Doctoral Thesis

The nucleosome core particle in vitro assembly and X-ray structure

Author(s):
Mäder, Armin W.

Publication Date:
1997

Permanent Link:
https://doi.org/10.3929/ethz-a-001763378

Rights / License:
In Copyright - Non-Commercial Use Permitted
The Nucleosome Core Particle: 
*in vitro* Assembly 
and X-ray Structure

A dissertation submitted to the 
Swiss Federal Institute of Technology Zürich

for the degree of 
Doctor of Natural Science

presented by Armin W. Mäder 
Dipl. Natw. ETH 
born December 12, 1967 
from Galgenen (SZ)

Prof. Dr. T. J. Richmond, examiner 
Dr. A. A. Travers, coexaminer 
Zürich, 1997
To my parents
Acknowledgments

Special thanks go to Prof. T. J. Richmond for giving me the opportunity to work on an exciting project in his laboratory. I had the chance of learning a wide range of methods in the field.

I am grateful to Dr. Karolin Luger for introduction into the laboratory techniques as well as to Dr. Song Tan for providing for many helpful technical discussions.

Dr. Dave Sargent deserves a very special thank. As the helping hand behind the crystallographic experiments, he was always open for discussions and his steady encouragement and enthusiasm was of great importance for my work. Thanks for running weeks of synchrotron night shifts without ever complaining.

My special gratitude goes to the F9 laboratory chief Dr. Thomas Rechsteiner who was responsible for the good working atmosphere. There were so many fruitful discussions about science and many other things that I only want to mention the Friday afternoon sessions. I am also grateful to Dr. Andrew Flaus for all discussions about nucleosomes and his collaboration with respect to several experiments.

Thank goes to Robin Richmond for her help in the purification of histone proteins, Yvonne Hunziker for oligonucleotide synthesis and Ritwa Stenzel for never complaining about the mess in F9 and for washing all those beakers, preparing media and sterilizing equipment.

I want to acknowledge the remaining members of the Richmond group and the support staff in the institute for providing an excellent working environment.

I want to express my thank to the people at the E.S.R.F. in Grenoble. We were supported by the groups of C. Riekel at BL-1/ID-13, P. Pattison at the SNBL and M. Wulff at BL-3/ID-9. I am also grateful to R. Bosch and his colleges at Dornier Friedrichshafen and K. Fuhrman (ESA) who were in charge of the organization of a microgravity experiment on Space Shuttle flight STS-78.

I am grateful and indebted to my parents. They gave me the freedom to study whatever I wanted and always supported me as best as possible.

This work was in part supported by the Swiss Science Foundation.
Curriculum vitae

Personal data: Dipl. Natw. ETH
Mäder, Armin Walter Lukas
born on the 2.12.1967, Swiss citizen

Education:
1974-1980 Primary school at Lachen (Switzerland)
1980-1987 Secondary school

Academic Education:
1991-1992: Semester-thesis in the group of Prof. Dr. P. Dimroth at the Institute of Microbiology (ETHZ) with work related to the anaerobic citrate metabolism of the bacteria Clostridium sphenoides.

1992: Diploma-thesis in the group of Prof. Dr. D. Arigoni at the Laboratory of Organic Chemistry (ETHZ) with work related to the biosynthesis of the toxin precursor circinatin from the fungus Periconia circinata.

1992: Diploma in Biochemistry, Molecular Biology, Immunology, Synthetic Organic Chemistry and Bioorganic Chemistry.

1993-1997: PhD thesis in the group of Prof. Dr. T. J. Richmond at the Institut for Molecular Biology and Biophysics.
Abbreviations

aa: Amino acid
ATP: Adenosine triphosphate
ATPase: ATP hydrolase
bp: base pairs
BSA: Bovine serum albumin
CIA: Chloroform-Isoamylalcohol
CPM: Coumarin derivative
DCC: Dicyclohexylcarbodiimide
DEAE: Diethylaminoethyl
DMF: Dimethylformamide
DMSO: Dimethyl sulfoxide
DNase I: Deoxyribonuclease I
ds: Double-stranded
DTT: Dithiothreitol
EDC: DCC derivative (see methods)
EDTA: Ethylene diamine tetraacetic acid
E.S.R.F.: European Synchrotron Radiation Facility
FLU: Fluorescein derivative
FRET: Fluorescence energy transfer
H1, H2A, H2B, H3, H4: Histone proteins
HMG: High mobility group protein
HPLC: High performance liquid chromatography
IAF: Fluorescein derivative
IPTG: Isopropyl-thiogalactopyranosid
kD: kilo Dalton
kJ: kilo Joules
MDa: mega Dalton
mol: moles
MPD: Methyl pentane diol
MW: Molecular weight
MWCO: Molecular weight cut off
NCP: Nucleosome core particle
NMR: Nuclear magnetic resonance
ODG: Octyl-β-D-glucopyranoside
PEG: Polyethylene glycol
PCR: Polymerase chain reaction
PMSF: Phenylmethanesulfonylfluoride
RHO: Rhodamin derivative
RMS: Root mean square
RP: Reverse phase
rpm: Revolutions per minute
RT: Room temperature (~20 °C)
SDS: Sodium dodecyl sulfate
ss: Single-stranded
TBE: Tris-Borate-EDTA
TE: Tris-EDTA
TFA: Trifluoroacetic acid
Summary

The DNA of higher organisms is organized in a tight nucleoprotein complex called chromatin, which is visible in the cell as chromosomes. The smallest repetitive unit present in all organisms is the nucleosome core particle (NCP). The NCP comprises the histone octamer which is composed of two copies of each of the histone proteins H2A, H2B, H3 and H4. Around this protein core, which has the form of a short cylinder, 146 base pairs of DNA are wrapped in approximately 1.8 left handed superhelical turns. The NCP has a molecular weight of 206 kilo-daltons which is equally distributed between the histone octamer and the DNA.

The X-ray structure determination of the NCP has been the subject of many years of research. However, the limited quality of the crystals did not allow the determination of the X-ray structure at high resolution. Part of this thesis was the cloning of DNA fragments which yielded crystals diffracting to 2.0 Å resolution. This made it possible to solve the crystal structure of the NCP at 2.8 Å resolution using heavy atom derivatives with the method of multiple isomorphous replacement. The work presented here focuses on the refinement of this first X-ray structure to 2.0 Å and the structure determination of a core particle made from a second DNA sequence.

The X-ray structures of the two NCP provide the basis to understand the protein-DNA interactions in chromatin. In the NCP, the DNA minor groove contacts the octamer at 14 positions, once approximately every 10 base pairs. The protein-DNA contacts are predominately between DNA phosphates and histone main chain amide atoms, surrounded by positively charged lysine and arginine side chains. The nucleosomal DNA appears to be highly accessible and shows a large degree of positional flexibility. The two X-ray structures provide a huge amount of new information about the helical parameters of nucleosomal DNA as the two 146 base pair long DNA fragments are by far the longest pieces of DNA ever crystallized.

In parallel, a biochemical system to monitor the dissociation of DNA from the histone octamer was established using the method of Fluorescence Resonance Energy Transfer (FRET). The measurements show that DNA dissociation from the NCP is a stepwise process, which takes place between 0.2 and 1 M NaCl. The dissociation of the DNA termini is clearly separated from the dissociation of the DNA center.
Zusammenfassung


Parallel dazu wurde mit Hilfe der Technik des Fluoreszenz Resonanz Energie Transfers (FRET) ein System entwickelt, mit dem die Dissoziation der DNS vom Histonoktamer untersucht werden konnte. Es konnte dabei eindeutig gezeigt werden, dass die Ablösung der DNS ein Vorgang ist, der schrittweise über einen ausgedehnten Salzbereich erfolgt, und dass verschiedene intermediäre Zustände existieren.
Summary of contents

1. Introduction ...................................................................................................................... 1

2. Materials and Methods ..................................................................................................... 20

3. Fluorescence energy transfer ............................................................................................ 56

4. Crystallization of the NCP ................................................................................................. 95

5. Refinement of the NCP structure ....................................................................................... 121

6. Structure of the α6s 146 NCP ......................................................................................... 133

7. HMG-14/17 and NCP ....................................................................................................... 172

8. Past goals and future perspectives .................................................................................... 180

9. References ......................................................................................................................... 185

Appendix ................................................................................................................................. 198
List of figures

Figure 1.1: NMR structure of the HMfB protein ........................................... 17

Figure 2.1: Schematic drawing of the doubling strategy ..................................... 36

Figure 3.1: Dependence of the transfer efficiency on the distance between probes ................................................................. 61
Figure 3.2: Reaction scheme for 5' phosphate labeling ........................................ 63
Figure 3.3: DNA gel of labeled asymmetric DNA (as146) ..................................... 65
Figure 3.4: Labeling of symmetric DNA (s146) near the dyad ................................ 67
Figure 3.5: Center-labeled symmetric DNA .......................................................... 68
Figure 3.6: SDS PAGE of labeled NCP ............................................................... 69
Figure 3.7: Emission spectra of the CPM/FLU labeled DNA .................................. 72
Figure 3.8: Salt dependent donor fluorescence for double-labeled DNA ................... 73
Figure 3.9: Donor emission of chicken NCP reconstituted with end-labeled DNA .................. 74
Figure 3.10: Donor quenching with different octamers ........................................... 76
Figure 3.11: Acetylation of NCP ............................................................. 78
Figure 3.12: Overlay of donor quenching curves of recombinant- and chicken NCP ................................................................. 79
Figure 3.13: Emission scan of NCP labeled on H4 and the DNA ends .................. 81
Figure 3.14: Donor quenching curves for dimer-tetramer transfer ............................ 82
Figure 3.15: Overlay of donor quenching curves .................................................. 84
Figure 3.16: Positions of labels on the NCP surface ............................................. 87

Figure 4.1: Mobility shift gel of NCP with as146 DNA ........................................... 100
Figure 4.2: Back shifting of NCP in increasing ionic strength .............................. 102
Figure 4.3: Mobility shift gel of NCP made from symmetric 5S sequences ............... 103
Figure 4.4: Construction of human alpha satellite sequences ............................... 108
Figure 4.5: Electrophoretic mobility of symmetric DNA fragments ........................ 110
Figure 4.6: Loops for mounting crystals along the crystallographic b-axis .................. 115

Figure 5.1: The model of the α3s NCP ............................................................... 128
Figure 5.2: Surface potential representation of the histone octamer. .............................................130

Figure 6.1: The α6s DNA sequence. ........................................................................................................133
Figure 6.2: Schematic drawing of the interactions between histones and DNA. ..........................139
Figure 6.3: Superposition of the α6s and α3s NCP structure .........................................................140
Figure 6.4: Crystal contacts in the NCP structures ...............................................................................142
Figure 6.5: Twist and global axis curvature. .......................................................................................146
Figure 6.6: Manganese binding sites .................................................................................................148
Figure 6.7: Deviation in the phosphate backbone positions ...............................................................151
Figure 6.8: B-factors of the DNA backbone .......................................................................................154
Figure 6.9: DNA superposition .........................................................................................................154
Figure 6.10: DNA minor groove dimensions ......................................................................................157

