with the nucleosomal DNA wrapping along a superhelical pathway around the histone core.

This chapter reports the superhelical properties of the nucleosomal DNA from the best-resolved structure (PDB ID: PD0287) in the cylindrical frame. Far from the ideal superhelical pathway, the ramp of the nucleosomal DNA follows a zig-zag route with notable kinks. Instead of a perfect circular shape, there is a large difference of radii (up to 3.7 Å) of the two DNA strands with the same angular location.

The origins of the middle frames of each base-pair step from the best-resolved nucleosome structure (PDB ID: PD0287) are projected on the *X, Y*-plane and the *X, Z*-plane. The central position of nucleosomal DNA (SHL −4 to +4) is relative round and the two ends of DNA (SHL ±6 to ±7) deviate most from the circle. The projection of the second half of DNA on the *X, Y*-plane and the *X, Z*-plane, when rotated around the *X*-axis by 180 °, overlaps the first half of DNA well, except for the difference at SHL ±2. This result shows that the two halves of the DNA (base-pair steps 1 to 73 and 76 to 146) are quite symmetrical.

The contribution of Slide to the superhelical pitch accumulates from 0 to 35 Å along the nucleosomal DNA. There are several dimer steps toward the two end of DNA (SHL −5.5, −4.5, −3.4, 3.5, and 5.6) with large contributions of Slide (up to 2.45 Å) to the superhelical pitch. The five of these dimer steps of which are CA/TG or TG/CA steps with large positive Slide, negative Roll, and increased Twist. Compared to Slide, the contribution of Shift and Rise are smaller and cancel each other so that the accumulation

curves fluctuate around zero. However, the contribution of Shift at the central position (SHL −3 to +3) is relatively large (between 1.0 Å to 1.5 Å or −1.0 Å to −1.5 Å).

The largest contributions of Slide along the radial direction occur at SHL −4.5, −0.3, +0.4, +4.6, and +5.8, all of which are CA/TG or TG/CA steps with large positive Slide, negative Roll, and increased Twist. The increased radius at the two ends of DNA is coupled with the increased contribution of Rise along the radial direction as the two ends of DNA deviate from a circular shape.

The superhelical properties of the SIN-mutant nucleosomes are compared to the wild-type structure. The difference in the *Z*-coordinates in all the SIN mutants is not significant compared the wild type. However, there are some dimer steps with large differences in the Shift and Slide contributions to the *Z*-coordinates. In case study of the R45H SIN mutant, the Shift and Slide contributions to the superhelical pitch are compensated by the variation of Tilt and Twist since the short axis and the normal axis of the dimer steps are approximately perpendicular to the superhelical axis (the global *Z*-axis).

Remarkably, an increase of superhelical radius occurs in the SIN mutants with high sliding rates in heat-shifting experiments. The R116H H3, T118H H3, T118I H3, R45E H4, and R45H H4 SIN mutants have considerable increases of radii at neighboring or successive dimer steps toward one end of nucleosomal DNA. Except for the R116H H3 SIN mutant, all the other mutants mentioned above show high sliding rates in heat-shifting experiments [5]. Notably, rather than an increase of radius in neighboring or successive dimer steps near the end of the nucleosomal DNA, the R45C H4 mutant has increased radii at the three successive dimer steps from SHL −2.8 to −2.5, near the

dimerization interface of histones H3H4 and H2AH2B. The R45C H4 mutant also has a high sliding rate in the heat-shifting experiment. The other mutants, R116A H3, T118A H3, V43A H4 and V43I H4 mutants do not show any increased radii in the neighboring or successive dimer steps. Consistently, they have low sliding rates in heat-shifting experiments. To some degree, the increase in radius at successive or neighboring dimer steps in the SIN mutants, either toward the end of nucleosomal DNA or near the histone dimerization interface, is related to the high sliding rates in heat-shift experiments. This suggests potential ways of nucleosome remodeling, such as nucleosomal DNA peeling off the histone cores from one end or bulging out near the histone dimerization interface. Further more, in almost all the SIN mutants, dimer step 58 at SHL −1.5 has a larger radius compared to the wild type. These results suggest that DNA may loop away from the protein core in the vicinity of the dyad during the remodeling of nuclesomes.

One end of the DNA (SHL −6 to −4.5) and the central position (SHL −2 to 0, especially at SHL −1.5) appear to be hot spots of changes of superhelical properties in the SIN mutants compared to the wild type (compare Figure 5.7 (b), (c) and Figure 5.8 (b), (c)). A decreased Shift contribution in the direction of the radial vectors occurs at dimer step 12 CC/GG (SHL −6), which is accompanied by a difference in the Shift contribution to the *Z*-coordinates at neighboring steps 14 TG/CA and 15 GC/GC. The increased radius in dimer step 58 GG/CC at SHL −1.5 is coupled with a decreased contribution of Slide along the *Z*-axis at the same step and an increased Shift contribution along the *Z*-axis at neighboring steps 59 GC/GC and 60 AT/AT. This result also suggests that the two regions, one end of the nucleosomal DNA and the site near SHL −1.5, may undergo changes of superhelical shape during the nucleosome remodeling.

1. Colasanti, A.V., *Conformational states of double helix DNA*, Ph. D. Thesis. 2006, Rutgers, the State University of New Jersey, New Brusnwick.
2. Lu, X.J., and Olson, W. K., *3DNA: a software package for the analysis rebuilding and visualization of three-dimensional nucleic acid structures*. Nucleic Acids Research, 2003. 31: p. 5108-5121.
3. Suto, R.K., Edayathumangalam, R. S., White, C. L., Melander, C., Gottesfeld, J. M., Dervan, P. B., and Luger, K., *Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands*. J. Mol. Biol., 2002. 326: p. 317-380.
4. Tolstorukov, M.Y., Colasanti, A. V., McCandlish, D., Olson, W. K., and Zhurkin, V. B., *A novel 'roll-and-slide' mechanism of DNA folding in chromatin. Imaplication for nucleosome positioning*. 2006.
5. Muthuranjan, U.M., Bao, Y., Forsberg, J., Edayathumagalam, R. S., Dyer, P. N., White, C. L., and Luger, K., *Crystal structures of histone Sin mutant nucleosomes reveal altered protein-DNA interactions*. EMBO J., 2004. 23: p. 260-271.

# Chapter 6 Threading

## 6.1 Methods

### 6.1.1 Energy function of a DNA dimer step

Knowledge-based elastic potentials provide a way to assess the deformation of nucleosomal DNA. The deformation energy reflects the conformational deviations of a DNA dimer step from its rest state. It has been presumed that the average values of each dimer step parameter over a large data set of protein-bound DNA molecules reveal its natural conformational characteristics. The force imposed on DNA by specific proteins are thought to cancel each other out [1].

The 16 possible dimer steps fall into 10 unique groups in terms of complementary dimer steps. That is to say, the complementary base pairs CA and TG, AG and CT, GG and CC, AA and TT, GA and TC, AC and GT fall in the same groups. The values of the six step parameters of the complementary dimer steps are the same except for different signs of Shift and Tilt. The four remaining base-pair steps are self-complementary, *i.e.*, have the same sequence in both strands, so that the mean values of Shift and Tilt are zero.

The values of dimer-step parameters have been collected from 173 high-resolution protein-DNA crystal structures in a non-redundant data set [2]. A culling procedure has been used iteratively to remove the outliers, *i.e.* the values of parameters greater than three standard deviations from the mean values [1]. The standard deviation is computed as $\sigma = (\langle\theta^2\rangle - \langle\theta\rangle^2)^{\frac{1}{2}}$, where $\theta$ is a given parameter and the brackets represent the average of all values. The average values and standard deviations of dimer-step parameters of the protein-DNA data set are summarized in Table 6.1.

Table 6.1. Average base-pair parameter values and standard deviations of parameters in the reference the protein-DNA data set.

| Step | Twist (°) | Tilt (°) | Roll (°) | Shift (Å) | Slide (Å) | Rise (Å) |
|---|---|---|---|---|---|---|
| CG | 34.40 | 0.00 | 5.47 | 0.00 | 0.36 | 3.41 |
| | (3.65) | (3.89) | (5.65) | (0.78) | (0.57) | (0.22) |
| CA | 35.04 | 0.11 | 5.07 | –0.05 | 0.22 | 3.38 |
| | (4.89) | (3.12) | (5.02) | (0.67) | (0.90) | (0.26) |
| TG | 35.04 | –0.11 | 5.07 | 0.05 | 0.22 | 3.38 |
| | (4.89) | (3.12) | (5.02) | (0.67) | (0.90) | (0.26) |
| TA | 37.42 | 0 | 2.49 | 0 | 0.37 | 3.32 |
| | (7.31) | (2.71) | (5.77) | (0.58) | (0.93) | (0.21) |
| AG | 32.47 | –1.05 | 4.13 | 0.19 | –0.32 | 3.35 |
| | (4.59) | (2.53) | (3.70) | (0.64) | (0.45) | (0.23) |
| CT | 32.47 | 1.05 | 4.13 | –0.19 | –0.32 | 3.35 |
| | (4.59) | (2.53) | (3.70) | (0.64) | (0.45) | (0.23) |
| GG | 33.27 | 0.66 | 4.95 | 0.02 | –0.52 | 3.45 |
| | (4.49) | (3.54) | (4.54) | (0.67) | (0.62) | (0.24) |
| CC | 33.27 | –0.66 | 4.95 | –0.02 | –0.52 | 3.45 |
| | (4.49) | (3.54) | (4.54) | (0.67) | (0.62) | (0.24) |
| AA | 35.09 | –1.08 | 0.74 | 0.08 | –0.16 | 3.25 |
| | (3.88) | (2.45) | (4.45) | (0.35) | (0.33) | (0.17) |
| TT | 35.09 | 1.08 | 0.74 | –0.08 | –0.16 | 3.25 |
| | (3.88) | (2.45) | (4.45) | (0.35) | (0.33) | (0.17) |
| GA | 35.54 | –1.50 | 1.86 | –0.21 | –0.12 | 3.32 |
| | (4.22) | (2.82) | (5.38) | (0.49) | (0.53) | (0.20) |
| TC | 35.54 | 1.5 | 1.86 | 0.21 | –0.12 | 3.32 |
| | (4.22) | (2.82) | (5.38) | (0.49) | (0.53) | (0.20) |
| AT | 29.84 | 0.00 | 1.02 | 0.00 | –0.66 | 3.24 |
| | (4.01) | (2.43) | (3.66) | (0.54) | (0.33) | (0.17) |
| AC | 31.66 | 0.61 | 1.56 | 0.22 | –0.63 | 3.27 |
| | (3.67) | (2.86) | (3.32) | (0.55) | (0.32) | (0,20) |
| GT | 31.66 | –0.61 | 1.56 | –0.22 | –0.63 | 3.27 |
| | (3.67) | (2.86) | (3.32) | (0.55) | (0.32) | (0,20) |
| GC | 33.660 | 0.000 | 1.160 | 0.000 | –0.330 | 3.360 |
| | (5.01) | (3.63) | (4.63) | (0.69) | (0.54) | (0.24) |

Values taken from Table 3.2 of the thesis of Dr. Sreekala Balasubramanian; standard deviation are noted in parenthesis.

The deformation energy of a DNA dimer step is approximated as a harmonic function in terms of the six dimer-step parameters [1]:

$$E = E_0 + \tfrac{1}{2}\sum_{i=1}^{6}\sum_{j=1}^{6} f_{ij}(\theta_i - \theta_{i,0})(\theta_j - \theta_{j,0}). \quad (1)$$

Here, $\theta_{i,0}$ represents the value of one of the six step parameters in the rest state, the average value for a given type of base-pair step collected from the DNA data set mentioned above. $\theta_i$ corresponds to the value of the dimer step in the crystal structure where the DNA sequence is threaded. $f_{ij}$ is a force constant that reflects the spread of structural samples in the dataset and is evaluated with an "inverse harmonic analysis" [1]. That is, the elements of the inverse covariance matrix $F^{-1}$ are computed as $\langle \Delta\theta_i \Delta\theta_j \rangle = \langle \theta_i \theta_j \rangle - \langle \theta_i \rangle \langle \theta_j \rangle$. And $f_{ij}$ is an element of the covariance matrix $F$ obtained by the matrix inversion of $F^{-1}$. The complementary steps share the same set of force constants except for different signs of the force constants linking Shift and Tilt to other parameters (Table 6.2). The signs of the constants linking Shift and Tilt, however, are the same for the two complementary dimer steps, since the product (−1) (−1) is +1.

Table 6.2. Force constants impeding deformations of protein-bound DNA base-pair steps using six parameters

| | CG | CA | TA | AG | GG | AA | GA | AT | AC | GC |
|---|---|---|---|---|---|---|---|---|---|---|
| Twist-Twist | 0.0906 | 0.0637 | 0.0418 | 0.0825 | 0.0686 | 0.0959 | 0.0937 | 0.0744 | 0.0933 | 0.0744 |
| Tilt-Tilt | 0.1081 | 0.1456 | 0.146 | 0.2267 | 0.141 | 0.2008 | 0.1713 | 0.1764 | 0.1404 | 0.1055 |
| Roll-Roll | 0.0342 | 0.0474 | 0.0566 | 0.0818 | 0.0538 | 0.071 | 0.0512 | 0.0854 | 0.111 | 0.0897 |
| Shift-Shift | 2.6647 | 3.259 | 3.1874 | 3.4716 | 3.5001 | 10.334 | 6.1417 | 3.6118 | 3.8202 | 2.8819 |
| Slide-Slide | 4.0303 | 2.0467 | 1.3232 | 6.0451 | 3.3325 | 11.1975 | 5.5592 | 10.0581 | 11.4583 | 6.2131 |
| Rise-Rise | 28.9759 | 18.6505 | 32.4385 | 34.7647 | 27.2818 | 42.6806 | 34.0138 | 42.8782 | 30.9656 | 27.5661 |
| Twist-Tilt | 0 | –0.02 | 0 | 0.0231 | 0.0025 | 0.0101 | 0.0291 | 0 | 0.0206 | 0 |
| Twist-Roll | 0.0147 | 0.0143 | 0.0322 | 0.0197 | 0.0102 | 0.0378 | 0.0308 | 0.0209 | 0.0292 | 0.0271 |
| Twist-Shift | 0 | 0.0109 | 0 | 0.1726 | 0.1408 | 0.2543 | –0.0924 | 0 | –0.0046 | 0 |
| Twist-Slide | –0.0976 | –0.1663 | –0.0337 | 0.0154 | –0.1128 | –0.2225 | –0.2829 | –0.0954 | –0.0851 | –0.1882 |
| Twist-Rise | –0.5387 | –0.2491 | –0.4875 | –0.9122 | –0.5149 | –0.497 | –0.5299 | –0.484 | –0.3944 | –0.6228 |
| Tilt-Roll | 0 | –0.0002 | 0 | 0.0133 | –0.0066 | 0.023 | –0.0044 | 0 | 0.0204 | 0 |
| Tilt-Shift | –0.3357 | –0.3321 | –0.1863 | –0.3317 | -0.3631 | –0.3859 | –0.3785 | –0.1692 | –0.1428 | –0.2939 |
| Tilt-Slide | 0 | 0.0426 | 0 | –0.0386 | 0.0475 | –0.0894 | –0.0262 | 0 | 0.1735 | 0 |
| Tilt-Rise | 0 | 0.3943 | 0 | –1.046 | –0.772 | –0.8929 | –0.7878 | 0 | 0.2848 | 0 |
| Roll-Shift | 0 | 0.0516 | 0 | 0.0365 | 0.0674 | 0.0246 | –0.0494 | 0 | 0.039 | 0 |
| Roll-Slide | –0.0497 | 0.0432 | 0.0076 | –0.154 | 0.0867 | –0.3051 | –0.1618 | 0.1402 | 0.0861 | 0.3491 |
| Roll-Rise | –0.0334 | 0.0554 | –0.1696 | –0.3195 | 0.066 | –0.3582 | 0.1911 | 0.24 | 0.4315 | 0.381 |
| Shift-Slide | 0 | 0.5194 | 0 | 0.5817 | –0.1108 | 1.8953 | 2.4161 | 0 | 1.6503 | 0 |
| Shift-Rise | 0 | –1.0555 | 0 | –0.3897 | 0.5324 | 2.5053 | 4.1598 | 0 | –0.3752 | 0 |
| Slide-Rise | 5.2257 | 2.4422 | 2.2621 | 3.9998 | 3.0822 | 4.2136 | 3.7106 | 5.3999 | 4.9325 | 6.6481 |

This table is obtained from the thesis of Dr. Sreekala Balasubramanian.

### 6.1.2 Threading method

A three-dimensional template is constructured from the set of values of the six dimer-step parameters along a centered DNA fragment of a known nucleosomal crystal structure. As the length of the central DNA fragment varies from 61 bp (60 dimer steps) to 131 bp (130 dimer steps), more and more DNA interactions with histone subunits are included. For example, whereas a 60 dimer step template contains the interactions between two copies of histones H3 and H4 with DNA, almost of all the interactions between DNA and the histone octamer (two copies of H2A, H2B, H3, and H4) are included in a 130 dimer step template. Using the deformation energy function (Equation (1)), the energy cost to deform any template-length DNA sequence from the rest state to a given three-dimensional template can be calculated. The template-length reading frame slides, one base pair at a time, along the DNA, the sequence of which is about 146 bp. The energy cost to deform each base-pair step in a given reading frame along the nucleosomal DNA sequence is collected and referred to as the threading score (Figure 6.1).

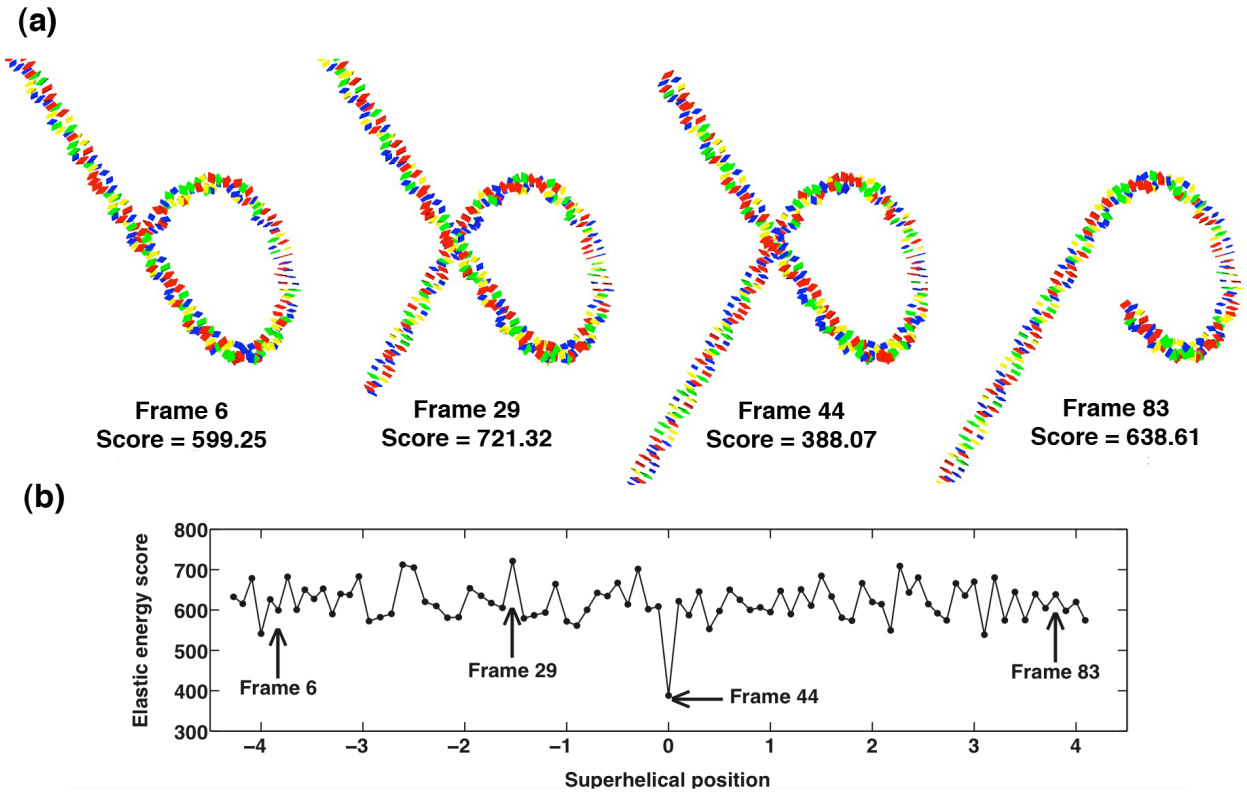



Figure 6.1. A diagram of the threading procedure. (a) Snapshots of the settings of a 146 bp nucleosomal DNA sequence from the wild-type structure (DNA2) on a 60 base-pair step nucleosomal template from the same structure. The three-dimensional template slides one base pair at a time along the DNA. The frame number (6, 29, 44, or 83) indicates where the first base-pair step of the template is positioned in the nucleosomal DNA sequence. (b) Threading score of each reading frame vs. superhelical position along the nucleosomal DNA. The arrows point to the threading scores of the reading frames shown in (a). Frame 44 corresponds to the sequence in its natural (observed) setting on the nucleosome structure.

## 6.2 Results

### 6.2.1 Threading templates of different lengths on the same nucleosomal DNA sequence.

Three-dimensional templates, made of the central 60, 80, 100, or120 dimer steps of the currently best-resolved (1.9 Å) nucleosome core-paritcle structure (PDB ID: 1KX5, NDB ID: PD0287) [3] were threaded on the α-satellite DNA sequence from the same crystal structure. As expected from the added number of base-pair steps, the average threading score increases in values, from 547 to 1056, as the template length increases from 60 to 120 dimer steps. However, the preferred position of the templates on the sequence, showing the lowest threading score compared to any other position of the template along the sequence, remains at SHL 0 for all templates (Figure 6.2). Deformation scores of the base-pair steps in the central setting will be analyzed in detail in the next section, revealing why the templates at the central position have the lowest threading scores. This result suggests that DNA is actually located in the best-fitting position in the crystal structure.

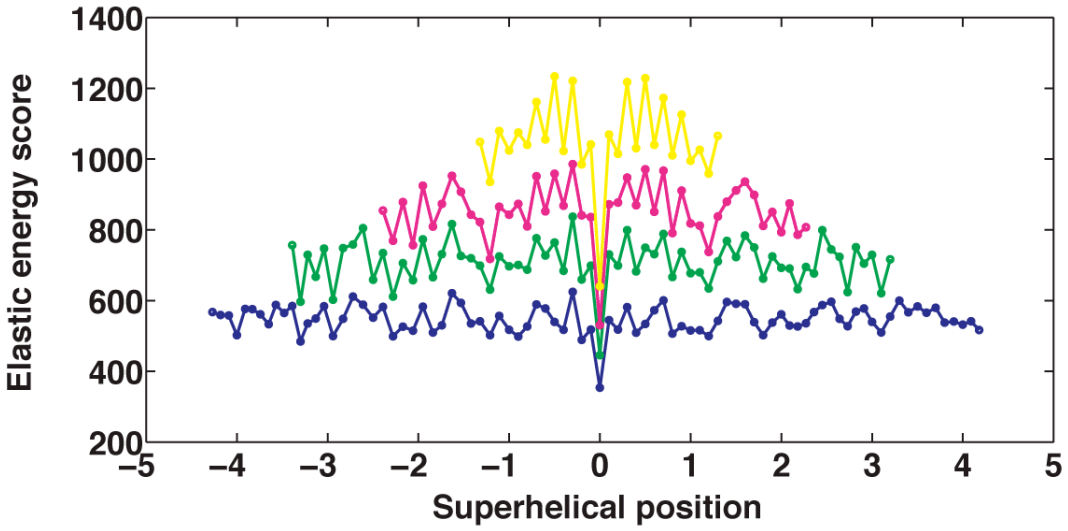



Figure 6.2. Threading score vs. the superhelical position of the template reading frame along the DNA sequence from the best-resolved nucleosome core particle structure (PDB ID: 1KX5) using templates of different length from the same structure. The curves from the bottom to top in the figure correspond to templates of 60, 80, 100, 120 base-pair steps (blue, green, magenta, yellow symbols and lines).

### 6.2.2 Deformation energy score of each base-pair step in every threading frame of the template

The total deformation score of each dimer step in every threading frame is mapped along the DNA sequence from the best-resolved nucleosome (PDB ID: 1KX5) as the 120 base-pair step template slides along the DNA sequence at one base-pair increments (Figure 6.3). Except for a few highly deformed steps at the center of the template (SHL ±0.5, ±1.5, ±2.0), the steps found at the ends of the template (SHL ±5.6, ±4.6, –4.0, ±3.5) contribute most to the high deformation scores when the template is displaced from its natural position on DNA. The steps, which make the largest contribution to the high average deformation scores (larger than 15), are summarized in Table 6.3. The total

score of these nine steps are larger than 15 in 16 of the 26 possible threading frames (60% of the threading frames). Most of the steps, which contribute to the extremely high deformation scores, occur at the ends of the template. The contributions of the six parameters to the deformation scores of these high-energy steps are also listed in Table 6.3. The contributions of Slide and Roll dominate the total deformation scores compared to the other parameters, in agreement with the results of Tolstorukov, *et al*. [4].

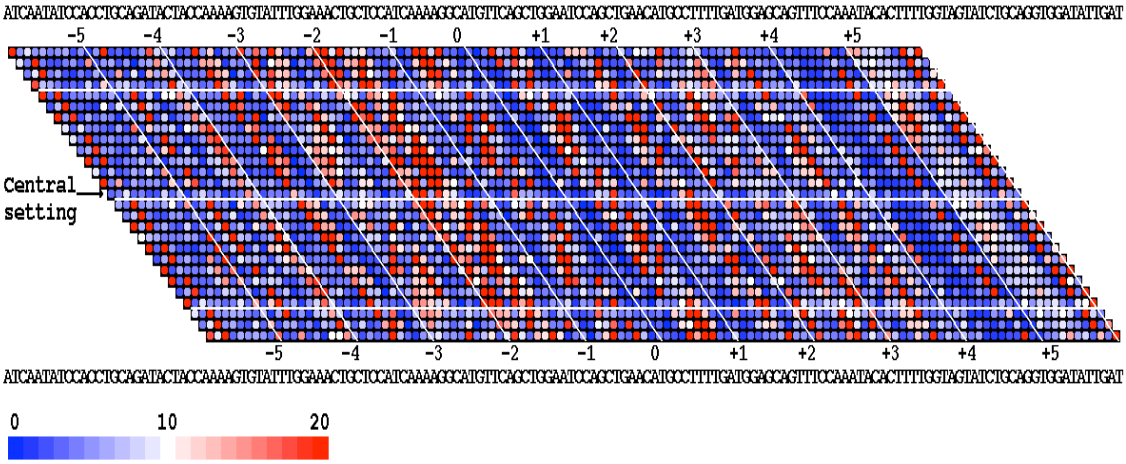



Figure 6.3. Deformation scores of each step in every threading frame are mapped along the DNA sequence from the best-resolved nucleosome structure (PDB ID: 1KX5). The DNA sequence, shown at the top and bottom of the figure, is threaded on the 120 base-pair step template extracted from the same nucleosome structure. The values of the deformation scores are shown in shades of blue (low values), white (medium values), and red (high values). Scores greater than 20 are colored as if they had a score of 20. Note the displacement of the line of scores, i.e., the frame, as the setting of the template is threaded along the DNA sequence. The set of scores associated with the natural central setting (frame 14) of the template on the sequence is denoted with an arrow. Settings displaced by 10 bp with respect of the natural central setting are illustrated by horizontal white lines.

Interestingly, the highly deformed steps listed in Table 6.3 produce relatively low deformation scores when the three-dimensional template is superimposed at the central position of the DNA sequence, the position of the template found in the crystal structure. Except for the AA/TT dimer at step 54, all of the highly deformed steps are taken up by CA/TG dimers in the crystal structure (Table 6.4). Because CA /TG steps are very flexible in terms of the low deformation force constants of the knowledge-based elastic function (Table 6.2), it is easy for those CA/TG dimers to assume the large negative Roll and large positive Slide of the most highly deformed steps of the template in the central setting (Table 6.4).

Table 6.3. Average total scores and the average contribution, over all settings of the 120 base pair step template from the best-resolved nucleosome structure (PDB ID: 1KX5), of individual step parameters to the high deformation scores (larger than 15) consistently found (in 16 out of 26 threading frames) at nine specific steps.

| Step Number | SHL | Contribution of step parameters (diagonal terms only) | | | | | | Total Score |
|---|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist | |
| 16 | -5.6 | 0.03 | 25.17 | 0.27 | 0.63 | 6.58 | 6.82 | 31.25 |
| 26 | -4.6 | 0.24 | 26.26 | 1.59 | 0.58 | 4.41 | 3.37 | 30.55 |
| 32 | -4.0 | 0.35 | 0.33 | 0.13 | 15.14 | 0.54 | 0.64 | 17.82 |
| 38 | -3.6 | 0.55 | 25.62 | 0.13 | 0.46 | 15.29 | 10.69 | 44.33 |
| 54 | -2.0 | 1.44 | 6.27 | 0.09 | 1.04 | 25.84 | 2.69 | 23.04 |
| 94 | 2.0 | 0.05 | 0.40 | 4.07 | 0.44 | 12.71 | 3.80 | 18.01 |
| 109 | 3.5 | 1.81 | 24.57 | 0.07 | 0.42 | 11.44 | 7.26 | 36.28 |
| 121 | 4.6 | 0.04 | 23.20 | 0.12 | 0.75 | 6.38 | 2.62 | 28.65 |
| 131 | 5.6 | 0.19 | 26.42 | 0.35 | 0.79 | 12.31 | 6.31 | 37.56 |

Table 6.4. (a) Values and (b) 'energy' contribution of individual step parameters of the highly deformed steps in the natural central setting of the central 120 base-pair step template from the best-resolved nucleosome (PDB ID: 1KX5).

(a)

| Step Number | Step | SHL | Step Parameters | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Shift (Å) | Slide (Å) | Rise (Å) | Tilt (°) | Roll (°) | Twist (°) |
| 16 | CA | -5.6 | 0.07 | 2.47 | 3.20 | -2.79 | -11.14 | 46.93 |
| 26 | CA | -4.6 | 0.29 | 2.46 | 3.00 | -2.62 | -8.45 | 43.36 |
| 32 | TG | -4.0 | 0.35 | -0.47 | 3.23 | -13.42 | 5.74 | 30.22 |
| 38 | TG | -3.6 | -0.43 | 2.58 | 3.25 | -2.30 | -18.43 | 50.04 |
| 54 | AA | -2.0 | 0.78 | 1.05 | 3.35 | 3.35 | 23.8 | 25.92 |
| 94 | TG | 2.0 | 0.03 | -0.53 | 3.81 | 2.11 | 21.63 | 24.50 |
| 109 | CA | 3.5 | 0.80 | 2.32 | 3.30 | 2.18 | -15.6 | 47.15 |
| 121 | TG | 4.6 | -0.04 | 2.29 | 3.25 | -2.76 | -10.71 | 41.68 |
| 131 | TG | 5.6 | -0.30 | 2.54 | 3.18 | 3.12 | -16.22 | 46.41 |

(b)

| Step Number | Step | SHL | Contribution of step parameters (diagonal terms only) | | | | | | Total Score |
|---|---|---|---|---|---|---|---|---|---|
| | | | Shift | Slide | Rise | Tilt | Roll | Twist | |
| 16 | CA | -5.6 | 0.02 | 5.18 | 0.30 | 0.61 | 6.23 | 4.50 | 8.58 |
| 26 | CA | -4.6 | 0.19 | 5.13 | 1.35 | 0.54 | 4.33 | 2.20 | 7.96 |
| 32 | TG | -4.0 | 0.15 | 0.49 | 0.21 | 12.90 | 0.01 | 0.74 | 15.43 |
| 38 | TG | -3.6 | 0.38 | 5.70 | 0.16 | 0.35 | 13.09 | 7.17 | 12.68 |
| 54 | AA | -2.0 | 2.53 | 8.20 | 0.21 | 1.97 | 18.88 | 4.03 | 22.34 |
| 94 | TG | 2.0 | 0.00 | 0.58 | 1.72 | 0.36 | 6.50 | 3.54 | 8.33 |
| 109 | CA | 3.5 | 1.18 | 4.51 | 0.06 | 0.31 | 10.13 | 4.67 | 10.34 |
| 121 | TG | 4.6 | 0.01 | 4.38 | 0.16 | 0.51 | 5.90 | 1.40 | 6.57 |
| 131 | TG | 5.6 | 0.20 | 5.51 | 0.37 | 0.76 | 10.74 | 4.12 | 12.57 |

The ‘energy’contributions of the six parameters (Shift, Slide, Rise, Tilt, Roll, and Twist) of each dimer step in every threading frame have been mapped along the DNA sequence. Slide and Roll contribute most to the threading scores of the 120 base-pair step template . The contribution of Slide of each dimer step in every threading frame has a pattern very similar to that of the total deformation score (Figure 6.3 and Figure 6.4). That is, the steps with highest total deformation scores and largest contribution of Slide are located not

only at the central 60 steps of the template but also at the two ends of the template. The steps with largest contribution of Roll primarily occur in the same regions, such as SHL ±6, −5, ±3.5, ±2, −1.5, where DNA bends into its major groove or minor groove (Figure 6.5). The contributions of the other step parameters namely Shift, Rise, Tilt, and Twist, are very small compared to that of Slide and Roll (data not shown).

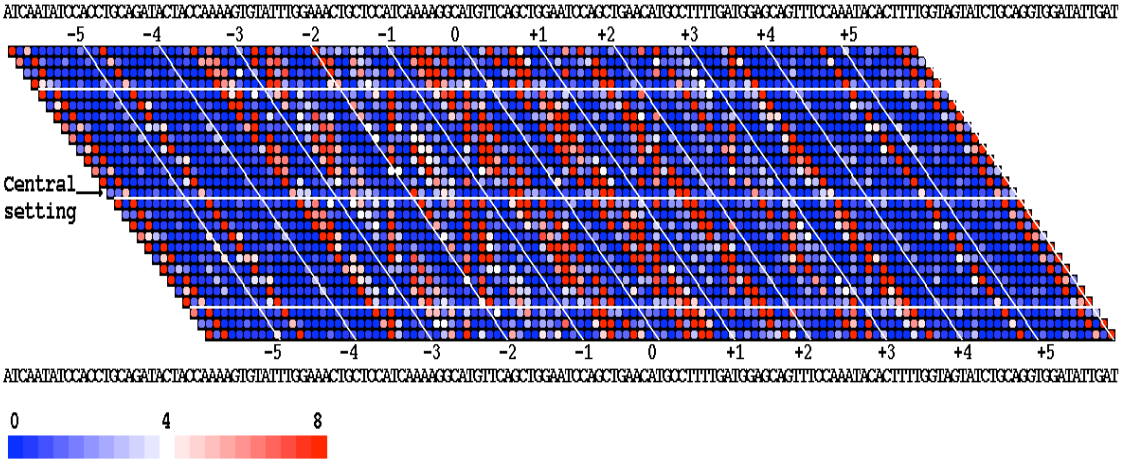



Figure 6.4. Contribution of Slide of each step in every threading frame mapped along the DNA sequence from the best-resolved nucleosome structure (PDB ID: 1KX5). The DNA sequence is threaded on the 120 base-pair step template extracted from the same nucleosome structure. The values of the deformation scores are shown in shades of blue (low values), white (medium values), and red (high values). Scores greater than 8 are colored as if they had a score of 8. See legend in Figure 6.3.

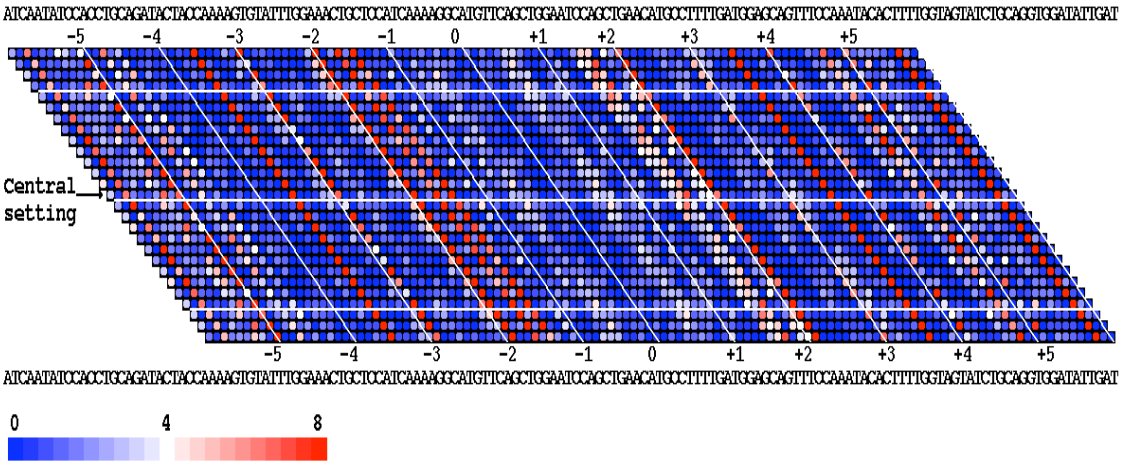



Figure 6.5. Contribution of Roll of each step in every threading frame mapped along the DNA sequence from the best-resolved nucleosome structure (PDB ID: 1KX5). The DNA sequence is threaded on the 120 base-pair step template extracted from the same nucleosome structure. The values of the deformation scores are shown in shades of blue (low values), white (medium values), and red (high values). Scores greater than 8 are colored as if they had a score of 8. See legend in Figure 6.3.

### 6.2.3 Threading the templates extracted from mutant and wild-type nucleosome structures on the same DNA sequence.

All SIN-mutant structures share exactly the same DNA sequence, which has only two base-pair differences from the wild-type (Figure 6.6). The templates extracted from the wild-type and mutant structures have been threaded on the wild-type DNA sequence (DNA2) and the DNA sequence shared by all the mutant structures, respectively. As there are only two base-pair differences between the SIN-mutant DNA sequence and the wild-type, the threading scores of the same template on the two sequences are very similar (Figure 6.7). This chapter reports the results obtained from threading templates on the wild-type DNA sequence. The corresponding results for the SIN-mutant DNA sequence are included in the appendix.

```
   DNA2   5' ATCAATATCC ACCTGCAGAT TCTACCAAAA GTGTATTTGG 3'
SIN mutants 5' ATCAATATCC ACCTGCAGAT TCTACCAAAA GTGTATTTGG 3'

   DNA2   5' AAACTGCTCC ATCAAAAGGC ATGTTCAGCT GAATTCAGCT 3'
SIN mutants 5' AAACTGCTCC ATCAAAAGGC ATGTTCAGCG GAATTCCGCT 3'

   DNA2   5' GAACATGCCT TTTGATGGAG CAGTTTCCAA ATACACTTTT 3'
SIN mutants 5' GAACATGCCT TTTGATGGAG CAGTTTCCAA ATACACTTTT 3'

   DNA2   5' GGTAGAATCT GCAGGTGGAT ATTGAT 3'
SIN mutants 5' GGTAGAATCT GCAGGTGGAT ATTGAT 3'
```

Figure 6.6. Literal alignment of the wild-type (DNA2) and the SIN-mutant nucleosomal DNA sequences. The bases that differ in the two sequences are highlighted in red.