Figure 7.1: Sequence alignment of human HMG-14/17. .................................................................173
Figure 7.2: HMG-17. ............................................................................................................................176

List of tables

Table 2.1: Large scale restriction enzyme digestion. ...........................................................................31
Table 2.2: E. coli cell lines and plasmids. ............................................................................................32
Table 2.3: Sequencing primers. ............................................................................................................36
Table 2.4: Restriction enzymes used for doubling. .............................................................................37
Table 2.5: Plasmid names of the multicopy plasmids. .........................................................................38
Table 2.6: Crystal soaking. ..................................................................................................................53
Table 2.7: Summary of used software. ................................................................................................55

Table 3.1: List of chromophores. .........................................................................................................63
Table 3.2: Summary of labeled NCP. ..................................................................................................70
Table 3.3: Distances between labeled sites on the NCP. .....................................................................87

Table 4.1: Table of 5S rDNA crystals. ...............................................................................................104
Table 4.2: Crothers sequence NCP crystals .................................................................106
Table 4.3a: Human alpha satellite repeat 8 NCP ......................................................109
Table 4.3b: Consensus and repeat 1 human alpha satellite NCP ..............................111
Table 4.4: Maximum oscillation angles ..................................................................115
Table 4.5: APCF space experiments .......................................................................117
Table 4.6: Diffraction of space grown crystals .......................................................118

Table 5.1: Data collection summary ......................................................................122
Table 5.2: Statistics of the α3s NCP data set ..........................................................124
Table 5.3: List of residues in the α3s NCP 2.3 Å structure ......................................125
Table 5.4: Manganese ions in the α3s NCP 2.3 Å structure .....................................126

Table 6.1: Data collection with α6s NCP ...............................................................134
Table 6.2: Summary of MR parameters ................................................................136
Table 6.3: Statistics of α6s data set .......................................................................138
Table 6.4: Summary of residues in the α6s NCP structure .....................................138
Table 6.5: Manganese ions in the α6s NCP 2.6 Å structure ....................................147

Table 7.1: Sampled crystallization conditions for NCP/HMG-17 complexes ..........177

Table 8.1: Potential labeling sites for FRET measurements .................................181
# Full table of contents

1. Introduction .........................................................................................................................1
   1.1 DNA organization in eukaryotic cells .................................................................................1
   1.2 Types of chromatin ...........................................................................................................4
      1.2.1 General .........................................................................................................................4
      1.2.2 Heterochromatin .........................................................................................................4
         Telomere ............................................................................................................................6
         Centromere .......................................................................................................................6
      1.2.3 Euchromatin ...............................................................................................................7
   1.3 Chromatin in transcription and replication .........................................................................9
      1.3.1 Factor access in transcription ......................................................................................9
      1.3.2 Replication and chromatin assembly ..........................................................................11
   1.4 The nucleosome core particle ..........................................................................................13
      1.4.1 The DNA in nucleosome core particles ....................................................................13
      1.4.2 Conformational transitions in the NCP ....................................................................15
   1.5 Histone homologies ...........................................................................................................16
   1.6 Objectives of the work .....................................................................................................19

2. Materials and Methods .......................................................................................................20
   2.1 Materials, Apparatus ........................................................................................................20
      2.1.1 Sources of chemicals .................................................................................................20
      2.1.2 Sources of enzymes and proteins .............................................................................23
      2.1.3 Gel electrophoresis, HPLC and FPLC equipment .....................................................23
      2.1.4 Centrifuges ...............................................................................................................23
      2.1.5 Concentration, lyophilisation ....................................................................................24
      2.1.6 Optical spectrometers ..............................................................................................24
      2.1.7 Crystallization, X-ray detection, X-ray sources ........................................................24
   2.2 General methods ...............................................................................................................25
      2.2.1 General buffers and bacterial media .........................................................................25
   2.3 Preparation and purification of DNA .................................................................................25
      2.3.1 Isolation of plasmid DNA from E. coli ....................................................................25
         Small scale alkaline lysis (miniprep) .............................................................................25
         Large scale alkaline lysis ...............................................................................................26
2.3.2 Preparative purification of DNA
- Ethanol precipitation of DNA fragments
- Spun column purification
- Purification of synthetic oligonucleotides
  - Denaturing gel method
  - RP-HPLC method
- Agarose gel purification
- HPLC purification methods
  - Fragment purification using DEAE 5PW
  - DNA strand separation using DEAE 5PW
- Reconstitution of ds DNA from ss DNA

2.3.3 Large scale isolation of multiple inserts from plasmids
- Releasing the insert with EcoRV
- Preparation of symmetric DNA fragments
- Large scale ligation of sticky end restriction fragments

2.4 Cloning methods

2.4.1 Transformation of competent cells
- Competent cell preparation
- Plasmid transformation

2.4.2 In vitro DNA manipulation
- Restriction digestion
- Ligation
- PCR cloning

2.4.3 DNA sequencing

2.4.4 Construction of plasmids with multiple direct repeats
- Doubling strategy

2.5 Protein methods

2.5.1 Chicken histone octamer purification

2.5.2 Recombinant histone octamer
- Recombinant histone proteins
- Recombinant histone octamer

2.5.3 Denaturing gel electrophoresis of histones
- SDS polyacrylamide gel electrophoresis
- Coomassie staining
- Triton-Urea-Acetic Acid gel electrophoresis

2.5.4 HMG-17

2.6 Fluorescence and chemical methods
2.6.1 DNA labeling
- Modification of the 5'-phosphate
- Coupling the modified DNA to the fluorescence dye
- Labeling of symmetric 146mer DNA near the center
  - Preparation of the labeled oligonucleotide
  - Ligation of the full length DNA

2.6.2 Octamer labeling
2.6.3 NCP acetylation

2.7 Chromatin (-characterization) methods
- Salt dialysis reconstitution
- Native gel electrophoresis
- Fluorescence measurements

2.8 Crystallization, X-ray work
- General
- Crystallization of the Nucleosome Core Particle
- Crystal Handling
- Data collection

2.9 Computing

3. Fluorescence energy transfer
- Introduction
- The theory of FRET
- Results
  - Labeling of DNA and histone proteins
    - The chromophores
    - Labeling the DNA termini
    - Labeling symmetric DNA near the dyad
    - Labeling of histone proteins
    - Characterization of labeled NCP
  - Characterization of the energy transfer system
  - Transfer between the DNA ends
  - Transfer between the DNA ends and H4 (H2A)
  - Dissociation of the dimer from the tetramer-DNA complex
  - Dissociation of the DNA center from H4

3.4 Discussion
- General
- Dissociation of the DNA from the octamer
The influence of the parameter Ro ........................................... 88
Mechanism of the DNA release ............................................... 89
Effects of modification of the basic histone tails on the release of the DNA ends ....................................................... 91
Dimer dissociation and the complete release of the DNA .......... 92

4. Crystallization of the NCP .................................................. 95
   4.1 Introduction ..................................................................... 95
      4.1.1 History of NCP crystallization .................................. 95
      4.1.2 Crystallization of all recombinant NCP .................... 97
      4.1.3 Nucleosomal DNA and crystallization .................... 97
      4.1.4 Objectives of the work ........................................... 99
   4.2 Results ......................................................................... 99
      4.2.1 Translational position of the DNA in the NCP .......... 99
      4.2.2 5S rDNA sequences ............................................. 103
      4.2.3 Crothers sequence .............................................. 105
      4.2.4 Human alpha satellite sequences ............................ 107
         General ........................................................................ 107
         Repeat 8 derived sequences .................................... 107
         Repeat 1 and consensus derived sequences .............. 110
      4.2.5 Optimization of crystal growth and handling .......... 111
         Crystallization ................................................................ 112
         Crystal soaking ...................................................... 112
         Crystal mounting and cryocooling ............................. 113
      4.2.6 Crystallization in microgravity ............................... 116
   4.3 Discussion .................................................................... 119

5. Refinement of the NCP structure ........................................ 121
   5.1 Introduction .................................................................... 121
   5.2 Results ......................................................................... 122
      5.2.1 Data collection .................................................... 122
      5.2.2 Data processing .................................................. 123
      5.2.3 Water molecules and manganese ions .................. 126
   5.3 Discussion ..................................................................... 127
      5.3.1 General view of the NCP ....................................... 128
      5.3.2 The DNA in the α3s NCP ..................................... 131
Chapter 1  Introduction

1. Introduction

1.1 DNA organization in eukaryotic cells

The DNA of all eukaryotic organisms is tightly packed in nucleoprotein complexes with histone proteins comprising the majority of proteins involved. The overall compaction is achieved in several steps with the fundamental, structurally repetitive unit common to all organisms being the nucleosome core (reviewed by van Holde, 1989; Wolffe, 1995). The nucleosome core particle (NCP) consists of a complex of two copies of each of the four core histone proteins H2A, H2B, H3 and H4 which form together a histone octamer (Kornberg, 1977). The histone octamer can be divided in two stable sub-complexes and exhibits a tripartite organization. Two copies of H3/H4 form the histone tetramer with two binding sites for histone H2A/H2B dimers. Thus, the octamer has a potential twofold symmetric structure. Around this cylindrical complex is wrapped 146 base pairs of DNA in approximately 1.8 left handed helical turns. The core particle has a molecular weight of 206 kD nearly equally distributed between protein and DNA.

The nucleosome core particle is a sub-fragment of the nucleosome, both of which are intermediates in the micrococcal digestion of chromatin (Noll & Kornberg, 1977). In the nucleosome, the DNA is approximately 165 base pairs long and a linker histone (such as H1, or H5 in chicken erythrocytes) is bound in the region of the incoming and the exiting DNA strands (reviewed by van Holde, 1989; Wolffe, 1995).

The four core histones H2A, H2B, H3 and H4 can be found associated with DNA in somatic cells of all eukaryotic organisms whereas the linker histone H1 is absent from yeast, although potential sequence homology has recently been identified (Landsman, 1996). The proteins have a MW of 11-15 kD (> 20 kD for H1/H5) with
more than 20% Lys/Arg content. Consequently, they carry a substantial positive charge at physiological pH which makes them ideal DNA binding proteins. The individual histones are composed of two distinct domains. A C-terminal (central) domain which forms a globular core when dimerizing, and a N-terminal basic tail which remains unstructured (at least in the nucleosome). For histone H2A and the linker histones H1/H5, both N- and C-termini remain unstructured. The actual core of the nucleosome contains the globular domains of the four core histones whereas the basic tails are thought to bind on the outer side of the DNA or to linker DNA between nucleosomes. It is interesting to note that the core histones are highly conserved in all eukaryotes from yeast to man both in length and primary sequence. Thereby, histones H3 and H4 are almost invariant, whereas H2A/H2B show a lower degree of conservation. H1 is the least conserved histone protein (reviewed by van Holde, 1989). The primary sequence conservation is not particularly high between the core histones, even though the histones share a common structural motif called the histone-fold.

The entire DNA is covered with nucleosomes with a repeat length of 160 bp to 240 bp per NCP depending on the species and tissue, the presence of linker histones, the DNA sequence and other yet uncharacterized factors (Blank & Becker, 1995; Chao et al., 1979; Crippa et al., 1993; van Holde, 1989). However, the packing of DNA into nucleosomes only results in compaction by a factor of 5, which is far to little with respect to the amount of DNA that needs to be organized in the cell nucleus. In eukaryotic cells, the nucleosomal array is most likely compacted in a 30 nm fiber or solenoid (Thoma et al., 1979). This structure consists of a higher order helix made from adjacent nucleosomes. Approximately 6-8 nucleosomes make one turn with the entrance/exit points of the DNA oriented towards the center of the solenoid. The handedness of the higher order helix is not clear and the nucleosomes are oriented with their flat surfaces parallel to the helix axis (McGhee et al., 1983a; McGhee et al., 1980). In the solenoid, the DNA is compacted by a factor of 50. When the solenoid is visualized with an electron microscope under low salt conditions, it adopts a 'beads on a
string' appearance (Thoma et al., 1979). This structure, although a preparation artifact, is also known as the 100 Å fiber and was one of the first indications for a more or less regular arrangement of nucleosomes on DNA.

In a next level of chromatin compaction, the solenoid can fold back on itself or coil into a fiber which forms loops of 40-80 kbp in length emerging from the nuclear skeleton (scaffold proteins) which defines the overall shape of the chromosome (Gasser & Laemmli, 1986). The main components of the chromatin scaffold are lamins (McKeon et al., 1986) and scaffold proteins I-III (Sc I-III) (Lewis & Laemmli, 1982). The lamins are filament-like proteins which are needed for the attachment of the chromosomes to the nuclear envelope, and they interact with the basic histone tails. This interaction is sensitive to cleavage of the histone tails with trypsin. The mode of interaction between chromosomes and the nuclear matrix may well be a regulative element in gene activity. Sc I has been identified as topoisomerase II (Earnshaw et al., 1985), can be localized throughout the chromosome and is indispensable for chromosome assembly and steady unraveling of DNA strands. Topoisomerase II serves no structural function as chromosomes depleted of the enzyme retain their overall shape. This role is played by Sc II, the Xenopus chromosome associated protein (XCAP-C/E), which is a heterodimeric coiled coil protein that seems to be responsible for chromosome condensation and structure (Hirano & Mitchison, 1994). ScIII has not yet been characterized.

DNA condensation with the help of histones is not the only mechanism used in higher eukaryotes. For example, during spermatogenesis the chromatin structure is completely remodeled and all histones are replaced by protamines. Protamines are small and highly basic proteins of a MW of 3-5 kD. The substitution of histones by protamines seems to result in an even tighter packing, probably making the DNA more resistant to physical and chemical perturbations. The removal of histones could also act as a negative regulator for all transcriptional processes that need the presence of histones (reviewed by van Holde, 1989; Wolffe, 1995).
1.2 Types of chromatin

1.2.1 General

The higher order structure of chromatin is not uniform throughout the chromosome. To adapt to the different cell cycles and to the specific requirements during cell development and differentiation, chromatin is subject to modification and compartmentalization. Chromatin was originally divided into euchromatin (decondensed) and heterochromatin (condensed) based on its appearance in light microscopy (Pardue & Hennig, 1990). Euchromatin is thought to represent the transcribed part of the genome, while heterochromatin is synonymous with the silenced regions of chromatin. On a molecular level the two types of chromatin differ with respect to histone and DNA modifications as well as by the occurrence of distinct proteins.

1.2.2 Heterochromatin

Heterochromatin is highly condensed chromatin. It comprises as much as 30% of the Drosophila genome, and can inactivate entire chromosomes such as one of the two X-chromosomes in female mammals. Regions with specialized roles in chromosome function such as centromeres and telomeres are also heterochromatic. In general, heterochromatic regions contain relatively few genes and are replicated late in the cell cycle. Exceptions to this are certain homeotic genes involved in cell differentiation and some ribosomal genes. Heterochromatic regions are characterized by the presence of distinct proteins and special histone protein modification.

Several proteins typical of heterochromatin have been characterized. For example, heterochromatin protein 1 (HP1) from Drosophila is a 19 kD protein with two domains
conserved from insects to mammals (Saunders et al., 1993). The N-terminal domain is approximately 50 amino acids in length, is also found in the polycomb protein (Pc) and has been termed the chromodomain, although its actual function is not clear. HP1 and related proteins are crucial for position effect variegation (PEV), a term describing the inactivation of genes, which are active in an euchromatic context. The polycomb group (PcG) of proteins is non-randomly distributed over the genome and is important in repression of many homeotic genes, presumably by condensation of chromatin (Zink & Paro, 1989). However, the proteins do not bind DNA directly and seem to be responsible only for the maintenance of repressed chromatin and not for its formation. It has been hypothesized that formation of heterochromatin relies on the cooperative formation of multisubunit complexes involving both protein-protein and protein-DNA interactions. After nucleation at a specific site, heterochromatin formation would proceed until some sort of termination signal is encountered (Hecht et al., 1995).

Histone proteins are generally subject to modifications, the best known example being acetylation. Special histone acetyl transferases and deacetylases regulate the acetylation state of lysine residues located in the basic tails of the core histone proteins (Kuo et al., 1996; Lewis et al., 1988; Parthun et al., 1996). Each acetylation removes one positive charge, which is likely to affect the DNA binding potential of the tails. Acetylation of histone tails seems to leave the core particle unaffected but probably destabilizes the higher order structure of chromatin by charge neutralization. Although the mechanisms of targeting the modifying enzymes to the chromosomal regions is not understood, heterochromatin is hypoacetylated (Jeppesen & Turner, 1993) in contrast to transcriptionally active chromatin which is generally hyperacetylated (Hebbes et al., 1994; Vettese-Dadey et al., 1996).

Another feature of heterochromatic regions in mammals (though not in Drosophila) is the increased methylation of cytosine bases at the N7 position of the pyrimidine ring in CpG dinucleotides (Bird, 1986). DNA methylation is known to down-regulate transcription, so active stretches of the chromosome are usually
unmethylated. The mechanism may involve either binding of special repressive factors recognizing the methylated DNA or exclusion of regulatory proteins because of the methylation.

**Telomere**

Telomeric regions are located at the ends of eukaryotic chromosomes and are indispensable for chromosome integrity and replication. They are also thought to be involved in the control of the number of cell divisions a cell can undertake (Feng et al., 1995). The DNA in this regions consists of long stretches of a tandemly repeated DNA sequence (e.g. TTAGGG in mammals). In yeast, the protein RAP1 binds over the entire length of the telomere and is needed for maintenance of telomere length. The X-ray structure of the DNA binding domain of RAP1 in complex with DNA has been solved recently and shows two tandemly repeated helix-turn-helix motifs (König et al., 1996). RAP1 binds to the proteins SIR3/4 (silent information regulators) which have been shown to interact with the N-termini of H3/H4 (Hecht et al., 1995). A complex between these proteins and nucleosomes is suggested to form the heterochromatic regions at the telomeres, and possibly also to mediate the attachment of the chromosomes to the nuclear matrix.

**Centromere**

The centromere is the part of the mitotic chromosome that acts as the point of spindle apparatus attachment, hence allows chromatid separation during cell division. It can be divided into three sub-structures. Firstly, there is the pairing domain where the inner centromere proteins (INCENP), responsible for sister chromatid association, are located. Secondly, a central domain containing very dense constitutive heterochromatin is attached to the pairing domain. Thirdly, the kinetochore as the actual site of spindle apparatus attachment is bound to the central domain. The DNA in the central domain is known as alpha satellite DNA after the tandemly repeated sequences first found in
centromeres of the African green monkey (see chapter 4). In addition to these specific sequences, proteins CENP-A, B, C have been specifically localized to the centromere region. CENP-A is a histone H3 variant differing mainly in the N-terminus (Sullivan et al., 1994). CENP-C is a hydrophilic protein located directly at the kinetochore-central domain interface and which binds directly to DNA (Saitoh et al., 1992). CENP-B is an 80 kD protein which binds a 17 bp long recognition site on the alpha-repeats of the centromere (Masumoto et al., 1989; Yoda et al., 1996). The N-terminus is the DNA binding domain whereas the C-terminus is thought to homodimerize and mediate the binding of other proteins. CENP-B could be involved in heterochromatin condensation and maintenance of the condensed state. Nucleosomes on alpha-satellite DNA seem to be rotationally positioned, whereas the translational position is not so well defined.

1.2.3 Euchromatin

In contrast to the highly condensed chromatin, transcriptionally active chromatin is less compacted. Several features are typical for active chromatin. Firstly, compared to heterochromatic histones, core histone proteins in actively transcribed chromatin are generally hyperacetylated in the tail regions (O'Neill & Turner, 1995). The reduction in charge will most likely weaken the higher order structure of chromatin although a direct effect on the binding of regulatory proteins to the histone tails can not be ruled out (Roth & Allis, 1996). Secondly, another modification typical for active chromatin is phosphorylation of serine residues in the basic histone tails, mainly of histone H1 (Dimitrov et al., 1994). Linker histone phosphorylation in non-mitotic systems leads to a decondensation of chromatin. Finally, high mobility group proteins 14/17 (HMG14/17) are among the most abundant non-histone proteins associated with chromatin. These 10-12 kD proteins are generally associated with transcriptionally active chromatin (Bustin et al., 1990).
The borders between hetero and euchromatic regions are very sharp and marked by the presence of distinct proteins. These chromosomal regions are called insulators (or boundaries), because they are able to block interaction between an enhancer with its corresponding promoter. Domain boundaries are generally assumed to be extremely stable nucleoprotein complexes attached to some supranuclear structure (e.g. nuclear matrix, chromosome scaffold). Transfer of an otherwise active gene into a heterochromatin region leads to a complete repression of the gene if it is not isolated from the heterochromatin by insulator sequences on both sides. Recently a boundary element-associated factor (32 kD, BEAF32) from a Drosophila heat shock promoter has been isolated (Zhao et al., 1995). The protein binds to a palindromic sequence present on both sides of the boundary element scs'. The two binding sites are separated by 200 bp of DNA which are highly DNaseI resistant although the flanking regions are not, so this DNA between the binding sites may be organized in a nucleosome.