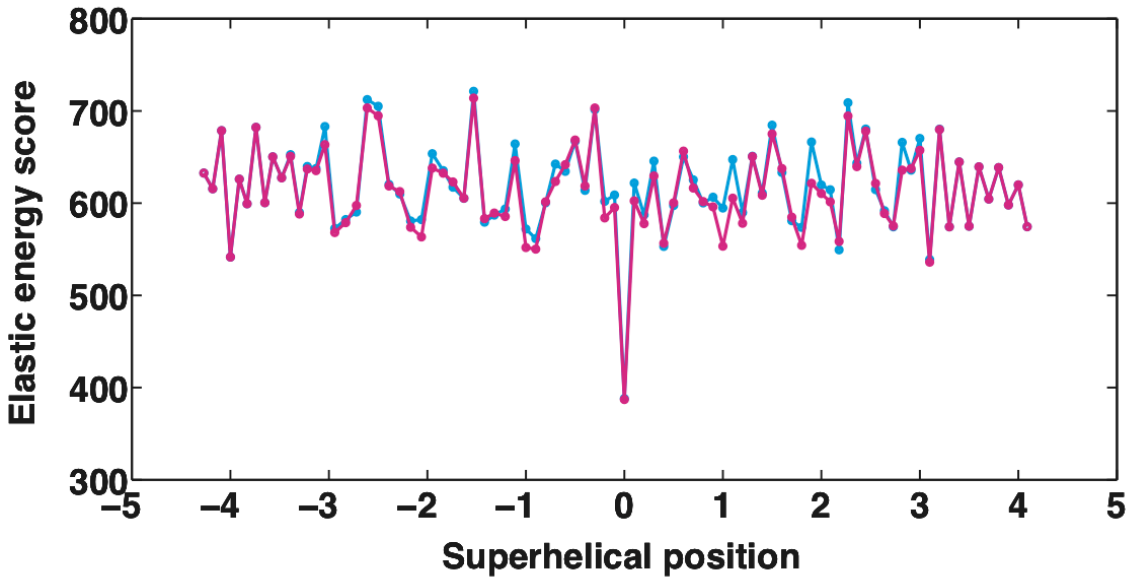



Figure 6.7. The threading scores of the wild-type template threaded on the DNA sequence from the wild-type nucleosome structure (shown in cyan) and the sequence from the SIN-mutant structures (shown in magenta). The threading score is plotted against the superhelical position of each template reading frame along the nucleosomal DNA. The template is obtained from the central 60 base-pair steps of the wild-type structure (DNA2).

Like the results obtained above for the best-resolved (147 bp) nucleosome core-particle structure (PDB ID: 1KX5), the original positioning of the (146 bp) wild-type and mutant

sequences on the wild-type X-ray crystal structure (the central position) has the lowest threading score.

The mutant structures fall into three groups according to the average threading scores. The first group, which includes the V43I H4 and R116A H3 mutants and the wild-type (DNA2) structures, have low average threading scores and are very close to each other (blue curves in (Figure 6.8 and Appendix Figure A6.1). The second group with intermediate threading scores (green curves in Figure 6.8 and Appendix Figure A1) includes the R116H H3, R45E H4, V43A H4, and T118A H3 mutants. The third group (red curves in Figure 6.8 and Appendix Figure A1), which includes T118A H3, T118I H3, R45A H4, R45C H4, and R45H H4 mutants, has much higher threading scores than the wild type. Except for R45E, all the mutants with high sliding rates in the heat-shifting experiments (Table 2.2 in Chapter 2), T118H, T118I, R45A, R45C, and R45C, belong to the high-threading-score group. This trend remains the same when the template length increases from 60 dimer steps to 130 dimer steps (Figure 6.9 and Figure 6.10 and Appendix Figure A6.2).

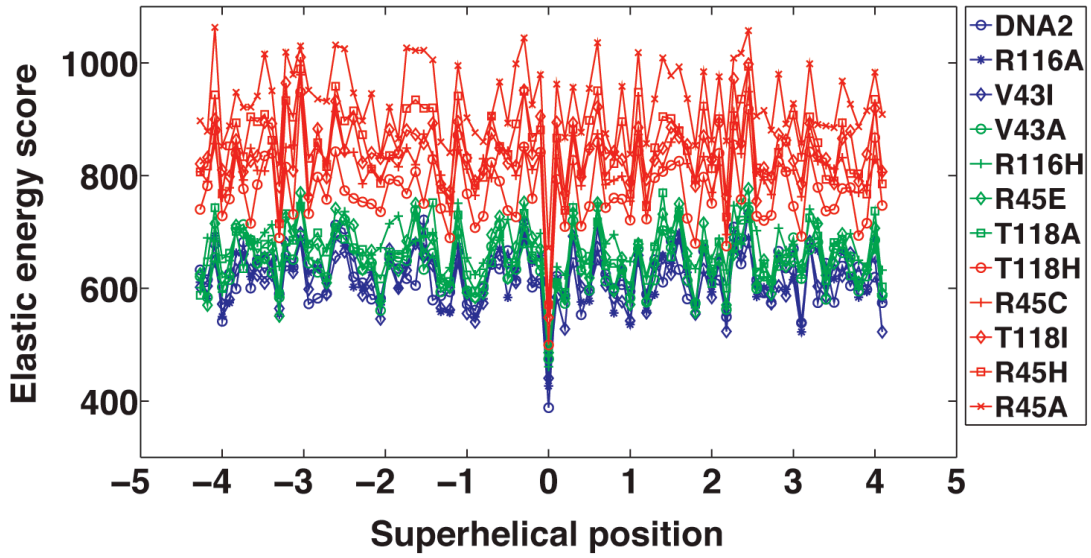



Figure 6.8. Threading score vs. superhelical location of the template reading frame along the DNA sequence from the wild-type nucleosome structure (DNA2). The templates are extracted from the wild-type and SIN-mutant nucleosome structures, and designated by the point-mutations in the legend.

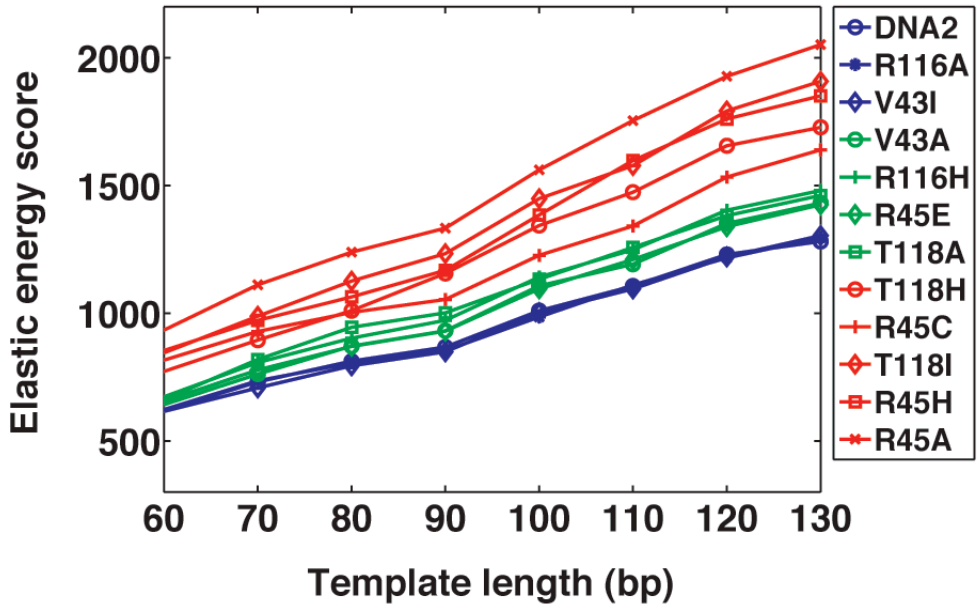



Figure 6.9. Average threading scores of different structural templates on the wild-type (DNA2) sequence vs. the length of the templates. The templates are extracted from the wild-type and SIN-mutant nucleosome structures, and designated by the point-mutations in the legend.

The threading scores, obtained with the 60-dimer-step templates from the wild-type and SIN-mutant structures, are plotted for each reading-frame position in Figure 6.8 and Appendix Figure A2. The curves of the high-scoring group, T118H, T118I, R45A, R45C, and R45C, are consistently shifted above the curve for the wild-type structure (DNA2), mimicking an excited energy state. Interestingly, the lowest threading scores at the central positions of this high-scoring group are comparable to the threading scores of the poorly positioned reading frames of the wild-type nucleosome (Figure 6.9 and Figure 6.10).

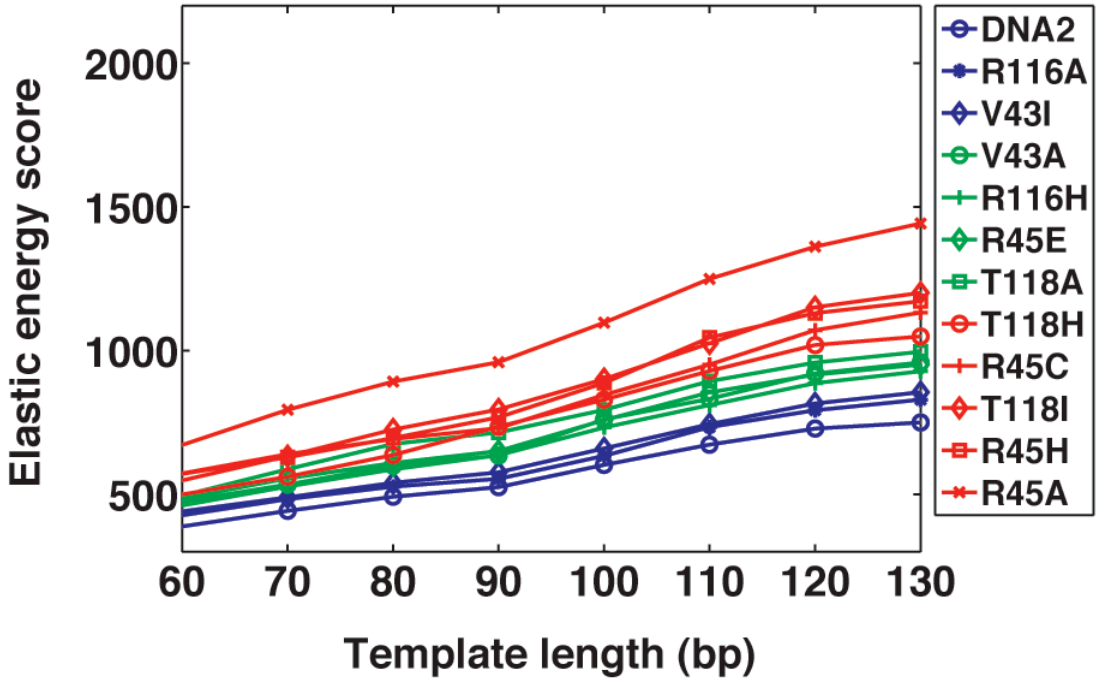



Figure 6.10. Minimal threading scores of different structural templates on the wild-type (DNA2) sequence vs. the length of the templates. The templates are extracted from the wild-type and SIN-mutant nucleosome structures, and designated by the point-mutations in the legend.

This feature may explain why the high-scoring group-T118H, T118I, R45A, R45C, and R45C-exhibits high sliding rates in heat-shifting experiments [5]. That is, even at the

best-fitting position (the central position) the SIN-mutant structures in the high-scoring group are not as well positioned as the wild-type and thus may be prone to slide along DNA at higher rates.

### 6.2.4 Contribution of the six step parameters to the threading scores of the SIN-mutant structures compared with the wild-type structure.

The average total threading energy scores are broken down in Table 6.5 into the different average contributions from individual step parameters. For example, the average contribution of Slide includes the cost of Slide alone (a diagonal term in the elastic function) and one half of the terms where Slide is coupled to the other five step parameters (cross terms in the elastic function). The average contribution of each step parameter from all reading frames of the central 60 dimer step-template threaded along the wild-type DNA sequence is listed in Table 6.5 (a). In all the structures, including the mutants and wild type, the deformation scores related to Slide are much higher than those associated with the other five parameters. The Slide terms are also relatively similar in all structures. The deformations in Roll and Rise also contribute appreciably to the total threading scores. At the same time the values of the Roll and Rise terms are clearly higher in the high-scoring group (T118H, T118I, R45A, R45C, and R45H) than in the low-threading-score group (DNA2, R116A, and V43I). The threading contribution of Tilt shows the most significant difference among threaded nucleosomes. The values of the Tilt terms are doubled or even tripled in the high-threading-score group compared to the low-threading-score group. The contributions of each step parameter alone (the diagonal terms) are reported in Table 6.5 (b). The diagonal terms of Slide, Roll, and Rise clearly play a primary role compared to the cross terms because of the similarity of the values in

Table 6.5 (a) and (b). Although the diagonal terms of Tilt, Shift and Twist are relatively large, the total contributions of Tilt, Shift and Twist are lowered by the negative cross terms. That is to say, the coupling of Tilt, Shift and Twist with other step parameters reduces the cost of deformation. Similar scoring trends occur with longer templates of 80, 100, and 120 dimer steps (Appendix Tables A6.1 to A6.3). The contribution of the cross terms and diagonal terms to the threading scores of the SIN-mutant templates on the SIN-mutant DNA sequence are reported in Appendix Tables A6.4 to A6.7.

Table 6.5. Average contribution of individual base-pair step parameters to the threading of the central 60 dimer steps from the different nucleosome structures on the DNA sequence from the wild-type nucleosome structure (DNA2). (a) The sum of the diagonal terms (deformations of step parameters alone) and the cross terms (deformations of a step parameter coupled with the other step parameters). (b) Diagonal terms only.

(a)

| Group | Structure ID | Contribution of the step parameters | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist | |
| Low | DNA2 | 59.45 | 179.97 | 119.94 | 40.88 | 139.07 | -2.57 | 619.03 |
| | R116A | 55.80 | 209.57 | 94.62 | 54.23 | 125.78 | 14.87 | 622.71 |
| | V43I | 54.16 | 221.46 | 99.85 | 57.33 | 103.07 | -10.62 | 615.98 |
| Medium | R116H | 68.68 | 183.45 | 108.16 | 82.24 | 145.04 | 21.63 | 672.54 |
| | T118A | 58.93 | 174.95 | 96.52 | 51.19 | 136.20 | 20.43 | 661.76 |
| | R45E | 37.54 | 224.35 | 118.73 | 75.26 | 113.88 | 11.89 | 656.41 |
| | V43A | 84.24 | 196.59 | 103.28 | 46.89 | 119.93 | 9.54 | 641.93 |
| High | T118H | 62.72 | 234.15 | 154.19 | 91.98 | 160.19 | -0.91 | 772.09 |
| | T118I | 69.34 | 236.28 | 177.56 | 123.59 | 153.71 | 22.33 | 845.02 |
| | R45A | 56.33 | 187.10 | 192.07 | 161.35 | 236.56 | 37.44 | 934.09 |
| | R45C | 70.19 | 250.05 | 172.85 | 107.07 | 155.04 | 18.06 | 816.34 |
| | R45H | 55.51 | 237.17 | 129.39 | 123.06 | 228.31 | -7.15 | 854.08 |

(b)

| Group | Structure ID | Contribution of the step parameters (diagonal terms only) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist |
| Low | DNA2 | 90.65 | 186.76 | 114.6 | 74.92 | 163.71 | 70.66 |
| | R116A | 94.55 | 208.22 | 80.48 | 94.68 | 140.34 | 72.28 |
| | V43I | 98.33 | 220.25 | 93.01 | 102.85 | 123.71 | 68.57 |
| Medium | R116H | 98.31 | 188.86 | 105.11 | 114.14 | 153.6 | 75.88 |
| | T118A | 123.15 | 190.11 | 90.81 | 118.45 | 163.27 | 99.52 |
| | R45E | 80.69 | 220.09 | 100.55 | 119.27 | 131.48 | 79.09 |
| | V43A | 118.92 | 199.05 | 102.16 | 82.48 | 135.52 | 85.27 |
| High | T118H | 101.89 | 226.24 | 149.99 | 132.39 | 170.99 | 60.36 |
| | T118I | 106.9 | 232.02 | 156.77 | 162.82 | 166.94 | 81.78 |
| | R45A | 84.77 | 199.85 | 191.71 | 191.94 | 242.6 | 86.46 |
| | R45C | 74.02 | 252.37 | 169.76 | 112.49 | 162.99 | 87.8 |
| | R45H | 96.65 | 228.91 | 124.58 | 166.42 | 247.7 | 77.61 |

6.2.5 Contributions of Roll, Tilt, and Rise of each dimer step in the templates from the SIN mutants compared to the wild type (DNA2).

The DNA sequence from the wild-type structure (DNA2) was threaded on the central 120 base-pair templates from one of the SIN-mutant structures (structure ID: R45H, PDB ID: 1P3I) and that from the wild-type, respectively. The difference in total deformation scores at each base-pair step in the two templates (the R45H H4 mutant compared to the wild-type structure), are mapped on each step, as the frames of the templates slide along the DNA sequence (Figure 6.11). The color-coded values describe lines in Figure 6.11. That is to say, wherever the template is positioned along the sequence, some steps in the R45H H4 mutant template always raise the deformation score compared to the corresponding steps in the wild-type template. The base-pair steps in the R45H H4 mutant with higher deformation scores are clustered at the two ends of the sequence (SHL −6 to −5 and around +5) and near the dyad (SHL −1.5).

The differences in the total scores are decomposed further into the differences in the contributions of Rise, Roll, Tilt (Figure 6.12, 6.13, and 6.14), and other step parameters (data not shown). There are three successive steps near SHL −5 in the H4 R45H mutant with values of Rise that cause much higher deformation scores than the corresponding steps in the wild type. Similarly, the values of Tilt of several steps near SHL −5 in the R45H H4 mutant have higher deformation scores than the wild type.

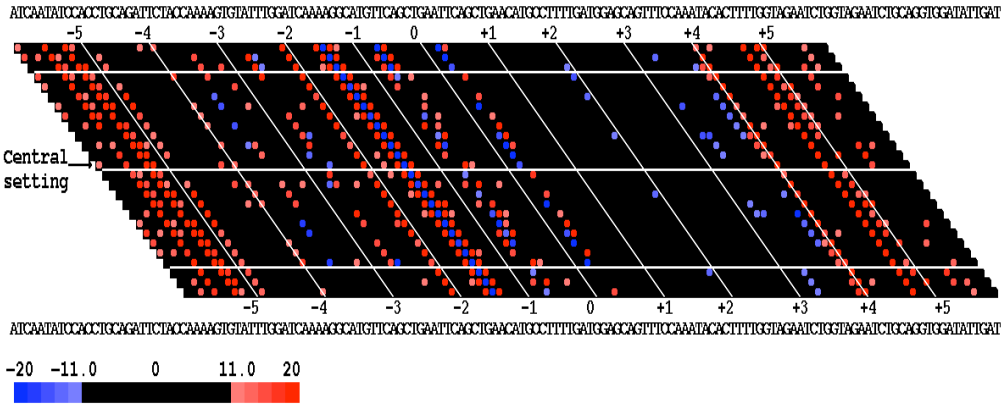



Figure 6.11. Difference of total deformation scores of a central 120 base-pair step structure from a SIN-mutant (structure ID: R45H, PDB ID: 1P3I) compared to the wild-type structure (DNA2) are mapped on each base-pair step along the wild-type DNA (DNA2). The templates slide along the DNA sequence from the wild type structure (DNA2) one base pair at a time. The difference in value of the mutant vs. the values of wild type is shown in shades of blue (negative) and red (positive). Differences from −11 to +11 are shown in black. The natural central setting (frame 13) is denoted by an arrow. See legend to Figure 6.3.

Except for the highly deformed steps at SHL −5, the steps with high contributions of Rise arer located at SHL −2, −1.5, +4, +4.5, +5.5. Steps with high Tilt contributions also occur at SHL −3 to −2, −1.5 and +4.5 to +5. The contribution of Roll of some base-pair steps in the R45H H4 mutant template are either much lower or much higher than the corresponding steps in the wild type. These steps are located approximately at sites where the DNA minor groove or major groove bending inward (SHL −5, −4.5, −3, −2.5, −1.5, −0.5, +3, +4, +5). Noticeably, the steps with decreased contributions of Roll compared to the wild type are primarily located in the first half of the DNA template (SHL −6 to 0). It seems that the Tilt and Rise values of the steps at the two ends of DNA (SHL −6 to −5 and +5 to +6) contribute most to the total threading scores of the R45H H4 mutant compared to the wild type.

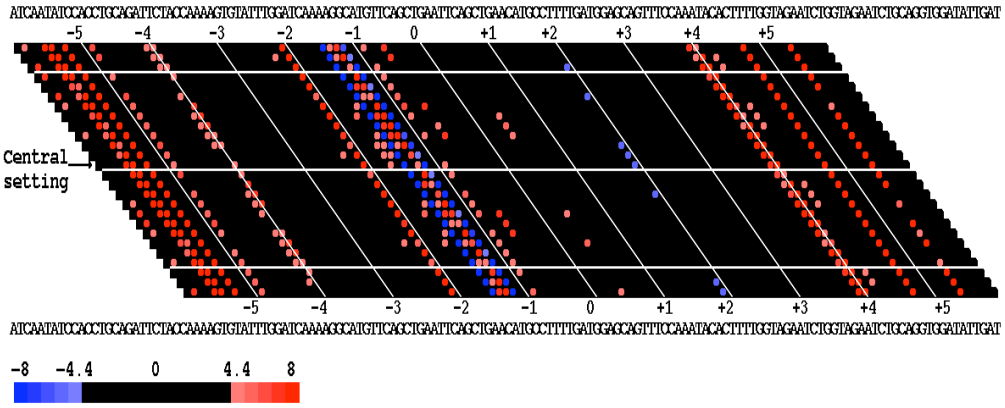



Figure 6.12. Difference of the contribution of Rise of a central 120 base-pair step template from a SIN-mutant (structure ID: R45H, PDB ID: 1P3I) compared to the wild-type structure (DNA2). Values are mapped on each base-pair step along the DNA sequence from the wild-type (DNA2). The templates slide along the DNA sequence from the wild type structure (DNA2) one base pair at a time. The difference in values of the mutant vs. the values of wild type is shown in shades of blue (negative) and red (positive). Difference from −11 to +11 are shown in black. The natural central setting (frame 13) is denoted by an arrow. See legend to Figure 6.3.

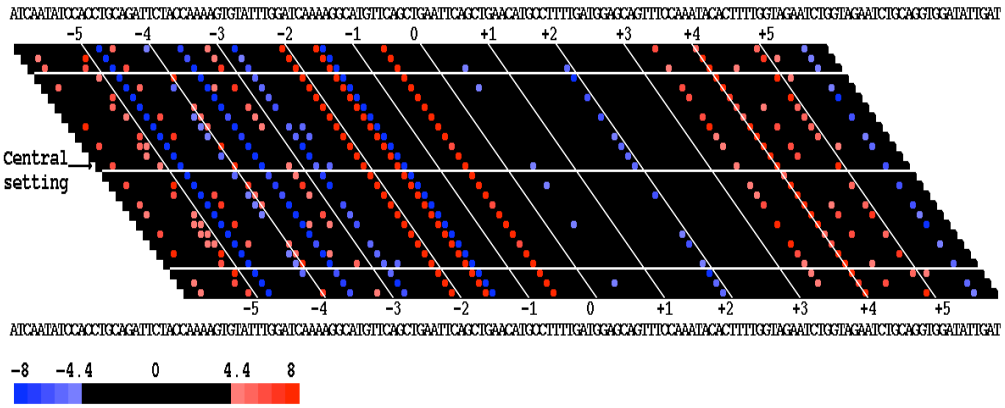



Figure 6.13. Difference of the contribution of Roll of a central 120 base-pair step template from a SIN-mutant (structure ID: R45H, PDB ID: 1P3I) compared to the wild-type structure (DNA2). Values are mapped on each base-pair step along the DNA sequence from the wild-type (DNA2). The templates slide along the DNA sequence from the wild type structure (DNA2) one base pair at a time. The difference in values of the mutant vs. the values of wild type is shown in shades of blue (negative) and red (positive). Difference from −11 to +11 are shown in black. The natural central setting (frame 13) is denoted by an arrow. See legend to Figure 6.3.

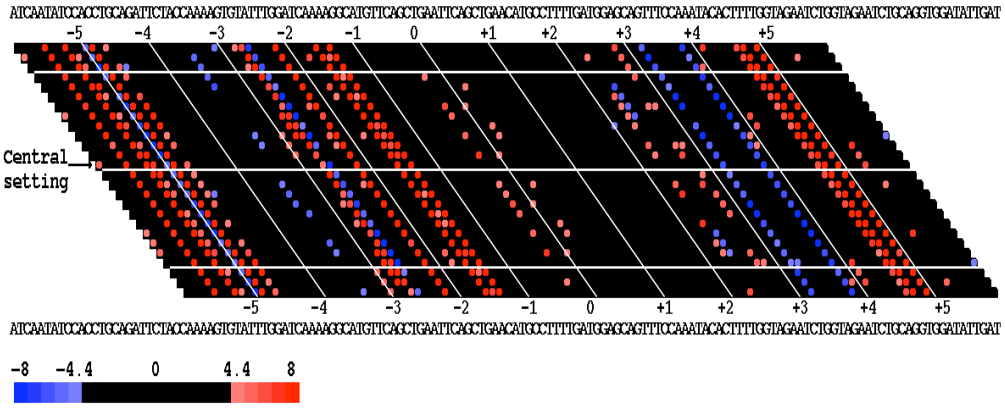



Figure 6.14. Difference of the contribution of Tilt of a central 120 base-pair step template from a SIN-mutant (structure ID: R45H, PDB ID: 1P3I) compared to the wild-type structure (DNA2). Values are mapped on each base-pair step along the DNA sequence from the wild-type (DNA2). The templates slide along the DNA sequence from the wild type structure (DNA2) one base pair at a time. The difference in values of the mutant vs. the values of wild type is shown in shades of blue (negative) and red (positive). Difference from −11 to +11 are shown in black. The natural central setting (frame 13) is denoted by an arrow. See legend to Figure 6.3.

The steps on the different SIN-mutant templates, which always have much higher or lower average deformation scores than the corresponding steps in the wild-type template (where the difference is larger than 11 or less than –11, in 25 out of 26 threading frames on the 120 bp template), are summarized in Figure 6.15. The steps with high deformation scores are located primarily in two regions, near one end of DNA (SHL –6 to –4) and near the dyad (SHL –2 to –1). The steps with higher deformation scores are asymmetrically distributed in the first half of the DNA (SHL –6 to 0).

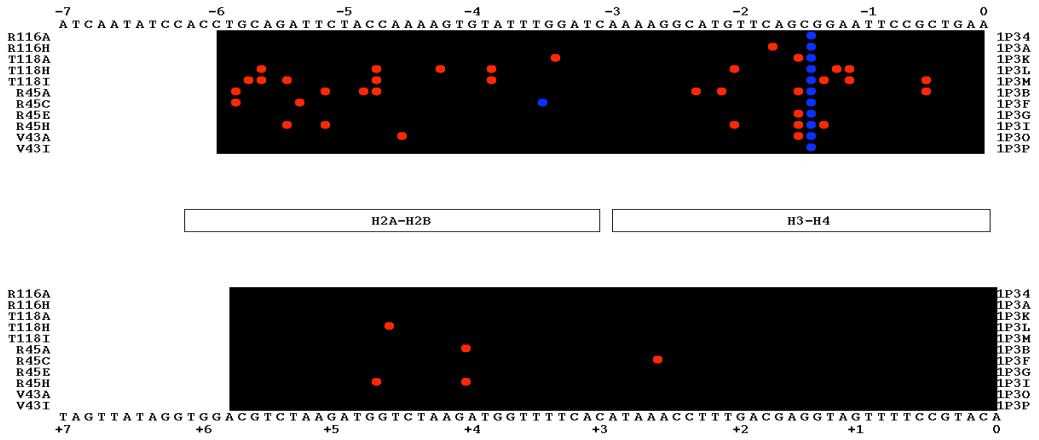



Figure 6.15. The base-pair steps on the central 120 dimer step templates from SIN mutant DNA structure, which almost always (25 out of 26 threading frames) introduce large difference in the deformation scores compared to that obtained with the wild-type structure (DNA2). The locations of these steps in the template are mapped along the DNA sequence. The threaded DNA sequence is from the wild type (DNA2). The PDB IDs of SIN mutants are shown at the right edge of the diagram. The structure ID is designated by point-mutations shown at the left edge. The steps with large differencess of deformation scores are shown in red (11 or larger) and blue (−11 or less), respectively.

### 6.2.6 Remodel the wild-type nucleosomal DNA according the values of step parameters of the highly deformed steps found in the SIN mutant structure.

The values of individual step parameters and the average contributions of the steps, which have high deformation scores on the template from a SIN-mutant structure (R45H H4 SIN mutant) (PDB ID: 1P3I), are listed in Table 6.6. We use these steps to simulate the remodeling of the wild-type nucleosomal DNA. The step parameters, which contribute most the the high deformation scores, are used to remodel the wild-type nucleosomal DNA. That is, we introduce the specific step parameters from the R45H H4 SIN-mutant structure in place of the steps in the wild-type structure and generate a computed model of the modified nucleosome. In this way, it is easy to see the global impact of the variation of the specific step parameters on nucleosomal DNA geometry.

We first change the Tilt and Rise of dimer step 18 GA/TC at SHL −5.4 in wild-type DNA from their original values (2.67 ° and 3.42 Å) to the values of the corresponding step from the R45H H4 SIN mutant (10.83 ° and 4.15 Å) (Figure 6.16 (a) and Table 6.7). The large positive Tilt coupled with the increase in Rise makes one end of the DNA peel off the histone octamer and flip out, increasing the root-mean-square deviation of the base pairs at one end of the remodeled DNA compared to the wild-type structure (Figure 6.16 (b)). As the changed step just affects one end of the remodeled DNA, the RMSD value averaged over the146 bp remodeled DNA compared to the wild-type DNA is small (Table 6.7 column 5). The whole remodeled DNA has also been fitted to the wild-type structure with a least-squares fitting procedure [6] to optimize the RMSD value. The optimized RMSD value is similar to the un-optimized value (compare Table 6.7 columns 5 and 6), suggesting the major part of the remodeled DNA is similar to the wild-type structure except for the large deviation at the end of DNA (up to 8.00 Å in Figure 6.16 (b)).

The large increases in Roll and Rise of dimer step 53 CA/TG, which is located at SHL −2.1 in the R45H H4 mutant, increase the DNA bending and appear to stretch one end of the DNA upon remodeling (Figure 6.17 (a) and Table 6.7). The remodeled DNA differs substantially from the wild-type structure with large RMSD values of 6.6 Å and deviates as much as 13 Å between base pairs at SHL −7 to −5 ((Figure 6.17 (b)). The RMSD decreases to 4.6 Å if the remodeled DNA as a whole is fitted to the wild-type structure. This result suggests that the general shape of the superhelical turns of the remodeled DNA changes a lot compared to the wild-type.

The extremely negative Roll and Tilt at SHL −1.5 (dimer step 58 GG/CC) change the DNA superhelical pathway even more dramatically, with the deviation of some base pairs as high as 51.08 Å (Figure 6.18). The simulated changes, however, introduce steric clashes with the histone core. The wild-type nucleosomal DNA has also been remodeled according to the other highly deformed steps (dimer step 21, 60, 121) in Table 6.7, the results of which are shown in Appendix Figure A6.3 to A6.5. The increased Roll and Rise at dimer step 21 at SHL −5.1 makes one end (SHL −5 to −7) closer to the second half of nucleosomal DNA. The increase Tilt at SHL −1.3 alone makes the first half of DNA bulge out. The decreased Twist and Rise at SHL + 4.7 have similar impact as the increased Tilt and Rise at SHL −5.1on nucleosomal DNA (compared Figure 6.16 to Appendix Figure A6.5).

The six step parameters of all five highly deformed steps in the first half (SHL −7 to 0) of the R45H H4 SIN mutant have been used together to remodel the wild-type nucleosome DNA. Such a model helps to see the total impact of these highly deformed steps on the wild-type nuclesomal DNA. The combined variation of base-pair step parameters at SHL −5.4, −5.1 and −2.1 and −1.3 with the extremely negative Roll and Tilt at SHL −1.5 brings more distortion of the nucleosomal DNA (24.25 Å RMSD) than the changes in Roll and Tilt at SHL −1.5 alone (19.8 Å RMSD) (Figure 6.19 and Table 6.8). The remodeled DNA, which has been fitted to the wild-type by least-squares fitting, is superimposed onto the wild-type DNA in Figure 6.20 (a). After least-squares fitting, the base pairs with the largest deviation (around 24 Å) occur at SHL −6 to −5 and +2 to +1 rather than SHL −5 found for the remodeled DNA without least-square fitting (compare

Figure 19 (b) and Figure 20 (b)). The largest deviation of each base pair in the remodeled DNA decreases from 51.08 Å to 25.00 Å. Correspondingly, the total RMSD of the remodeled DNA decreases by one third after the least-squares-fitting procedure (24.25 Å vs. 16.80 Å). The RMSD (16.80 Å) of the remodeled DNA with all the highly deformed steps is less than that of the DNA subject to negative Roll and Tilt at SHL −1.5 (19.84 Å) but more than that found after least-squares fitting of that model (13.65 Å). This result suggests that the variation of base-pair step parameters at other steps may compensate for the extreme value of Roll and prevent steric clashes of the histone core. Interestingly, after the least-squares fitting of the remodeled DNA to the wild-type structure, the base pairs with exceptionally low RMSD (around 2.5 Å) occur at SHL −1.5 and +6, suggesting that these steps may serve as anchor points of the histone core to the nucleosomal DNA (Figure 6.20).

The superhelical properties of the remodeled DNA compared to the wild-type DNA are examined. The wild-type DNA is transformed into a cylindrical frame using the procedure described in Chapter 5. The remodeled DNA, which has already fitted to the wild type, are translated and rotated using the same matrices obtained from the analysis of the wild-type DNA. In this way, the change of superhelical shape of the remodel DNA is shown clearly compared to the wild-type DNA. The superhelices of the part of the remodeled DNA (SHL −7 to +1) change from left-hand to right-hand (Figure 6.21). The superhelical pitch of the first half of the remodeled DNA almost diminishes (Figure 6.22, left part). In concert with the flattened superhelices of the first half of DNA, the central part of the remodeled DNA (SHL 0 to +2) lifts up (Figure 6.22, right part). One end of the remodeled DNA (SHL −7 to −5) stretches away from the core histone with increased

radii (Figure 6.20, Figure 6.23, Figure 6.24, left part). Two regions (SHL −5 to −2 and −1 to +2) of the remodeled DNA shrink inward (Figure 6.24, left part) with decreased radii and one region (SHL +2 to +6) bulges out (Figure 6.24, right part).

The simulated DNA remodeling suggests that DNA conformational changes near SHL −5 and −1.5 in the SIN mutants may alter the ramp of DNA wrapping around the histone octamer.

Table 6.6. (a) Values and (b) average contribution, over all settings of the 120 base pair step template from a SIN-mutant structure (PDB ID: 1P3I, structure ID: R45H) on DNA, of individual step parameters to the high deformation scores (larger than 11) consistently found (in 25 out of 26 threading frames) at specific steps.

(a)

| Step Number | Step | SHL | Step Parameters | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Twist (°) | Tilt (°) | Roll (°) | Shift (Å) | Slide (Å) | Rise (Å) |
| 18 | GA | -5.4 | 36.27 | 10.83 | -2.40 | -0.95 | -0.77 | 4.15 |
| 21 | TC | -5.1 | 23.17 | 20.06 | 9.73 | 1.27 | 0.75 | 1.87 |
| 53 | CA | -2.1 | 37.64 | -3.84 | 27.02 | -0.49 | 0.37 | 4.33 |
| 58 | GG | -1.5 | 39.05 | -17.13 | -57.21 | 0.22 | 1.90 | 3.73 |
| 60 | CA | -1.3 | 46.83 | 16.82 | 11.60 | -0.89 | 0.92 | 3.89 |
| 114 | CA | 4.0 | 43.73 | -3.44 | 21.85 | 0.57 | 0.06 | 4.18 |
| 121 | GG | 4.7 | 21.52 | -19.80 | -0.86 | -0.01 | 0.54 | 2.07 |

(b)

| Step Number | SHL | Contribution of step parameters (diagonal terms only) | | | | | | Total Score |
|---|---|---|---|---|---|---|---|---|
| | | Twist | Tilt | Roll | Shift | Slide | Rise | |
| 18 | -5.4 | 1.54 | 10.34 | 0.50 | 1.63 | 0.06 | 17.75 | 29.10 |
| 21 | -5.1 | 7.17 | 29.50 | 1.59 | 3.45 | 2.10 | 35.76 | 44.59 |
| 53 | -2.1 | 0.21 | 0.47 | 16.21 | 0.24 | 0.67 | 17.35 | 40.73 |
| 58 | -1.5 | 0.51 | 21.09 | 91.93 | 0.05 | 2.89 | 1.14 | 112.10 |
| 60 | -1.3 | 8.51 | 28.19 | 2.28 | 0.85 | 4.65 | 5.07 | 53.61 |
| 114 | 4.0 | 2.41 | 0.92 | 6.67 | 0.63 | 0.03 | 5.97 | 17.82 |
| 121 | 4.7 | 5.82 | 28.86 | 0.83 | 0.00 | 0.10 | 16.00 | 53.32 |

Table 6.7. The altered step parameters of the modified base-pair steps and root-mean-square deviations of different remodeled nucleosomal DNA compared to the wild-type nucleosomal DNA (DNA2) are listed.

| Base pair step | SHL | Models | Parameters | | RMSD [a] (Å) | RMSD [b] (Å) (optimized) | Figure[c] number |
|---|---|---|---|---|---|---|---|
| | | | Tilt (°) | Rise (Å) | | | |
| 18 GA/TC | −5.4 | Wild type | 2.76 | 3.42 | 1.72 | 1.60 | Figure 6.16 |
| | | Remodeled | 10.83 | 4.15 | | | |
| | | | Tilt (°) | Rise (Å) | | | |
| 21 TC/GA | −5.1 | Wild type | 9.36 | 2.73 | 2.51 | 2.06 | Appendix Figure A6.3 |
| | | Remodeled | 20.06 | 1.87 | | | |
| | | | Roll (°) | Rise (Å) | | | |
| 53 CA/TG | −2.1 | Wild type | 17.15 | 3.94 | 6.60 | 4.61 | Figure 6.17 |
| | | Remodeled | 27.02 | 4.33 | | | |
| | | | Tilt (°) | Roll (°) | | | |
| 58 GG/CC | −1.5 | Wild type | -9.06 | -31.16 | 19.84 | 13.65 | Figure 6.18 |
| | | Remodeled | -17.13 | -57.21 | | | |
| | | | Tilt (°) | | | | |
| 60 CA/TG | −1.3 | Wild type | 8.96 | | 5.83 | 3.93 | Appendix Figure A6.4 |
| | | Remodeled | 16.82 | | | | |
| | | | Twist (°) | Rise (Å) | | | |
| 121 GG/CC | 4.7 | Wild type | 26.70 | 2.61 | 1.26 | 1.05 | Appendix Figure A6.5 |
| | | Remodeled | 21.52 | 2.07 | | | |

a. Root-mean-square deviation (RMSD) of the remodeled structure compared to the wild type structure (DNA2). The 5′-terminus of the first five remodeled structures and the 3′-terminus of the last remodeled structure are displaced from the wild-type structure. The RMSD calculation is performed over all 146 bp steps.

b. The optimized RMSD obtained by using a least-squares fitting procedure to fit the entire remodeled DNA to its original state found in the wild-type nucleosome structure (DNA2)

c. Three-dimensional image of the different remodeled nucleosomal DNA structures are shown in the figures listed in this column.