The factors and modifications discussed so far are very descriptive rather than providing an explanation of their role in the process. For example, the observation that transcriptionally active chromatin is hyperacetylated does not explain the mechanism by which acetylation affects transcription. It is still questionable, whether acetylation is a prerequisite for chromatin disruption or whether histone acetylation is a consequence of chromatin disruption. The observation that transcription regulator proteins such as GCN5p have histone acetyl transferase activity could provide a solution to the question of targeting (Kuo et al., 1996; Roth & Allis, 1996; Wolff & Pruss, 1996). Based on this idea, it had been speculated that targeting of modifying activities could be by specific interactions to either distinct DNA sequences or to regulatory proteins.
1.3 Chromatin in transcription and replication

1.3.1 Factor access in transcription

Over many years, chromatin formation and transcription have been regarded as two completely unrelated topics, because the high degree of compaction observed for the eukaryotic genome at first sight seems to be an obstacle for transcription. Although a simple one dimensional diffusion process can explain site recognition on DNA in prokaryotes, it is difficult to imagine how such a simple mechanism could work on eukaryotic DNA packed into nucleoprotein complexes. Only recently has it become evident that the two processes are linked in many ways (Tijan & Maniatis, 1994; Wolffe, 1994). Therefore, nucleosomes can be visualized as both, obstacles when blocking the binding sites, or enhancers when bringing recognition sequences in close proximity. For example in the MMTV promoter, binding of all relevant transcription factors may only be possible when the promoter is associated with a nucleosome (Truss et al., 1995).

More important is the mechanism by which binding sites covered by the octamer are made available for recognition. Recently, an interesting dynamic model for chromatin has been proposed based on restriction enzyme cutting of nucleosomal DNA (Polach & Widom, 1995). Nucleosomal DNA was cut in every position within the nucleosome, with the cutting efficiency decreasing with decreasing distance of the recognition site from the nucleosomal dyad. For this to occur, the NCP must be a dynamic structure, temporary exposing stretches of its DNA. Such processes would allow non-histone proteins (e.g. transcription factors) to gain access to the DNA. Supporting evidence for this model comes from experiments with the glucocorticoid receptor which is found to bind even to a GRE positioned only 40 bp away from the nucleosomal dyad oriented toward the histone octamer (Li & Wrange, 1995). Any modification of the histone proteins increasing this dynamic behavior would therefore
facilitate factor binding. Hyperacetylation of histone tails in transcriptionally active chromatin could account for such an effect. Studies that investigated factor binding to globular nucleosomes (N-terminus removed by trypsin treatment) showed indeed enhanced binding without significant distortion of the NCP (Ausio & Holde, 1986; Vettese-Dadey et al., 1996; Vettese-Dadey et al., 1994).

However, considering the high level of compaction found in the eukaryotic genome, it seems highly unlikely that transcription could be started only by passive invasion of regulatory proteins. That is why the existence of chromatin remodeling activities had been postulated. In the recent past, several such activities were characterized and partially purified from eukaryotic cells. The SWI/SNF complex from yeast is the best characterized activity at present (Carlson & Laurent, 1994; Peterson & Tamkun, 1995; Struhl, 1996). SWI/SNF was originally identified as a positive regulator of the two genes HO (endonuclease gene involved in mating-type switch) and SUC2 (enzyme of the sucrose metabolism) in yeast and subsequently found to be important for the induction of other genes. SWI/SNF is a 2 MDa complex made from at least 5 polypeptides (Peterson et al., 1994), which acts as a general activator in the transcription of various genes. The complex has no specific DNA binding activity, although some of the polypeptides associated in the complex contain activator-like domains (Glu, Pro rich; acidic regions). In the SWI2/SNF2 subunit of the complex a 500 residue domain was has sequence homology to the members of the family of DNA dependent ATPases. The function of the SWI/SNF complex is indeed ATP dependent and mutations in the ATPase domain abolish the functionality of the complex. Similar complexes have been shown to exist in higher organisms such as Drosophila, mouse and humans (Chiba et al., 1994). The transcription induction by mammalian GR expressed in yeast, and induction in Drosophila nuclear cell extracts, were both found to be SWI/SNF dependent. Importantly, SWI/SNF enhances the binding of transcription factors and activators to their binding sites (upstream activator sites, UAS) compacted in nucleosomes (Côté et al., 1994; Imbalzano et al., 1994; Kwon et al., 1994) in a
strictly ATP dependent manner. Several histone mutants suppressing SWI/SNF mutations have been characterized, but it is not clear whether these findings indicate specific interactions between histones and the SWI/SNF complex, or whether the point mutations destabilize nucleosomes in a way that the SWI/SNF activity is no longer needed (Lenfant et al., 1996). The mechanism of activation is therefore unclear, but an alteration of the chromatin structure seems to be crucial. Recently SWI/SNF has been identified as part of the proteins associated with the C-terminal domain (CTD) of RNA polymerase II holoenzyme (Wilson et al., 1996). This finding suggests an elegant way of targeting SWI/SNF to genes, as part of the RNApol II it would always be present and could modulate chromatin structure.

1.3.2 Replication and chromatin assembly

Apart from transcription initiation, there are other processes in which the interaction of chromatin with additional protein factors is important. Enzyme complexes involved in replication (and transcription elongation) must also have developed a way to deal with omnipresent nucleosomes on the DNA templates. Taking into account the size of the replication/transcription machinery and the speed at which polymerization runs, the chromatin remodeling factors involved must be very effective. The torsional stress introduced in the DNA template is an example for the problems encountered here. Assuming no rotation of the polymerase around the DNA template, one positive supercoil per 10 bp is introduced in front of the polymerase (destabilizing NCP) while the same degree of negative supercoiling occurs behind the polymerase (Patterton & von Holt, 1993). Dissociation of the histone octamer appears to be inevitable under such circumstances. Acceptor proteins may help to overcome the energetically unfavorable dissociation of histones from the DNA.

Two proteins, that could act as a histone sink, both from Xenopus oocytes, have been characterized in the recent past. Nucleoplasm is a 22 kD protein which is
Chapter 1

Introduction

essential for chromatin assembly in oocytes (Chen et al., 1994; Laskey & Earnshaw, 1980). It is an acidic (20% Asp/Glu) and thermostable protein, which forms homopentamers. Nucleoplasm specifically binds H2A/H2B dimers and is supposed to work as a core histone dimer storage protein, promoting NCP assembly when present in excess over histones. Likewise the protein N1/N2 has been found to form specific complexes with the H3/H4 tetramer. N1/N2 is a 64.8 kD protein which also contains polyacidic tracts (Kleinschmidt et al., 1986). However, the exact mode of action for these proteins remains to be elucidated.

Xenopus oocytes and Drosophila embryo extracts have been used extensively for the characterization of the chromatin assembly process under near physiological conditions. The advantage of such systems is the presence of large quantities of histone proteins and a high replication rate, consequently generating a high rate of chromatin assembly. Recently, an ATP dependent nucleosome assembly activity has been characterized in a cell free extract isolated from Drosophila embryos (Becker & Wu, 1992; Blank & Becker, 1995; Blank & Becker, 1996; Varga-Weisz et al., 1995). The activity is capable of spacing nucleosomes on DNA similar to the in vivo spacing, with repeat length and positioning mainly depend on the ionic strength. DNA primary sequence constraints seem to be responsible only for the fine tuning of the positioning (Blank & Becker, 1996).

In contrast, recent in vitro work using Sp6 phage RNA polymerase (96 kD; 5 times smaller than the eukaryotic enzyme) suggested a direct mechanism (Studitsky et al., 1994; Studitsky et al., 1995), in which the octamer can step around the transcribing polymerase by direct transfer from the front of the enzyme to the DNA behind the polymerase by loop formation. Generally, dissociation of the histone octamer into dimer and tetramer does not seem to be crucial for transcription, as crosslinked octamer is not an obstacle in assays using T7 RNA polymerase. H1 does seem to slow down transcription (reviewed by Lewin, 1994).
1.4 The nucleosome core particle

The nucleosome core particle (NCP) is an assembly of two independently unstable structures. Under physiological conditions, the high content of positive charges in the histone proteins forces the octamer to dissociates into one tetramer and two dimers. Likewise free DNA is unlikely to bend as much as it does in the superhelix of the nucleosome core particle. Nevertheless, together a remarkably stable and conserved complex is formed.

1.4.1 The DNA in nucleosome core particles

The period of the nucleosomal DNA as probed with DNaseI and hydroxyl radicals as well as based on AT/GC sequence periodicity analysis is approximately 10.15-10.35 bp/turn as compared to 10.5 bp/turn found for B-form DNA in solution (Drew & Travers, 1985; Hayes et al., 1990; Hayes et al., 1991a). It was proposed that the DNA period is effectively 10.0 bp/turn over most of the core particle surface with an underwinding to 10.7 bp/turn in the central three turns (Hayes et al., 1990; Wolffe, 1995). The overall overwinding of nucleosomal DNA is in good agreement with results from linking number studies, where it would account for the discrepancy between the number of negative superhelical turns (1.8) and the experimentally observed reduction in linking number of 1-1.25 introduced by one core particle. This difference has also been known as the linking number paradox.

The histone octamer should be expected to bind every DNA sequence equally well, as the entire genomic pool of sequences has to be organized in the nucleoprotein complex. On the other hand, DNA is known to adopt primary sequence dependent bends and kinks especially when associated with proteins. For example, AT base pairs preferentially bend in the minor groove, where in GC base pairs this bend could be
inhibited due to the presence of a third Watson-Crick hydrogen bond forming between bases. However, alternative proposals say that AT stretches have a high propeller twist and are rather straight-bending only occurs at the borders to GC sequences. To complicate matters, bending characteristics seem to be dependent on the presence of bivalent cations and additives such as MPD (Dlakic & Harrington, 1996; Dlakic et al., 1996; Goodsell & Dickerson, 1994; Haran et al., 1994). It is likely that such constraints also affect DNA binding to the octamer surface.

Analysis of nucleosomal DNA revealed an increased probability of AT base pairs where the DNA minor groove faces the octamer, except for the 2 central turns where the pattern was inverted (Satchwell et al., 1986). Likewise, GC base pairs were enriched in positions where the minor groove faces away from the histone. Longer AT stretches were preferentially located at the DNA ends near the entry and exit points in the NCP. Sequences made from tandem repeats of the base triplet CTG have been shown to be preferentially incorporated into nucleosomes as compared to other sequences (Godde & Wolffe, 1996; Wang et al., 1994b). This result is in contradiction to gel mobility experiments where the anomalous electrophoretic behavior of CTG sequences was interpreted as the result of decreased bendability as compared to random sequences (Chastain et al., 1995).

It is unlikely that the DNA primary sequence is the only determining factor for positioning of nucleosomes. Together with the DNA length, the primary sequence is certainly the crucial factor in in vitro high salt reconstitutions but it seems to be difficult to find a common pattern. Perhaps related to this weak effect of sequence, NCP positioning experiments in more native like systems suggest that the positioning influence of the DNA primary sequence is of minor importance compared to other factors. Positioning in nuclear extracts appears to depend mainly on charge neutralization by cations and the DNA topology (Blank & Becker, 1996). In the presence of repressors or transcription factors, NCP positioning may also depend on
additional proteins bound to the DNA as shown for the albumin transcriptional enhancer (McPherson et al., 1996; McPherson et al., 1993).

1.4.2 Conformational transitions in the NCP

Since histone-DNA interactions are largely electrostatic in nature, the effect of ionic strength on the NCP has provided the basis of numerous experiments. Most of these early characterizations on core particles were done using sedimentation coefficient-, circular dichroism- (CD) and intrinsic tyrosine fluorescence measurements (Ausio et al., 1984; van Holde, 1989).

Generally, low concentrations of multivalent cations such as Mg$^{2+}$ stabilize the NCP (Pennings et al., 1991), whereas higher concentrations of monovalent cations disrupt DNA-protein interactions by competition. At very low ionic strength (< 20 mM), the NCP undergoes a structural transition during which the complex is significantly disrupted. However, there is confusion over the nature and degree of the change, as different methods resulted in highly contradictory results. It seems to be generally agreed that there is a relatively sharp transition around 1 mM salt (Brown et al., 1993). The DNA-protein interactions are unlikely to be broken under these conditions, because the low ionic environment would favor histone-DNA binding. It is more likely that the particle is disrupted by the charge repulsion between adjacent superhelical turns of the DNA while the histone proteins remain associated with the DNA. In intermediate ionic strength (20-300 mM) no large changes of the overall core particle structure were detected. At salt concentrations between 300 and 800 mM (high salt transition) a general loosening of DNA-protein interactions was detected. Above 800 mM ionic strength, histone proteins start to dissociate from the DNA.

Another method frequently used to characterize (nucleosomal) DNA was thermal denaturation. Starting with native NCP at 20°C, CD measurements at 40°C suggested
Chapter 1

Introduction

the dissociation of 20-25 bp at each DNA end. These findings are consistent with a stop site in the micrococcal digestion of nucleosomal DNA of 105 bp (at 37°C). At 60°C the dissociated DNA ends are molten followed by complete disintegration of the particle at 80°C.

Most of the work presented above was done at neutral pH. Lowering the pH to values below 4.2 leads to precipitation of the NCP. Between pH 5 and 6 and around pH 7 some minor transitions of unknown nature were recorded. As expected, urea concentrations up to 4 M resulted only in a general swelling of the particle. At higher concentrations, histone-histone interactions are broken first followed by the protein-DNA interactions at 5-6 M urea.

1.5 Histone homologies

The histone-fold motif has been identified as the central structural feature common to the globular part of all four core histones (Arents et al., 1991). The motif resembles an extended helix-strand/loop-helix motif. A short helix (helix I, ~10 aa) is followed by a loop/β-strand segment and a long helix (helix II, ~27 aa) which is again followed by a loop/β-strand segment and another short helix (helix III, ~10 aa). Figure 1.1 shows the NMR structure of a typical example, the histone-like HMfB homodimer. The orientation of the histone folds in a head-to-tail association results in the formation of a 'handshake motif'. The core of both the H2A/H2B dimer as well as the H3/H4 'half tetramer' are found to contain this orientation of secondary structure elements. The tight packing of the two histone-fold motifs is most likely responsible for the very high stability of the dimeric assemblies.

Subsequently, the histone-fold motif was identified by primary sequence homology in bacteria, Archaea and in many different proteins from eukaryotic organisms (Baxevanis et al., 1991). Most of these proteins are involved in protein-
protein or protein-DNA interactions. In eukaryotic organisms, several transcription factors and activators are members of the histone-fold super family (see below).

A family of small basic DNA-binding proteins, that share up to 50% primary sequences homology with eukaryotic histone proteins and compact 90 to 150 bp of DNA into nucleosome like structures, has been found in archae bacteria. The best characterized members of this family are the HMf proteins from *Methanothermus fervidus*. The solution NMR structure of a HMfB homodimer (7.7 kD, 69 aa; see figure 1.1) from this organism was solved recently and showed a secondary structure as expected from sequence alignments (Starich *et al.*, 1996). The overall fold of the two histone fold motifs is very similar to H3/H4 as seen in the chicken octamer structure (Arents *et al.*, 1991; Arents & Moudrianakis, 1993; Arents & Moudrianakis, 1995). The similarities between the proteins suggest that the archaeal protein evolved from the same ancestor as the eukaryal histones. It is interesting to note that gene regulation of archaeal bacteria is generally more closely related to eukaryotic gene regulation, whereas metabolic pathways seem to be more bacteria-like. However, the exact mode of action of HMfB with respect to DNA compaction remains to be determined.

![Figure 1.1: NMR structure of the HMfB protein.](image)

NMR structure of the histone fold protein HMfB from *Methanothermus fervidus*. The two monomers are shown in red and blue, respectively.
Histone like proteins have also been found in the general transcription factor TFIID. This factor is a major component of the basal transcription initiation complex involved in RNA polymerase II transcription in eukaryotic cells. It consists of the TATA-box binding protein (TBP) and approximately 13 (human) TBP-associated factors (TAFs). A subset of these TAFs has now been identified to contain the histone fold motif. Distinct domains of the human proteins TAF20/15 (20/15 kD), TAF31 and TAF80 were found to be homologous to histones H2B, H3 and H4, respectively (Hoffmann et al., 1996; Oelgeschläger et al., 1996). The X-ray structure of the amino terminal portions of dTAF42 and dTAF62 (Drosophila homologues of hTAF31 and hTAF80) has recently been solved to 2.0 Å resolution (Xie et al., 1996). The proteins form a heterotetramer very similar to the histone H3/H4 tetramer. In the transcription initiation process, TAFs interact with and mediate the functions of activators. The presence of a histone octamer-like structure in a transcription initiation complex offers interesting solutions for the problem of histone disruption during transcription. However, in vitro the N-termini have not yet been shown to bind DNA, which could be explained by the observation that the histone homologue domains of hTAF31/80 as a complex have no residual positive charge (pI=6; T. Rechsteiner, personal communication).
1.6 Objectives of the work

The characterization of large, multisubunit complexes generally involves biochemical as well as structural methods in order to obtain both dynamic and structural information. The study of the nucleosome core particle has a long history with respect to both methodologies. In this thesis, a biochemical characterization of NCP unfolding in increasing ionic strength is presented along with a high resolution X-ray crystallographic analysis.

Many details of the unfolding of the nucleosome core particle, such as the dissociation of the DNA from the histone proteins, have remained unclear. Using the technique of Fluorescence Resonance Energy Transfer (FRET), a highly specific method for mapping DNA movements in the core particle was established with the goal to investigate the dynamic behavior of the NCP with increasing monovalent cation concentration. Thereby, the dissociation of specific regions of the nucleosomal DNA from the histone octamer was characterized to develop an image of the dissociation steps. Together with high resolution structural data, this biochemical work provides the basis for a detailed interpretation of histone-DNA interactions.

The second part of the work dealt with the X-ray structure determination of the NCP. X-ray crystallography is the only method available to obtain structural information of large macromolecular complexes such as the NCP (206 kD) at atomic resolution. Crystals of the NCP diffracting to a maximum resolution of 4 Å were available in the laboratory when I began my thesis studies, so solving the X-ray structure at higher resolution required improved crystals. Optimization of NCP crystals eventually yielded crystals diffracting to 2.0 Å resolution. As a result the crystal structure of the NCP could be solved at 2.3 Å.
2. Materials and Methods

2.1 Materials, Apparatus

2.1.1 Sources of chemicals

Unless stated otherwise, chemicals of the highest obtainable quality (Fluka puriss.) were used.

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetic acid</td>
<td>Fluka</td>
</tr>
<tr>
<td>Acrylamide</td>
<td>Fluka</td>
</tr>
<tr>
<td>Adenosine 5’-triphosphate, disodium</td>
<td>Pharmacia</td>
</tr>
<tr>
<td>Agar</td>
<td>Sigma</td>
</tr>
<tr>
<td>Agarose (type II, A-6877)</td>
<td>FMC</td>
</tr>
<tr>
<td>Agarose (ultrapure grade)</td>
<td>FMC</td>
</tr>
<tr>
<td>α-[35S]-dATP</td>
<td>Amersham</td>
</tr>
<tr>
<td>α-[33P]-dATP</td>
<td>Amersham</td>
</tr>
<tr>
<td>γ-[32P]-ATP</td>
<td>Amersham</td>
</tr>
<tr>
<td>Amberlite MB-3</td>
<td>BDH</td>
</tr>
<tr>
<td>Ammonium persulfate</td>
<td>Fluka, Sigma</td>
</tr>
<tr>
<td>Ampicillin</td>
<td>Sigma</td>
</tr>
<tr>
<td>Bis-Tris (bis-(2’-hydroxyethyl)amino-tris(hydroxymethyl)methane)</td>
<td>Fluka</td>
</tr>
<tr>
<td>Boric acid</td>
<td>Fluka</td>
</tr>
<tr>
<td>Bromophenol blue</td>
<td>Fluka</td>
</tr>
<tr>
<td>Butan-2-ol</td>
<td>Fluka</td>
</tr>
<tr>
<td>Cacodylic acid</td>
<td>Fluka</td>
</tr>
<tr>
<td>Calcium chloride</td>
<td>Fluka</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>Fluka</td>
</tr>
<tr>
<td>Chloroform</td>
<td>Fluka</td>
</tr>
<tr>
<td>Coomassie brilliant blue R250 (BDH Electran grade)</td>
<td>Fluka</td>
</tr>
<tr>
<td>Cystamine dihydrochloride</td>
<td>Fluka</td>
</tr>
<tr>
<td>2’-deoxyadenosine 5’-tripophosphate, sodium</td>
<td>Pharmacia</td>
</tr>
<tr>
<td>2’-deoxycytidine 5’-tripophosphate, sodium</td>
<td>Pharmacia</td>
</tr>
<tr>
<td>2’-deoxyguanosine 5’-tripophosphate, sodium</td>
<td>Pharmacia</td>
</tr>
<tr>
<td>2’-deoxythymidine 5’-tripophosphate, sodium</td>
<td>Pharmacia</td>
</tr>
</tbody>
</table>
Chapter 2