Table 6.8. The altered step parameters of the modified base-pair steps and root-mean-square deviations (RMSD) of the remodeled DNA compared to the original state found in the wild-type nucleosome structure (DNA2). Three-dimensional representations of this remodeled DNA, whose view points are obtained without or with least-square fitting, are shown in Figures 6.19 and 6.20, respectively.

| Step Number | Step | SHL | Models | Step parameters | | | | | | RMSD [a] (Å) | RMSD [b] (optimized) (Å) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Twist (°) | Tilt (°) | Roll (°) | Shift (Å) | Slide (Å) | Rise (Å) | | |
| 18 | GA/TC | −5.4 | Wild type | 36.72 | 2.76 | −4.88 | −0.28 | −0.81 | 3.42 | 24.25 | 16.80 |
| | | | Remodeled | 36.27 | 10.83 | −2.40 | −0.95 | −0.77 | 4.15 | | |
| 21 | TC/GA | −5.1 | Wild type | 26.15 | 9.36 | 9.82 | −0.22 | −0.42 | 2.73 | | |
| | | | Remodeled | 23.17 | 20.06 | 9.73 | 1.27 | 0.75 | 1.87 | | |
| 53 | CA/TG | −2.1 | Wild type | 32.58 | 1.88 | 17.15 | −0.56 | −0.03 | 3.94 | | |
| | | | Remodeled | 37.64 | −3.84 | 27.02 | −0.49 | 0.37 | 4.33 | | |
| 58 | GG/CC | −1.5 | Wild type | 34.36 | −9.06 | −31.2 | −0.22 | 0.32 | 3.28 | | |
| | | | Remodeled | 39.05 | −17.13 | −57.21 | 0.22 | 1.90 | 3.73 | | |
| 60 | CA/TG | −1.3 | Wild type | 38.95 | 8.96 | 10.9 | 0.03 | 0.65 | 3.52 | | |
| | | | Remodeled | 46.83 | 16.82 | 11.60 | −0.89 | 0.92 | 3.89 | | |

a. Root-mean-square deviation (RMSD) of the remodeled structure compared to the wild type structure (DNA2).
b. The optimized RMSD obtained by using least-square fitting procedure to fit the remodeled DNA to its original state found in the wild-type nucleosome structure (DNA2).

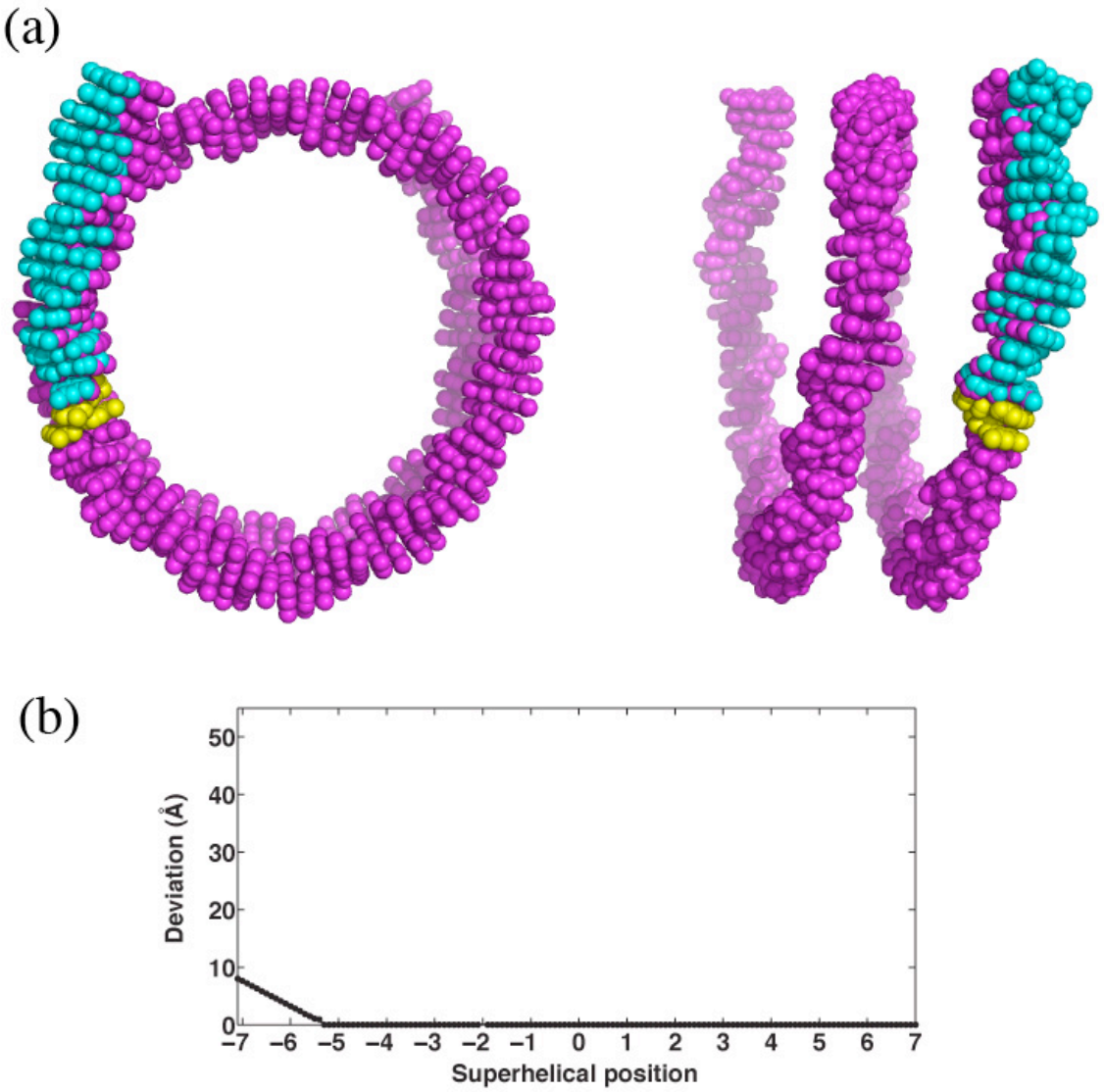



Figure 6.16. (a) Top view (on the left) and side view (on the right) of the remodeled DNA of the wild-type nucleosome structure (DNA2). (b) Root-mean-square deviation of the base atoms at each base pair of the remodeled DNA compared to the wild-type nucleosomal DNA (DNA2). Note: The altered step parameters of the modified base-pair step are listed in Table 6.7. The wild type DNA and the part of the remodeled DNA, which overlap the wild type DNA, are shown in magenta. The changed pathway of the remodeled DNA is shown in cyan. The remodeled base pair step, the GA/TC dimer at base-pair step 18, is highlighted in yellow.

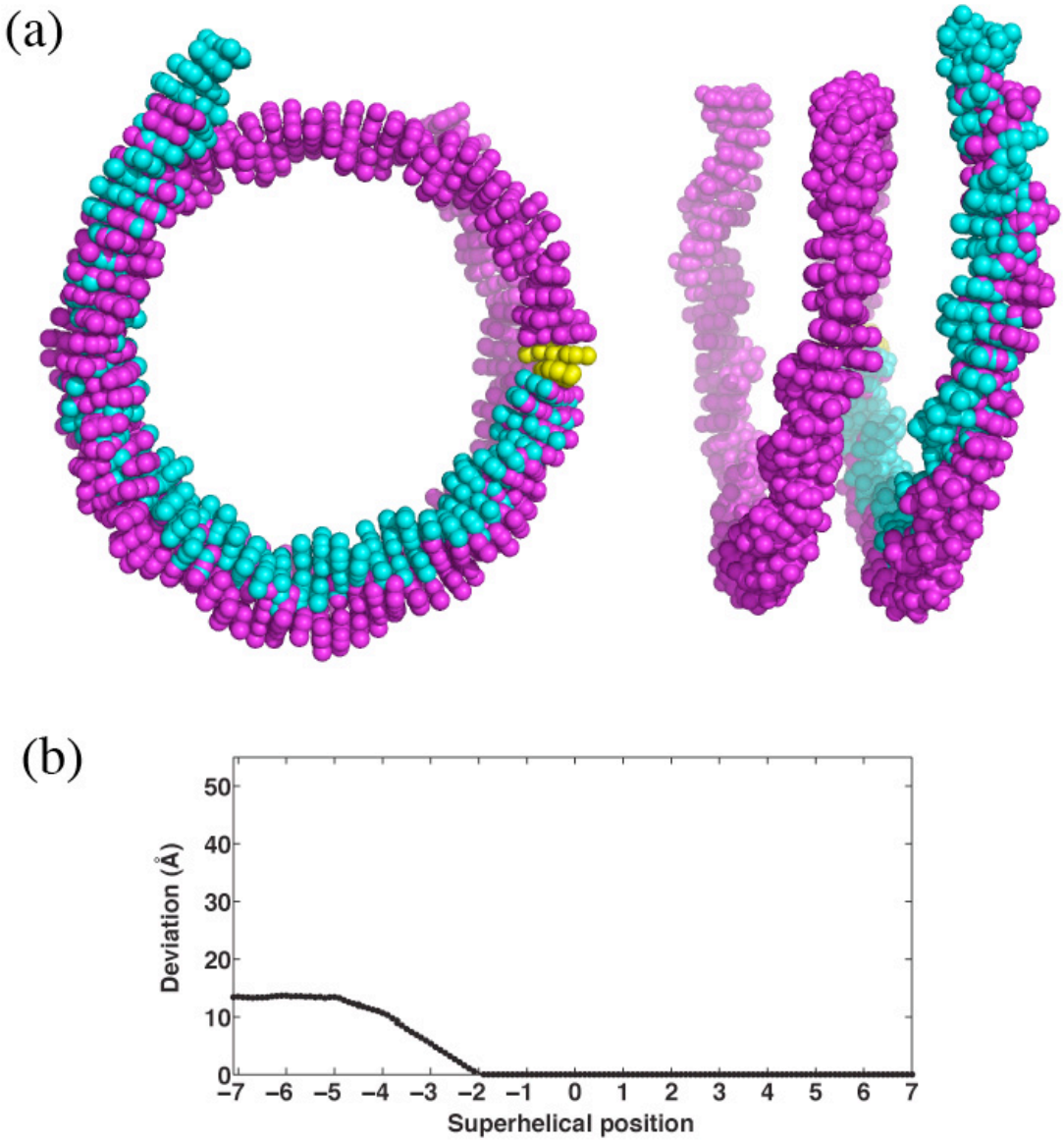



Figure 6.17. (a) Top view (on the left) and side view (on the right) of the remodeled DNA of the wild-type nucleosome structure (DNA2). (b) Root-mean-square deviation of the base atoms at each base pair of the remodeled DNA compared to the wild-type nucleosomal DNA (DNA2). Note: The altered step parameters of the modified base-pair step are listed in Table 6.7. The wild type DNA and the part of the remodeled DNA, which overlaps the wild type DNA, are shown in magenta. The changed pathway of the remodeled DNA is shown in cyan. The remodeled base pair step, the CA/TG dimer at the base-pair step 53, is highlighted in yellow.

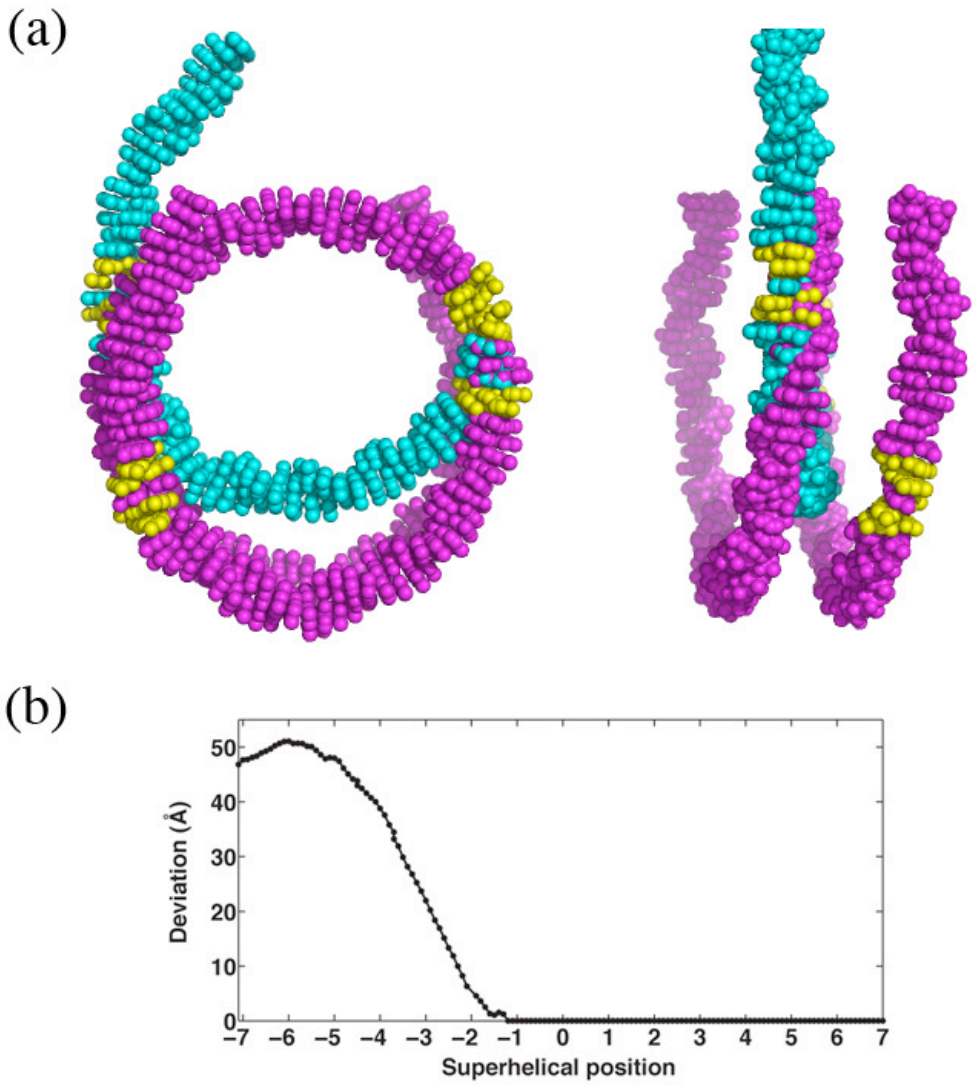



Figure 6.18. (a) Top view (on the left) and side view (on the right) of the remodeled DNA of the wild-type nucleosome structure (DNA2). (b) Root-mean-square deviation of the base atoms at each base pair of the remodeled DNA compared to the wild-type nucleosomal DNA (DNA2). Note: The altered step parameters of the modified base-pair steps are listed in Table 6.8. The wild type DNA and the part of the remodeled DNA, which overlaps the wild type DNA, are shown in magenta. The changed pathway of the remodeled DNA is shown in cyan. The remodeled base pair steps, the GG/CC dimer at base-pair step 58, are highlighted in yellow.

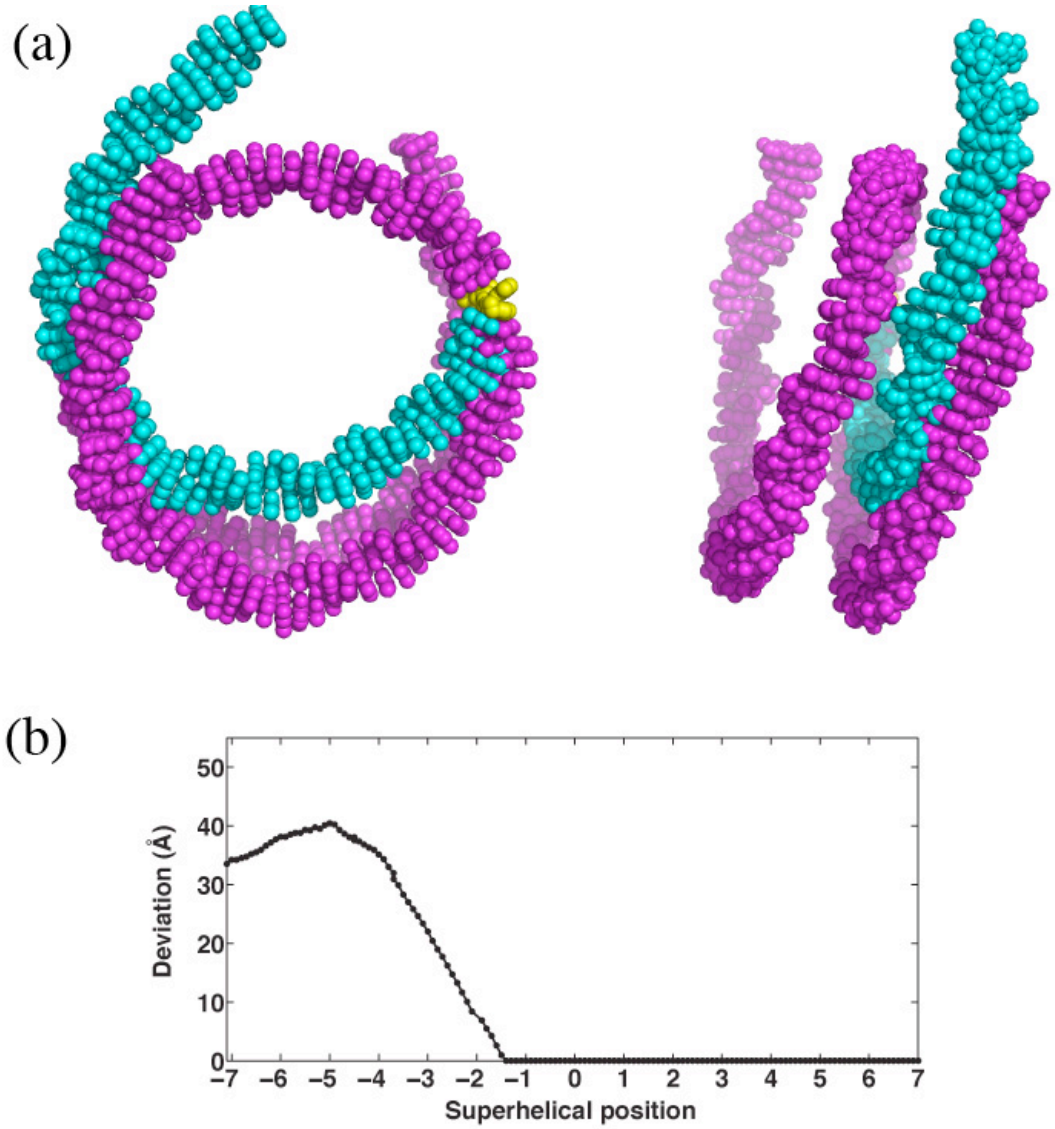



Figure 6.19. (a) Top view (on the left) and side view (on the right) of the remodeled DNA of the wild-type nucleosome structure (DNA2). (b) Root-mean-square deviation of the base atoms at each base pair of the remodeled DNA compared to the wild-type nucleosomal DNA (DNA2). Note: The altered step parameters of the modified base-pair step are listed in Table 6.7. The wild type DNA and the part of the remodeled DNA, which overlaps the wild type DNA, are shown in magenta. The changed pathway of the remodeled DNA is shown in cyan. The remodeled base pair steps —base-pair step 18 GA/TC, 21 TC/GA, 53 CA/TG, 58 GG/CC, and 60 CA/TG — are highlighted in yellow.

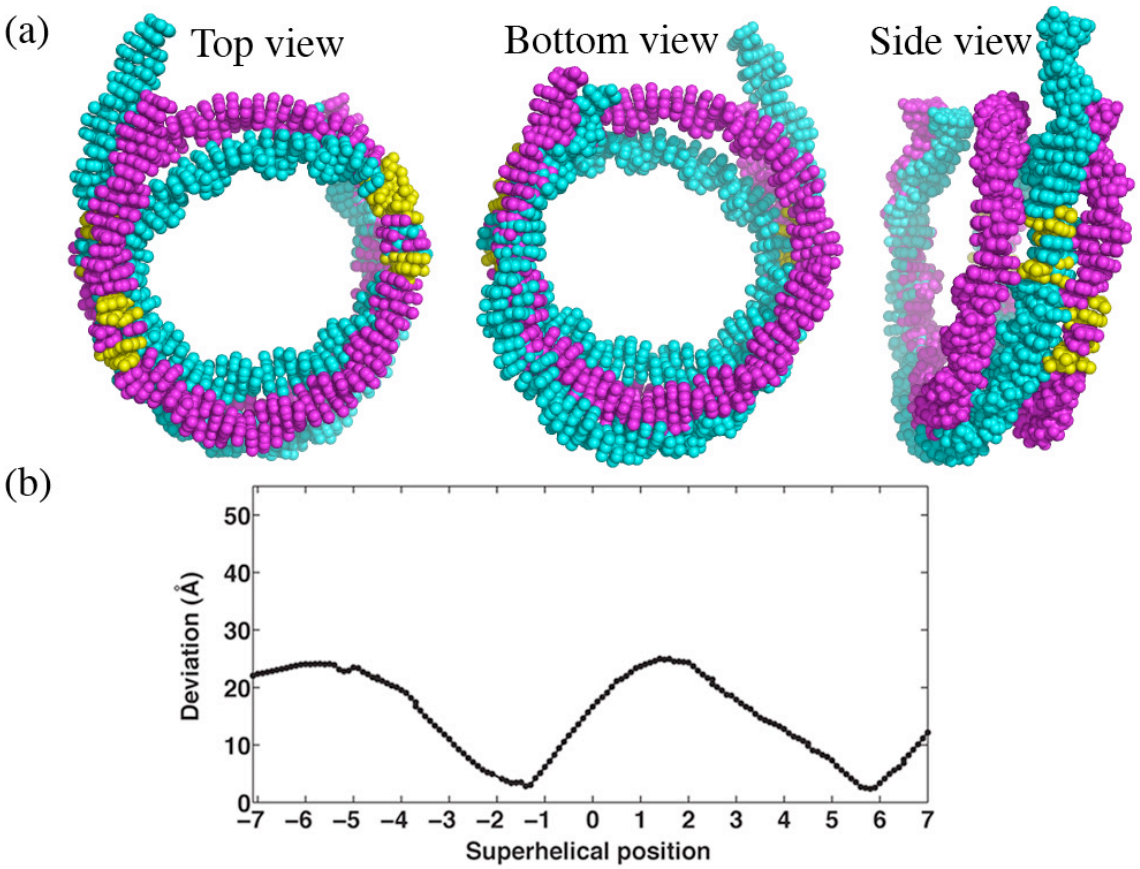



Figure 6.20. (a) Top view (on the left), bottom view (in the middle), and side view (on the right) of the remodeled DNA of the wild-type nucleosome structure (DNA2), whose root-mean-square deviation is optimized with least-square fitting. (b) Root-mean-square deviation of the base atoms at each base pair of the remodeled DNA compared to the wild-type nucleosomal DNA. Note: The altered step parameters of the modified base-pair steps are listed in Table 6.8. The wild type DNA and the part of the remodeled DNA, which overlaps the wild type DNA, are shown in magenta. The changed pathway of the remodeled DNA is shown in cyan. The remodeled base-pair steps are highlighted in yellow.

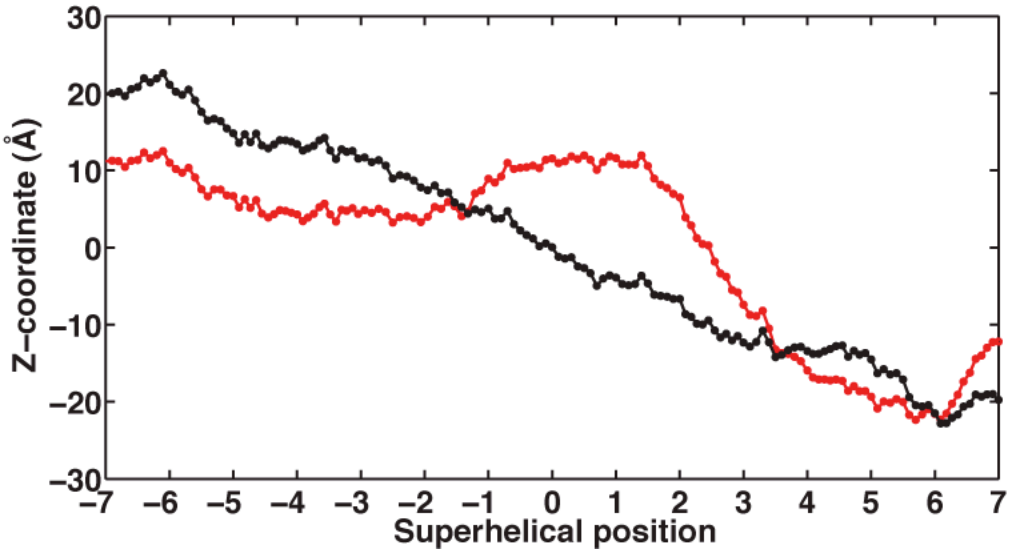



Figure 6.21. Z-coordinates vs. superhelical position of each base pair along the remodeled DNA in Figure 6.20 compared to the wild-type nucleosomal DNA. The curves of the remodeled and the wild-type DNA are shown in red and black, respectively.

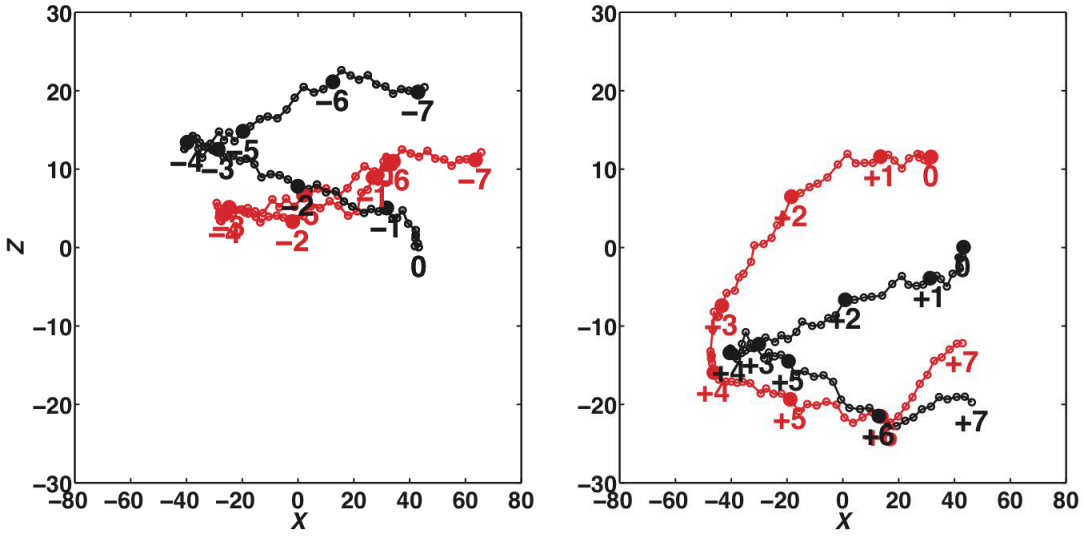



Figure 6.22. *X*-coordinate vs. *Z*-coordinate of the origin of each base pair of the remodeled DNA in Figure 6.20 (red curves) compared to the wild type (black curves). The first halves (base pair 1 to 73) and the second halves (base pair 73 to 146) are shown in the left and the right of the figure, respectively. The origins of the base pairs with integral superhelical numbers are highlighted with large markers and labeled with the corresponding superhelical numbers.

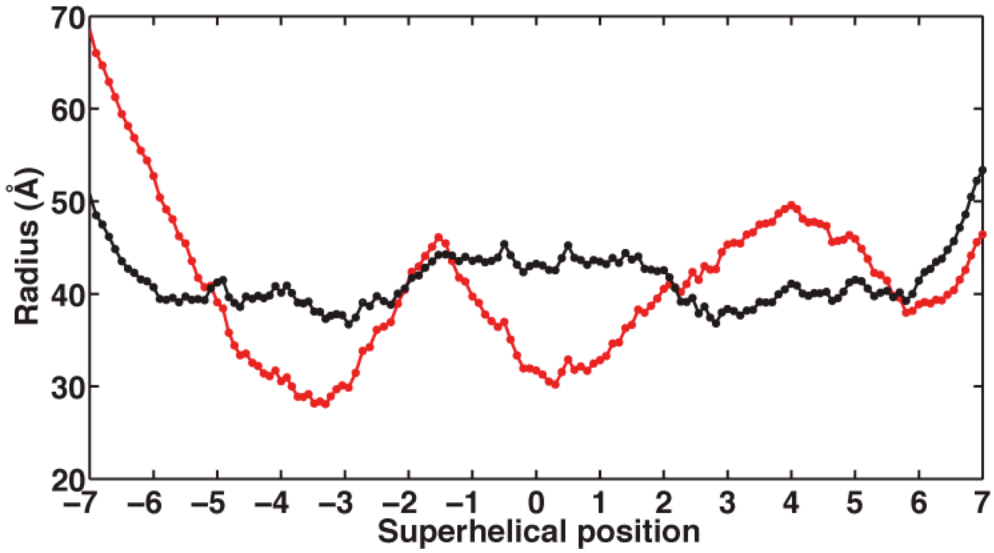



Figure 6.23. Radius vs. superhelical position of each base pair along the remodeled DNA in Figure 6.20 compared to the wild-type nucleosomal DNA. The curves of the remodeled and the wild-type DNA are shown in red and black, respectively.

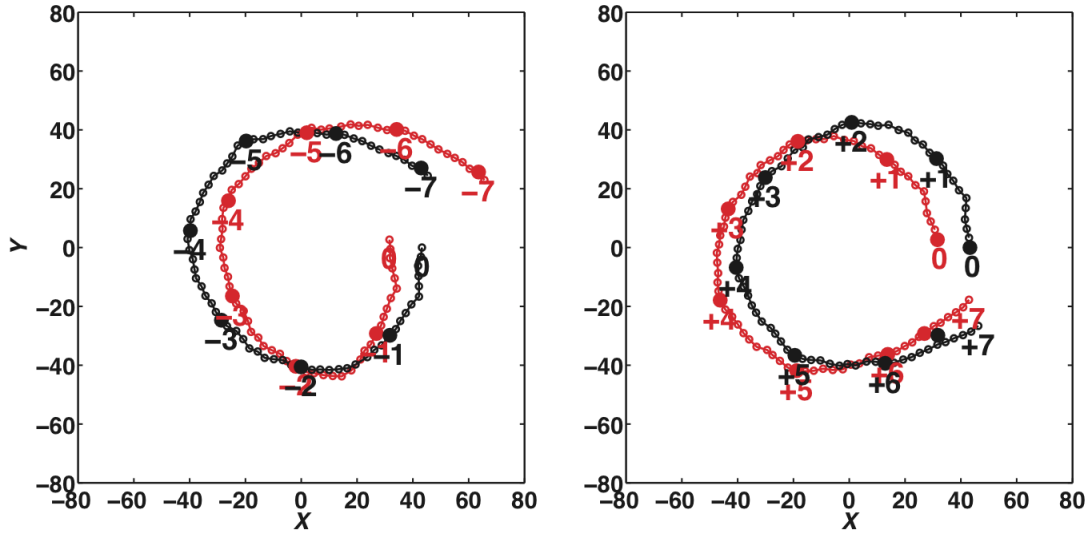



Figure 6.24. *X*-coordinate vs. *Y*-coordinate of the origin of each base pair of the remodeled DNA in Figure 6.20 (red curves) compared to the wild type (black curves). The first halves (base pair 1 to 73) and the second halves (base pair 73 to 146) are shown in the left and the right of the figure, respectively. The origins of the base pairs with integral superhelical numbers are highlighted with large markers and labeled with the corresponding superhelical numbers.

For the other SIN-mutant structures in the high-scoring group, the values and average contribution of step parameters are reported in the appendix (Table A6.8 – A6.9).

6.2.6 Deformation scores of each dimer step of nucleosomal DNA threaded on its natural template in wild-type and SI-mutant structures

The energy of deforming each dimer step from the rest state to the state found in the mutant and wild-type structures, *i. e.* the natural setting has been calculated and aligned along either the wild-type nucleosomal DNA sequence (Figure 6.25) or the SIN-mutant nucleosomal DNA sequence (Appendix Figure A6).

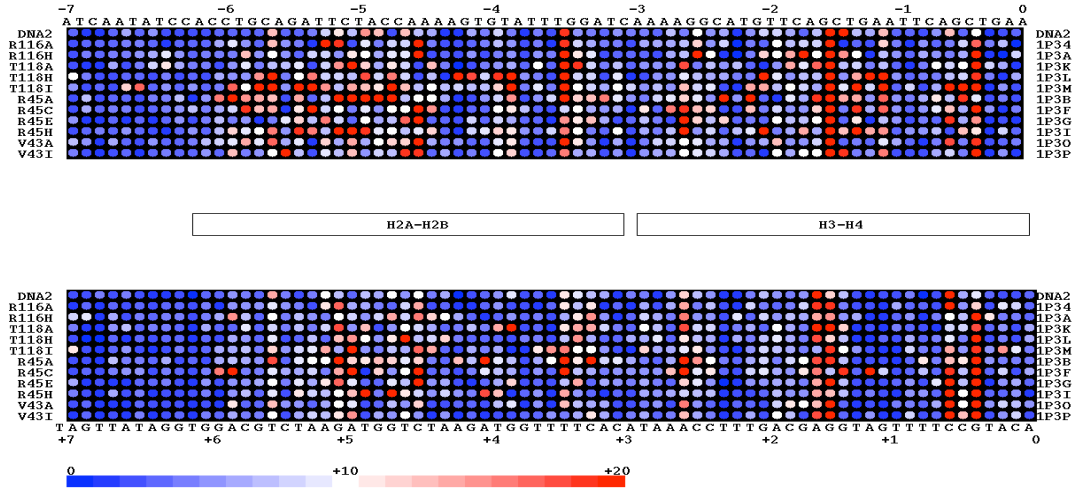



Figure 6.25. Values of the deformation scores of SIN mutants and the wild type (DNA2), are mapped on each base-pair step along the DNA sequence from the wild-type (DNA2). The PDB IDs, except for DNA2, are shown at the right edge of the diagram. The corresponding structure IDs, except for DNA2, designated by point-mutations are shown at the left edge. The values of deformation scores are shown in shades of blue (small values), white (medium values), and red (large values). Scores greater than 20 are colored as if they had a score of 20.

The average value of the deformation energy over all dimer steps in the natural settings of the wild-type and SIN-mutant structures is 7.22 (standard deviation is 2.45), *i.e.*, all boxes in Figure 6.25. The deformation energy scores of each dimer step along nucleosomal DNA do not change in a symmetrical manner as in the best-resolved nucleosome core-particle structure (PDB ID: 1KX5) since the first half of DNA (1 to 72 bp) is one base pair shorter than the second half (73 to 146 bp). (The dyad axis is located at base pair 73.) Besides the dimer steps at SHL ±1.5 and ±0.5, dimer step 38 at SHL −3.5 also has a high deformation score (average deformation scores of the single steps over all the SIN-mutant and wild-type structure is 19.91). The negative Roll (average −33.4˚ and standard deviation 6.36 ˚) in dimer step 38 TG/CA contributes to the high deformation energy scores (average 21.79 for Roll alone vs. average total score of 19.91). Other dimer steps with high deformation scores are scattered in the region from SHL −4.5 to −5.5 (Figure 6.25).

The deformation energy differences of each base-pair step in the SIN mutants (difference between scores in the boxes in the first row compared to the boxes in the other rows of Figure 6.25 are mapped along the wild-type nucleosomal DNA sequence in Figure 6.26 and along the SIN-mutant DNA sequence in Appendix Figure A7. The increase in the deformation score at SHL −6 to −4.5 is related to the conformational changes in the same region of the SIN-mutant nucleosome structure (Figure 5.1). Note that many of the steps that have high deformation scores in the natural template are sites of high threading scores compare dots in Figure 6.25 with Figure 6.15). Except for the different templates used, other factors, such as the force constants, the rest state, and the sequence, are the same in the computation of the deformation scores. That is to say, it is energetically

costly for the nucleosomal DNA sequence at SHL –6 to –4.5 to bend into the states found in the crystallized structures of the SIN mutants. As the nucleosomal DNA at SHL –6 to –4.5 is primarily contacted by histones H2A and H2B (Figure 1.2 in Chapter 1), the deformation-score increase at those regions suggests that the interactions of histones H2A and H2B and nucleosomal DNA may be destabilized in the SIN-mutant structures.

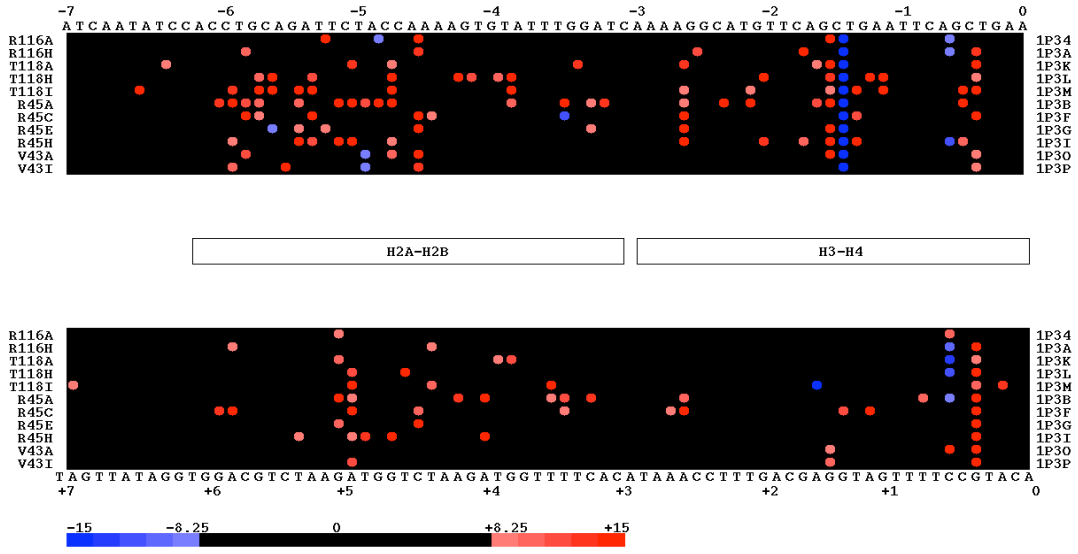



Figure 6.26. Difference in deformation scores of the SIN mutants compared to the wild-type (DNA2), are mapped on each base-pair step along the DNA sequence from the wild type (DNA2). The PDB IDs, except for DNA2, are shown at the right edge of the diagram. The corresponding structure IDs, except for DNA2, designated by point-mutations, are shown at the left edge. The difference in values of the mutant vs. the values of wild type is shown in shades of blue (negative) and red (positive). Differences of –7.5 to +7.5 are shown in black.

## 6.3 Discussion

The central natural settings of templates with different length (from central 60 to 130 base-pair steps) from the best-resolved nucleosome structure (PDB ID: 1KX5) always have the lowest threading scores compared to other settings, when these templates slide

along the DNA sequence from the same structure (Figure 6.2). In the natural setting, most of the highly deformed steps are CA/TG dimers, which are naturally flexible in terms of their low force constants. When the template moves to other locations along the DNA sequence, dimer steps other than CA/TG steps must take up these highly deformed conformation, causing an increase in the threading scores. Like the best-resolved 147-bp nucleosome structure, the central natural settings of the 146-bp SIN-mutant and wild-type nucleosome structure always have the lowest threading scores (Figure 6.8).

Most surprisingly, the threading scores can discriminate among different SIN-mutant structures, although the global structures of the SIN-mutant structures and the wild-type nucleosome are very similar (except for the subtle DNA conformational changes mentioned above). Roll, Rise, and Tilt contribute to the differences in the threading scores between the highest-scoring SIN-mutant structures and the wild-type structure (Table 6.5). The contribution of Slide, which are responsible for superhelical pitch of nulcoesomal DNA, do not change much in the different structures [4]. Both Roll and Tilt contribute to the DNA bending (http://dimax.rutgers.edu/~davidm/movies/movies.htm). The increase of the 'energy' contribution of Roll and Tilt in the SIN mutants suggests that it is more difficult for the DNA sequences to bend along the pathways around histone octamers bearing SIN mutations. The variation in Rise in the SIN mutants appears to stretch out nucleosomal DNA (Figure 6.17).

The most highly deformed steps in terms of large deformation scores in the SIN mutant structures are located primarily at SHL −6 to −4 and SHL −2 to −1. Remodeling the wild-type nucleosomal DNA with the step parameters of the highly deformed steps in the R45H H4 SIN mutant, which helps to see the global impact of these conformational

change on nucleosomal DNA structure. The extremely negative Roll at SHL −1.5 cause the dramatically change of the wild-type nucleosomal DNA, which may be compensated, to some degree, by the step parameters of the other steps. The remodeled DNA, which includes the changes of all highly deformed steps at SHL –6 to –5 and –2 to -1 in the R45H H4 SIN mutant, distort the general cylindrical shape of nucleosomal DNA (Figure 6.20 (a)). The pathway of the remodeled DNA deviates appreciably much from the wild-type one at SHL −6 to −5 and +2 to +1(Figure 6.20 (b)). Interestingly, the remodeled DNA remains similar to the wild-type at SHL −1.5 and +6. This similarity suggests ways, in which the nucleosomal DNA may loop out or peel off during the nucleosome remodeling.

The natural setting of the natural templates from the SIN-mutant and wild-type nucleosomes are also examined. Most of the steps with higher deformation scores in the SIN mutants compared the wild type are located at SHL −6 to −4, −2.5, −1.5, and −0.5 (Figure 6.26). This result is consistent with the results from the threading. That is, the steps in the SIN-mutant templates, which always cause the high deformation scores compared to the wild-type, are concentrated at SHL −6 to −4 and SHL −2 to −1 (Figure 6.15).

The threading scores reported in this chapter measure the relative ease of deforming a DNA sequence on the surface of a nucleosome core particle. That is to say, it is energetically more costly for the nucleosomal DNA sequences (from the wild-type and the SIN-mutant nucleosome structures) to adopt the conformations, especially the highly deformed conformations at SHL −6 to −4, found in the SIN-mutant structures.

Consistently, the SIN-mutant structures with high threading scores have high sliding rates in the heat-shifting experiments.

The threading scores reported in this chapter measure the relative ease of deforming a DNA sequence on the surface of a nucleosome core particle. It is energetically more costly for the nucleosomal DNA sequences (from the wild-type and the SIN mutant nucleosome structures) to adopt the conformations found in the SIN-mutant structures than in the wild-type structure.

Consistently, the SIN-mutant structures, which have high threading scores, also have high sliding rates in heat-shifting experiments. When histone protein is reconstitute at 4 °C with 146 bp nucleosomal DNA, the wild-type and the SIN-mutant structures, which have low- or intermediate scoring in threading-score computation, are positioned at two off-center positions of the 146-bp nculeosomal DNA sequence at 4 °C in the heat-shifing experiment. When heated at 37 °C, the wild-type, the low-scoring, and intermediate-scoring SIN-mutant nucleosomes slide to the central position on the 146-bp nucleosomal DNA. By contrast, the SIN mutants in the high-scoring group of the threading-score computation only form nucleosomes at the central position even at 4 °C in the heat-shifting experiments. Those nucleosomes appear to slide from the two-off center positions on the 146-bp nucleosomal DNA to the central position too quickly to measure the sliding rates. The high sliding rates and the high threading scores of the SIN-mutant structures may be related to each other.

It may be very difficult for the SIN-mutants in the high scoring group to accommodate the histone protein cores at the off-center positions because of the sky-rocking high threading scores at those positions. The interactions between the SIN-mutant histone

proteins and DNA are optimized at the central position because of the ease of deforming DNA, *i.e.* the lowest threading score occurs at the centeral position rather than the off-center positions. This may be the reason why the SIN-mutants in the high-scoring group only form at the central position in heat-shifting experiments at 4 °C. Interestingly, the lowest threading scores of the SIN mutants in the high-scoring group are comparable to the average threading scores of all the reading frames of the wild-type and the SIN-mutants in the low- and intermediate-scoring groups. Since the SIN mutants in the high-scoring group can form nucleosome structures at that level of threading scores, it may be relatively easy for the wild-type and the SIN-mutants in the low- and intermediate-scoring groups to form nucleosomes at off-center positions. This may be the reason why the wild-type and the SIN-mutants in the low-scoring and intermediate scoring groups can be positioned at the two off-center positions when the histone proteins and DNA are reconstituted to form nucleosomes at 4 °C. However, the nucleosomes at the two off-center positions are not optimally stable. When the wild-type and the SIN-mutant nucleosomes with low- and intermediate threading scores are heated at 37 °C in the heat-shifting experiments, the nucleosomes slide to the central position where the DNA deformation is optimized in terms of the lowest threading scores.

1. Olson, K.W., Gorin, A. A., Lu, X. J., Hock, L. M., and Zhurkin, V. B. , *DNA sequence-dependent deformability deduced from protein-DNA crystal complexes*. Proc. Natl. Acad. Sci. USA, 1998. 96: p. 11163-11168.
2. Balasubramanian, S., Xu, F., Olson, W. K., *DNA sequence-directed organization of chromatin: knowledge-based potentials for the analysis of positioning sequences*. 2007.
3. Luger, K., Maedar, W., Richmond, R. K., Sargent, D. F., and T. J. Richmond, *Crystal structure of the nucleosome core paricle at 2.8 Å resolution*. Nature, 1997. 389: p. 251-259.

4. Tolstorukov, M.Y., Colasanti, A. V., McCandlish, D., Olson, W. K., and Zhurkin, V. B., *A novel 'roll-and-slide' mechanism of DNA folding in chromatin. Imaplication for nucleosome positioning*. J. Mol. Biol., 2007. 371: p. 725-738.
5. Muthuranjan, U.M., Bao, Y., Forsberg, J., Edayathumagalam, R. S., Dyer, P. N., White, C. L., and Luger, K., *Crystal structures of histone Sin mutant nucleosomes reveal altered protein-DNA interactions*. EMBO J., 2004. 23: p. 260-271.
6. Horn, B.K.P., *Closed-form solution of absolute orientation using unit quaternions*. J.Opt. Soc. Am. A, 1987. 4: p. 629-642.

# Chapter 7 Nucleosome Structures with different DNA sequences

## 7.1 Introduction

Two nucleosome structures (PDB ID: 1KX3 and 1KX4) have identical histone cores but different 146-bp DNA sequences [1] (Figure 7.1).The best-resolved nucleosome structure (PDB ID: 1KX5) has a DNA sequence similar to that of one of the 146-bp nucleosome structures (PDB ID: 1KX3), except for one extra base pair near the dyad (147 bp in 1KX5 vs. 146 bp in 1KX3) [1] (Figure 7.1). These three structures are made up of recombinant *X. laevis* histones and fragments of human α-satellite DNA [1].

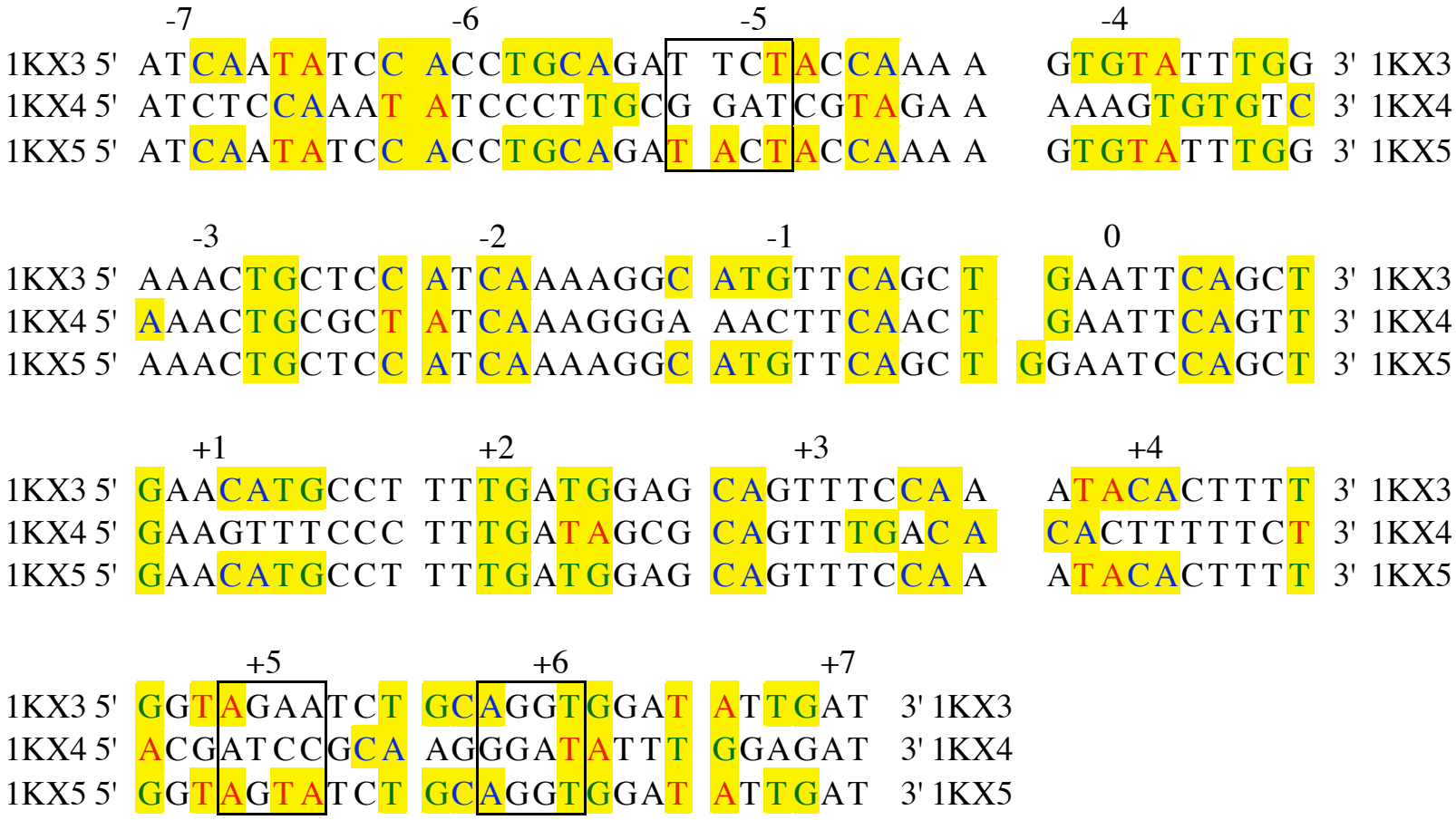



Figure 7.1. Literal alignment of the DNA sequences from the nucleosome structures, 1KX3, 1KX4, and 1KX5. Flexible CA, TG, and TA dimer steps are highlighted in blue, green, and red, respectively. The base pairs with the integral superhelical location numbers are labeled along the DNA sequences. The base pairs are assigned the superhelical location numbers obtained from the structure 1KX3. Boxed sites have distinctly different superhelical properties.

In this chapter, the superhelical shapes and contacts of the histone proteins to DNA in these three nucleosome structures are compared with each other, regardless of the differences in nucleotide composition and length of DNA. The DNA sequences are also threaded on the three-dimensional templates from the three nucleosome structures. These analyses may help to understand how the nucleosome adapts to different DNA sequences. For convenience, we refer to the three structures by their PDB identities as 1KX3, 1KX4, and 1KX5.

## 7.2 Methods

We use the methods introduced respectively in Chapters 5 and 6 to compute the superhelical properties of DNA and perform threading calculations. As described below, we check the approach used to determine the superhelical structure and the rotational phase of DNA on the sufaces of the three nucleosome.

### 7.2.1 A preliminary step to define the superhelical axis

Because the two ends of nucleosomal DNA deviate from an ideal superhelical ramp, the ends of DNA are removed in increments of five base pairs in the determination of superhelical parameters. Using the procedures mentioned in Chapter 5, the superhelical axes are determined from the coordinates of the nucleosomal DNA in each structure, with 0, 10, 15, 20, 30 bp removed from each of the two ends. The superhelical pitch and radii are then computed from the superhelical axes defined by the different lengths of the truncated nucleosomal DNA (Figure 7.2, Appendix Figure A7.1, and Figure 5.3 in [2]). For the three structures (1KX3, 1KX4, and 1KX5), the computed superhelical radius and the local variation in superhelical pitch are more uniform when 15 bp are removed from

each end of the nucleosomal DNA. We therefore use DNA fragments of 116 and 117 base pairs in the determination of the cylindrical axes used to analyze the superhelical folding of the three nucleosomal structures.

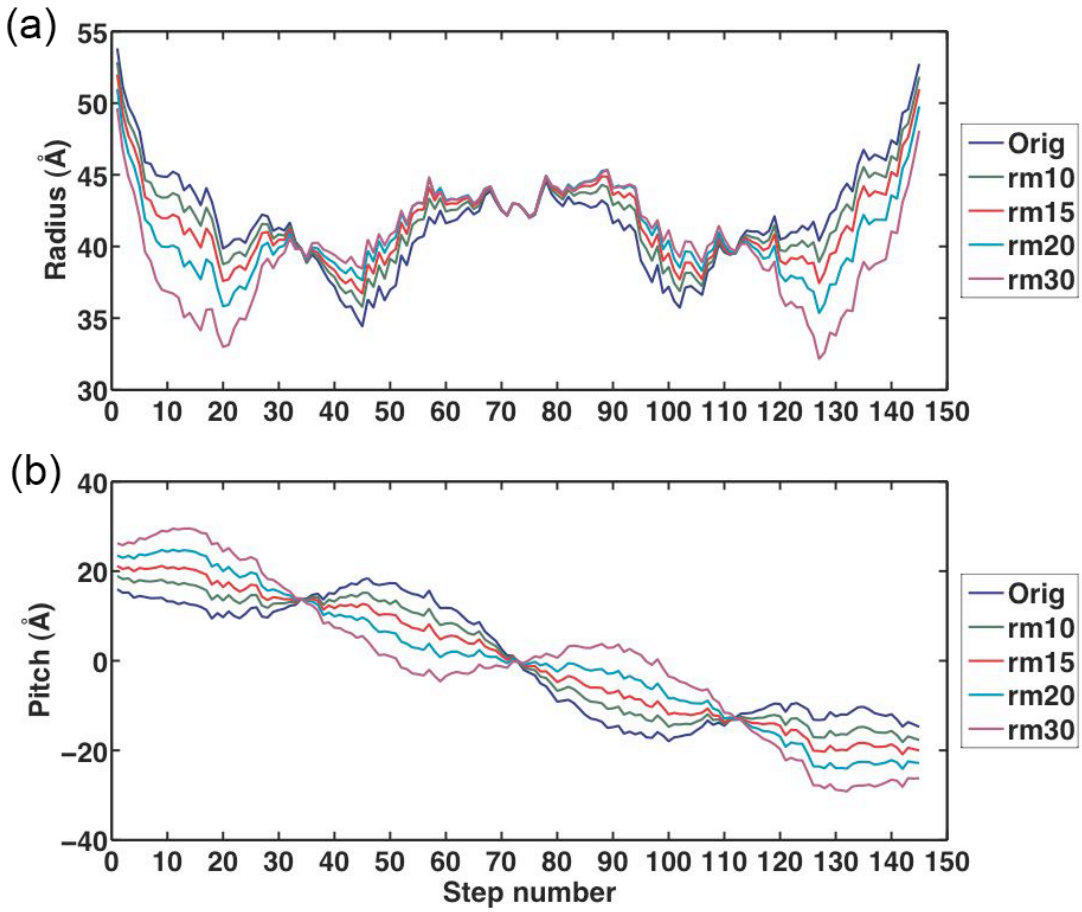



Figure 7.2. Effects of the nucleosomal DNA chain length on (a) superhelical radius and (b) *Z*-coordinates of each base-pair step in the structure (PDB ID: 1KX4). Orig represents the full length 146 bp DNA, rm10, rm15, rm20, and rm30 correspond to chains of 126, 116, 106, and 86 bp where the 10, 15, 20, and 30 base pairs have been removed from each end of DNA.

### 7.2.2 Superhelical location

The superhelical pitch and radius, which are discussed in the later sections, are determined from the origins of the local middle frames of base-pair steps or dimer steps.

The superhelical locations of dimer steps are approximated by the superhelical locations of the first base pairs of the dimer steps. The superhelical location of each base pair of nucleosomal DNA is determined with the method described below.

Once the three structures (1KX3, 1KX4, and 1KX5) are transformed into their respective cylindrical frames, the superhelical locations of each base pair can be computed. The superhelical location shows the relative position of a base pair with respect to the global *Z*-axis. For each base pair along nucleosomal DNA, the projection of the long axis of the base-pair frame along the global *Z*-axis, $z_{SHL}$, is computed from the scalar product:

$$z_{SHL} = \vec{v}_{long} \cdot \vec{n} ,$$

where $\vec{v}_{long}$ is a unit vector along the long axis of the base-pair frame, and $\vec{n} = (0, 0, 1)$ is a unit vector along the global *Z*-axis.

The base pair at the dyad position of nucleosomal DNA, where $z_{SHL} \approx 1$, is defined as SHL 0. Sequentially, the other base pairs with local maxima in $z_{SHL}$ are assigned integral superhelical numbers from SHL −7 to +7. The superhelical numbers of the $m^{th}$ base pair relative to the preceding base pair with integral superhelical numbers, $SHL_{\text{int}}$, are approximated as:

$$SHL = SHL_{\text{int}} + m/N_p$$

where $N_P$ is the number of base pair between the two base pairs with local maxima in $z_{SHL}$ on either side of the base pair of interest.

## 7.3 Results

### 7.3.1 Dimer-step composition of nucleosomal DNA sequences

The number of occurrences of the 16 different dimer steps in the three nucleosome structures is summarized in Table 7.1. As the structures 1KX3 and 1KX4 have almost identical sequencyies, the frequency of occurrence of the 16 dimer steps are almost the same. Surprisingly, except for the large difference in the number of occurrences of CA, TG, and TA dimer steps between 1KX3 and 1KX4, the populations of other dimer steps in 1KX3 and 1KX4 are very similar to one another. The CA, TG, and TA steps are flexible in terms of the low force constants of the knowledge-based elastic function (Chapter 6 Table 6.2). The locations of these flexible dimer steps play an important role in nucleosomal DNA bending and in formation of the superhelical pitch [3].

The highly deformed CA, TG, TA dimer steps at SHL −5.6, −4.0, −3.6, +3.5, and +5.6, in 1KX5 (Chapter 6, Table 6.4 (b)) are replaced by stiffer dimer steps in 1KX4 with higher force constants (Chapter 6, Table 6.2). The locations of CA, TG, and TA dimer steps in the three structures 1KX3, 1KX4 and 1KX5 are highlighted in color in Figure 7.1. The CA dimer steps at SHL −6.9 and +1.1 in the DNA sequences of the 1KX3 and 1KX5 structures disappear from in the 1KX4 structure and are replaced by other stiffer dimer steps. The CA dimer step at SHL −5.6 in 1KX3 disappears and another flexbile dimer step, TG dimer step appears at SHL −5.5 in 1KX4. The CA dimer step at SHL +3.4 in 1KX3 is shifted to SHL +3.5 in 1KX4. A TG dimer step also appears at SHL +3.2 in 1KX4. The CA dimer step at SHL +4 in 1KX3 is shifted to SHL +3.7 in 1KX4. The CA dimer at SHL +5.7 in 1KX3 is shifted to SHL +5.4 in 1KX4. At SHL −4.6, the CA dimer

step in 1KX3 and 1KX5 changes to a flexible TA dimer step in 1KX4. The TG dimer steps in 1KX3 and 1KX4 at SHL −5.8, −1.1, +1.3, +5.5, and +6.8 are replaced by other stiffer dimer steps. The TG dimer steps at SHL −4.2 and −3.5 in 1KX3 are shifted, respectively, to SHL −3.9 and −3.4 in 1KX4. At SHL +2.4, +4.6, and +6.1, the TG dimer step in 1KX3 and 1KX4 are replaced by another flexible dimer step TA. The TA dimer step at SHL +4.8 in 1KX3 and 1KX5 is replaced by other stiffer dimer steps, such as TC, GT, and GA, in 1KX4. The TA dimer step at SHL −4.9 in 1KX3 is shifted to SHL −4.6 in 1KX4. TA dimer step at SHL −3.9 in 1KX3 is shifted and replaced by the TG dimer step at SHL −3.8 in 1KX4. The TA dimer step at SHL +3.8 in 1KX3 disappears and a CA dimer step occurs at SHL +3.7 in 1KX4.

Table 7.1. Counts of dimer steps occurring in the DNA sequences from the structures 1KX3, 1KX4, 1KX5.

| Step | PDB ID | | |
|---|---|---|---|
| | 1KX3 | 1KX4 | 1KX5 |
| CG | 0 | 6 | 0 |
| CA | 14 | 9 | 14 |
| TG | 14 | 9 | 14 |
| TA | 6 | 6 | 8 |
| AG | 9 | 9 | 9 |
| CT | 9 | 9 | 9 |
| GG | 6 | 6 | 7 |
| CC | 6 | 6 | 7 |
| AA | 14 | 16 | 13 |
| TT | 14 | 16 | 12 |
| GA | 9 | 11 | 8 |
| TC | 9 | 11 | 8 |
| AT | 15 | 11 | 15 |
| AC | 6 | 7 | 7 |
| GT | 6 | 7 | 7 |
| GC | 8 | 6 | 8 |

### 7.3.2 Superhelical location

The base pairs with integral superhelical locations are reported in Table 7.2. According to these data, the base pairs found in 1KX3 and 1KX4 differ in superhelical locations at SHL −4, −3, +2, and +5. Base pairs 33 at SHL −4 and 43 at SHL −3 in 1KX3 change to base pairs 32 and 42, respectively, in 1KX4. Over the region from SHL −5 to -4, there are 11 base pairs in 1KX3 vs. 10 base pairs in 1KX4. Thus, the DNA from SHL −5 to −4 in 1KX3 is under-twisted compared to that in 1KX4. The average value of Twist of the base-pair steps from SHL −5 to −4 is 34.05 ° in 1KX3 vs. 36.74 ° in 1KX3. Similarly, over the region from SHL −3 to −2, there are 10 base pairs in 1KX3 vs. 11 base pairs in 1KX4. The average value of Twist from SHL −3 to −2 is 36.36 °in 1KX3 vs. 34.25 ° in 1KX4.

Although the base pair located at SHL +2 changes from base pair 93 in 1KX3 to base pair 94 in 1KX4, this change does not affect the average value of Twist of the base pairs from SHL +1 to +3. The average Twist from SHL +1 to +2 is 34.36 ° in 1KX3 vs. 33.89 ° in 1KX4, and the average Twist from SHL +2 to +3 is 34.51 ° in 1KX3 vs. 34.65 ° in 1KX4.

The change of base pair 125 at SHL +5 in 1KX3 to base pair 124 in 1KX4, however, changes the average Twist of base pairs from SHL +4 to +6. The value of average Twist from SHL +4 to +5 is 33.25 ° in 1KX3 vs. 35.46 ° in 1KX4, and from SHL +5 to +6 is 35.48 ° in 1KX3 vs. 34.06 ° in 1KX4. This difference in Twist is the main cause of the difference in superhelical pitch at SHL −4 to −6 mentioned below in Section 7.3.4.

As noted above, the DNA sequences in 1KX3 and 1KX4 have the same number of base pairs but different nucleotide composition. The histone core takes up these two DNA sequences differently, changing the identities and numbers of base pairs between integral superhelical locations.

Table 7.2. Base pairs assigned integral superhelical locations (left) and the corresponding superhelical phase angle (right) in the three structures (PDB ID: 1KX3, 1KX4, and 1KX5).

| | Base-pair number of structures | | | Phase angle (°) of base pairs of structures | | |
|---|---|---|---|---|---|---|
| SHL | 1KX3 | 1KX4 | 1KX5 | 1KX3 | 1KX4 | 1KX5 |
| −7 | 2 | 2 | 2 | 32.41 | 31.14 | 32.08 |
| −6 | 12 | 12 | 12 | 72.21 | 72.33 | 71.74 |
| −5 | 22 | 22 | 22 | 118.57 | 120.12 | 118.81 |
| −4 | 33 | 32 | 33 | 171.84 | 168.5 | 170.67 |
| −3 | 43 | 42 | 43 | 221.08 | 218.13 | 219.74 |
| −2 | 53 | 53 | 53 | 269.7 | 270.7 | 267.26 |
| −1 | 63 | 63 | 64 | 316.62 | 315.36 | 315.98 |
| 0 | 73 | 73 | 74 | 0 | 0 | 0 |
| 1 | 83 | 83 | 84 | 43.98 | 43.52 | 43.84 |
| 2 | 93 | 94 | 94 | 88.92 | 93.9 | 88.33 |
| 3 | 104 | 104 | 105 | 141.77 | 139.07 | 140.3 |
| 4 | 114 | 114 | 115 | 189.58 | 189.47 | 189.13 |
| 5 | 125 | 124 | 126 | 241.93 | 236.57 | 240.87 |
| 6 | 135 | 135 | 136 | 288.15 | 286.57 | 287.63 |
| 7 | 145 | 145 | 146 | 326.83 | 326.47 | 326.83 |

Although the regions containing 11 bp differ in 1KX4 compared to 1KX3, the phase angles of the 15 sites at the integral superhelical locations are very close to each other (Table 7.2, right part). This suggest that, regardless of the specific nucleotide content, the angular locations of these 15 sites may serve as fixed anchor points in the superhelical ramp of nucleosomal DNA.

Comparison of the superhelical locations and phase angles of the two structures, 1KX3 and 1KX5, with identical DNA sequences shows that the extra base pair lies in the region from SHL −2 to −1, where the DNA periodicity is 11 bp in 1KX5 vs. 10 bp in 1KX3 (Table 7.2). Surprisingly, all other regions of the nucleosomal DNA have the same periodicities in 1KX3 and 1KX5. The angular locations of the base pairs with integral superhelical locations are also very similar to each other in the two structures, 1KX3 and 1KX5.

### 7.3.3 Superhelical radii

The superhelical radii of all base-pair steps in the three nucleosome structures (1KX3, 1KX4, and 1KX5) are plotted against the phase angle in Figure 7.3.

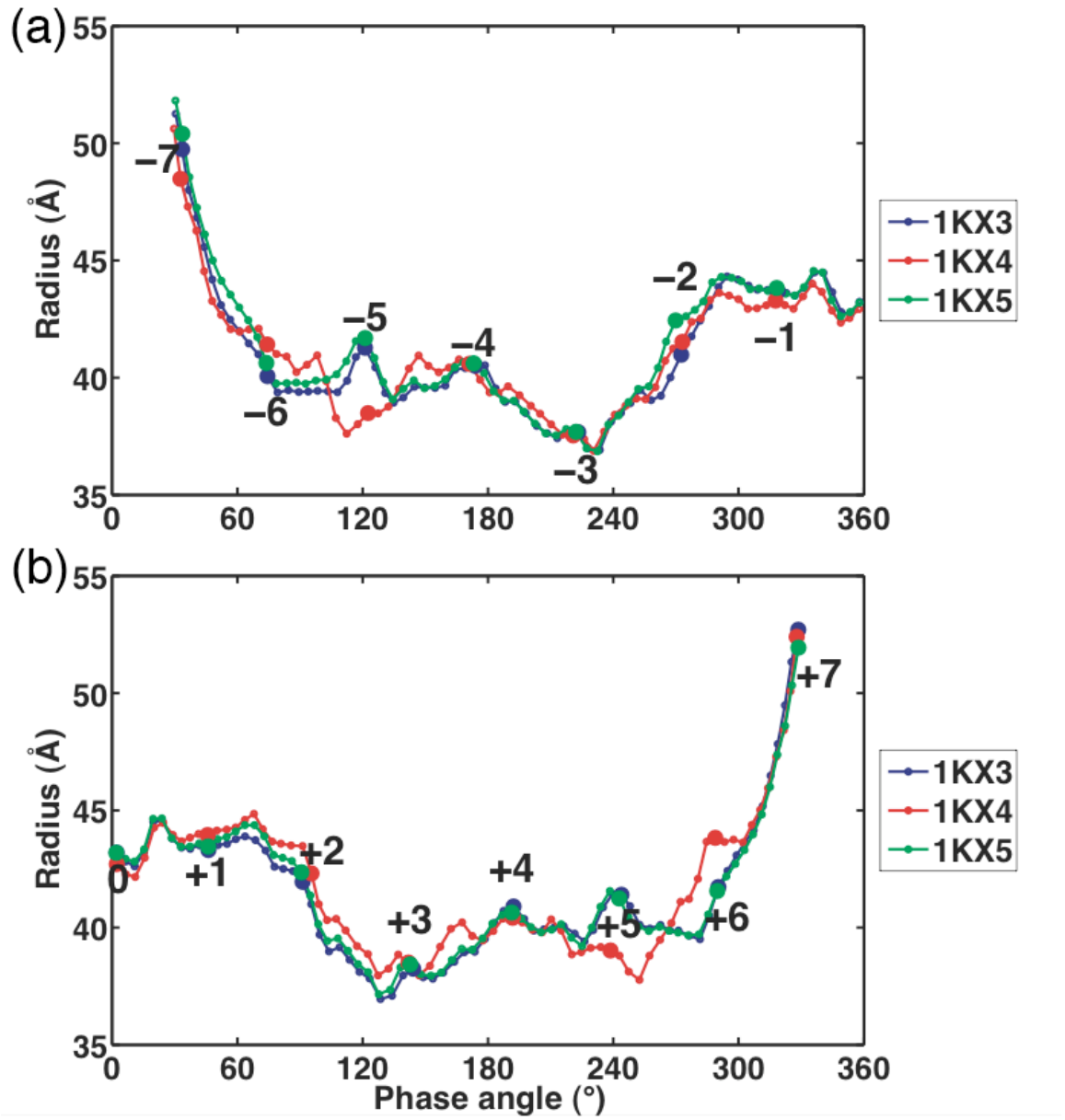



Figure 7.3. Radius vs. phase angle of each base-pair step in (a) the first half and (b) the second half of the nucleosomal DNA from, 1KX3 (blue), 1KX4 (red), 1KX5 (green). The first half of nucleosomal DNA contains base-pair steps 1 to 73 in 1KX3 and 1KX4, and base-pair steps 1 to 74 in 1KX5. The second half includes base-pair step 74 to 145 in 1KX3 and 1KX4, and base-pair steps 75 to 146 in 1KX5. The origins of the base-pair steps with integral superhelical numbers are highlighted with larger markers and labeled with the corresponding superhelical numbers. The superhelical number of each dimer step is assigned that of the first base pair in the dimer step.

Although the two structures 1KX3 and 1KX5 have different lengths of nucleosomal DNA, the superhelical radii are very similar. By contrast, in the two structures, 1KX3 and 1KX4, with different nucleotide sequences but the same length of DNA, there are large differences in radii in the regions from SHL −5.5 to −5 and +5 to +6 (Figure 7.3).

The base-pair steps of 1KX4 with large differences (greater than 2 Å or more negative than −2 Å) in superhelical radius compared the corresponding steps of 1KX3 and the differences in the contributions of Shift, Slide, and Rise to the superhelical radius at these sites are reported in Table 7.3 (a). Slide makes a large contribution (−1.66 Å, 1.93 Å, and 1.09 Å) to the difference in superhelical radius at base-pair steps 20, 127 and 133 of 1KX4 vs. 1KX3. The difference of the radius at dimer step 127 is negative (−2.35 Å), while the difference in the Slide contribution to the radius at base-pair step 127 is positive (1.93 Å). Slide at base-pair step 127 may be related to the large kink of superhelical pathway at SHL +5.2 (note the direction of the large kink at SHL +5.2 along the red curve in Figure 7.3). Beyond the sharp kink in the superhelical radius at SHl +5.2, the difference of superhelical radius in the following three dimer steps from 128 to 130 changes gradually from around −2 Å to zero.

The small differences in the Shift, Slide, and Rise contributions to the difference in radius at most base-pair steps in Table 7.3 (a), show that the difference in radius of 1KX4 compared to 1KX3 may be caused by differences in the rotational step parameters, Tilt, Roll, and Twist.

Table 7.3. (a) Differences in the contributions of Shift, Slide, and Rise to the supehelical radius at the dimer steps of the structure 1KX4 with large differences (more than 2.0 Å, or less than −2.0 Å) in superhelical radius, compared to 1KX3. (b) The values of step parameters of the corresponding dimer steps in the two structures.

(a)

| Step number | SHL (of 1KX4) | Difference in contribution of | | | Difference of radius (Å) |
|---|---|---|---|---|---|
| | | Shift (Å) | Slide (Å) | Rise (Å) | |
| 20 | −5.2 | 0.49 | −1.66 | 0.03 | −2.26 |
| 21 | −5.1 | −0.08 | 0.25 | −0.17 | −2.86 |
| 22 | −5 | −0.13 | −0.1 | −0.2 | −2.79 |
| 124 | 5 | 0.06 | −0.1 | −0.46 | −2.41 |
| 125 | 5.09 | −0.76 | −0.62 | −0.04 | −2.6 |
| 126 | 5.18 | −0.38 | −0.7 | −0.22 | −2.8 |
| 127 | 5.27 | 0.26 | 1.93 | −0.47 | −2.35 |
| 133 | 5.81 | 0.83 | 1.09 | −0.19 | 2.58 |
| 134 | 5.9 | −0.05 | −0.05 | −0.11 | 3.15 |
| 135 | 6 | 0.62 | −0.03 | 0.22 | 2.1 |

(b)

| PDB ID | Step number | Step | SHL | Step parameters | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Shift (Å) | Slide (Å) | Rise (Å) | Tilt (º) | Roll (º) | Twist (º) |
| 1KX3 | 20 | TT/AA | −5.2 | 0.01 | −1.38 | 3.4 | 3.43 | 4.23 | 29.58 |
| 1KX4 | | GG/CC | | −0.62 | 0.62 | 3.72 | 3.06 | 2.32 | 29.9 |
| 1KX3 | 21 | TC/GA | −5.1 | −0.36 | −0.33 | 2.72 | 3.84 | 9.76 | 27.11 |
| 1KX4 | | GA/TC | | −0.29 | −1.23 | 3.39 | 2.16 | 19.68 | 31.22 |
| 1KX3 | 22 | CT/AG | −5 | −0.24 | −0.41 | 3.81 | 1.15 | 11.7 | 28.65 |
| 1KX4 | | AT/AT | | −0.61 | −0.46 | 3.32 | −6.53 | 1.93 | 31.74 |
| 1KX3 | 124 | AG/CT | 5 | 0.32 | 0.08 | 3.29 | 0.59 | 10.65 | 29.83 |
| 1KX4 | | AT/AT | | 0.3 | −0.73 | 2.94 | −0.73 | 5.13 | 25.68 |
| 1KX3 | 125 | GA/TC | 5.09 | −0.42 | −0.87 | 3.08 | −5.11 | 14.41 | 24.28 |
| 1KX4 | | TC/GA | | −1.52 | −1.62 | 3.28 | 4.62 | 0.48 | 34.31 |
| 1KX3 | 126 | AA/TT | 5.18 | 1 | −0.45 | 3.06 | −7.34 | 2.14 | 36.94 |
| 1KX4 | | CC/GG | | 0.48 | −1.13 | 3.08 | 5.29 | 13.2 | 31.59 |
| 1KX3 | 127 | AT/AT | 5.27 | −0.48 | −1.06 | 3.32 | −0.08 | −10.34 | 29.96 |
| 1KX4 | | CG/CG | | −0.23 | 0.91 | 3.59 | −4.12 | 3.35 | 33.24 |
| 1KX3 | 133 | AG/CT | 5.81 | 0.35 | 0.84 | 2.83 | −5.84 | 8.44 | 28.39 |
| 1KX4 | | GG/CC | | −0.78 | −0.77 | 3.89 | 9.99 | 6.98 | 32.48 |
| 1KX3 | 134 | GG/CC | 5.9 | −0.43 | −0.41 | 3.58 | −0.09 | 17.27 | 32.09 |
| 1KX4 | | GA/TC | | −0.14 | −0.06 | 3.08 | −1.86 | 7.19 | 35.76 |
| 1KX3 | 135 | GT/AC | 6 | −0.6 | −0.54 | 3.32 | −2.01 | 6 | 33.02 |
| 1KX4 | | AT/AT | | 1.38 | −0.88 | 3.2 | −2.68 | 7.72 | 30.44 |

The step parameters of the dimers with the large difference in superhelical radii in 1KX4 compared to 1KX3 are reported in Table 7.3 (b). Large differences in Roll at five base-pair steps (dimer steps 21, 22 near SHL −5, dimer steps 124, 125 near SHL 5, and dimer step 134 near SHL 6) contribute to the large changes in superhelical radius in the two structures. For example, Roll increases from 1.93° at dimer step 22 of 1KX3 to 11.7° in 1KX4. At the same time, the long axis of the dimer step 22, given by the unit vector (–0.19, 0.13, 0.97), of dimer step 22 is roughly parallel to the global *Z*-axis. That is to say, the change of Roll about this axis in dimer step 22 affects the superhelical radius. The large differences in superhelical radii at dimer steps 126 and 133 are related to large changes in Tilt in 1KX4 compared to 1KX3 at these steps. The short axis of dimer step 133, along the unit vector (0.15, 0.63, –0.76), is roughly anti-parallel to the global *Z*-axis. The change in Tilt about this axis at dimer step 133, from 9.99° in 1KX3 to –5.84° in 1KX4 is large enough to affect the superhelical radius. There are no significant changes in Tilt, Roll, and Twist in dimer step 135. Dimer step 135 lies near SHL +6, where the difference in radii between 1KX3 and 1KX4 results from conformational changes at dimer steps 133 and 134 (see the large kink in the red curve around SHL +6 in Figure 7.3 (b)). Beyond dimer step 135, the difference of radii of the structure 1KX4 compared to 1KX3 decreases gradually to zero (see the red curve and blue curves around SHL +6 in Figure 7.3 (b)).

### 7.3.4 Superhelical pitch

The superhelical pitch of each base-pair step along the nucleosomal DNA in the three structures 1KX3, 1KX4, and 1KX5 is plotted against the superhelical phase angle in

Figure 7.4. The superhelical pitch of 1KX3 and 1KX5 are very similar (compare the blue and green curves in Figure 7.4). The nucleosomal DNA sequences in these two structures are very similar, except for the one extra base pair in 1KX5.

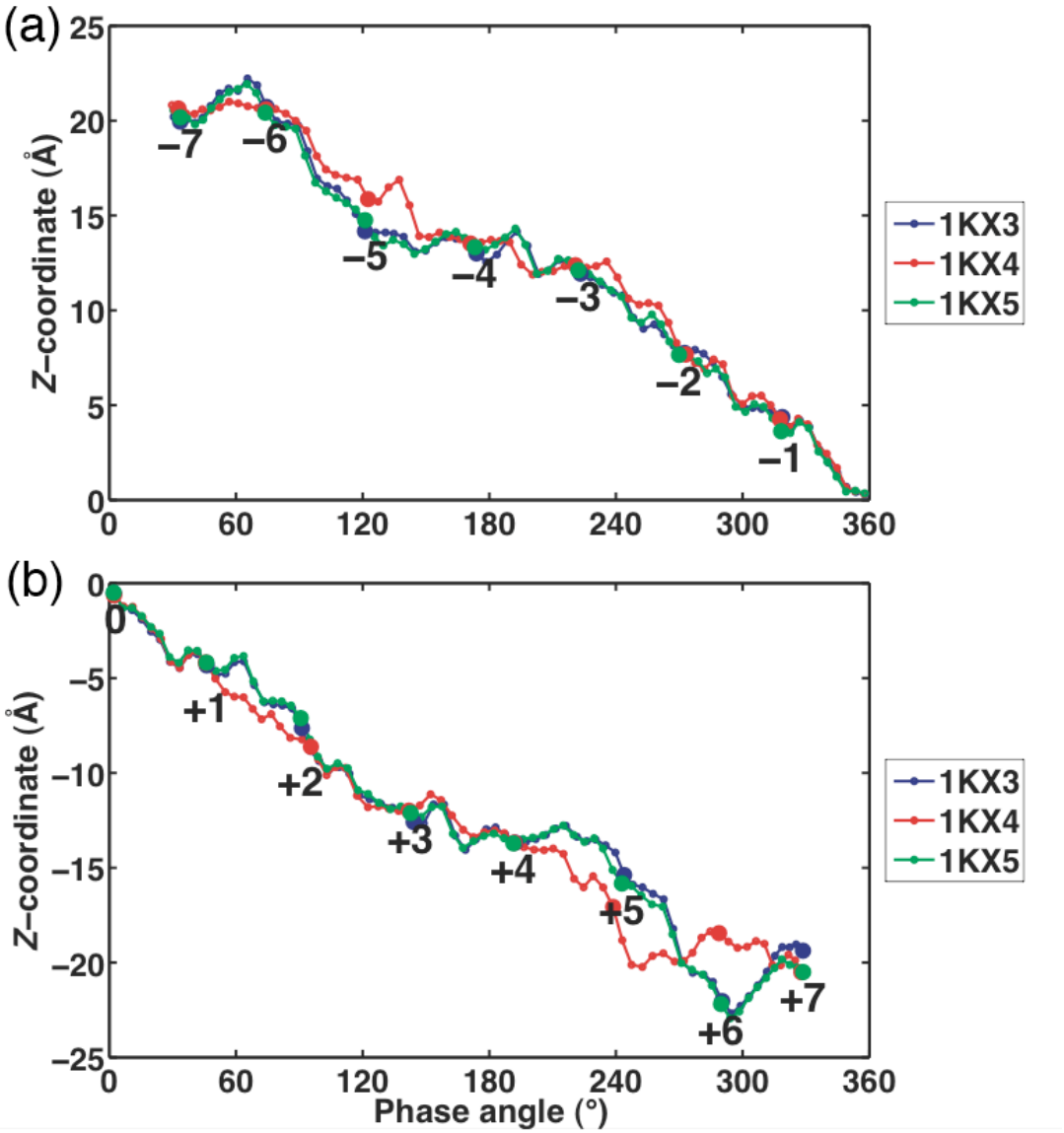



Figure 7.4. Z-coordinate vs. phase angle of each base-pair step in (a) the first half and (b) the second half of the nucleosomal DNA from 1KX3 (blue), 1KX4 (red), 1KX5 (green). The origins of the base-pair steps with integral superhelical numbers are highlighted with larger markers and labeled with the corresponding superhelical numbers. See the legend to Figure 7.7.3 for the base-steps in the first and second halves of the nucleosomal DNA and the assignment of superhelical numbers to dimer steps.