Materials and Methods

Dicyclohexychlorodiimide (DCC) Fluka
2'3'-dideoxyadenosine 5'-triphosphate, sodium Pharmacia
2'3'-dideoxycytidine 5'-triphosphate, sodium Pharmacia
2'3'-dideoxyguanosine 5'-triphosphate, sodium Pharmacia
2'3'-dideoxythymidine 5'-triphosphate, sodium Pharmacia
DIEA (diisopropylethylamine) Fluka
N-(4-(-(Diethylamino)-4-methylcoumarin-3-yl)phenyl)maleimide (CPM) Molecular Probes
Disodium hydrogen phosphate Fluka
Dimethyldichlorosilane Fluka
N,N-Dimethyl dodecylamine-N-oxide Fluka
Dimethylformamide Fluka
Dimethylsulfoxide Fluka
Dipotassium hydrogen phosphate Fluka
DMF (dimethylformamide) Fluka
DMS (dimethylsulfate) Fluka
DTT (1,4-dithiothreitol) Fluka
EDTA (ethylenediaminetetraacetic acid disodium/dipotassium salt Fluka
Ethidium bromide Sigma
1-Ethyl-3,3-dimethylaminopropylcarbodiimide (EDC) Fluka
Fluorescein-5-maleimide (FLU) Molecular Probes
Formamide Fluka
Glucose Merck
Glycerol Fluka
Glycine Fluka
Guanidine hydrochloride (Ultra grade) Gibco BRL, Fluka
Hydrochloric acid Fluka
N-Hydroxysulfo succimide sodium salt Fluka
5-Iodoactamidofluorescein (5-IAF) Molecular Probes
Isopentyl alcohol (isoamyl alcohol) Fluka
Isopropanol Fluka
Isopropyl-β-D-1-thiogalactopyranosid (IPTG) Fluka
Magnesium acetate, tetrahydrate Fluka
Magnesium sulfate Fluka
β-mercaptoethanol Fluka
Methanol Fluka
2-Methyl-2,4-pentane diol Fluka
N,N'-methylenebisacrylamide (bisacrylamide) Fluka
1-Methylimidazole Fluka
Mineral oil Sigma
NP-40 (Nonidet P40) Fluka
Octadecylamine Fluka
Octyl-β-D-glucopyranoside Fluka
Paraffin oil Fluka
PEG 400, 600, 1000, 2000, 4000, 6000, 8000 (Polyethylene glycol) Fluka
Phenol Fluka
Phosphoric acid Fluka
Potassium acetate Fluka
Potassium chloride Fluka
Potassium dihydrogen phosphate Fluka
Propane 35 (Reingas) Carbagas
Q-Sepharose Fast Flow Pharmacia
Sephadex G-25 Medium Pharmacia
Silicone oil DC 200, 12 (100) mPa.S Fluka
Silicone oil AR 200, 235 mPa.S Fluka
Sodium acetate, anhydrous Fluka
Sodium azide Fluka
Sodium chloride Fluka
Sodium dihydrogen phosphate Fluka
Sodium dodecyl sulphate Fluka
Sodium hydroxide Fluka
Spermidine Fluka
Sulphuric acid Fluka
Sucrose [D(+) - saccharose] Fluka
Tetramethylrhodamine-5( and -6) maleimide (RHO) Molecular Probes
D(+)-Trehalose Dihydrate Fluka
1,1,1-Trichloroethane Fluka
Tris(hydroxymethyl) amino methane (Trizma base) Sigma
Triton X-100 Fluka
Tryptone Difco
Urea Fluka
Xylene cyanol Fluka
Xylitol Fluka
Yeast extract Difco
Chapter 2  

Materials and Methods

2.1.2 Sources of enzymes and proteins

Alkaline phosphatase, calf intestinal (molecular biology grade)  
Bactotryptone  
Bovine serum albumin  
EcoRV restriction enzyme  
Restriction endonucleases  
T4 DNA ligase  
T4 DNA polynucleotide kinase  
T4 DNA polymerase  
T7 DNA Polymerase  
VentR® DNA polymerase  

Boehringer  
Difco  
Boehringer, NEB  
gift Dr S. Halford  
NEB, Fermentas  
NEB  
NEB  
Pharmacia  
NEB

2.1.3 Gel electrophoresis, HPLC and FPLC equipment

Preparative Gel Apparatus : Biorad Model 491 Prep Cell  
HPLC: Gilson with Rainin Dynamix software running on a Macintosh computer; columns from Machery Nagel and Toyosoda  
FPLC: Gilson pump, detector, fraction collector and Pharmacia chromatography columns

2.1.4 Centrifuges

benchtop Heraeus, Eppendorf (both fixed angle, 13 krpm))  
tabletop Heraeus swinging bucket (4 krpm)  
large Heraeus Suprafuge 22, Sorvall RC 26 Plus(SS-34, HFA 14290, HFA 12500 rotors)
2.1.5 Concentration, lyophilisation

**Speedvac**
"Speedvac plus", model SC110A (Savant).

**Lyophilisation**
Lyovac GT3 (Leybold-Heraeus).

**Concentration**
Sartorius 13200E ultrathimbles; Centricon 10 from Amicon

2.1.6 Optical spectrometers

**UV/VIS**
Varian Cary 1E

**Fluorescence**
SPEX FL112 fluorimeter with a HAAKE C thermostat and connected to a 286 PC running dm3000 software

2.1.7 Crystallization, X-ray detection, X-ray sources

**Multiwell plates**
Falcon 347, Becton Dickinson

**Sitting drops**
Micro-Bridges®, Crystal microsystems

**Silicon grease**
Bayer medium viscosity

**Glass capillaries**
Astrophysics Research

**Film**
CEA 25 mounted on a Huber Precession camera

**Imaging plate**
MARresearch (18 or 30 cm diameter) mounted on a Huber 4 circle diffractometer (laboratory) or on the original "MAR-base" during synchrotron runs

**Rotating anodes**
Rigaku FR-C or RU-200 generators with copper target and double focusing mirrors (\(\lambda = 1.54\text{Å}\))

**Synchrotron radiation**
European Synchrotron Radiation Facility (E.S.R.F.) beamlines: BL-1, SNBL and BL-3

**Crystal freezing and storage**
Locally developed equipment as described elsewhere

**Crystal cooling**
cold gas stream, FTS systems
2.2 General methods

2.2.1 General buffers and bacterial media

TBE buffer: 89 mM Tris-borate
2 mM EDTA, usually made as a 10x stock.

TE(10:1) pH 8: the numbers in parentheses stand for the millimolar concentrations of TrisCl and EDTA respectively

DNA gel loading buffer (6x): 30% glycerol, 60 mM EDTA pH 8.0, 2.5 mg/ml bromophenol blue, 2.5 mg/ml xylene cyanol

2TY media: 1.6% bactotryptone, 1.0% yeast extract, 0.5% NaCl

TB media: 1.2% bactotryptone, 2.4% yeast extract, 0.5% glycerol; 17/72 mM KH₂PO₄/K₂HPO₄ (sterilized separately)

Agar plates: 1.0% bactotryptone, 0.5% yeast extract, 0.8% NaCl, 1.5% agar.

All media were sterilized in an autoclave for 20 min at 121° C. 50 µg/ml ampicillin or/and 25 µg/ml chloramphenicol was added depending on the type of cells used.

2.3 Preparation and purification of DNA

2.3.1 Isolation of plasmid DNA from E. coli

Small scale alkaline lysis (miniprep)

Overnight cultures of E. coli (5 ml 2TY media) harboring the desired plasmid were centrifuged for 5 min (3000 rpm, RT, tabletop centrifuge) in 15 ml polypropylene tubes (Falcon; Becton Dickinson). The pelleted cells were resuspended in 100 µl lysis buffer (25 mM TrisCl pH 8.0, 50 mM glucose, 10 mM EDTA) and transferred to 1.5 ml Eppendorf tubes. 200 µl NaOH/SDS solution (0.2M NaOH, 1% SDS) was then added, and the tubes were shaken vigorously prior to incubation on ice for 5 min. 300 µl of cold 4/6M potassium/acetate solution was added to each, the tubes shaken briefly and then incubated on ice for 5 min. After centrifugation of the samples (5 min, 13000 rpm,
Chapter 2

Materials and Methods

benchtop centrifuge) the supernatant was removed and then extracted with 500 µl phenol/CIA and 500 µl CIA. After ethanol precipitation the DNA pellet was resuspended in 50 µl TE(10;0.1) pH 8 containing 40 µg/ml DNase free RNaseA and incubated for 30 min at 37°C (procedure adapted from Sambrook et al., 1989).

Large scale alkaline lysis

For preparative isolation of DNA fragments, 6-12 liters of cells harboring the desired plasmid were grown for 16-20 hours from 100 ml starter cultures in TB media. The cells were pelleted in 4-6 500 ml centrifuge bottles by centrifugation of the media (6 min, 6000 rpm, 4°C, HFA 12500 rotor) and resuspended in 70 ml lysis buffer per bottle. 140 ml NaOH/SDS solution were added, the bottles shaken very rigorously and then incubated on ice for 10 min prior to the addition of 210 ml cold 4/6 M potassium/acetate. The mixture was shaken gently, incubated on ice for 20 min, and centrifuged (5 min, 10000 rpm, 4°C, HFA 12500 rotor) to remove the chromosomal DNA and cell debris. The supernatant was filtered through miracloth (Calbiochem) and the plasmid DNA was precipitated by the addition of 0.52 vol. isopropanol. After centrifugation (30 min, 10000 rpm, RT, HFA 12500 rotor) the pellets were air dried for 1 h and subsequently transferred into 40 ml of TE(10,50) pH 8/40ug/ml RNaseA in two 50 ml polypropylene tubes. The DNA in the pellets was dissolved by incubation at 37°C (approx. 3h). The RNase digest was extracted 3 times with an equal volume of TE equilibrated phenol (centrifugation 5 min, 20000 rpm, RT, SS-34), and once with CIA. In these steps, poorly defined interface regions were transferred to Eppendorf tubes and re-centrifuged to obtain a clear separation. The DNA was PEG precipitated by addition of NaCl and PEG 6000 to 0.5M and 10%, respectively. After incubation on ice for 30 min the precipitated plasmid DNA was collected in 50 ml polypropylene tubes (10 min, 15000 rpm, 4°C, SS-34). The pellet was resuspended in 20 ml TE(10, 0.1) pH 8, extracted 3 times with an equal volume CIA. The DNA was ethanol
precipitated (adapted from Sambrook et al., 1989). The DNA preparation was free of chromosomal DNA and suitable for enzymatic digests.

2.3.2 Preparative purification of DNA

Ethanol precipitation of DNA fragments

DNA fragments 30 to 6000 bp in length were precipitated by adding 2.5 volumes of absolute ethanol to the DNA containing 0.5 M NaCl. The precipitated DNA was harvested by 10-20 min centrifugation at the maximum allowed speed for the used centrifuge system (benchtop 13 krpm, tabletop 3.8 krpm, large 8 krpm). The DNA pellets were resuspended in TE(10, 0.1) pH 8 to give a final DNA concentration of 0.5-10 mg/ml.

Spun column purification

Spun columns were constructed by placing siliconized glass wool in the base of a 1 ml plastic syringe barrel, filling the barrel with Sephadex G25 medium, equilibrated in the appropriate buffer. The syringe was placed inside a 15 ml Falcon polypropylene tube, and pre-equilibrating by centrifugation for 10 min (2000 rpm, RT, tabletop centrifuge). The sample (<120 µl) was applied to the top of the Sephadex G25 column, the assembly placed in a fresh 15 ml tube containing a Eppendorf tube without cap, and eluted by centrifugation as above.