The superhelical pitch of 1KX3 increases significantly over that of 1KX4 at the regions from SHL −5 to −4.5 and from +5.5 to +6.5, and decreases considerably at the regions from SHL +4.5 to +5.5.

The large kinks along the superhelical ramp at dimer steps 24 and 25 (see red curve near SHL −5 in Figure 7.4 (a)) reflect large differences in Tilt and Twist in both steps together with a large contribution of Shift at dimer step 24 (Table 7.4 (a) and (b)). The Shift contribution to superhelical pitch at dimer step 24 is 1.47 Å. At the same time, the short axis (–0.63, 0.77, 0.13) and normal vector (–0.75, –0.64, 0.14) of dimer step 24 are approximately perpendicular to the global *Z*-axis. Thus, the large differences of Tilt and Twist in the structure 1KX4 affect the superhelical pitch compared to 1KX3 (Tilt is 8.04° in 1KX4 vs. −12.23° in 1KX3, and Twist is 38.88° in 1KX4 vs. 30.85° in 1KX3). As mentioned in Section 7.2.1, the Twist of 1KX4 differs from that in 1KX3 in the region from SHL +4 to +6, with 11 bp from SHL +4 to +5 and 10 bp from SHL+5 to +6 in 1KX4 vs. 10 bp and 11 bp, respectively, in 1KX3 (Table 7.1). The change of Twist, in combination with differences in other step parameters, such as Tilt and Slide, cause the large difference (up to –4.22 Å) in superhelical pitch at SHL +4 to +6. That is, the large difference in the *Z*-coordinates at dimer steps 120 to 135 in Table 7.4 (a) mainly reflect the change of Twist in 1KX4 compared to 1KX3. In addition to Twist, Slide in dimer step 123 affects the superhelical pitch, making a contribution of 0.93 Å. The change of Tilt, from −5.75° in dimer step 129 of 1KX3 to 9° of 1KX4 in the same step, leads to the difference of the *Z*-coordinates, as the short axis (–0.37, –0.93, 0.03) is roughly perpendicular to the global *Z*-axis. There are small differences in Shift, Slide, and Rise

contributions to superhelical pitch at dimer steps 135 to 139. At these steps, the differences of the rotational parameters, Tilt, Roll, and Twist, are also not significant. The conformational changes, especially in Twist, at dimer steps 114 to 134 (SHL +4 to +6), in 1KX4 cause the large increase in the *Z*-coordinates at dimer steps 135 to 139 compared to 1KX3 (compare red and blue curves at SHL +6 to +6.5 in Figure 7.4 (b)). Beyond dimer step 139, the large positive difference in the *Z*-coordinates of 1KX4 gradually diminishes to zero, compared to 1KX3.

Table 7.4. (a) Differences in the contributions of Shift, Slide, and Rise to the superhelical pitch at the dimer steps of 1KX4 with large difference (more than 2.0 Å, or less than −2.0 Å) in *Z*-coordinates, compared to 1KX3. (b) The values of step parameters of the corresponding dimer steps in the two structures.

(a)

| Step number | SHL (of 1KX4) | Difference in Contribution of | | | Difference of pitch (Å) |
|---|---|---|---|---|---|
| | | Shift (Å) | Slide (Å) | Rise (Å) | |
| 24 | −4.8 | 1.47 | 0.16 | 0.46 | 2.39 |
| 25 | −4.7 | −0.96 | −0.27 | −0.01 | 2.84 |
| 120 | 4.6 | −1.25 | 0.92 | −0.52 | −2.21 |
| 121 | 4.7 | 0.55 | 0.04 | −0.18 | −2.43 |
| 123 | 4.9 | −0.46 | −0.93 | 0.21 | −2.22 |
| 124 | 5 | 0.06 | −0.8 | 0.63 | −2.87 |
| 125 | 5.09 | −0.76 | −0.49 | 0.23 | −3.44 |
| 126 | 5.18 | −0.38 | −0.11 | 0 | −4.22 |
| 127 | 5.27 | 0.26 | −0.24 | 0.52 | −4.2 |
| 128 | 5.36 | 0.45 | −0.09 | 0.94 | −3.27 |
| 129 | 5.45 | 0.23 | −2.24 | 1.52 | −2.85 |
| 134 | 5.9 | −0.05 | 0.34 | 0.6 | 2.66 |
| 135 | 6 | 0.62 | −0.34 | 0.68 | 3.58 |
| 136 | 6.1 | −0.54 | −0.55 | 0.53 | 3.79 |
| 137 | 6.2 | −1.12 | −0.13 | 0.36 | 3.06 |
| 138 | 6.3 | −0.19 | 0.07 | 0.1 | 2.6 |
| 139 | 6.4 | 0.34 | −0.83 | −0.04 | 2.33 |

(b)

| PDB ID | Step number | Step | SHL | Step parameters | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Shift (Å) | Slide (Å) | Rise (Å) | Tilt (°) | Roll (°) | Twist (°) |
| 1KX3 | 24 | AC/GT | −4.8 | −1.23 | −0.17 | 2.93 | −12.23 | 3.63 | 30.85 |
| 1KX4 | | CG/CG | | 0.34 | 0.79 | 3.52 | 8.04 | −1.82 | 38.88 |
| 1KX3 | 25 | CC/GG | −4.7 | 0.59 | −0.23 | 2.71 | 13.89 | 7.8 | 29.59 |
| 1KX4 | | GT/AC | | −0.44 | 0.48 | 3.48 | −8.32 | 6.97 | 36.53 |
| 1KX3 | 120 | TG/CA | 4.6 | 0.02 | 2.49 | 3.27 | 1.66 | −3.9 | 40.97 |
| 1KX4 | | TA/TA | | 1.57 | 1.3 | 2.87 | 5.96 | −4.64 | 42.37 |
| 1KX3 | 121 | GG/CC | 4.7 | −0.53 | 0.21 | 2.67 | −12.71 | 2.13 | 26.09 |
| 1KX4 | | AC/GT | | −1.06 | −0.25 | 3.34 | 0.79 | −1.86 | 31.71 |
| 1KX3 | 123 | TA/TA | 4.9 | −0.71 | 0.22 | 3.62 | 3.35 | 14.72 | 35.23 |
| 1KX4 | | GA/TC | | 0.06 | −0.83 | 3.64 | −0.55 | 7.07 | 35.15 |
| 1KX3 | 124 | AG/CT | 5 | 0.32 | 0.08 | 3.29 | 0.59 | 10.65 | 29.83 |
| 1KX4 | | AT/AT | | 0.3 | −0.73 | 2.94 | −0.73 | 5.13 | 25.68 |
| 1KX3 | 125 | GA/TC | 5.09 | −0.42 | −0.87 | 3.08 | −5.11 | 14.41 | 24.28 |
| 1KX4 | | TC/GA | | −1.52 | −1.62 | 3.28 | 4.62 | 0.48 | 34.31 |
| 1KX3 | 126 | AA/TT | 5.18 | 1 | −0.45 | 3.06 | −7.34 | 2.14 | 36.94 |
| 1KX4 | | CC/GG | | 0.48 | −1.13 | 3.08 | 5.29 | 13.2 | 31.59 |
| 1KX3 | 127 | AT/AT | 5.27 | −0.48 | −1.06 | 3.32 | −0.08 | −10.34 | 29.96 |
| 1KX4 | | CG/CG | | −0.23 | 0.91 | 3.59 | −4.12 | 3.35 | 33.24 |
| 1KX3 | 128 | TC/GA | 5.36 | 0.38 | −0.59 | 3.48 | 3.71 | −3.28 | 40.82 |
| 1KX4 | | GC/GC | | 1.17 | −0.33 | 3.14 | −4.08 | 9.27 | 36.65 |
| 1KX3 | 129 | CT/AG | 5.45 | −0.91 | −0.4 | 3.43 | −5.75 | 5.2 | 32.02 |
| 1KX4 | | CA/TG | | −0.92 | 1.91 | 3.81 | 9 | −14.04 | 41.08 |
| 1KX3 | 134 | GG/CC | 5.9 | −0.43 | −0.41 | 3.58 | −0.09 | 17.27 | 32.09 |
| 1KX4 | | GA/TC | | −0.14 | −0.06 | 3.08 | −1.86 | 7.19 | 35.76 |
| 1KX3 | 135 | GT/AC | 6 | −0.6 | −0.54 | 3.32 | −2.01 | 6 | 33.02 |
| 1KX4 | | AT/AT | | 1.38 | −0.88 | 3.2 | −2.68 | 7.72 | 30.44 |
| 1KX3 | 136 | TG/CA | 6.1 | 0.09 | −0.01 | 3.46 | 1.78 | 9.35 | 33.6 |
| 1KX4 | | TA/TA | | −0.65 | −0.83 | 3.03 | 8.63 | 10.13 | 31.79 |
| 1KX3 | 137 | GG/CC | 6.2 | 0.67 | 0.35 | 3.13 | −0.09 | 5.92 | 28.63 |
| 1KX4 | | AT/AT | | −0.46 | −0.74 | 3.43 | −5.28 | 2.71 | 37.1 |
| 1KX3 | 138 | GA/TC | 6.3 | −0.53 | −0.4 | 3.17 | −3.94 | 3.07 | 39.28 |
| 1KX4 | | TT/AA | | −0.74 | −0.56 | 3.18 | 3.09 | −7.49 | 39.18 |
| 1KX3 | 139 | AT/AT | 6.4 | −0.16 | −0.79 | 3.27 | −0.81 | −3.26 | 30.02 |
| 1KX4 | | TT/AA | | 0.68 | 0.12 | 3.53 | −2.72 | 2.19 | 36.37 |

The cumulative contributions of Shift, Slide, and Rise to superhelical pitch are consistent with these results. The differences in the accumulation of the contributions Shift and Rise to the superhelical pitch are negative and positive, respectively, in 1KX4 compared to 1KX3 (Figure 7.5 (a) and (c)). The net effect on the difference of *Z*-coordinates is nearly zero, except that positive contribution of Rise exceeds the negative contribution of Shift at SHL +6 to +7. The contribution of Slide is more negative at SHL +5.5 to +7 in 1KX4 compared to 1KX3 (compare the red curve with the blue curve in Figure 7.5 (b)). That is, the difference in the Slide contribution of 1KX4 is negative compared to 1KX3. However, the difference in the *Z*-coordinates at SHL +5.5 to +7 in 1KX4 is positive compared to 1KX5 (Figure 7.4 (b)). Hence, the increase of the Z-coordinates at SHL +5.5 to +7 in 1KX4 vs. 1KX3 is not caused by Slide but rather by Twist, as shown in the previous paragraph.

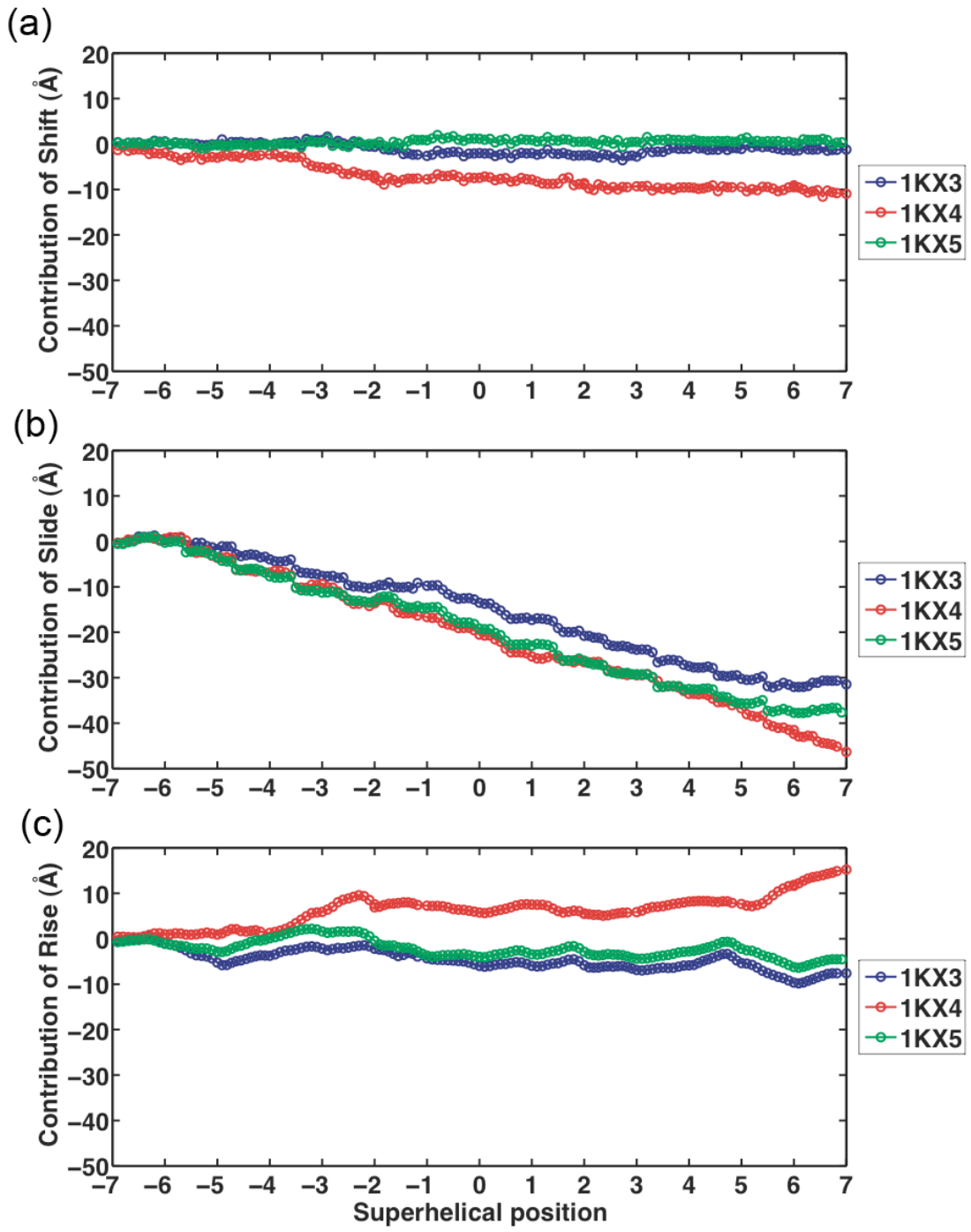



Figure 7.5. The accumulation of the contributions of (a) Shift, (b) Slide, and (c) Rise in the direction of the Z-axis vs. superhelical position of each base-pair step along DNA from the three structures, 1KX3, 1KX4, and 1KX5.

### 7.3.5 Thread the different DNA sequences on structural templates from 1KX3

The DNA sequence from 1KX3 is threaded on the structural templates from 1KX3 (see Section 6.1 in Chapter 6 for the method). The structural templates are made of the central 60, 80, 100, or120 dimer steps of 1KX3. Similar to the best-resolved nucleosome structures (PDB ID: 1KX5) and the SIN-mutant structures, the lowest threading score of all the reading frames along the DNA sequence from 1KX3 remains at SHL 0 for the different length 1KX3 templates (note the deep 'V' shape in the blue curves in Figure 7.6 (a) and (b) for the 60 and 80 base-pair step template; the data of the template length of 100, 120 are not shown).

By contrast, there is no deep global minimum in the threading scores when the 1KX4 sequence is threaded on the 1KX3 templates. The lowest threading score of the 1KX4 sequence on the 60-bp template from 1KX3 occurs at the central position, but differs little from the scores of other threading frames (552 in central position vs. 576 in threading frame 31 at SHL −1.3) (Figure 7.6 (a)). For the 1KX3 template of 80 dimer steps, the reading frame 21 of 1KX4 at SHL −1.3 has the lowest threading score, but this score is very close to the scores of threading frames 34 at SHL 0 and 55 at SHL +2 (812 for reading frame 21 vs. 816 for reading frame 34, and 825 for reading frame 55) (Figure 7.6 (b)).

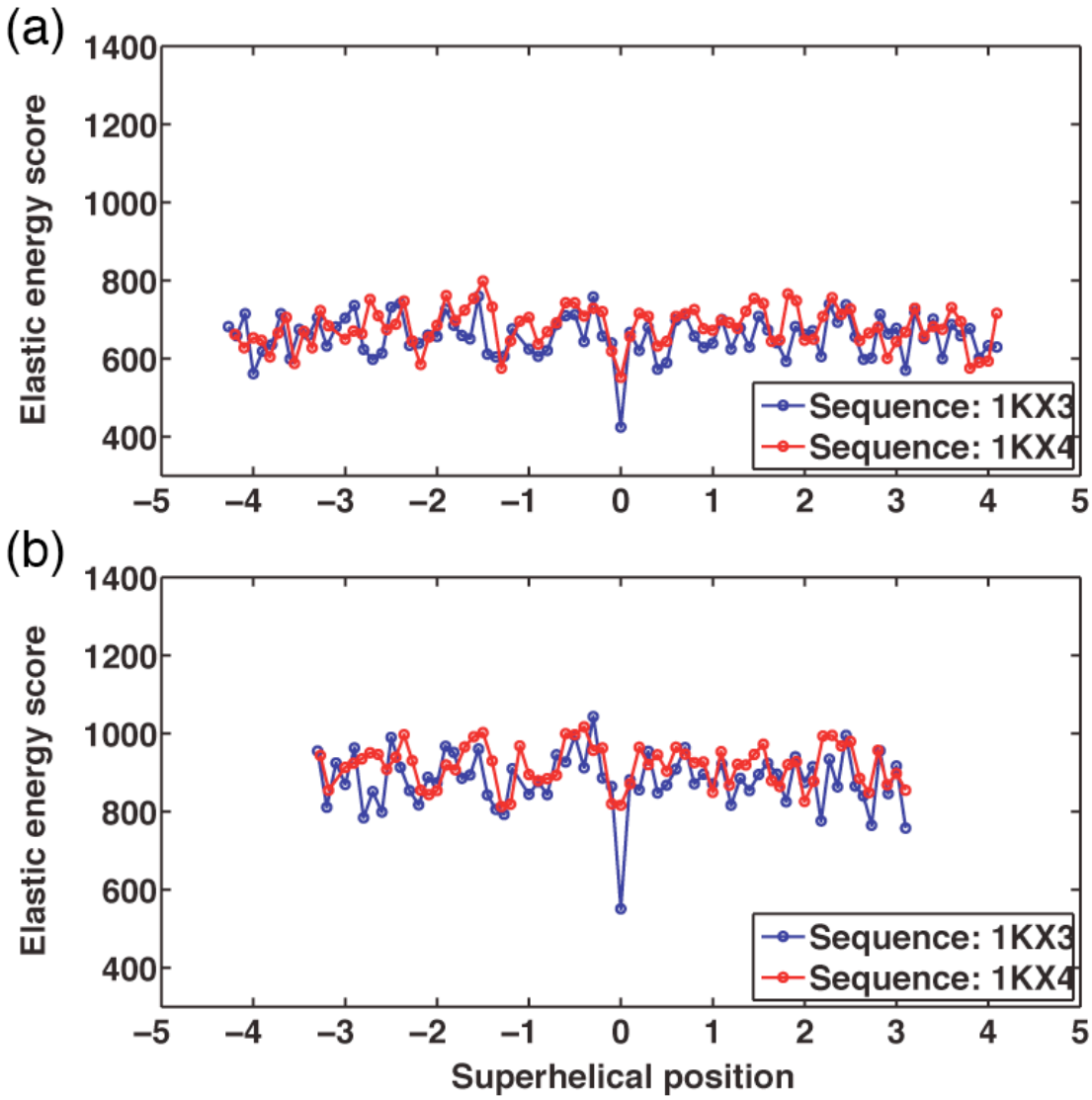



Figure 7.6. Threading score vs. superhelical location of the (a) central 60 dimer-step and (b) central 80 dimer-step template reading frames along the DNA sequences from 1KX3 (blue) and 1KX4 (red). The templates are extracted from 1KX3.

Threading scores of the 1KX4 sequence on the 1KX3 templates are higher than those of the 1KX3 sequence on the 1KX3 templates (for the 120 dimer-step template, the optimum threading score of the 1KX3 sequence is 819 vs. 1224 of 1KX4 sequence). As the same structural templates and force constants are used to compute threading scores, the increased threading scores reflect the capability of the different dimer steps in the

1KX4 sequence, compared to the 1KX3 sequence, to take up the same structural templates. The deformation scores of each base-pair step in the central setting of the 1KX3 sequence threaded on the 120 dimer step template are compared to those of the 1KX4 sequence in the same setting (Figure 7.7 (a)). Significant increases in the deformation score occur at SHL −5.6, −3.5, −1.5 to −1, +1.1, +3.5 and +5.5, where dimer steps CA, TG and TA in the 1KX4 sequence are shifted compared to the flexible dimmers in the 1KX3 sequence (Figure 7.1). The dimer steps of the 1KX4 sequence, which have much large deformation scores than the corresponding dimer steps in the 1KX3 sequence, are listed in Table 7.5 (a). Surprisingly, except for dimer steps 59 and 78 (both of which are GC steps), CA and TG dimer steps in the 1KX3 sequence are replaced by the other dimer steps such as TT, GT, and AC at the high-scoring sites (Table 7.5 (a)). This replacement of flexible CA and TG dimer steps by the stiffer dimer steps with larger force constants causes the increased threading scores of the 1KX4 sequence compared to the 1KX3 sequence. Notably, the difference in the contribution of Slide is the main cause of the increased deformation scores (Figure 7.7 (b) and Table 7.5 (a) column 6). The six step parameters of the dimers with large deformation scores in the 1KX4 vs. 1KX3 sequence threading are listed in Table 7.5 (b). Seven out of 9 dimer steps from the 1KX3 template have large positive Slide values, which are coupled with negative Roll and increased Twist in 5 of the steps. That is to say, in the central (natural) setting of the 1KX3 sequence on the 1KX3 template, the highly deformed steps (large positive Slide, negative Roll and increased Twist) in the 1KX3 structural template are taken up by the flexible steps such as CA and TG. As the dimer steps CA and TG are shifted in the 1KX4 sequence, other stiffer dimer steps adopt this highly deformed conformation and cause

the dramatically increased deformation scores in the central (non-natural) setting of the 1KX4 sequence.

Table 7.5. (a) The total deformation score and contribution of six step parameters at steps with large differences (greater than 15 or less than -15) in total score for the two central settings of 1KX3 and 1KX4 on the 120 dimer-step template of 1KX3. (b) The six step parameters of the corresponding dimer steps in the structural template of 1KX3.

(a)

| Step number | SHL | DNA sequence | Dimer | Contributions of step parameters | | | | | | Total score |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Shift | Slide | Rise | Tilt | Roll | Twist | |
| 16 | | 1KX4 | TT | 0.06 | 18.96 | 1.03 | 0.47 | 17.62 | 1.86 | 42.31 |
| | −5.6 | 1KX3 | CA | 0.01 | 2.18 | 1.14 | 0.71 | 16.78 | 1.26 | 15.61 |
| 38 | | 1KX4 | GT | 0.81 | 48.85 | 0.13 | 0.27 | 29.1 | 11.63 | 72.74 |
| | −3.5 | 1KX3 | TG | 1.38 | 4.38 | 0.37 | 0.15 | 16.53 | 4.91 | 17.12 |
| 59 | | 1KX4 | GA | 8.06 | 0.04 | 25.73 | 0.27 | 12.97 | 0.42 | 46.68 |
| | −1.45 | 1KX3 | GC | 2.86 | 0.34 | 19.52 | 0 | 21.33 | 0.89 | 28.23 |
| 60 | | 1KX4 | AA | 0.27 | 5.71 | 0 | 25.54 | 1.52 | 2.65 | 36.5 |
| | −1.36 | 1KX3 | CA | 0.02 | 0.41 | 0.16 | 15.86 | 0.12 | 1.79 | 15.75 |
| 62 | | 1KX4 | AC | 0.11 | 21.12 | 10.41 | 0.11 | 2.47 | 4.15 | 44.96 |
| | −1.18 | 1KX3 | TG | 0.01 | 1.17 | 4.7 | 0.02 | 0.24 | 1.17 | 7.65 |
| 78 | | 1KX4 | GT | 1.48 | 15.98 | 0.22 | 0.42 | 4.14 | 2.04 | 23.1 |
| | 0.5 | 1KX3 | GC | 1.74 | 5.83 | 0.01 | 0.49 | 3.05 | 0.79 | 4.85 |
| 84 | | 1KX4 | GT | 0.1 | 7.98 | 2.48 | 0.06 | 5.51 | 0.62 | 21.89 |
| | 1.1 | 1KX3 | CA | 0.26 | 0.11 | 0.78 | 0 | 0.99 | 0 | 2.53 |
| 108 | | 1KX4 | AC | 0.06 | 48.52 | 0.04 | 3.8 | 25.49 | 12.85 | 74.37 |
| | 3.4 | 1KX3 | CA | 0.32 | 4.34 | 0.24 | 4.5 | 14.74 | 5.57 | 15.23 |
| 130 | | 1KX4 | AA | 0.96 | 33.88 | 1.92 | 0.9 | 11.6 | 5.54 | 48.74 |
| | 5.5 | 1KX3 | TG | 0.26 | 4.43 | 1.72 | 0.3 | 11.9 | 3.71 | 13.68 |

(b)

| Step | SHL | Step parameters | | | | | |
|---|---|---|---|---|---|---|---|
| | | Shift (Å) | Slide (Å) | Rise (Å) | Tilt (º) | Roll (º) | Twist (º) |
| 16 CA | −5.6 | 0.03 | 1.68 | 3.03 | 3.24 | −21.54 | 41.32 |
| 38 TG | −3.5 | −0.87 | 2.29 | 3.18 | 1.34 | −21.34 | 47.45 |
| 59 TG | −1.45 | 1.41 | 0 | 4.55 | 0.28 | −20.65 | 38.54 |
| 60 CA | −1.36 | −0.15 | 0.85 | 3.25 | 14.87 | 7.29 | 42.53 |
| 62 TG | −1.18 | −0.02 | 1.29 | 4.09 | −0.62 | 8.23 | 41.09 |
| 78 GC | 0.5 | −1.1 | 1.04 | 3.39 | −3.06 | −7.08 | 38.27 |
| 84 CA | 1.1 | −0.45 | 0.55 | 3.67 | 0.35 | 11.52 | 35.3 |
| 108 CA | 3.4 | 0.39 | 2.28 | 3.22 | 7.97 | −19.87 | 48.26 |
| 130 TG | 5.5 | −0.35 | 2.3 | 2.95 | 1.91 | −17.34 | 45.84 |

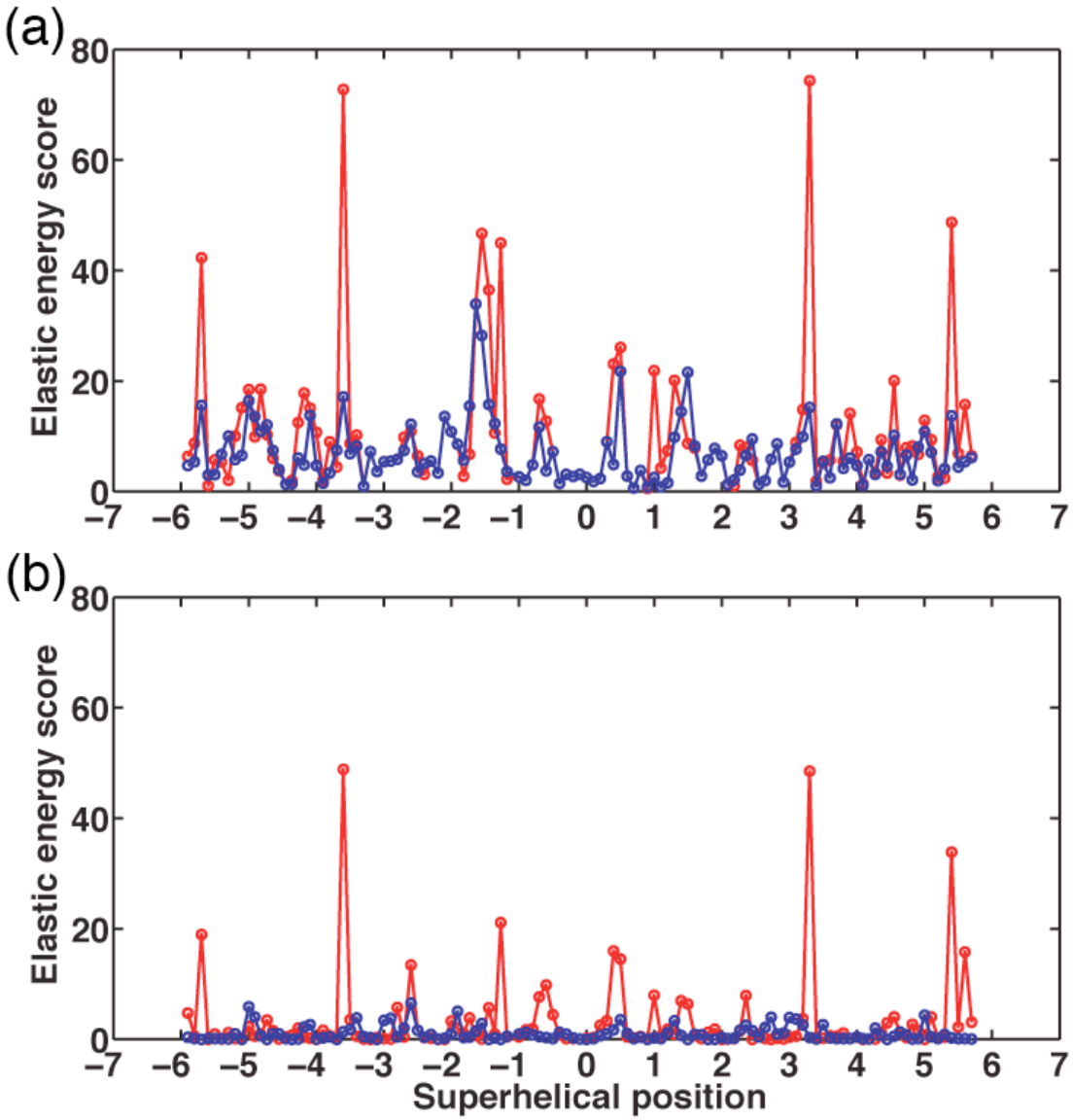



Figure 7.7. (a) Deformation score and (b) contribution of Slide of each dimer step in the central setting of the 1KX4 (red) and 1KX3 (blue) sequences threaded on the 120 dimer-step template from 1KX3.

### 7.3.6 Threading the 1KX4 and 1KX3 sequences on the 1KX4 template

A similar situation happens when the 1KX4 and 1KX3 sequences are threaded on the central 60 to 120 dimer-step templates from the 1KX4 structure. The natural central

setting of the 1KX4 templates has the lowest threading score of all the threading frames along the nucleosomal DNA from 1KX4. The deep global minimum in the threading scores is lost when the 1KX3 sequence is threaded on one of the 1KX4 templates (Appendix Figure A7.2).

The 1KX3 sequence threaded on the varios 1KX4 templates has higher threading scores than the 1KX4 sequence. Table 7.6 (a) lists the deformation scores of each dimer step of the central (natural) setting of the 1KX4 sequence compared to those the central (non-natural) setting of the 1KX3 sequence on the 120 dimer-step 1KX4 template. The dimer steps with large difference (greater than 15 or less than in −15) in deformation scores occur around SHL ±5.5, −4.8, ±3.5, +0.5, and +1.6 (Figure 7.8 (a) and Table 7.6 (a)). Three dimer steps (16, 78, and 108) have lower deformation scores for the 1KX3 sequence than the 1KX4 sequence. The other 10 dimer steps (17, 19, 24, 37, 40, 89, 109, 126, 127, and 129) show the higher deformation scores in the 1KX3 sequence than in the 1KX4 sequence (Table 7.6 (a)). The Slide contribution is again the main cause of the increased deformation scores (Figure 7.8 (b) and Table 7.6 (a), column 6). In dimer step 89, Tilt also contributes a lot to the increased deformation score; and in dimer step 129, Roll and Tilt-Rise coupling (6.28 in 1KX3 vs. −1.39 in 1KX4) in dimer step 129 add to the increased deformation scores. Except for the large positive Roll in dimer steps 40 and 129, the other dimer steps have negatively Roll. Four dimer steps (37, 89, 108, 129) have highly deformed conformation — large positive Slide, increased Twist, and large negative Roll (Table 7.6 (b)). Five of these ten dimer steps with large positive Slide are CA or TG steps, which adopt the highly deformed conformation with the lower deformation scores in the central natural setting of the 1KX4 sequence. The replacement

of these CA and TG dimer steps by other dimer steps in the central setting of the 1KX3 sequence causes the increased deformation scores in the same template.

Table 7.6. (a) The total deformation score and contribution of six step parameters at steps with large difference (greater than 15 or less than −15) in total score for the two central settings of 1KX4 and 1KX3 on the 120 dimer-step template of 1KX4. (b) The six step parameters of the corresponding dimer steps in the structural template of 1KX4.

(a)

| Step number | SHL | DNA sequence | Dimer | Contribution of step parameter | | | | | | Total score |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Shift | Slide | Rise | Tilt | Roll | Twist | |
| 16 | −5.6 | 1KX3 | CA | 2.31 | 0.45 | 1.97 | 0.45 | 1.31 | 0.19 | 5.03 |
| | | 1KX4 | TT | 7.69 | 6.06 | 7.43 | 0.23 | 0.34 | 0.28 | 20.16 |
| 17 | −5.5 | 1KX3 | AG | 0.18 | 19.81 | 0.21 | 3.11 | 2.52 | 3.79 | 31.48 |
| | | 1KX4 | TG | 0.05 | 4.18 | 0.18 | 1.35 | 1.83 | 1.57 | 5.65 |
| 19 | −5.3 | 1KX3 | AT | 5.04 | 17.4 | 0.17 | 0.01 | 0.58 | 4 | 23.31 |
| | | 1KX4 | CG | 3.72 | 1.42 | 0.98 | 0 | 1.13 | 1.53 | 7.71 |
| 24 | −4.8 | 1KX3 | AC | 0.03 | 11.55 | 0.97 | 3.88 | 0.63 | 2.43 | 21.26 |
| | | 1KX4 | CG | 0.15 | 0.37 | 0.18 | 3.49 | 0.91 | 0.91 | 4.59 |
| 37 | −3.55 | 1KX3 | TT | 0.03 | 49.39 | 4.13 | 2.4 | 10.65 | 10.35 | 73.33 |
| | | 1KX4 | TG | 0.01 | 6.86 | 0.9 | 1 | 11.11 | 6.92 | 13.64 |
| 40 | −3.2 | 1KX3 | GA | 3.19 | 5.53 | 6.75 | 1.4 | 7.52 | 1.03 | 34.23 |
| | | 1KX4 | CA | 1.21 | 1.17 | 3.03 | 2.33 | 4.6 | 0.86 | 16.47 |
| 78 | 0.5 | 1KX3 | GC | 0.83 | 10.86 | 0.02 | 1.48 | 2.58 | 1.17 | 7.93 |
| | | 1KX4 | GT | 0.56 | 26.98 | 0.26 | 1.54 | 3.54 | 2.69 | 34.72 |
| 89 | 1.6 | 1KX3 | CT | 1.99 | 6.18 | 0.03 | 11.29 | 16.07 | 0.63 | 34.04 |
| | | 1KX4 | CC | 2.69 | 4.43 | 0.05 | 4.82 | 11.46 | 0.33 | 16.56 |
| 108 | 3.4 | 1KX3 | CA | 6.2 | 0.33 | 0.73 | 0.79 | 6.34 | 1.09 | 9.28 |
| | | 1KX4 | AC | 5.39 | 11.55 | 2.35 | 0.55 | 9.16 | 3.97 | 28.88 |
| 109 | 3.5 | 1KX3 | AA | 2.05 | 13.45 | 3.41 | 0.14 | 2.67 | 0.06 | 30.2 |
| | | 1KX4 | CA | 0.94 | 1.4 | 0.68 | 0 | 4.01 | 0.04 | 6.21 |
| 126 | 5.2 | 1KX3 | AA | 0.79 | 5.27 | 0.62 | 4.06 | 5.51 | 0.59 | 20.31 |
| | | 1KX4 | CC | 0.42 | 0.62 | 1.87 | 2.49 | 1.83 | 0.1 | 4.48 |
| 127 | 5.3 | 1KX3 | AT | 0.1 | 12.4 | 2.63 | 1.5 | 0.23 | 0.43 | 19.89 |
| | | 1KX4 | CG | 0.07 | 0.61 | 0.47 | 0.92 | 0.08 | 0.06 | 2.69 |
| 129 | 5.5 | 1KX3 | CT | 0.93 | 15.03 | 3.84 | 7.18 | 13.49 | 3.06 | 58.51 |
| | | 1KX4 | CA | 1.23 | 2.92 | 1.81 | 5.77 | 8.65 | 1.16 | 21.63 |

(b)

| Step | SHL | Step parameters | | | | | |
|---|---|---|---|---|---|---|---|
| | | Shift (Å) | Slide (Å) | Rise (Å) | Tilt (°) | Roll (°) | Twist (°) |
| 16 TT | −5.6 | 1.14 | 0.88 | 3.84 | 2.59 | −2.37 | 37.5 |
| 17 TG | −5.5 | −0.13 | 2.24 | 3.24 | 4.19 | −3.72 | 42.05 |
| 19 CG | −5.3 | 1.67 | 1.2 | 3.15 | −0.3 | −2.67 | 40.21 |
| 24 CG | −4.8 | 0.34 | 0.79 | 3.52 | 8.04 | −1.81 | 38.88 |
| 37 TG | −3.55 | −0.01 | 2.81 | 3.69 | −3.81 | −16.58 | 49.78 |
| 40 CA | −3.2 | 0.81 | 1.29 | 3.95 | −5.55 | 19 | 40.24 |
| 78 GT | 0.5 | −0.76 | 1.54 | 3.4 | −5.3 | −6.43 | 39.26 |
| 89 CC | 1.6 | −1.26 | 1.11 | 3.39 | −8.93 | −15.69 | 36.37 |
| 108 AC | 3.4 | 1.9 | 0.79 | 3.66 | 3.41 | −11.29 | 40.89 |
| 109 CA | 3.5 | 0.71 | 1.39 | 3.65 | 0.09 | −7.93 | 36.22 |
| 126 CC | 5.2 | 0.47 | −1.13 | 3.08 | 5.28 | 13.2 | 31.59 |
| 127 CG | 5.3 | −0.23 | 0.91 | 3.59 | −4.12 | 3.35 | 33.24 |
| 129 CA | 5.5 | −0.92 | 1.91 | 3.82 | 9.01 | −14.03 | 41.08 |

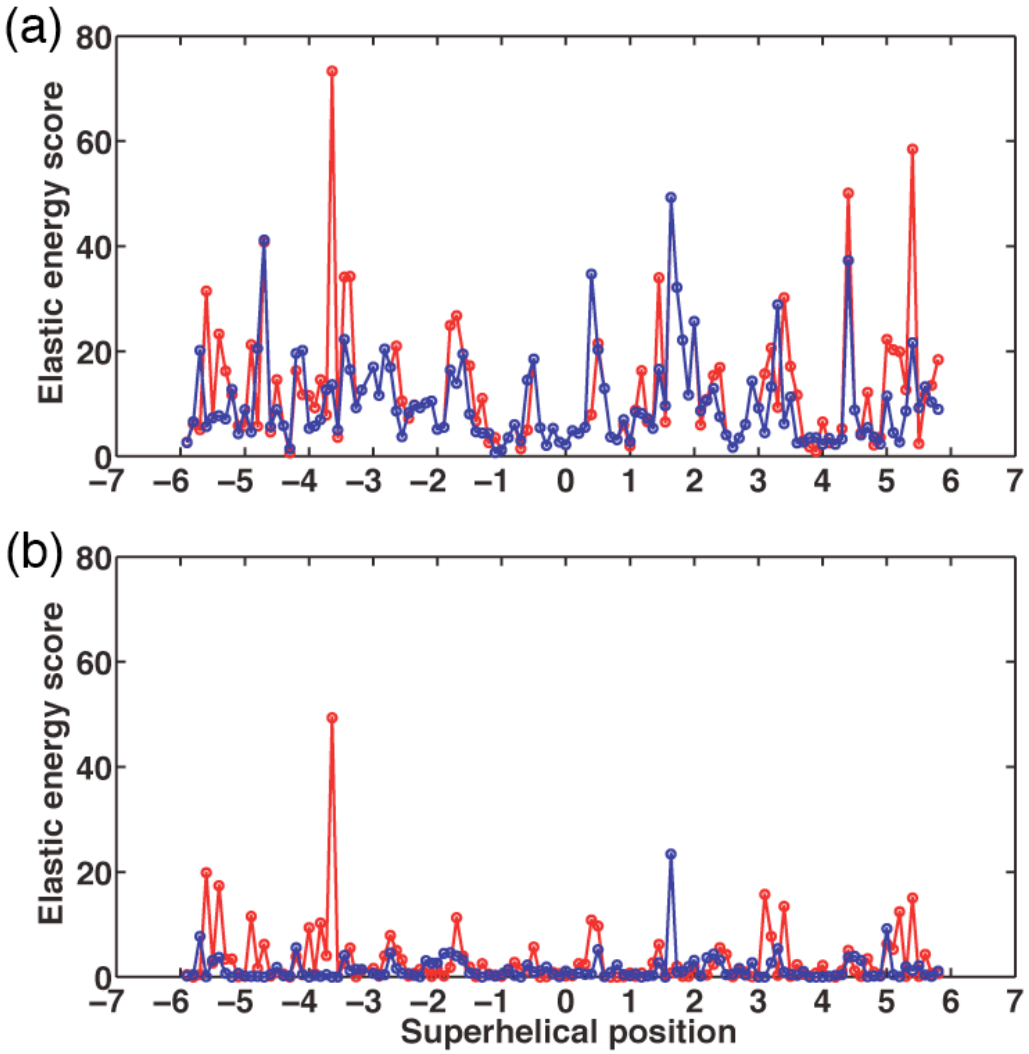



Figure 7.8. (a) Deformation score and (b) contribution of Slide of each dimer step in the central setting of the 1KX3 (red) and 1KX4 (blue) sequences threaded on the 120 dimer-step template from 1KX4.