Purification of synthetic oligonucleotides

Denaturing gel method

Oligonucleotides were synthesized at 0.2 µmol scale on an ABI 308B DNA synthesizer (Applied Biosystems) and de-protected by incubation at 55° C for 14 hours. The de-protected synthesis product was dried in a Speedvac and resuspended in 100 µl H₂O. 25 µl formamide dye was added to an equal volume of the crude oligonucleotide
Chapter 2  

Materials and Methods

stock. The solution was boiled for 3 min and loaded onto a 10-15% (1:20 bisacrylamide: acrylamide), 7M urea, 1 x TBE gel (60 cm x 20 cm x 1 mm). The gel was run at 40W for 5-6 hours, removed from the plates and wrapped in Saran wrap (Dow). Bands were visualized by their absorbance against a fluorescent TLC plate (Merck) under UV illumination at 254 nm, and thus could be excised. The gel fragments were placed in a 12-15000 MWCO dialysis bag together with 1 ml H₂O and dialyzed against two changes of a large volume of H₂O. The liquid in the bag was isolated by filtering through a 0.45 μm filter (Sartorius) and the DNA was dried in a Speedvac. The oligonucleotide was resuspended in 100 μl TE(10,0.1).

RP-HPLC method

Oligonucleotides were synthesized at 0.2/1 μmol scale with trityl groups left on, then de-protected by incubation at 55° C for 14 hours. Using a method adapted by S. Tan (Zon, 1990), Tris base (100 μl, 1M) was added to the de-protected reaction product and the solution was evaporated to dryness without heating in a Speedvac. The remaining solid was resuspended in 400 μl 100 mM triethylammonium acetate pH 6.7 and loaded onto a C8 reverse phase column (Nucleosil 300-5, 4 mm x 250 mm; Machery-Nagel) which was equilibrated in the same buffer. The column was washed with 12 ml of buffer at a flow rate of 1 ml/min. A 3 ml pulse of 0.5% trifluoroacetic acid was injected which cleaved the trityl groups and allowed elution of the oligonucleotide with a gradient of 10-16% acetonitrile in 100 mM triethylammonium acetate pH 6.7 over 16 min. The eluted material, as detected by its absorbance at 260 nm, was concentrated to a small volume in a Speedvac and dissolved in TE (10,0.1) pH 8.

Agarose gel purification

DNA (<2 pmol/10mm lane) in 1 x DNA gel loading buffer was applied to 0.8% agarose minigels (8 cm x 8 cm x 0.5 cm, SeaChem ME grade, 1 x TBE, 1 μl 10 mg/ml Ethidium bromide in the gel), and electrophoresed at 70V for 40 min. The desired bands
were excised while being visualized on a UV transilluminator. The agarose slices were then placed in a 0.5 ml safe lock Eppendorf tube prepared with a 0.5 mm hole pierced in the bottom and siliconized glass wool immediately above this. The 0.5 ml tube was placed in an open 1.5 ml Eppendorf tube and centrifuged for 5 min (7000 rpm, RT, benchtop centrifuge). The DNA was used for cloning purposes without further purification.

**HPLC purification methods**

**Fragment purification using DEAE 5PW**

DNA fragments of 69-150bp in length were bound to a TSK DEAE-5PW ion exchange column (21.5 mm x 150 mm; Toyosoda) equilibrated in 20 mM TrisCl pH 7.5, 0.1 mM EDTA, 300 mM NaCl. The DNA was eluted by a gradient ranging from 400-500 mM NaCl in the same buffer (approx. 40 minutes) at 4 ml/min. The peak fractions as detected by absorbance at 260nm and polyacrylamide gel electrophoresis were pooled, and the DNA ethanol precipitated. The pellet was resuspended in TE(10,0.1) pH 8.

**DNA strand separation using DEAE 5PW**

Asymmetric 146mer was separated into single strands using a TSK DEAE 5PW (Toyo Soda) column at 1 ml/min at RT (as suggested by S. Tan). 1 mg of asymmetric 146mer (Richmond et al., 1988) in a total volume of 1.5 ml 10 mM NaOH/ 0.5 M NaCl was incubated for 15 min. at RT. The sample was loaded onto the column and eluted in a gradient of 2 mM NaCl/min. The first strand eluted at 0.61 M NaCl/10 mM NaOH, the second eluted 20 mM later. After adjusting the pH of the fractions to approximately 6 by the addition of 5 μl 3M NaAc pH 5.2 per 1 ml fraction, the DNA was directly ethanol precipitated and dissolved in TE (10, 0.1) pH 8.
Reconstitution of ds DNA from ss DNA

Equimolar amounts of the two complementary strands were heated to 90° C for 3 min. After cooling on ice for 5 min, the DNA was brought to 75° C for 15 min and allowed to cool to RT slowly (1.5 h). Remaining single stranded DNA was removed by preparative PAGE (10 W, RT, 15 ml of 8% 40:1 acrylamide:bisacrylamide gel, 0.5 TBE as running and elution buffer). The peak fractions were pooled, ethanol precipitated and dissolved in TE (10, 0.1) pH 8 at -20° C.

2.3.3 Large scale isolation of multiple inserts from plasmids

Releasing the insert with EcoRV

Plasmid-DNA containing the inserts (<200 bp) flanked by EcoRV restriction sites (Richmond et al., 1988) was digested as follows: 1 mg/ml plasmid in 50 mM TrisCl pH 7.6, 100 mM NaCl, 10 mM MgCl\textsubscript{2} and 1 mM DTT was digested with 30u EcoRV/nmol EcoRV sites. Incubation was at 37° C for 12 h. After checking the completeness of digestion by gel electrophoresis, 0.192 vol. 4M NaCl and 0.346 vol. PEG 6000 were added. The mixture stood on ice for 60 min and was centrifuged for 15 min (15000 rpm, 4° C, SS-34). The supernatant containing the insert-DNA was ethanol precipitated and the precipitate collected by centrifugation. The DNA pellet was resuspended in TE(10,0.1) buffer.

Preparation of symmetric DNA fragments

Dephosphorylation of blunt end fragments was performed as follows: the reaction mix contained 1 mg/ml DNA in 50 mM TrisCl pH 8, 0.1 mM EDTA, 10 mM MgCl\textsubscript{2} and 1u CIP/nmol DNA ends. After o/n incubation at 37° C the DNA was extracted twice with phenol/CIA, once with CIA, and then ethanol precipitated.
In order to create the sticky DNA end needed in the preparation of symmetric DNA fragments, the dephosphorylated DNA was dissolved in the buffer suitable for particular restriction enzyme used (table 2.1). The digest was performed according to the conditions listed. The large fragments were purified on DEAE HPLC as described previously.

<table>
<thead>
<tr>
<th></th>
<th>Reaction conditions</th>
<th>Enzyme concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>HindIII</td>
<td>0.1-0.2 mg DNA/ml in 50 mM TrisCl pH 8, 125 mM NaCl, 25 mM MgCl₂, 10 mM DTT</td>
<td>20 μ/nmol sites; incubation at 37°C for approx. 36 h</td>
</tr>
<tr>
<td>EcoRI</td>
<td>1 mg DNA/ml in 100 mM TrisCl pH 7.5, 50 mM NaCl, 10 mM MgCl₂, 0.025% Triton X-100</td>
<td>20-30 μ/nmol sites, incubation at 37°C for approx. 24 h</td>
</tr>
<tr>
<td>HinfI</td>
<td>1 mg DNA/ml in 10 mM TrisCl pH 8, 50 mM NaCl, 10 mM MgCl₂, 1 mM DTT</td>
<td>10-15 μ/mol sites, incubation at 37°C for approx. 12 h</td>
</tr>
</tbody>
</table>

Table 2.1: Large scale restriction enzyme digestion.

Large scale ligation of sticky end restriction fragments

1 mg/ml DNA was ligated in a buffer containing 20 mM TrisCl pH 7.6, 10 mM MgCl₂, 10 mM DTT, 2 mM rATP and 0.5 u T4 DNA ligase /21 pmol restriction ends to be ligated. Incubation was at RT for 12 h. Unligated DNA was removed by DEAE HPLC purification (conditions as described earlier).

2.4 Cloning methods

2.4.1 Transformation of competent cells

Competent cell preparation

The requisite *E. coli* strains (see table 2.2) were obtained from S. Tan and restreaked on TYE agar plates. Using colonies from this plate, a 250 ml culture (250 ml
Chapter 2

Materials and Methods

2TY media with 5.6 ml competent cell salts solution [0.44M MgCl₂, 0.44M MgSO₄, 0.11M KCl]) was grown to an optical density of 0.6 at 600 nm. The culture was chilled on ice for 10 minutes and then centrifuged (4000 rpm, HFA 12500, 10 min, 4°C). The pelleted cells were gently resuspended in 80 ml ice-cold transformation buffer (10 mM PIPES, 250 mM KCl, 15 mM CaCl₂, 55 mM MnCl₂) and the centrifugation was repeated. The cell pellet was re-suspended in 20 ml of ice-cold transformation buffer, and the suspension was transferred to a sterile polypropylene tube. DMSO (1.5 ml) was added with gentle swirling and the cells were incubated on ice for 10 minutes. Aliquots of 400 μl were dispensed into 1.5 ml Eppendorf tubes and flash frozen in liquid nitrogen. The cells were stored at -80°C and gave a transformation efficiency of 10⁶-10⁷ colonies/μg pUC19 (Inoue et al., 1991; and adapted by S. Tan).

Table 2.2: E. coli cell lines and plasmids.

<table>
<thead>
<tr>
<th>Cells</th>
<th>Genotype hsdΔ5, thi, supEΔ(lac-proA) F' [traD36 proA+B⁺ lacQZΔM15]; (Sambrook et al., 1989)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HG1</td>
<td>Genotype hsdS recA proA leu-6 ara14 galK2 lacY1 xyl-5 mtl-1 prsL20 thi-1 sup44; (Boyer &amp; Roulland-Dussoix, 1969)</td>
</tr>
<tr>
<td>BL21(DE)</td>
<td>F- ompT hsdSB (rB- mB-; an E. coli B strain with a λ prophage carrying the T7 RNA polymerase gene). pLysS is a chloramphenicol-resistant plasmid carrying a gene for lysozyme (Studier et al., 1990).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pUC19</td>
<td>General purpose E.coli cloning vector of 2686 bp carrying an ampicillin-resistance gene; (Yanisch-Perron et al., 1985)</td>
</tr>
<tr>
<td>pWM528</td>
<td>2050 bp long derivative of pUC19 in which noncoding regions and several restriction sites have been removed. (Mandecki et al., 1990)</td>
</tr>
<tr>
<td>pET3a</td>
<td>Component of an E. coli protein expression system carrying an ampicillin-resistance gene. These vectors have a T7 phage promoter which is recognized only by the T7 RNA polymerase but not by the host RNA polymerase (used with BL21 cells). (Studier et al., 1990).</td>
</tr>
</tbody>
</table>
Plasmid transformation

Plasmid DNA (<0.1 ng, or 2-5 μl of the ligation reaction product) was added to 200 μl competent cells which had been thawed whilst standing on ice. The cell suspension was mixed by tumbling (Spiromix) at 4°C for 40 min. The cells were heat-shocked at 42°C for 30 seconds and immediately chilled on ice for 30-120 seconds. 500 μl of 2TY media was added and the cells were incubated for 40 min at 37°C. The transformed cells (300 μl) were plated onto 2TY agar plates containing appropriate antibiotics.