7.3.7 Compare the highly deformed steps in the 1KX3 template and 1KX4 template

The dimer steps with the most highly deformed conformations of the 1KX3 and 1KX4 templates with large positive Slide, negative Roll, and increased Twist are listed in Table 7.5 (b) and Table 7.6 (b), respectively.

Except for the slight shifting of the locations of the highly deformed steps, the steps with large positive Slide, negative Roll, and increased Twist are located around SHL ±5.5,

±4.5, ±3.5 in both 1KX3 and 1KX4 structures (Figure 7.9). Notably, the TA dimer step 26 at SHL −4.6 has a huge Roll (−37.5°) accompanied by large positive Slide (2.68 Å) and increased Twist (45.06°) in 1KX4.

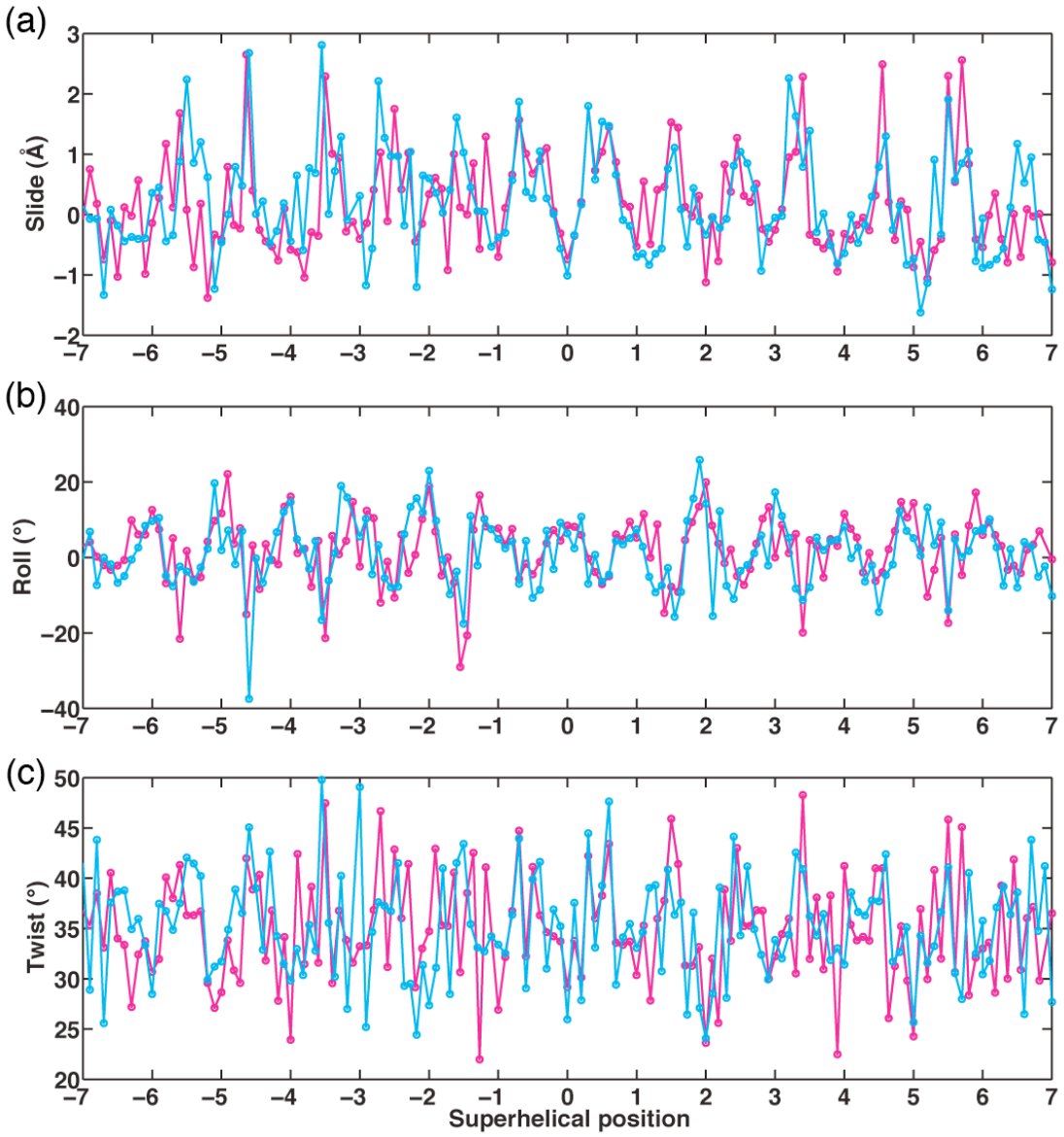



Figure 7.9. (a) Slide, (b) Roll, and (c) Twist vs. superhelical location of the structures 1KX3 (magenta) and 1KX4 (cyan).

The locations of the highly deformed steps are shifted one or two steps away near SHL ±5.5 and ±3.5. The highly deformed conformation of dimer step 16 CA at SHL −5.6 in 1KX3 shifts by one base pair to a deformed 17 TG dimer step at SHL −5.5 in 1KX4. These changes in structure are clear from the shifted peaks of Slide and Twist in Figure 7.9 (a) and (c). As Roll is negative at both dimer steps (−2.37° at 17 TG in 1KX3 vs. −21.54° at16 CA in 1KX3), there is no peak shift near SHL −5.5 in Figure 7.9 (b). The peaks of Slide, Roll, and Twist also shift at SHL −3.5 at the point of where the highly deformed steps 38 TG in 1KX3 shifts to the deformed step 37 TG in 1KX4. The symmetrically-related deformed dimer step 108 CA at SHL +3.4 in 1KX3 shifts to the deformed dimer step 109 CA at SHL +3.5 in 1KX4. A large Slide (2.26 Å) is also taken up by a nearby TG dimer (dimer step 106) at SHL +3.1 in 1KX4. This TG dimer in 1KX4 is replaced by a TC dimer step at the corresponding location in 1KX3, where Slide is small (2.26 Å in 1KX4 vs. 0.95 Å in 1KX3). However, Roll increases from −19.87° in dimer step 108 CA in 1KX3 to −7.93° at dimer step 109 CA in 1KX4. Hence the peak of Roll at SHL +3.5 diminishes rather than shifts. Like the peak in Roll, the peak of Twist also diminishes at SHL +3.5 because of the decrease of Twist in dimer step 109 in 1KX4 compared to dimer step 108 in 1KX3 (36.22° vs. 48.26°). The highly deformed conformation of dimer step 130 TG at SHL +5.5 in 1KX3 shifts to dimer step 129 CA at SHL +5.5 in 1KX4. As dimer steps 130 in 1KX3 and 129 in 1KX4 have the same superhelical location, the two peaks of Slide, Roll, and Twist overlap, respectively, at SHL +5.5 in Figure 7.9 (a), (b) and (c).

## 7.4 Summary and discussion

The two structures, 1KX3 and 1KX5, have the almost the same DNA sequence except for an extra base pair near the dyad in 1KX5 (146 bp in 1KX3 vs. 147 bp in 1KX4) and very similar superhelical properties. The extra base pair in 1KX5 is accommodated by over-twisting in the region from SHL −2 to −1 in 1KX5.

The two structures, 1KX3 and 1KX4, have different nucleosomal DNA sequences, but the length of these two DNA sequences is the same, 146 bp. Three regions of the nucleosomes, SHL −5 to −4, −3 to −2, and +4 to +6 in 1KX4, show notable rotational changes, *i.e*. there are change of both Twist and the number of base pairs between the integral superhelical locations. The change of Twist has a dramatic effect on the superhelical pitch of the region from SHL +4 to +6 in 1KX4 compared to 1KX3 (Figure 7.4 (b)). The region from SHL +5 to +6 also undergoes a significant change in superhelical radius (Figure 7.3 (b)), which is mainly caused by Roll and Tilt. Rise contributes to the increased *Z*-coordinates at SHL +6 to +7. On the other hand, the superhelical ramp near SHL −5.5 to −4.5 is kinked along both the *Z*-axis and the radial direction (Figure 7.4 (a) and Figure 7.3 (a)). The large kink in the superhelical pitch at SHL −4.8 comes from Shift, Tilt, and Twist. The considerable decrease of radii at SHL −5.2 to −5 is caused by Slide and Roll.

To summarize, the region near SHL −5.5 and −4.5 has a sharp and dramatic local change in both superhelical pitch and radius in 1KX3 vs. 1KX4. The change of superhelical pitch is caused by the difference in Tilt, Twist, and Shift of the dimer steps near SHL −4.8. The change of superhelical radius is caused by the change of Tilt, Roll, and Slide of the dimer

steps from SHL −5.2 to −5. The superhelical pathway of the region near SHL +4 to +7 has large-scale changes in superhelical pitch and radii. The change of superhelical pitch from SHL +4 to +7 is due to the change of Twist from SHL +4 to +6. The change of superhelical radius from SHL +5 to +6 is caused by Roll and Tilt.

The structural templates from 1KX3 have been threaded on the sequences from 1KX3 and 1KX4. The natural central settings of the 1KX3 sequence threaded on the 1KX3 always have the lowest threading scores. Highly deformed conformations — with large positive Slide, negative Roll, and increased Twist — are accommodated by flexible dimer steps such as CA and TG near SHL ±3.5 and ±5.5 so that the deformation score in low natural central setting. By contrast, in the central non-natural setting of the 1KX4 sequence, the same highly deformed conformations are costly for 1KX4 since the CA and TG dimer steps are replaced by other stiffer dimer steps such as TT, GA, and AA with higher force constants. A similar situation happens when the 1KX3 and 1KX4 sequences are threaded on the 1KX4 templates.

Interestingly, the locations of the dimer steps with the most highly deformed conformations (large positive Slide, negative Roll and increased Twist) shift to take account of the locations of the flexible CA and TG dimer steps in the difference nucleosomal sequence. Near SHL ±3.5 and ±5.5, the CA or TG dimer steps that take up large positive Slide, negative Roll, and increased Twist in 1KX3 shift by one or two base pairs and adopt similar conformation in 1KX4. However, the superhelical locations of these shifting dimer steps are very close or even the same (dimer step 108 in 1KX3 and 109 in 1KX4). This result suggests that the locations of SHL ±3.5 and ±5.5 are fixed spots along the superhelical ramps of nucleosomal DNA and at the same time

nucleosomal DNA adjust its conformation according to its sequence specificity, *i.e*. the locations of the flexible dimer steps such as CA and TG.

The comparison between 1KX3 and 1KX4 in aspects of the superhelical properties and the threading scores shows how the difference nucleosomal DNA sequences accommodate a histone protein core. Although different regions of DNA in the 1KX3 and 1KX4 structures are subjected to rotational changes such as over-twisted, the base pairs with integral superhelical locations in these two structures have very similar phase angles in the cylindrical frames. This suggest that, regardless of the specific nucleotide content, the angular locations of the integral superhelical locations may serves as fixed points in the superhelical ramp of nucleosomal DNA. The locations of SHL ±3.5 and ±5.5 may also serve as fixed spots, around which the flexible CA, TG or TA dimer steps take up the highly deformed conformation — large positive Slide, negative Roll, and increased Twist. This highly deformed conformation plays an important role in nucleosomal DNA bending and the formation of superhelical pitch [3].

1. Davey, C.A., Sargent, D. F., Luger, K., Maeder, A. W., and Richmond, T. J., *Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution*. J. Mol. Biol., 2002. 319: p. 1097-1113.
2. Colasanti, A.V., *Conformatioinal state of double helix DNA*, Ph.D. Thesis 2006, Rutgers, the State University of New Jersey, New Brusnwick.
3. Tolstorukov, M.Y., Colasanti, A. V., McCandlish, D., Olson, W. K., and Zhurkin, V. B., *A novel 'roll-and-slide' mechanism of DNA folding in chromatin. Implications for nucleosome positioning*. J. Mol. Biol., 2007. 371: p. 725-738.

# Chapter 8 Discussion

The analyses of the SIN-mutant and wild-type nucleosome structures in this thesis provides new insights into the mechanisms of nucleosome remodeling in three aspects: the locations of remodeling; the interactions between histone subunits involved in remodeling; and the potential pathways of nucleosomal DNA remodeling.

## 8.1 One end of nucleosomal DNA and the region from SHL −3 to −2 are important for the nucleosome remodeling.

Comparison of the SIN-mutant nucleosome structures with the wild-type structure reveals notable differences in protein-DNA interactions, step parameter values, superhelical shape, and deformatioin scores. Two regions at one end of the nucleosomal DNA (SHL −6 to −4.5) and at nucleotides two to three helical turns from the dyad (SHL −3 to −2), show major differences in the SIN-mutant structures compared to the wild-type one. DNA locations near the dimerization interface of histones H3H4 and H2AH2B (SHL +3 on strand I and SHL −3 on strand II) lose close contacts (< 3.0 Å) of protein and DNA. The regions toward the two ends of DNA tend to lose more distant contacts with larger cutoffs varying from 3.2 Å to 4.0 Å. The contacts of arginine and lysine residues to DNA in the histone proteins of the SIN mutants do not change much compared to those in the wild type. Rather than arginine and lysine, the contact losses at the two regions mentioned above involve serine residues.

The loss of contacts in the SIN-mutant structures compared to the wild-type is accompanied by appreciable conformational changes near one end of nucleosomal DNA, from SHL −5.5 to −4.5, and in a more central region, from SHL −2.8 to −1.5. The second

halves of the SIN-mutant nucleosomal DNA structures (base pairs 74 to 146 from SHL 0 to +7), however, resemble the wild-type structure. This similarity may be related to the pseudo-symmetry of the nucleosome structures, *i.e.* the first half of the structure spans 72 bp and the second half spans 73 bp (Chapter 3).

The superhelical radius increases consistently at neighboring base-pair steps toward one end of the DNA (SHL −6.5 to −4.5) in five SIN-mutant structures (R116H H3, T118H H3, T118I H3, R45E H4, and R45H H4 SIN mutants). The radius increases in the R45C H4 SIN mutant structure at three successive dimer steps, from SHL −2.8 to −2.5.

The threading of the common DNA sequence on the structural templates from the SIN-mutant and wild-type nucleosomes produces distinctly different scores. The same steps in the SIN-mutant templates, found at SHL −6 to −4 and SHL −2 to −1, always raise the deformation scores.

We propose that, in the SIN-mutant nucleosomes, the disruption of protein-DNA interactions at two regions, at one end of nucleosomal DNA bound by histones H2AH2B and at a location near the H3H4-H2AH2B dimerization interface (SHL −3 to −2), may help overcome the deficiency of the SWI/SNF complex.

Photochemical studies show consistent results [1, 2]. The position of the histone core on a nucleosomal DNA sequence can be mapped with high resolution, photochemical cross-linking approaches. For example, the modified amino acid produced by mutation of serine 47 in histone H4 to cysteine and subsequent attachment to a photoreactive (aryl azides) group, cross-links the proximal nucleotide under UV exposure. Alkaline cleavage of the DNA at the cross-linking site reveals the contacted nucleotide if the 5´ end of strand I or II of the DNA is radio-labeled. This method has mapped the translocation of

native histone octamers along a DNA sequence and detected the contacts lost between the modified H4 cysteine residue and specific locations of DNA [1, 2]. During the nucleosome remodeling of SWI/SNF complexes as much as 50 bp peels off one end of nucleosomal DNA, disrupting the interactions of histones H2AH2B with the DNA [1, 2]. Perturbation of the DNA sequence show that the regions from SHL +3 to +2 and SHL −3 are very important for the nucleosome remodeling [2]. Specifically, one-nucleotide gaps of nucleosomal DNA in the region from SHL +3 to +2 with respect to the observed position, prevents nucleosome sliding on DNA, *i.e.*, the percentage of nucleosome sliding decreases and the percentage of immobilized nucleosomes increases. Gaps of two- and four- nucleotides at SHL −3 also interfere with the nucleosome translocation in terms of a decreased percentage of sliding nucleosomes.

8.2 SIN mutations transmit their information through allosteric interactions between histone H3H4 and H2AH2B.

The H2AH2B histones, which bind the one end of nucleosomal DNA (from SHL −7 to −5), stack with the H3H4 histones bound in the vicinity of SHL +0.5, where the site-specific changes in protein sequence occur. The SIN mutations in histone H3 or H4 may affect the stacking between histones H2AH2B and histones H3H4, leading to the observed disrupting of protein-DNA interactions, in terms of contact loss and DNA conformational changes, at the end of nucleosomal DNA (from SHL −7 to −5).

8.3 DNA may peel off or loop away from the histone core during nucleosome remodeling.

The simulated remodeling of nucleosome DNA, based on the highly deformed base-pair steps clustered at SHL −5.5 to −5.1 and −2.1 to −1.3 in the R45H H4 SIN mutant, suggests how the mutations might perturb the structure of wild-type nucleosomal DNA. The remodeled DNA shows a stretched end (SHL −7 to −5), diminished superhelical pitch (from SHL −7 to +1), and a bulged-out region (from SHL +2 to +6). These features suggest that DNA may peel off or loop away from the histone core during nucleosome remodeling. Interestingly, the base pairs at SHL −1.5 and +6 in the remodeled DNA show exceptionally low RMSD values from the corresponding atoms in the wild-type structure even though the remodeled DNA undergoes dramatic changes compared to the wild-type DNA. The steps at SHL −1.5 and +6 may thus serve as anchor points of the histone core to the nucleosomal DNA.

8.4 The relationship between the heat-shifting experiments and the threading scores of the SIN mutant nucleosomal DNA.

The SIN-mutant structures, which have high threading scores, also exhibit high sliding rates from two off-center positions to the central position in heat-shifting experiments [3]. The extremely high threading scores suggest that the two off-center positions are not favored by the SIN-mutant nucleosomes in the high-scoring group. The low threading score at the central position appears to favor binding of the SIN-mutant nucleosomes to this site rather than to the off-center positions. Most of the SIN-mutants in the high-scoring group form nucleosomes at the central position in heat-shifting experiments at a

low temperature (4 °C). Only under the careful control of the reconstitution temperature and sample handling, is it possible for the SIN-mutant nucleosomes in the high-scoring group to be positioned, albeit briefly (under 5 min), at the two off-center positions [3]. Notably, the lowest threading scores of the SIN mutants in the high-scoring group are comparable to the average threading scores of the mutants in the low- and intermediate groups. Although the SIN mutants in the intermediate- and low-scoring groups can occupy the two off-center positions at low temperature (4 °C), these positions are not optimally stable. After heating at 37 °C, these nucleosomes only occupy the central position, the most favorable position in terms of the lowest threading scores. Furthermore, the central position appears to be more stable than the two off-center positions for the SIN-mutant nucleosomes in the high-scoring group, while the stability of the central position and the two off-center positions appears to be quite comparable at low temperature (4 °C) in the SIN-mutants in the low- and intermediate- scoring groups. Consistently, the well depth of the threading score curves (Figure 6.8 in Chapter 6) of the high-scoring group, *i.e*. the difference between the lowest threading scores at the central position and the average threading scores, is larger than that of the low- and intermediate-scoring groups. The threading scores obtained from the knowledge-based elastic function thus appear to be good measures of the relative ease of deforming a DNA sequence on the surface of a nucleosome core particle.

8.5 Locations of base-pair steps with highly deformed conformation are shifted in different nucleosomal DNA sequences.

The analysis of the two nucleosomal structures (PDB ID: 1KX3 and 1KX4), which have identical histone cores but different 146-bp DNA sequences, reveals how different DNA

sequences associate with the histone protein core. Consistent with the findings of M. Y. Tolstorukov *et.al*. [3], flexible dimer steps such as CA and TG adopt highly deformed conformations — with large positive Slide, negative Roll, and increased Twist —near SHL ±3.5 and ±5.5 in the nucleosome structures. Interestingly, the locations of the most highly deformed conformations shift along the DNA to take advantage of the locations of the flexible CA and TG dimer steps in the different nucleosomal sequences. This suggests that, regardless of the specific nucleotide content, this highly deformed conformation (large positive Slide, negative Roll, and increased Twist) may play an important role in nucleosomal DNA deformation and positioning.

1. Kassabov, S.R., Zhang, B., Persinger, J., and Bartholomew, B., *SWI/SNF unwraps, slides, and rewraps the nucleosome.* Mol. Cell, 2003. 11: p. 391-403.
2. Zofall, M., Persinger, J., Kassabov, S. R., and Bartholomew, B., *Chromatin remodeling by ISW2 and SWI/SNF requires DNA translocation inside the nucleosome.* Nat. Struct. Mol. Biol., 2006. 13: p. 339-346.
3. Tolstorukov, M.Y., Colasanti, A. V., McCandlish, D., Olson, W. K., and Zhurkin, V. B., *A novel 'roll-and-slide' mechanism of DNA folding in chromatin. Implications for nucleosome positioning.* J. Mol. Biol., 2007. 371: p. 725-738.

Appendix

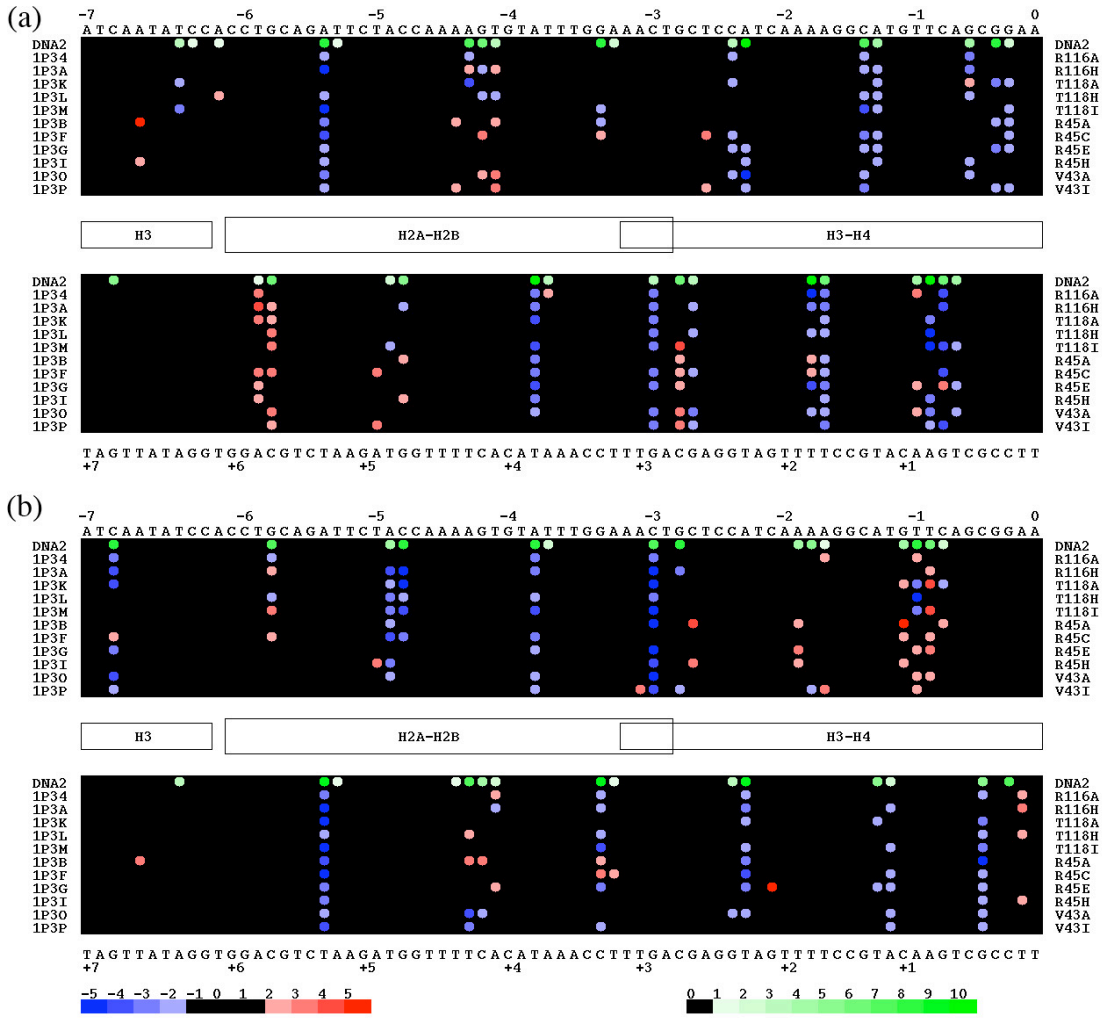



Figure A4.1. Difference in the numbers of close contacts within 3.6 Å to DNA phosphate atoms in (a) DNA strand I and (b) DNA strand II compared with the wild type. Data are mapped on each nucleotide along the nucleosomal DNA. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The mutation locations are shown at the right edge of the diagram. The number of contacts in the wild type is shown in different shades of green, and the difference in contacts to mutant vs. the wild type in shades of blue (negative) and red (positive). Differences of –1, 0, +1 are shown in black.

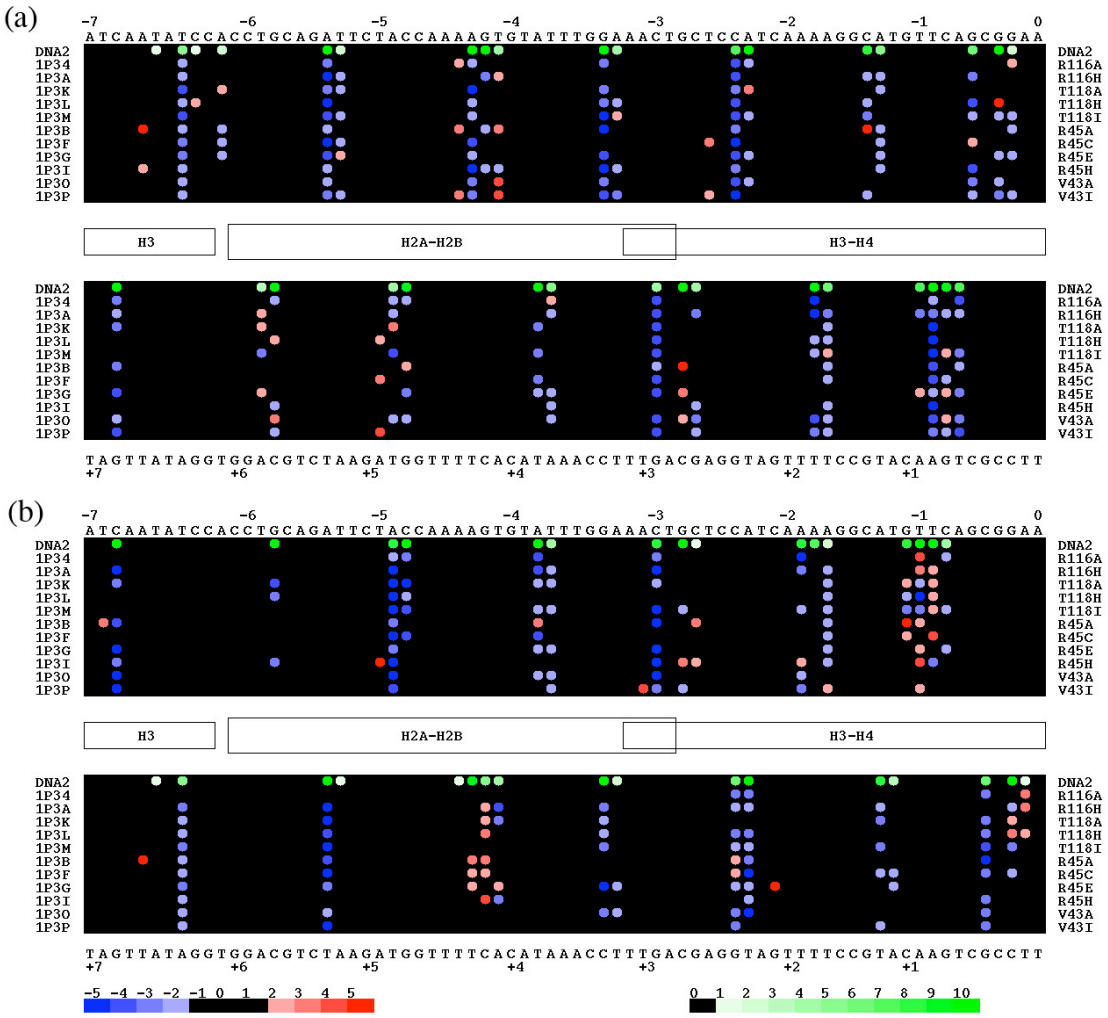



Figure A4.2. Difference in the numbers of close contacts within 3.8 Å to DNA phosphate atoms in (a) DNA strand I and (b) DNA strand II compared with the wild type. Data are mapped on each nucleotide along the nucleosomal DNA. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The mutation locations are shown at the right edge of the diagram. The number of contacts in the wild type is shown in different shades of green, and the difference in contacts to mutant vs. the wild type in shades of blue (negative) and red (positive). Differences of –1, 0, +1 are shown in black.

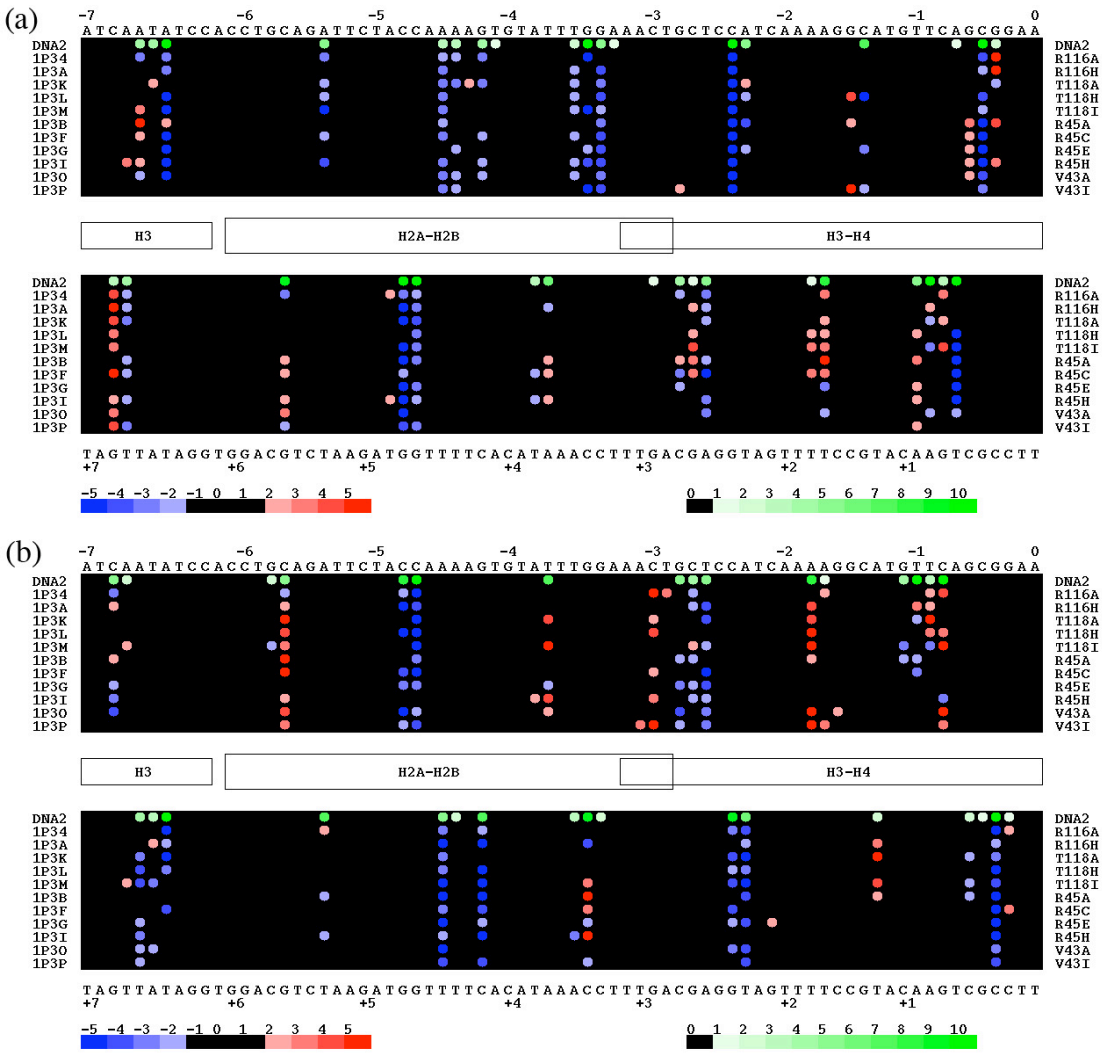



Figure A4.3. Difference in the numbers of close contacts within 4.0 Å to DNA phosphate atoms in (a) DNA strand I and (b) DNA strand II compared with the wild type. Data are mapped on each nucleotide along the nucleosomal DNA. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The mutation locations are shown at the right edge of the diagram. The number of contacts in the wild type is shown in different shades of green, and the difference in contacts to mutant vs. the wild type in shades of blue (negative) and red (positive). Differences of –1, 0, +1 are shown in black.

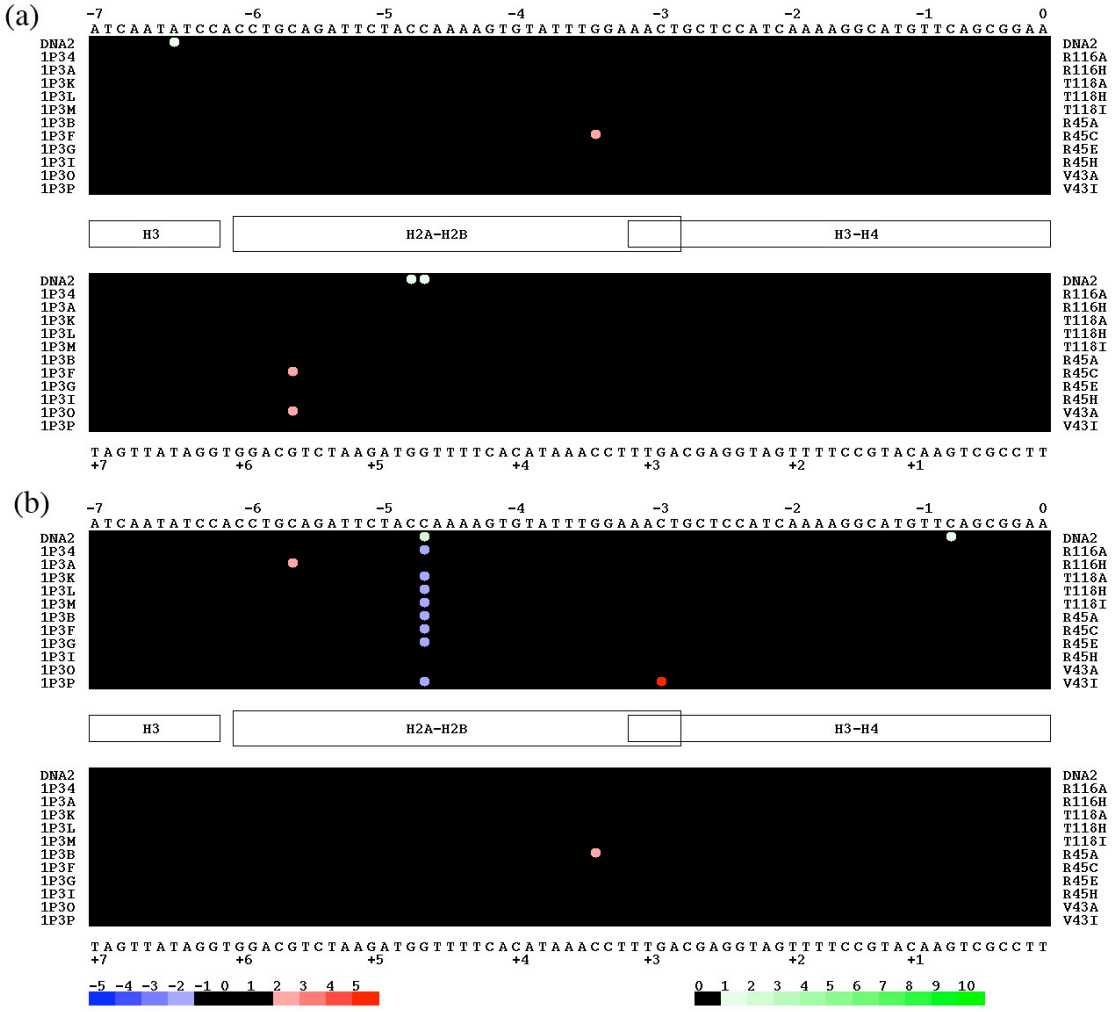



Figure A4.4. The difference in the numbers of close contacts within 3.0 Å to DNA sugar atoms in (a) DNA strand I, (b) DNA stand II compared with those of the wild type. Data are mapped on each nucleotide along the nucleosomal DNA. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The mutation locations are shown at the right edge of the diagram. The number of contacts in the wild type is shown in different shades of green, and the difference is contacts to mutant vs. contacts to wild type in shades of blue (negative) and red (positive). Differences of –1, 0, +1 are shown in black.

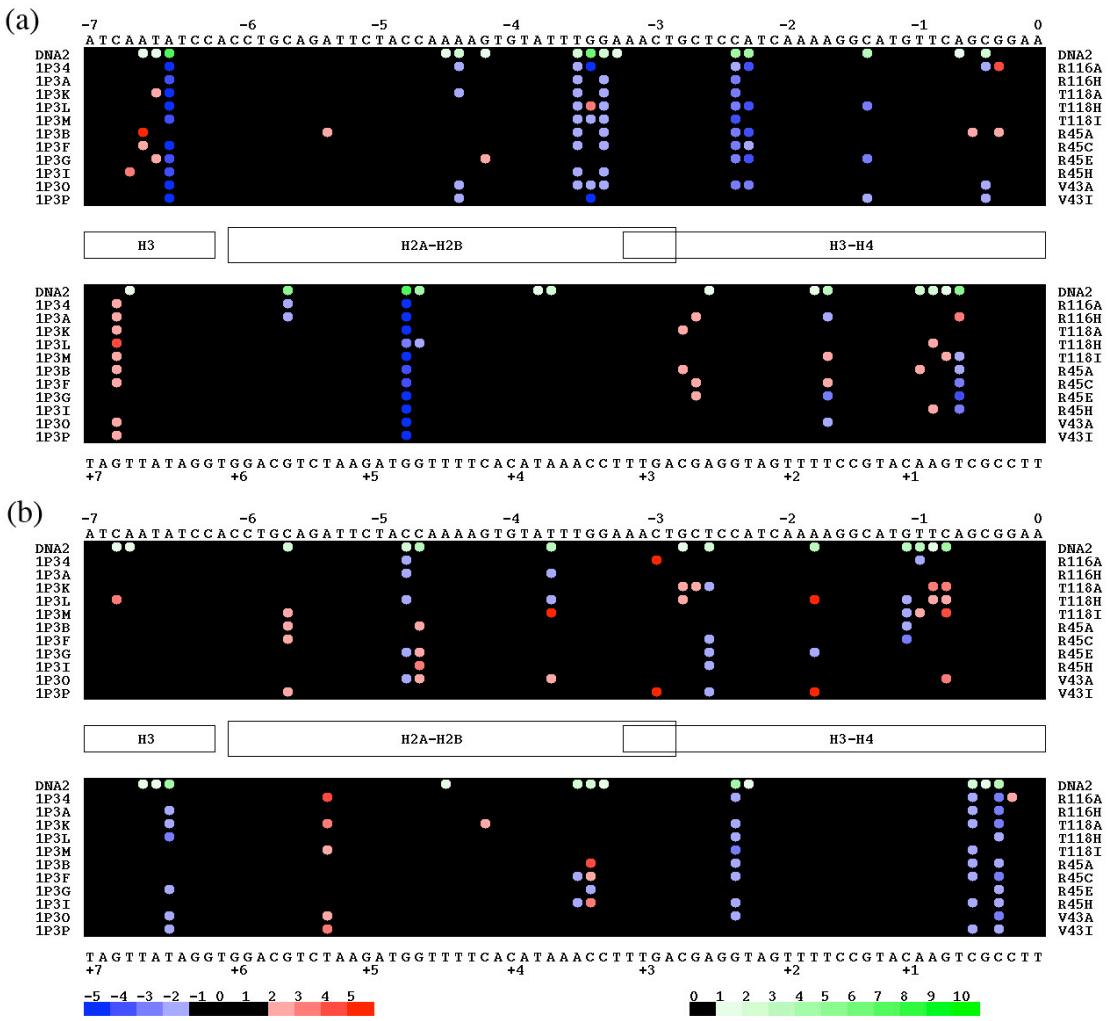



Figure A4.5. The difference in the numbers of close contacts within 3.6 Å to DNA sugar atoms in (a) DNA strand I, (b) DNA stand II compared with those of the wild type. Data are mapped on each nucleotide along the nucleosomal DNA. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The mutation locations are shown at the right edge of the diagram. The number of contacts in the wild type is shown in different shades of green, and the difference is contacts to mutant vs. contacts to wild type in shades of blue (negative) and red (positive). Differences of –1, 0, +1 are shown in black.

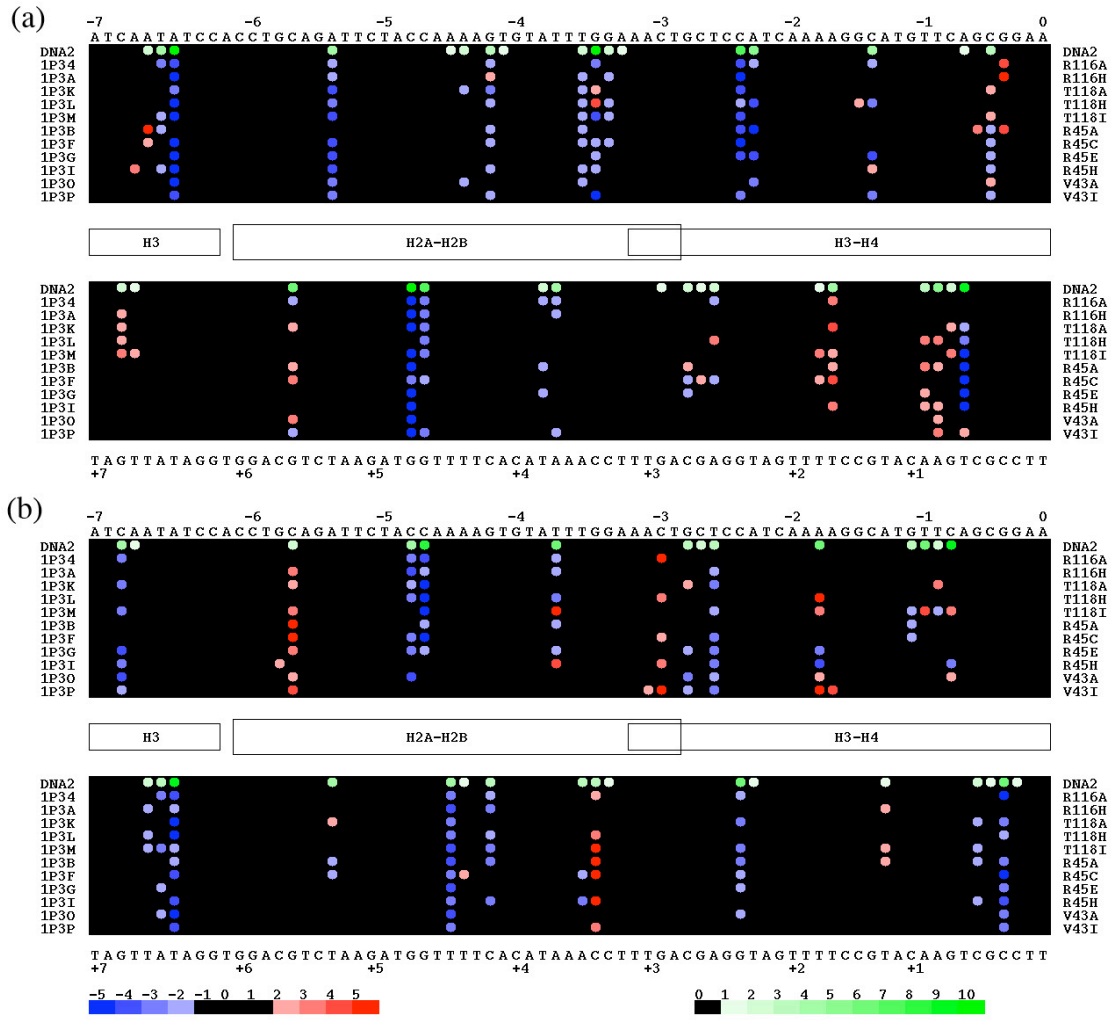



Figure A4.6. The difference in the numbers of close contacts within 3.8 Å to DNA sugar atoms in (a) DNA strand I, (b) DNA stand II compared with those of the wild type. Data are mapped on each nucleotide along the nucleosomal DNA. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The mutation locations are shown at the right edge of the diagram. The number of contacts in the wild type is shown in different shades of green, and the difference is contacts to mutant vs. contacts to wild type in shades of blue (negative) and red (positive). Differences of –1, 0, +1 are shown in black.

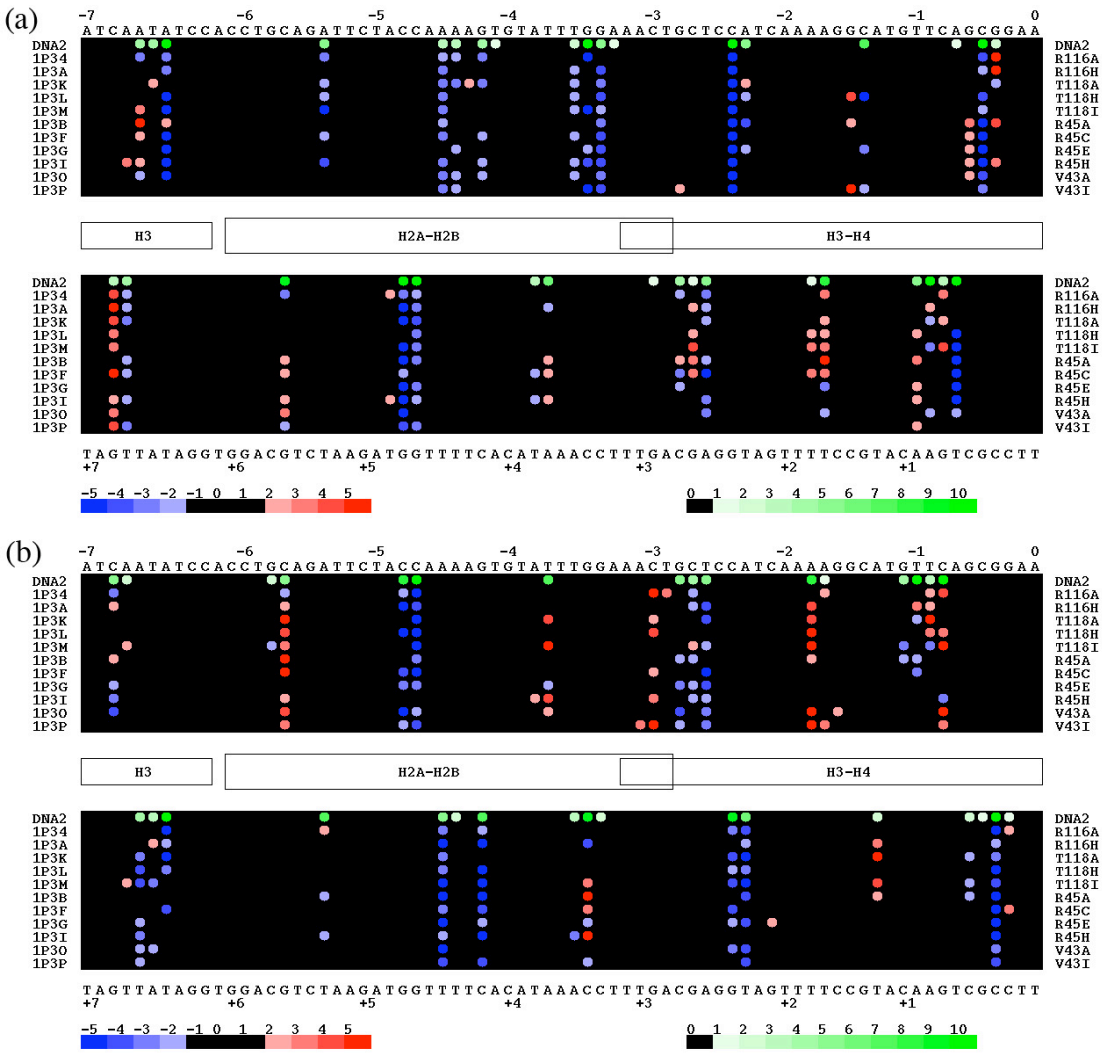



Figure A4.7. The difference in the numbers of close contacts within 4.0 Åto DNA sugar atoms in (a) DNA strand I, (b) DNA stand II compared with those of the wild type. Data are mapped on each nucleotide along the nucleosomal DNA. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The mutation locations are shown at the right edge of the diagram. The number of contacts in the wild type is shown in different shades of green, and the difference is contacts to mutant vs. contacts to wild type in shades of blue (negative) and red (positive). Differences of –1, 0, +1 are shown in black.

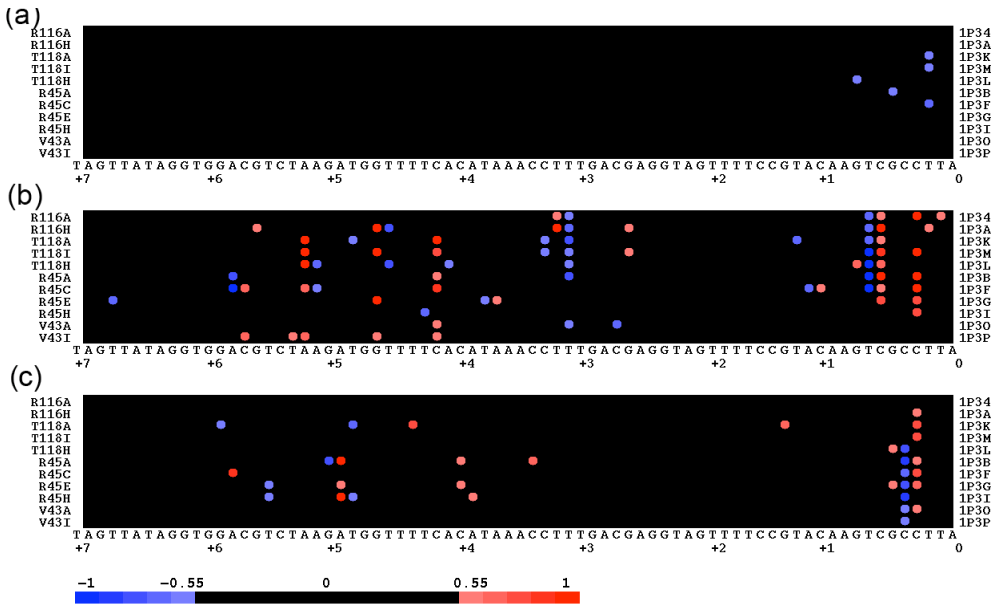



Figure A5.1. The difference in (a) *Z*-coordinates, (b) the contribution of Shift in the radius direction and (c) the contribution of Slide in the *Z*-axis of SIN mutant compared to the wild type (DNA2), are mapped on each base-pair step along the second half of the nucleosomal DNA from base-pair step 74 to 145. The PDB IDs, except for DNA2, are shown at the left edge of the diagram. The differences in values of the mutants vs. the values of wild type are shown in shades of blue (negative) and red (positive). Differences between –0.55 Å to 0.55 Å, are shown in black.

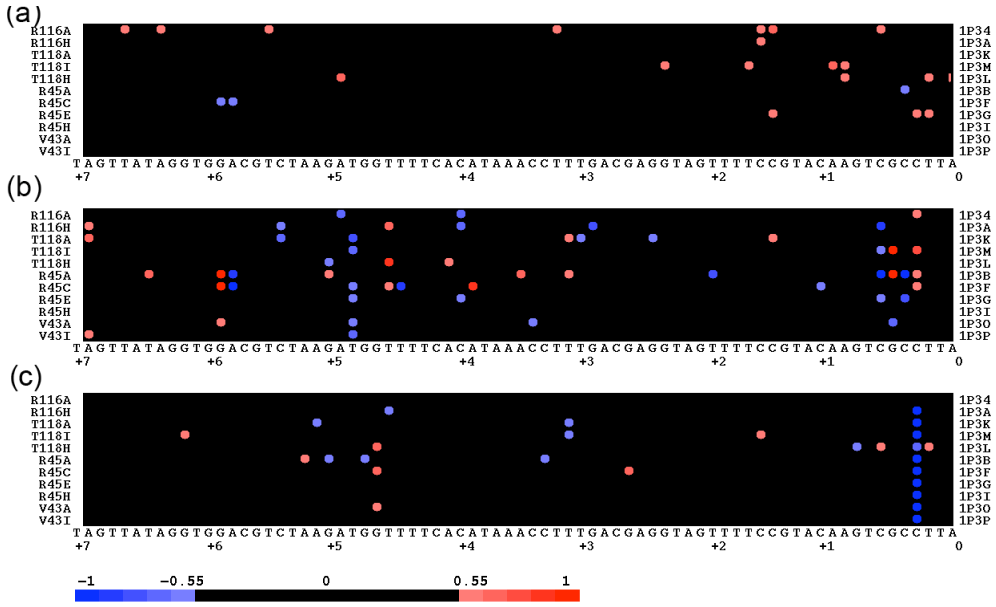



Figure A5.2. The difference in (a) radius, (b) the contribution of Shift in the radial direction and (c) the contribution of Slide in the radial direction of SIN mutants compared to the wild type (DNA2), are mapped on each base-pair step along the second half of the nucleosomeal DNA from base-pair step 74 to 145. See legend to Figure A5.1.

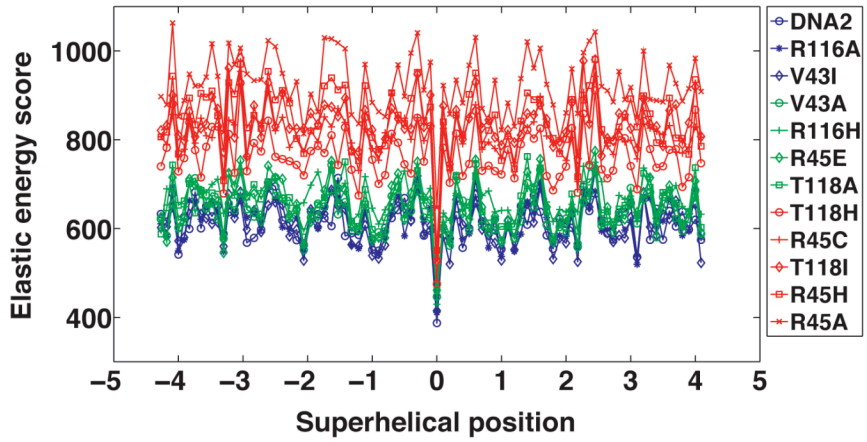



Figure A6.1. Threading score vs. superhelical location of the template reading frame along the DNA sequence from the SIN-mutant nucleosome structure. The templates are extracted from the wild-type and SIN mutant nucleosome structures, and designated by the point-mutations in the legend.

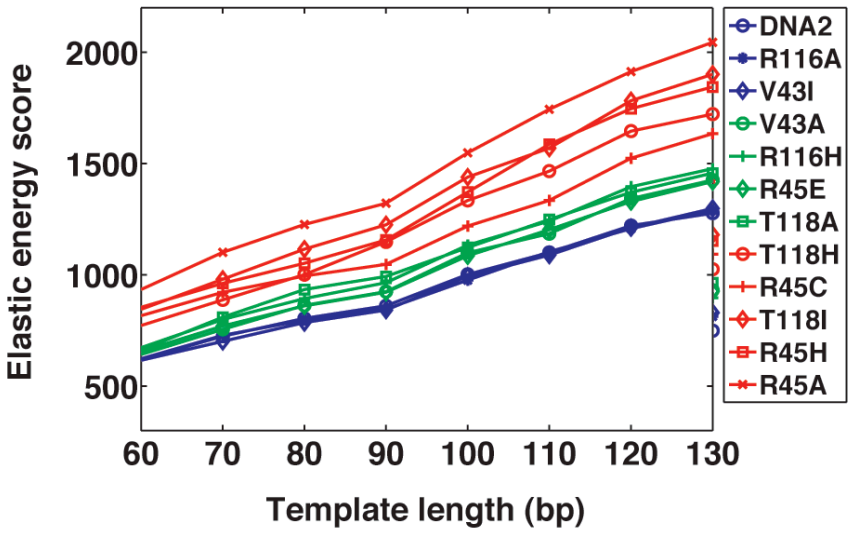



Figure A6.2. Average threading scores of different structural templates on the SIN mutant sequence vs. the length of the templates. The templates are extracted from the wild type and SIN mutant nucleosome structures, and designated by the point-mutations in the legend.

Table A6.1. Average contribution of individual base-pair step parameters to the threading of the central 80 dimer steps from the different nucleosome structures on the DNA sequence from the wild-type nucleosome structure (DNA2). (a) The sum of the diagonal terms (deformations of step parameters alone) and the cross terms (deformation of a step parameter coupled with the other step parameters). (b) Diagonal terms only.

(a)

| Group | Structure ID | Contribution of the step parameters | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist | |
| Low | DNA2 | 65.56 | 240.71 | 137.63 | 70.13 | 183.36 | -2.24 | 812.65 |
| | R116A | 71.29 | 268.03 | 111.34 | 70.68 | 167.84 | 17.38 | 805.65 |
| | V43I | 66.63 | 279.59 | 121.19 | 75.99 | 130.94 | 0.58 | 794.25 |
| Medium | R116H | 89.23 | 257.60 | 121.41 | 119.97 | 182.21 | 29.22 | 902.46 |
| | T118A | 64.72 | 230.84 | 129.59 | 137.41 | 190.46 | 25.53 | 945.33 |
| | R45E | 55.62 | 282.88 | 139.35 | 123.50 | 145.99 | 23.74 | 869.98 |
| | V43A | 92.89 | 266.65 | 120.69 | 80.53 | 168.77 | 11.85 | 871.65 |
| High | T118H | 87.45 | 307.94 | 185.04 | 130.24 | 205.89 | -2.46 | 1010.35 |
| | T118I | 86.20 | 301.28 | 212.45 | 167.23 | 220.58 | 30.13 | 1126.00 |
| | R45A | 88.19 | 240.05 | 235.19 | 216.24 | 319.87 | 54.39 | 1239.34 |
| | R45C | 93.55 | 302.10 | 188.23 | 130.21 | 196.54 | 23.60 | 1003.15 |
| | R45H | 80.33 | 299.73 | 151.38 | 140.79 | 286.48 | -8.22 | 1064.03 |

(b)

| Group | Structure ID | Contribution of the step parameters (diagonal terms only) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist |
| Low | DNA2 | 111.59 | 254.92 | 134.57 | 117.58 | 216.87 | 94.64 |
| | R116A | 125.06 | 269.93 | 98.09 | 126.12 | 188.94 | 96.59 |
| | V43I | 122.32 | 286.42 | 119.61 | 132.53 | 154.64 | 98.03 |
| Medium | R116H | 139.41 | 269.39 | 119.25 | 170.84 | 201.06 | 105.33 |
| | T118A | 153.51 | 252.87 | 119.31 | 228.26 | 230.37 | 127.79 |
| | R45E | 101.89 | 285.61 | 127.59 | 171.09 | 169.09 | 113.62 |
| | V43A | 140.53 | 275.24 | 121.39 | 129.54 | 204.01 | 131.22 |
| High | T118H | 127.14 | 305.93 | 190.95 | 170.25 | 223.14 | 89.19 |
| | T118I | 140.56 | 306.09 | 200.37 | 223.47 | 242.94 | 120.70 |
| | R45A | 113.91 | 260.51 | 242.19 | 243.71 | 336.12 | 128.31 |
| | R45C | 98.53 | 316.93 | 196.02 | 136.32 | 211.22 | 113.06 |
| | R45H | 126.96 | 295.05 | 154.77 | 188.53 | 310.49 | 101.78 |

Table A6.2. Average contribution of individual base-pair step parameters to the threading of the central 100 dimer steps from the different nucleosome structures on the DNA sequence from the wild-type nucleosome structure (DNA2). (a) The sum of the diagonal terms (deformations of step parameters alone) and the cross terms (deformation of a step parameter coupled with the other step parameters). (b) Diagonal terms only.

(a)

| Group | Structure ID | Contribution of the step parameters | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist | |
| Low | DNA2 | 72.66 | 303.50 | 158.07 | 105.14 | 227.70 | -6.55 | 1009.45 |
| | R116A | 94.65 | 328.68 | 121.34 | 100.24 | 192.07 | 22.93 | 985.56 |
| | V43I | 83.86 | 347.08 | 137.02 | 105.43 | 156.27 | 24.17 | 1002.30 |
| Medium | R116H | 117.33 | 332.59 | 134.65 | 171.99 | 203.96 | 32.13 | 1141.87 |
| | T118A | 76.78 | 276.51 | 155.40 | 167.23 | 221.74 | 33.53 | 1133.24 |
| | R45E | 91.37 | 345.78 | 148.73 | 134.20 | 191.78 | 43.90 | 1095.31 |
| | V43A | 124.90 | 314.77 | 147.00 | 118.41 | 200.52 | 13.17 | 1107.45 |
| High | T118H | 102.57 | 378.55 | 268.95 | 198.70 | 265.39 | -5.90 | 1343.34 |
| | T118I | 99.29 | 405.70 | 269.11 | 241.35 | 241.06 | 32.59 | 1448.62 |
| | R45A | 97.58 | 314.33 | 291.07 | 295.98 | 357.00 | 64.20 | 1561.51 |
| | R45C | 118.43 | 359.94 | 197.78 | 165.49 | 234.07 | 32.25 | 1226.95 |
| | R45H | 97.51 | 364.96 | 195.06 | 230.50 | 349.21 | -3.21 | 1385.04 |

(b)

| Group | Structure ID | Contribution of the step parameters (diagonal terms only) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist |
| Low | DNA2 | 128.38 | 322.36 | 164.83 | 162.82 | 264.42 | 115.57 |
| | R116A | 153.27 | 335.08 | 117.62 | 162.04 | 219.25 | 123.97 |
| | V43I | 151.22 | 358.87 | 139.47 | 174.02 | 186.10 | 141.09 |
| Medium | R116H | 175.24 | 359.14 | 149.46 | 229.58 | 230.07 | 147.61 |
| | T118A | 170.89 | 315.96 | 156.01 | 264.71 | 267.49 | 160.24 |
| | R45E | 146.60 | 357.43 | 147.86 | 192.19 | 226.49 | 164.29 |
| | V43A | 188.39 | 336.74 | 163.92 | 183.80 | 251.28 | 172.00 |
| High | T118H | 148.14 | 383.77 | 286.34 | 244.01 | 289.69 | 126.44 |
| | T118I | 175.74 | 420.10 | 266.29 | 316.49 | 270.66 | 158.88 |
| | R45A | 134.86 | 352.47 | 325.15 | 335.64 | 376.95 | 177.77 |
| | R45C | 136.44 | 389.85 | 212.00 | 182.48 | 265.60 | 159.58 |
| | R45H | 154.79 | 366.60 | 222.55 | 290.16 | 369.92 | 132.02 |

Table A6.3. Average contribution of individual base-pair step parameters to the threading of the central 120 dimer steps from the different nucleosome structures on the DNA sequence from the wild-type nucleosome structure (DNA2). (a) The sum of the diagonal terms (deformations of step parameters alone) and the cross terms (deformation of a step parameter coupled with the other step parameters). (b) Diagonal terms only.

(a)

| Group | Structure ID | Contribution of the step parameters | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | Rise | Roll | Shift | Slide | Tilt | Twist | |
| Low | DNA2 | 93.29 | 380.40 | 178.52 | 126.20 | 279.76 | -10.17 | 1230.01 |
| | R116A | 120.66 | 402.23 | 139.51 | 124.70 | 240.38 | 15.59 | 1216.80 |
| | V43I | 115.49 | 422.35 | 153.78 | 127.42 | 193.45 | 34.22 | 1220.60 |
| Medium | R116H | 143.02 | 412.31 | 156.69 | 220.45 | 250.15 | 28.07 | 1403.45 |
| | T118A | 95.87 | 346.87 | 183.60 | 199.71 | 271.35 | 44.23 | 1382.26 |
| | R45E | 123.87 | 431.32 | 154.80 | 174.74 | 220.77 | 54.53 | 1339.04 |
| | V43A | 150.68 | 392.64 | 159.57 | 149.17 | 261.67 | 10.86 | 1351.81 |
| High | T118H | 130.80 | 466.08 | 289.83 | 239.18 | 344.52 | -2.06 | 1655.97 |
| | T118I | 146.13 | 484.09 | 307.83 | 313.73 | 316.53 | 34.68 | 1792.39 |
| | R45A | 136.12 | 405.37 | 327.83 | 339.85 | 438.29 | 80.41 | 1927.65 |
| | R45C | 150.45 | 453.43 | 214.86 | 211.94 | 291.91 | 56.60 | 1533.33 |
| | R45H | 137.58 | 461.19 | 256.88 | 290.51 | 408.66 | -10.80 | 1760.70 |

(b)

| Group | Structure ID | Contribution of the step parameters (diagonal terms only) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Rise | Roll | Shift | Slide | Tilt | Twist |
| Low | DNA2 | 151.52 | 408.82 | 190.66 | 186.64 | 332.96 | 141.43 |
| | R116A | 180.63 | 419.83 | 144.15 | 192.24 | 292.75 | 160.95 |
| | V43I | 183.87 | 442.83 | 158.80 | 199.48 | 235.50 | 173.99 |
| Medium | R116H | 205.89 | 448.42 | 177.75 | 286.31 | 297.84 | 180.00 |
| | T118A | 196.43 | 403.04 | 194.00 | 303.58 | 329.62 | 196.20 |
| | R45E | 171.50 | 465.14 | 163.60 | 228.00 | 275.99 | 213.81 |
| | V43A | 212.81 | 428.81 | 181.72 | 216.27 | 334.91 | 204.51 |
| High | T118H | 183.68 | 483.60 | 322.08 | 295.17 | 392.74 | 166.32 |
| | T118I | 216.27 | 509.12 | 315.82 | 388.61 | 360.21 | 191.78 |
| | R45A | 184.85 | 468.88 | 381.30 | 392.31 | 474.24 | 225.87 |
| | R45C | 165.96 | 503.83 | 240.69 | 229.39 | 333.62 | 213.99 |
| | R45H | 204.41 | 483.12 | 309.47 | 362.61 | 452.20 | 165.57 |

Table A6.4. Average contribution of individual base-pair step parameters to the threading of the central 60 dimer steps from the different nucleosome structures on the DNA sequence from the SIN nucleosome structures. (a) The sum of the diagonal terms (deformations of step parameters alone) and the cross terms (deformation of a step parameter coupled with the other step parameters). (b) Diagonal terms only.

(a)

| Group | Structure ID | Contribution of the step parameters | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist | |
| Low | DNA2 | 59.18 | 177.67 | 120.13 | 39.33 | 136 | -2.49 | 613.59 |
| | R116A | 55.5 | 207.13 | 94.53 | 52.34 | 122.7 | 14.88 | 616.73 |
| | V43I | 53.84 | 218.93 | 99.84 | 55.29 | 100.2 | -10.6 | 610.07 |
| Medium | R116H | 68.39 | 181.21 | 108.38 | 80.05 | 142 | 21.66 | 666.53 |
| | T118A | 58.47 | 172.3 | 96.25 | 48.85 | 132.4 | 20.66 | 654.77 |
| | R45E | 37.26 | 221.81 | 118.68 | 72.94 | 110.8 | 11.94 | 650.12 |
| | V43A | 83.88 | 194.25 | 103.3 | 45.19 | 116.9 | 9.63 | 636.41 |
| High | T118H | 62.41 | 231.79 | 154.86 | 89.44 | 157.5 | -0.96 | 765.81 |
| | T118I | 69.01 | 234.07 | 178.09 | 120.56 | 150.9 | 22.3 | 838.27 |
| | R45A | 56.06 | 184.54 | 192.41 | 157.84 | 232 | 37.48 | 925.5 |
| | R45C | 70.04 | 247.47 | 173.51 | 104.91 | 152 | 18.03 | 810.45 |
| | R45H | 55.21 | 234.52 | 129.69 | 119.96 | 223.9 | -7.04 | 845.64 |

(b)

| Group | Structure ID | Contribution of the step parameters (diagonal terms only) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist |
| Low | DNA2 | 151.07 | 407.59 | 190.98 | 185.21 | 330.16 | 141.3 |
| | R116A | 180.28 | 418.4 | 144.23 | 189.97 | 289.92 | 160.87 |
| | V43I | 183.44 | 441.39 | 159.01 | 197.39 | 232.9 | 173.89 |
| Medium | R116H | 205.56 | 447.21 | 177.84 | 284.71 | 296.06 | 180.02 |
| | T118A | 195.96 | 402.11 | 194.02 | 300.92 | 326.3 | 196.44 |
| | R45E | 171.23 | 463.71 | 163.74 | 224.53 | 273.82 | 213.71 |
| | V43A | 212.32 | 427.51 | 181.94 | 214.34 | 332.08 | 204.64 |
| High | T118H | 183.29 | 482.14 | 322.26 | 292.11 | 390 | 166.19 |
| | T118I | 215.83 | 508.12 | 316.11 | 384.88 | 357.17 | 191.8 |
| | R45A | 184.56 | 467.91 | 381.62 | 387.4 | 469.62 | 225.78 |
| | R45C | 165.66 | 502.29 | 241.06 | 226.93 | 331.25 | 214.01 |
| | R45H | 204.06 | 481.62 | 309.62 | 359.24 | 447.21 | 165.55 |

Table A6.5. Average contribution of individual base-pair step parameters to the threading of the central 80 dimer steps from the different nucleosome structures on the DNA sequence from the SIN nucleosome structures. (a) The sum of the diagonal terms (deformations of step parameters alone) and the cross terms (deformation of a step parameter coupled with the other step parameters). (b) Diagonal terms only.

(a)

| Group | Structure ID | Contribution of the step parameters | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist | |
| Low | DNA2 | 65.03 | 237.03 | 137.81 | 67.53 | 178.83 | -2.31 | 804.10 |
| | R116A | 70.78 | 264.24 | 111.00 | 67.65 | 163.27 | 17.31 | 796.39 |
| | V43I | 66.11 | 275.69 | 121.06 | 72.78 | 126.91 | 0.59 | 785.38 |
| Medium | R116H | 88.59 | 254.02 | 121.54 | 116.42 | 177.80 | 29.09 | 892.96 |
| | T118A | 63.91 | 226.78 | 128.87 | 133.05 | 184.80 | 25.75 | 933.86 |
| | R45E | 55.15 | 278.90 | 139.01 | 119.75 | 141.54 | 23.73 | 860.24 |
| | V43A | 92.31 | 262.87 | 120.66 | 77.88 | 164.06 | 11.94 | 862.89 |
| High | T118H | 86.88 | 304.16 | 185.92 | 126.68 | 201.83 | -2.68 | 1000.96 |
| | T118I | 85.51 | 297.71 | 213.24 | 162.71 | 216.14 | 29.93 | 1115.42 |
| | R45A | 87.82 | 236.05 | 235.53 | 210.97 | 313.01 | 54.54 | 1226.35 |
| | R45C | 93.25 | 298.12 | 189.20 | 127.30 | 192.10 | 23.49 | 994.58 |
| | R45H | 79.76 | 295.60 | 151.81 | 136.60 | 280.13 | -8.17 | 1051.83 |

(b)

| Group | Structure ID | Contribution of the step parameters (diagonal terms only) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist |
| Low | DNA2 | 111.14 | 254.28 | 135.02 | 115.53 | 213.69 | 94.62 |
| | R116A | 124.58 | 269.20 | 98.31 | 123.68 | 186.17 | 96.57 |
| | V43I | 121.84 | 285.58 | 119.91 | 129.91 | 152.36 | 98.02 |
| Medium | R116H | 138.90 | 268.80 | 119.53 | 167.80 | 198.09 | 105.34 |
| | T118A | 152.93 | 252.44 | 119.52 | 224.74 | 227.00 | 127.91 |
| | R45E | 101.47 | 284.79 | 127.88 | 168.01 | 166.62 | 113.63 |
| | V43A | 139.97 | 274.53 | 121.78 | 127.34 | 201.15 | 131.31 |
| High | T118H | 126.63 | 304.88 | 191.44 | 167.14 | 219.89 | 89.14 |
| | T118I | 140.01 | 305.15 | 200.80 | 219.35 | 239.53 | 120.76 |
| | R45A | 113.44 | 259.99 | 242.84 | 239.02 | 331.12 | 128.36 |
| | R45C | 98.12 | 315.98 | 196.66 | 133.70 | 208.13 | 113.10 |
| | R45H | 126.47 | 294.09 | 155.10 | 184.70 | 305.75 | 101.81 |

Table A6.6. Average contribution of individual base-pair step parameters to the threading of the central 100 dimer steps from the different nucleosome structures on the DNA sequence from the SIN nucleosome structures. (a) The sum of the diagonal terms (deformations of step parameters alone) and the cross terms (deformation of a step parameter coupled with the other step parameters). (b) Diagonal terms only.

(a)

| Group | Structure ID | Contribution of the step parameters | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist | |
| Low | DNA2 | 72.43 | 300.04 | 158.43 | 103.21 | 222.79 | -6.19 | 1001.65 |
| | R116A | 94.44 | 324.86 | 120.89 | 97.61 | 186.89 | 23.22 | 976.34 |
| | V43I | 83.64 | 343.11 | 136.84 | 102.82 | 151.53 | 24.53 | 993.46 |
| Medium | R116H | 117.02 | 329.07 | 135.12 | 169.57 | 199.26 | 32.30 | 1133.52 |
| | T118A | 76.45 | 272.22 | 154.75 | 164.20 | 215.69 | 34.09 | 1122.79 |
| | R45E | 91.16 | 341.59 | 148.34 | 130.78 | 186.72 | 44.21 | 1085.32 |
| | V43A | 124.67 | 311.11 | 146.74 | 116.19 | 195.44 | 13.63 | 1099.10 |
| High | T118H | 102.31 | 374.91 | 269.74 | 194.95 | 261.04 | -5.80 | 1333.88 |
| | T118I | 98.95 | 402.39 | 269.99 | 237.50 | 236.45 | 32.82 | 1439.00 |
| | R45A | 97.42 | 310.65 | 291.38 | 290.84 | 349.27 | 64.56 | 1548.14 |
| | R45C | 118.49 | 356.09 | 198.80 | 162.96 | 229.56 | 32.48 | 1218.94 |
| | R45H | 97.44 | 360.63 | 195.23 | 226.36 | 341.59 | -2.66 | 1371.85 |

(b)

| Group | Structure ID | Contribution of the step parameters (diagonal terms only) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist |
| Low | DNA2 | 127.95 | 321.63 | 165.27 | 161.04 | 261.16 | 115.54 |
| | R116A | 152.86 | 334.18 | 117.72 | 159.79 | 216.22 | 124.00 |
| | V43I | 150.78 | 357.83 | 139.67 | 171.74 | 183.39 | 141.04 |
| Medium | R116H | 174.80 | 358.39 | 149.66 | 227.04 | 227.17 | 147.65 |
| | T118A | 170.37 | 315.33 | 156.16 | 262.04 | 263.86 | 160.41 |
| | R45E | 146.21 | 356.35 | 148.03 | 189.18 | 223.80 | 164.28 |
| | V43A | 187.89 | 335.91 | 164.26 | 181.93 | 248.31 | 172.11 |
| High | T118H | 147.71 | 382.62 | 286.79 | 240.79 | 286.30 | 126.38 |
| | T118I | 175.27 | 419.07 | 266.67 | 312.68 | 267.25 | 158.97 |
| | R45A | 134.48 | 351.91 | 325.67 | 330.76 | 371.53 | 177.81 |
| | R45C | 136.07 | 388.71 | 212.51 | 180.06 | 262.48 | 159.69 |
| | R45H | 154.39 | 365.42 | 222.75 | 286.17 | 364.30 | 132.06 |

Table A6.7. Average contribution of individual base-pair step parameters to the threading of the central 120 dimer steps from the different nucleosome structures on the DNA sequence from the SIN nucleosome structures. (a) The sum of the diagonal terms (deformations of step parameters alone) and the cross terms (deformation of a step parameter coupled with the other step parameters). (b) Diagonal terms only.

(a)

| Group | Structure ID | Contribution of the step parameters | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist | |
| Low | DNA2 | 92.80 | 376.15 | 177.88 | 124.15 | 274.70 | -10.12 | 1220.94 |
| | R116A | 120.52 | 397.63 | 139.37 | 122.49 | 235.22 | 15.66 | 1207.29 |
| | V43I | 115.25 | 417.55 | 154.23 | 125.69 | 188.64 | 34.29 | 1211.84 |
| Medium | R116H | 142.72 | 407.95 | 156.49 | 218.71 | 246.29 | 27.80 | 1395.69 |
| | T118A | 94.98 | 341.67 | 182.73 | 196.42 | 265.12 | 44.44 | 1370.56 |
| | R45E | 123.69 | 426.38 | 155.44 | 171.66 | 216.15 | 54.40 | 1329.23 |
| | V43A | 150.30 | 388.05 | 159.36 | 146.93 | 256.40 | 11.08 | 1342.47 |
| High | T118H | 130.15 | 461.59 | 289.89 | 235.34 | 340.29 | -2.41 | 1645.43 |
| | T118I | 145.44 | 480.15 | 308.84 | 310.45 | 311.46 | 34.62 | 1782.44 |
| | R45A | 135.46 | 400.94 | 327.59 | 334.81 | 430.51 | 80.33 | 1913.27 |
| | R45C | 149.81 | 448.75 | 215.31 | 209.15 | 287.11 | 56.41 | 1523.87 |
| | R45H | 137.10 | 455.78 | 255.82 | 286.49 | 401.07 | -10.65 | 1746.46 |

(b)

| Group | Structure ID | Contribution of the step parameters (diagonal terms only) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist |
| Low | DNA2 | 151.07 | 407.59 | 190.98 | 185.21 | 330.16 | 141.30 |
| | R116A | 180.28 | 418.40 | 144.23 | 189.97 | 289.92 | 160.87 |
| | V43I | 183.44 | 441.39 | 159.01 | 197.39 | 232.90 | 173.89 |
| Medium | R116H | 205.56 | 447.21 | 177.84 | 284.71 | 296.06 | 180.02 |
| | T118A | 195.96 | 402.11 | 194.02 | 300.92 | 326.30 | 196.44 |
| | R45E | 171.23 | 463.71 | 163.74 | 224.53 | 273.82 | 213.71 |
| | V43A | 212.32 | 427.51 | 181.94 | 214.34 | 332.08 | 204.64 |
| High | T118H | 183.29 | 482.14 | 322.26 | 292.11 | 390.00 | 166.19 |
| | T118I | 215.83 | 508.12 | 316.11 | 384.88 | 357.17 | 191.80 |
| | R45A | 184.56 | 467.91 | 381.62 | 387.40 | 469.62 | 225.78 |
| | R45C | 165.66 | 502.29 | 241.06 | 226.93 | 331.25 | 214.01 |
| | R45H | 204.06 | 481.62 | 309.62 | 359.24 | 447.21 | 165.55 |

Table A6.8. Average contribution of individual base-pair step parameters to the threading of the central 120 dimer steps from the different nucleosome structures on the DNA sequence from the SIN nucleosome structures. (a) The sum of the diagonal terms (deformations of step parameters alone) and the cross terms (deformation of a step parameter coupled with the other step parameters). (b) Diagonal terms only.

(a)

| Group | Structure ID | Contribution of the step parameters | | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist | |
| Low | DNA2 | 92.80 | 376.15 | 177.88 | 124.15 | 274.70 | -10.12 | 1220.94 |
| | R116A | 120.52 | 397.63 | 139.37 | 122.49 | 235.22 | 15.66 | 1207.29 |
| | V43I | 115.25 | 417.55 | 154.23 | 125.69 | 188.64 | 34.29 | 1211.84 |
| Medium | R116H | 142.72 | 407.95 | 156.49 | 218.71 | 246.29 | 27.80 | 1395.69 |
| | T118A | 94.98 | 341.67 | 182.73 | 196.42 | 265.12 | 44.44 | 1370.56 |
| | R45E | 123.69 | 426.38 | 155.44 | 171.66 | 216.15 | 54.40 | 1329.23 |
| | V43A | 150.30 | 388.05 | 159.36 | 146.93 | 256.40 | 11.08 | 1342.47 |
| High | T118H | 130.15 | 461.59 | 289.89 | 235.34 | 340.29 | -2.41 | 1645.43 |
| | T118I | 145.44 | 480.15 | 308.84 | 310.45 | 311.46 | 34.62 | 1782.44 |
| | R45A | 135.46 | 400.94 | 327.59 | 334.81 | 430.51 | 80.33 | 1913.27 |
| | R45C | 149.81 | 448.75 | 215.31 | 209.15 | 287.11 | 56.41 | 1523.87 |
| | R45H | 137.10 | 455.78 | 255.82 | 286.49 | 401.07 | -10.65 | 1746.46 |

(b)

| Group | Structure ID | Contribution of the step parameters (diagonal terms only) | | | | | |
|---|---|---|---|---|---|---|---|
| | | Shift | Slide | Rise | Tilt | Roll | Twist |
| Low | DNA2 | 151.07 | 407.59 | 190.98 | 185.21 | 330.16 | 141.30 |
| | R116A | 180.28 | 418.40 | 144.23 | 189.97 | 289.92 | 160.87 |
| | V43I | 183.44 | 441.39 | 159.01 | 197.39 | 232.90 | 173.89 |
| Medium | R116H | 205.56 | 447.21 | 177.84 | 284.71 | 296.06 | 180.02 |
| | T118A | 195.96 | 402.11 | 194.02 | 300.92 | 326.30 | 196.44 |
| | R45E | 171.23 | 463.71 | 163.74 | 224.53 | 273.82 | 213.71 |
| | V43A | 212.32 | 427.51 | 181.94 | 214.34 | 332.08 | 204.64 |
| High | T118H | 183.29 | 482.14 | 322.26 | 292.11 | 390.00 | 166.19 |
| | T118I | 215.83 | 508.12 | 316.11 | 384.88 | 357.17 | 191.80 |
| | R45A | 184.56 | 467.91 | 381.62 | 387.40 | 469.62 | 225.78 |
| | R45C | 165.66 | 502.29 | 241.06 | 226.93 | 331.25 | 214.01 |
| | R45H | 204.06 | 481.62 | 309.62 | 359.24 | 447.21 | 165.55 |

Table A6.9. (a) Values and (b) Average contribution, over all settings of the 120 base pair step template from a SIN mutant structure (PDB ID: 1P3M, structure ID: T118I) on DNA, of individual step parameters to the high deformation scores (larger than 11) consistently found (in 25 out of 26 threading frames) at specific steps.

(a)

| Step Number | Step | SHL | Step Parameters | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Twist (°) | Tilt (°) | Roll (°) | Shift (Å) | Slide (Å) | Rise (Å) |
| 15 | GC | -5.7 | 42.13 | -10.91 | 11.54 | 0.75 | -0.15 | 3.42 |
| 16 | CA | -5.6 | 38.01 | 6.12 | -31.93 | -0.19 | 1.83 | 2.95 |
| 18 | GA | -5.4 | 40.01 | 12.65 | 7.68 | -0.64 | -0.30 | 4.16 |
| 25 | CC | -4.7 | 23.75 | 20.45 | 16.90 | 1.53 | 0.86 | 2.58 |
| 34 | TA | -3.9 | 45.65 | -10.17 | 10.12 | 0.39 | -0.21 | 4.27 |
| 60 | CA | -1.3 | 41.87 | 10.68 | 13.13 | -1.14 | 1.51 | 3.97 |
| 62 | TG | -1.1 | 40.34 | -11.68 | 18.99 | -0.99 | 2.02 | 4.56 |
| 68 | GC | -0.5 | 29.68 | 7.44 | -8.57 | -0.96 | -0.03 | 4.14 |

(b)

| Step Number | SHL | Contribution of step parameters (diagonal terms only) | | | | | | Total Score |
|---|---|---|---|---|---|---|---|---|
| | | Twist | Tilt | Roll | Shift | Slide | Rise | |
| 15 | -5.7 | 2.03 | 13.19 | 2.40 | 0.90 | 0.00 | 0.17 | 24.41 |
| 16 | -5.6 | 0.28 | 2.63 | 32.45 | 0.03 | 2.65 | 1.72 | 33.77 |
| 18 | -5.4 | 3.85 | 14.11 | 1.89 | 0.74 | 0.65 | 18.15 | 40.90 |
| 25 | -4.7 | 3.11 | 31.42 | 3.84 | 4.20 | 3.17 | 10.32 | 25.67 |
| 34 | -3.9 | 4.79 | 11.66 | 1.75 | 0.10 | 0.02 | 15.35 | 27.39 |
| 60 | -1.3 | 3.64 | 10.51 | 3.31 | 1.57 | 10.12 | 6.68 | 42.01 |
| 62 | -1.1 | 1.71 | 8.56 | 5.30 | 1.65 | 10.75 | 16.81 | 41.90 |
| 68 | -0.5 | 1.40 | 7.29 | 3.08 | 5.59 | 0.09 | 16.90 | 35.98 |

Table A6.10. (a) Values and (b) Average contribution, over all settings of the 120 base pair step template from a SIN mutant structure (PDB ID: 1P3B, structure ID: R45A) on DNA, of individual step parameters to the high deformation scores (larger than 11) consistently found (in 25 out of 26 threading frames) at specific steps.

(a)

| Step Number | Step | SHL | Step Parameters | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Twist (°) | Tilt (°) | Roll (°) | Shift (Å) | Slide (Å) | Rise (Å) |
| 14 | TG | -5.8 | 28.99 | -6.76 | -16.67 | -1.42 | 1.47 | 3.38 |
| 21 | TC | -5.1 | 21.91 | 15.09 | 0.56 | 1.29 | 0.19 | 2.02 |
| 24 | AC | -4.8 | 26.33 | -16.40 | -2.14 | -0.43 | 0.45 | 2.64 |
| 25 | CC | -4.7 | 27.15 | 18.48 | 10.33 | 0.94 | -0.06 | 2.31 |
| 50 | CA | -2.4 | 44.99 | -12.48 | -0.36 | 0.44 | 1.32 | 2.51 |
| 52 | TC | -2.1 | 30.24 | -2.18 | 14.58 | 0.14 | -1.55 | 4.78 |
| 58 | GG | -1.5 | 26.59 | -19.68 | -56.06 | 0.78 | 0.88 | 3.84 |
| 68 | GC | -0.5 | 36.77 | 12.82 | -10.31 | 1.23 | 0.62 | 4.19 |
| 114 | CA | 4.0 | 42.37 | 7.71 | 21.26 | 0.57 | 0.11 | 4.19 |

(b)

| Step Number | SHL | Contribution of step parameters (diagonal terms only) | | | | | | Total Score |
|---|---|---|---|---|---|---|---|---|
| | | Twist | Tilt | Roll | Shift | Slide | Rise | |
| 14 | -5.8 | 0.03 | 4.03 | 13.36 | 3.64 | 22.82 | 0.42 | 38.96 |
| 21 | -5.1 | 8.70 | 15.82 | 0.04 | 3.58 | 0.27 | 28.74 | 37.94 |
| 24 | -4.8 | 1.33 | 20.31 | 0.76 | 0.81 | 6.68 | 6.15 | 33.29 |
| 25 | -4.7 | 1.28 | 25.83 | 0.78 | 1.61 | 0.35 | 17.73 | 20.31 |
| 50 | -2.4 | 4.70 | 18.46 | 0.04 | 1.40 | 12.26 | 11.69 | 53.23 |
| 52 | -2.1 | 0.31 | 0.57 | 2.49 | 0.03 | 1.77 | 24.13 | 29.99 |
| 58 | -1.5 | 2.27 | 27.88 | 88.56 | 0.87 | 0.45 | 1.97 | 143.59 |
| 68 | -0.5 | 0.14 | 19.40 | 4.33 | 6.83 | 3.41 | 18.86 | 43.13 |
| 114 | 4.0 | 1.71 | 4.20 | 6.21 | 0.63 | 0.01 | 6.12 | 19.36 |

Table A6.11. (a) Values and (b) Average contribution, over all settings of the 120 base pair step template from a SIN mutant structure (PDB ID: 1P3F, structure ID: R45C) on DNA, of individual step parameters to the high deformation scores (larger than 11) consistently found (in 25 out of 26 threading frames) at specific steps.

(a)

| Step Number | Step | SHL | Step Parameters | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Twist (°) | Tilt (°) | Roll (°) | Shift (Å) | Slide (Å) | Rise (Å) |
| 14 | TG | -5.8 | 38.79 | 8.03 | -17.95 | -0.37 | 1.87 | 3.18 |
| 19 | AT | -5.3 | 24.18 | -9.67 | -6.24 | -0.25 | -0.96 | 2.72 |
| 99 | AG | 2.5 | 31.38 | -9.66 | -18.19 | 1.43 | 0.23 | 2.81 |
| 133 | AG | 5.8 | 40.50 | -10.25 | -1.44 | 1.70 | 1.79 | 3.00 |

(b)

| Step Number | SHL | Contribution of step parameters (diagonal terms only) | | | | | | Total Score |
|---|---|---|---|---|---|---|---|---|
| | | Twist | Tilt | Roll | Shift | Slide | Rise | |
| 14 | -5.8 | 2.98 | 5.69 | 15.37 | 0.25 | 32.19 | 0.08 | 44.33 |
| 19 | -5.3 | 3.66 | 6.83 | 2.16 | 0.10 | 1.17 | 5.84 | 19.58 |
| 99 | 2.5 | 0.43 | 6.64 | 12.82 | 3.10 | 0.00 | 3.03 | 31.29 |
| 133 | 5.8 | 0.95 | 7.49 | 1.00 | 4.44 | 2.52 | 1.35 | 16.63 |

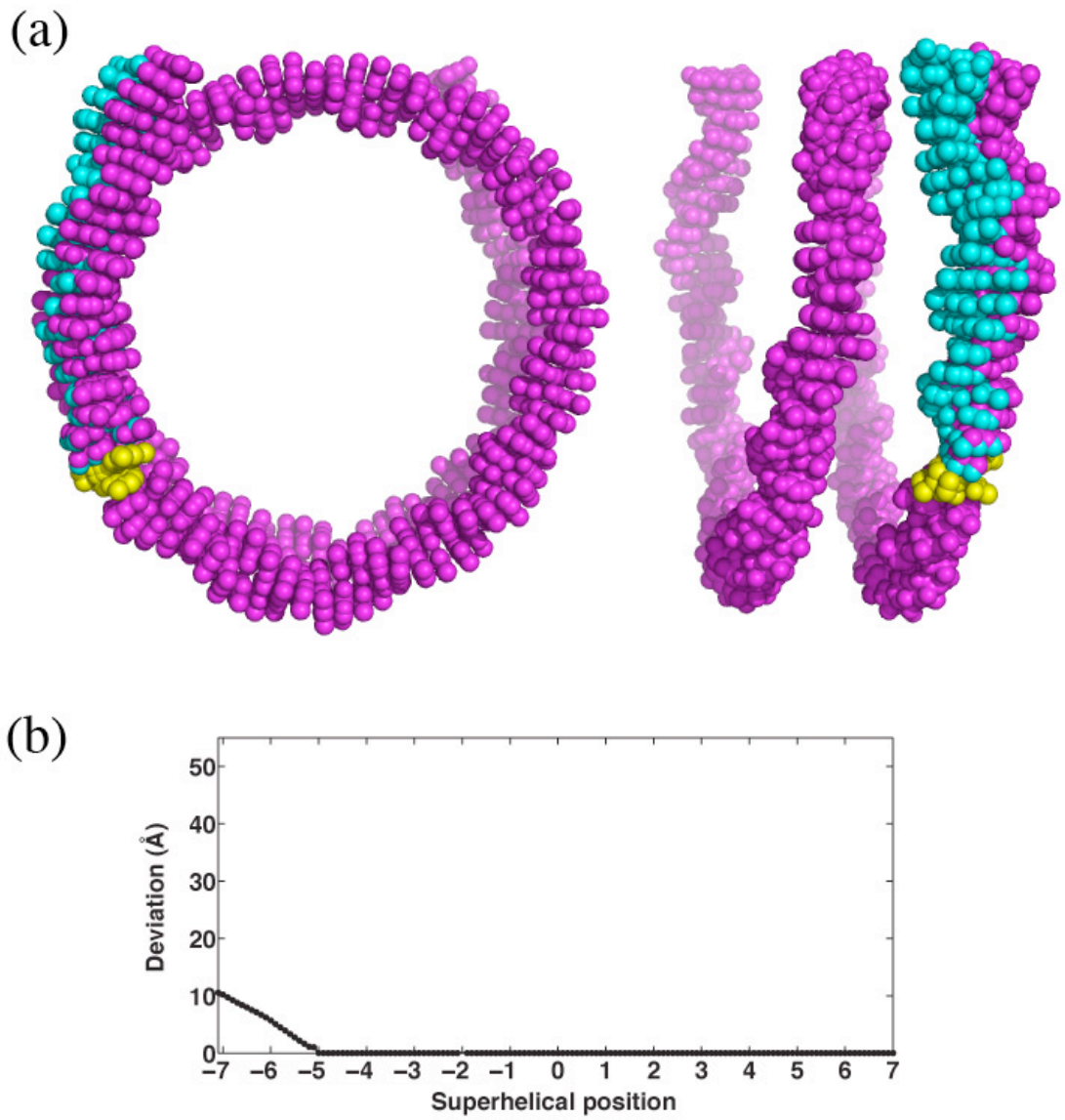



Figure A6.3. (a) Top view (on the left) and side view (on the right) of the remodeled DNA of the wild-type nucleosome structure (DNA2). (b) Root-mean-square deviation of the base atoms at each base pair of the remodeled DNA compared to the wild-type nucleosomal DNA (DNA2). Note: The altered step parameters of the modified base-pair step are listed in Table 6.7. The wild type DNA and the part of the remodeled DNA, which overlaps the wild type DNA, are shown in magenta. The changed pathway of the remodeled DNA is shown in cyan. The remodeled base pair step is highlighted in yellow.

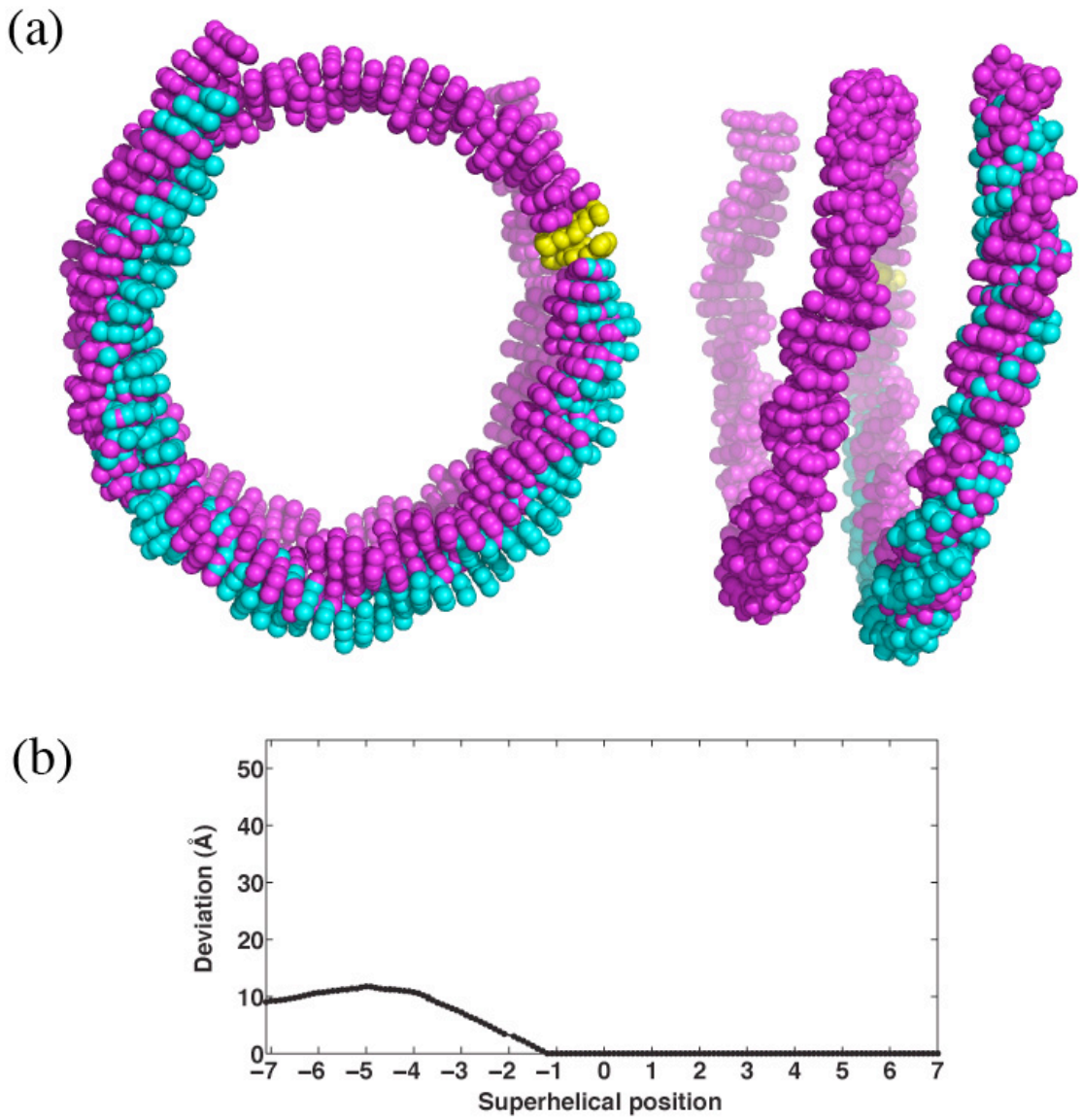



Figure A6.4. (a) Top view (on the left) and side view (on the right) of the remodeled DNA of the wild-type nucleosome structure (DNA2). (b) Root-mean-square deviation of the base atoms at each base pair of the remodeled DNA compared to the wild-type nucleosomal DNA (DNA2). Note: The altered step parameters of the modified base-pair step are listed in Table 6.7. The wild type DNA and the part of the remodeled DNA, which overlaps the wild type DNA, are shown in magenta. The changed pathway of the remodeled DNA is shown in cyan. The remodeled base pair step is highlighted in yellow.

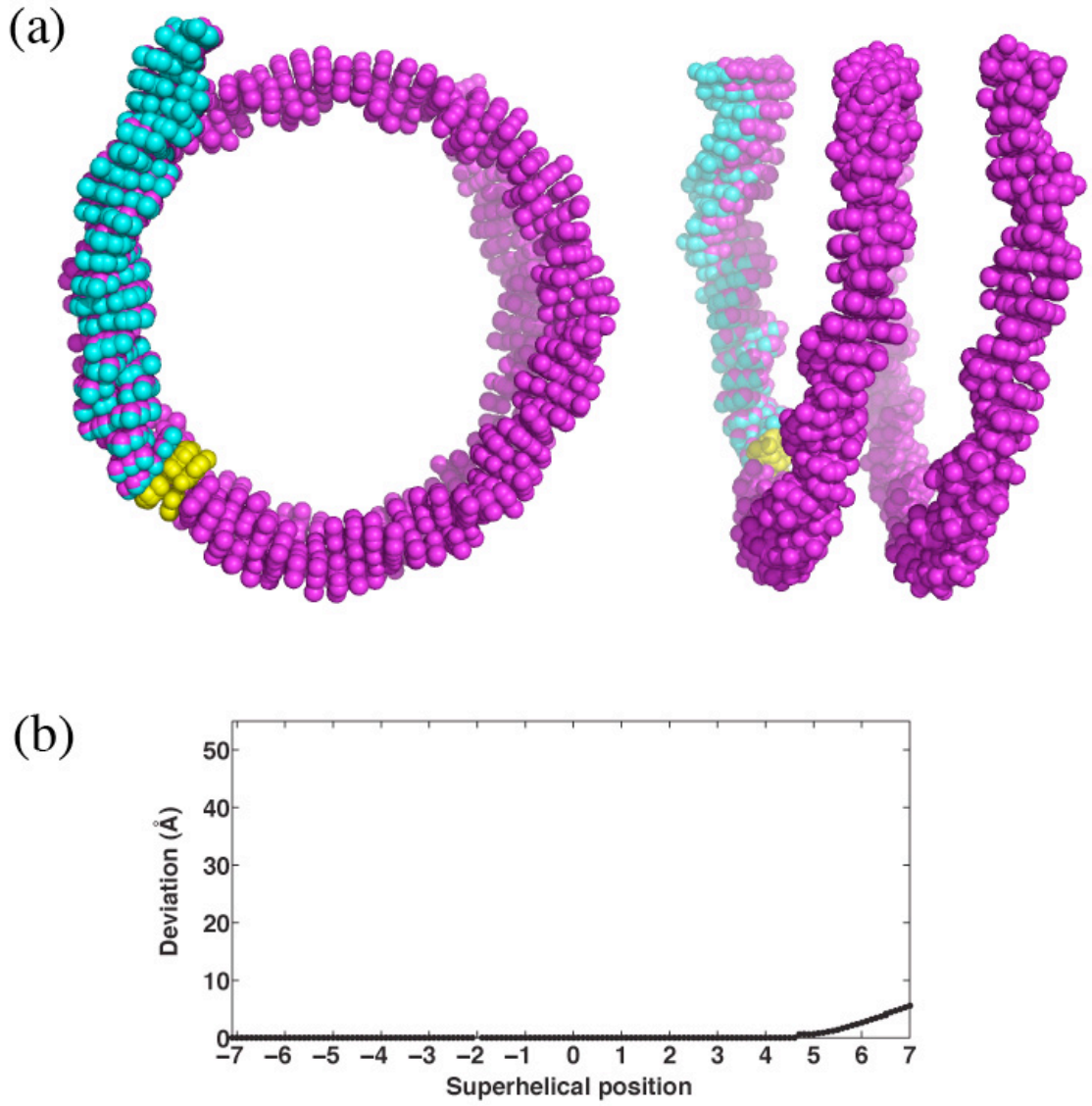



Figure A6.5. (a) Top view (on the left) and side view (on the right) of the remodeled DNA of the wild-type nucleosome structure (DNA2). (b) Root-mean-square deviation of the base atoms at each base pair of the remodeled DNA compared to the wild-type nucleosomal DNA (DNA2). Note: The altered step parameters of the modified base-pair step are listed in Table 6.7. The wild type DNA and the part of the remodeled DNA, which overlaps the wild type DNA, are shown in magenta. The changed pathway of the remodeled DNA is shown in cyan. The remodeled base pair step is highlighted in yellow.

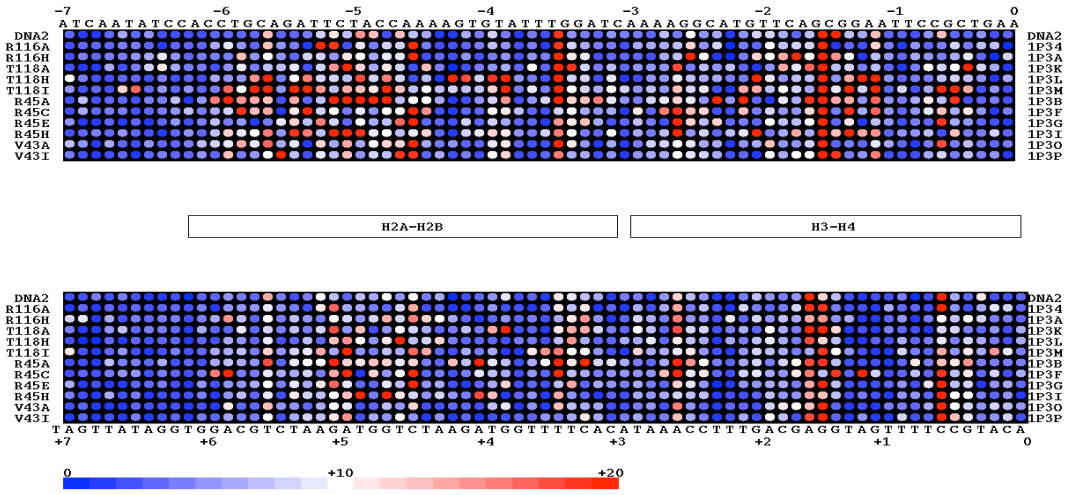



Figure A6.6. Values of the deformation score of SIN mutants and the wild type (dna2), are mapped on each base-pair step along the DNA sequence from the SIN mutants. The PDB IDs, except for DNA2, are shown at the right edge of the diagram. The corresponding structure IDs, except for DNA2, designated by the point-mutations are shown at the left edge. The values of deformation scores are shown in shades of blue (small values), white (medium values), and red (large values). The range is 0 to 20.

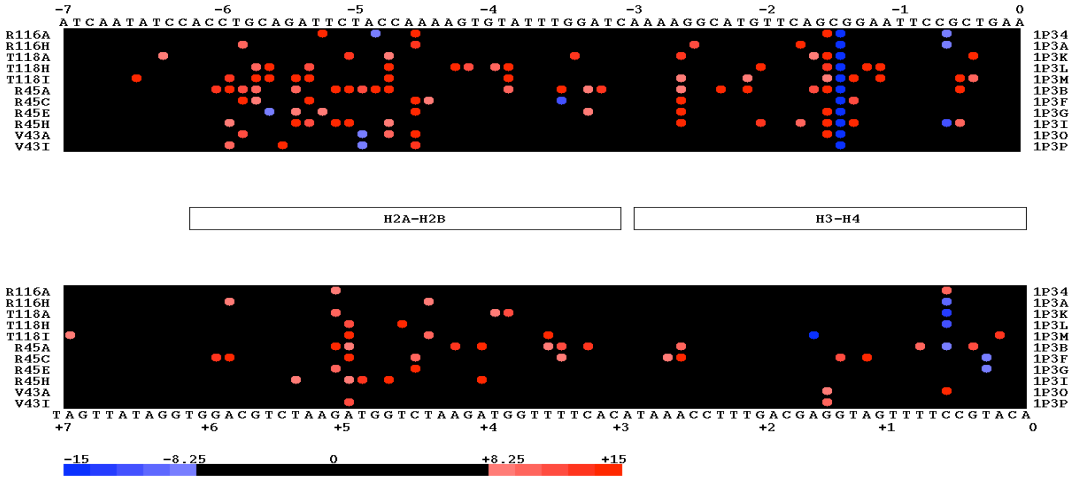



Figure A6.7. Difference in deformation scores of the SIN mutants compared to the wild type (DNA2), are mapped on each base-pair step along the DNA sequence from the SIN mutants. The PDB IDs, except for DNA2, are shown at the right edge of the diagram. The corresponding structure IDs, except for DNA2, designated by the point-mutations are shown at the left edge. The difference in values of the mutant vs. the values of wild type is shown in shades of blue (negative) and red (positive). Differences of −7.5 to +7.5 are shown in black.

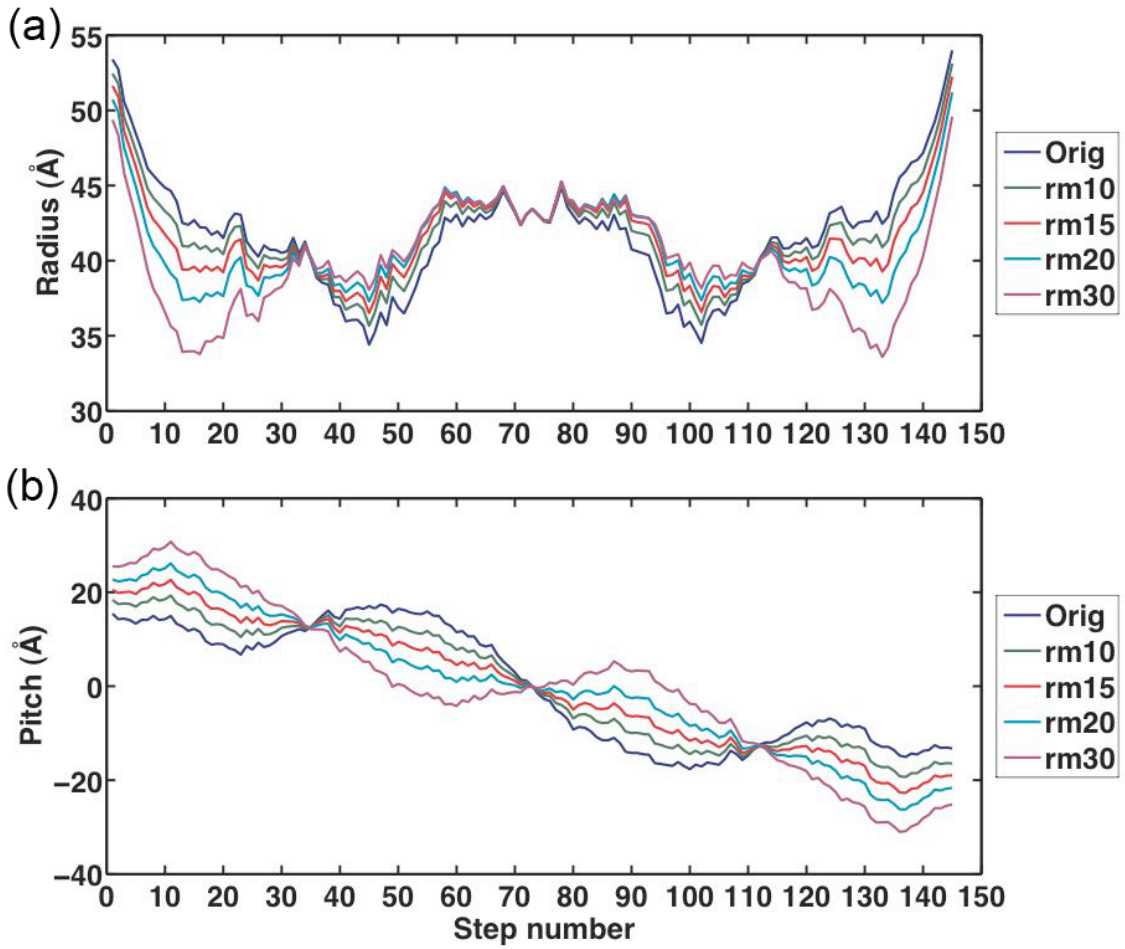



Figure A7.1. Effects of the nucleosomal DNA chain length on (a) superhelical radius and (b) *Z*-coordinates of each base-pair step in the structure (PDB ID: 1KX3). Orig represents the full length 146 bp DNA, rm10, rm15, rm20, and rm30 correspond to chains of 126, 116, 106, and 186 bp where the 10, 15, 20, and 30 base pairs have been removed from each end of DNA.

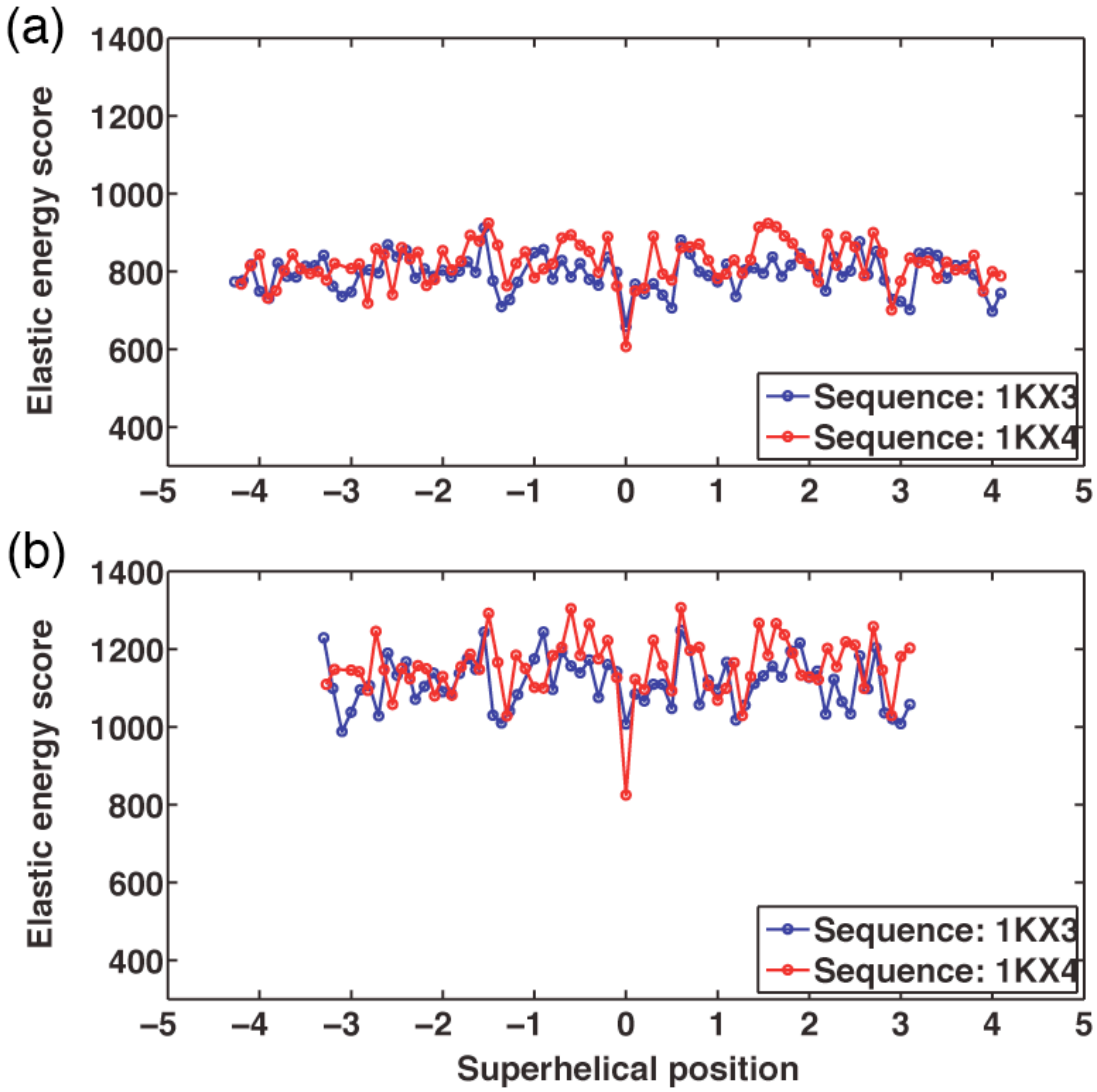



Figure A7.2. Threading score vs. superhelical location of the (a) central 60 dimer-step and (b) central 80 dimer-step template reading frame along the DNA sequences from 1KX4 (blue) and 1KX3 (red). The templates are extracted from 1KX4.

# Curriculum Vita

Fei Xu

| | |
|---|---|
| October 2007 | Ph.D. candidate, Biochemistry, Rutgers, the State University of New Jersey, New Brunswick, New Jersey |
| October 2006 | M.S. Food Science, Rutgers, the State University of New Jersey, New Brunswick, New Jersey |
| June 2000 | B.S., Biology Specialty of the Life Sciences Department, Wuhan University, Wuhan, Hubei, China |
| 2002-2007 | Graduate Assistant<br>Rutgers University, Biochemistry, New Brunswick, NJ |
| 2002-2004 | Programmer and Grader<br>Rutgers University, Waksman Challenge Program, New Brunswick, NJ |
| 2001-2002 | Teaching Assistant<br>Rutgers University, Chemistry, New Brunswick, NJ |
| 2005 | Zhurkin, Victor B., Tolstorukov, Michael Y., **Xu, Fei**, Colasanti, Andrew V., and Olson, Wilma K., "Sequence-dependent Variability of B DNA: An Update on Bending and Curvature," in DNA Conformation and Transcription, Takashi Ohyama, Ed., Landes Bioscience/Eurekah.com, Georgetown, TX (2005). |