2.4.2 In vitro DNA manipulation

Restriction digestion

Restriction digestions were carried out in volumes of 20-50 μl in restriction buffers supplied by New England Biolabs. The digestion mix contained 10-20 units of restriction enzyme and the solution was incubation at 37°C for 1-3 hours.

If dephosphorylation was required, 1u of Calf intestinal alkaline phosphatase (CIP) was added for the last 30 min of the digestion.

For analytical purposes, 10 μl of the digestion mix was added to 2 μl of 6 x DNA gel loading buffer, and the samples were applied to appropriate gels. To analyze DNA fragments in the range of 50-500 bp polyacrylamide gels (10% 1:60 bisacrylamide: acrylamide, 1 x TBE, 10 cm x 10 cm x 0.5 mm) were run at 12W for 20 min. The gels were stained briefly with Ethidium bromide. Longer fragments (>500 bp) were applied to agarose gels (0.8% SeaChem ME grade agarose (FMH), 1 x TBE, Ethidium bromide present in the gel) run at 70V for 40 min. For preparative purposes, the digestion mixes were extracted once with phenol/CIA and CIA respectively and electrophoresed as described for preparative gel methods above.
Ligation

For cohesive-ended ligation, approximately 200 ng of vector and insert were mixed in a volume of 10 µl ligation buffer containing 50 mM TrisCl pH 7.8, 10 mM MgCl2, 10 mM DTT, 25 µg/µl BSA, 1 mM ATP and 40 units of T4 DNA ligase (New England Biolabs). Ligation was for 2-3 hours at RT or overnight at 16° C. The ligation reaction was centrifuged briefly and then used for plasmid transformation.

PCR cloning

Amplification of DNA for cloning purposes was performed by the polymerase chain reaction (Saiki & Gelfand, 1988). In a reaction volume of 100 µl, 10 to 100 ng of plasmid DNA containing the target sequence to be amplified was incubated on ice in a 0.5 ml safe lock Eppendorf tube containing PCR reaction buffer (20 mM TrisCl pH 8.8, 10 mM KCl, 10 mM (NH4)2SO4, 2 mM MgSO4, 0.1% Triton X-100) with 0.2 mM dNTP and 0.2 µM primers. Vent DNA polymerase (2 units, New England Biolabs) was added, 60 µl of mineral oil (Sigma) was overlaid, and the samples were transferred to a thermocycler (Perkin Elmer) which had been pre-warmed to 95° C. The samples were incubated at 95° C for 2 min and then the amplification was automatically carried out using following the scheme: 4 cycles of 30 seconds at 95° C, 60 seconds at a temperature 5° C lower than the lowest melting temperature of the two primers, and extension at 75° C for 30 seconds followed by 25 cycles of 30 sec at 95° C, 30 sec at 45° C and 30 sec at 75° C. A final “polishing” incubation of 5 min at 75° C was used, followed by holding at 0° C. DNA products were analyzed by DNA PAGE as described earlier. The reaction products were phenol/CIA and CIA extracted. The DNA was ethanol precipitated and used for cloning purposes without further purification.
2.4.3 DNA sequencing

Sequencing of ds DNA was carried out by the dideoxy chain termination method (Sanger et al., 1977) using T7 DNA polymerase (Tabor & Richardson, 1987). 2 μg of plasmid DNA in 20 μl were alkali denatured at room temperature for 5 min by addition of 2 μl of 2M NaOH, 2 mM EDTA solution. After addition of 7 μl H2O, the denatured DNA was ethanol precipitated and the dried DNA pellet was resuspended in 14 μl of a mixture containing 10 pmol of sequencing primer and 2 μl of T7 anneal buffer (280 mM TrisCl pH 7.5, 0.35M NaCl, 100 mM MgCl2). The sample was incubated for 10 min at 37° C and 10 min at room temperature. T7 DNA polymerase mix (3 μl of 100 mM DTT, 2.0 μl of T7 label mix (2 μM dTTP, 2 μM CTP, 2 μM dGTP), 0.5 μl of T7 DNA polymerase dilution buffer (Pharmacia), 0.2 μl 10 μCi/μl α-[35S] or α-[33P]-dATP, and 0.5 μl of 2 units/μl T7 DNA polymerase (Pharmacia) was added to the primer-annealed template. The mixture was incubated at room temperature for 5 min. Aliquots (4.5 μl) were pipetted on the side of 4 wells of a microtitre plate (Falcon; Becton-Dickinson), each well containing 2.5 μl of one of the four termination mixes (40 mM TrisCl pH 7.5, 50 mM NaCl, 10 mM MgCl2, 150 μM dNTP, 10 μM ddXTP, where X=A,C,G,T). The microtitre plate was centrifuged briefly to mix the reagents, and incubated at 42° C for 5 min before 2 μl of formamide dyes was added to each well to stop the reaction. Samples were incubated for 15 min in a 80° C oven and loaded onto a denaturing polyacrylamide gel. Sequencing gels (6% (1:20 bisacrylamide:acrylamide, 7M urea, 0.5-5 x TBE gradient, 60 cm x 20 cm x 0.1 mm using siliconized glass plates), were run for approximately 150 min at 45W using 0.5 x TBE in the upper and lower buffer reservoirs. The gels were fixed in 10% ethanol/10% acetic acid for 15 min, dried onto blotting paper (Sihl Papier) in a vacuum dryer (BioRad) and exposed to X-ray film (Fuji) at room temperature for 3-18 hours. The sequencing primers used are summarized in table 2.2.
Table 2.3: Sequencing primers.

<table>
<thead>
<tr>
<th>Identity</th>
<th>Sequence (5' to 3')</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST02</td>
<td>G'ITTTCAGTCACGAC</td>
<td>primer for pUC19 based plasmids</td>
</tr>
<tr>
<td>ST0429</td>
<td>TTAAGTTGGGTAACGCG</td>
<td></td>
</tr>
<tr>
<td>ST03</td>
<td>CAGGAAACAGCTATGAC</td>
<td>reverse primer for pUC19 based plasmids</td>
</tr>
<tr>
<td>ST0430</td>
<td>GTTGTGTGGAATTGTGAG</td>
<td></td>
</tr>
<tr>
<td>ST036</td>
<td>TAAACAATTTCACACAGGA</td>
<td>primer for pWM528 based plasmids</td>
</tr>
<tr>
<td>ST0175</td>
<td>ATTTTCCAGTCTCGTC</td>
<td>reverse primer for pWM528 based plasmids</td>
</tr>
<tr>
<td>ST037</td>
<td>TAATACGACTCCTATA</td>
<td>primer for pET3a based plasmids</td>
</tr>
<tr>
<td>ST0127</td>
<td>CCCCTCAAGACCCGTTT</td>
<td>reverse primer for pET3a based plasmids</td>
</tr>
</tbody>
</table>

2.4.4 Construction of plasmids with multiple direct repeats

Doubling strategy

The construction of multiple direct repeats is most simply achieved in vectors where the insert (defined as the element between restriction sites A and B in figure 2.1) is located between two restriction sites generating compatible cohesive ends (D and E), which when ligated no longer form a restriction site. A third, unique restriction site C outside D or E is used in the following way: digestion of the plasmid at sites C/D creates a vector which can be ligated to the insert created by digestion with D/E. The number of inserts (A, B) is doubled with each cloning step (after the suggestion of Prof. T. Richmond). Different variants of this concept were used. Figure 2.1 shows a schematic drawing of the strategy.

![Figure 2.1: Schematic drawing of the doubling strategy.](image)

Overview of a one copy plasmid. The open rectangle is the insert defined as sequence between restriction sites A and B.
Table 2.4 gives a list of the 4 different enzyme combinations used for the cloning of the fragments.

<table>
<thead>
<tr>
<th>System</th>
<th>A</th>
<th>B</th>
<th>B’</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HindIII</td>
<td>EcoRV</td>
<td>EcoRV</td>
<td>EcoRI</td>
<td>BglII</td>
<td>BamHI</td>
</tr>
<tr>
<td>2</td>
<td>EcoRI</td>
<td>EcoRV</td>
<td>-</td>
<td>HindIII</td>
<td>SalI</td>
<td>Xhol</td>
</tr>
<tr>
<td>3</td>
<td>EcoRI</td>
<td>EcoRV</td>
<td>EcoRV</td>
<td>Kpnl</td>
<td>BglII</td>
<td>BamHI</td>
</tr>
<tr>
<td>4</td>
<td>Hinfl</td>
<td>EcoRV</td>
<td>EcoRV</td>
<td>Kpnl</td>
<td>BglII</td>
<td>BamHI</td>
</tr>
</tbody>
</table>

Table 2.4: Restriction enzymes used for doubling.

In table 2.5 all plasmids constructed to produce symmetric DNA fragments are summarized. When using the enzyme combination D/E for the first subcloning step often vectors with 2 or 3 inserts could be obtained in one step. Doubling was achieved by several cycles of digestion of the plasmid, isolation of vector and insert DNA, purification on agarose gels and ligation of vector and insert. In some cases inserts of previous doubling steps were used to obtained the desired copy number. *E. coli* HB101 was used for almost all subcloning, occasionally replaced by TG1 in the subcloning of plasmids with less than 4 inserts.
### Table 2.5: Plasmid names of the multicopy plasmids.

<table>
<thead>
<tr>
<th>plasmid</th>
<th>copy no.</th>
<th>frag. name</th>
<th>enzyme comb.</th>
<th>insert construction</th>
<th>first copy cloned into</th>
</tr>
</thead>
<tbody>
<tr>
<td>p32-71 (pUC19)</td>
<td>32</td>
<td>s142</td>
<td>1</td>
<td>PCR with TR51 and AMO4 from 5s RNA gene fragment</td>
<td>pUC19 EcoRI/BamHI</td>
</tr>
<tr>
<td>p32-69</td>
<td>32</td>
<td>s138</td>
<td>1</td>
<td>PCR with TR51 and AMO5 from 5s RNA gene fragment</td>
<td>pUC19 EcoRI/BamHI</td>
</tr>
<tr>
<td>p16-α2s (pWM528)</td>
<td>16</td>
<td>α2s 146</td>
<td>2</td>
<td>constructed from AMO7 and AMO8</td>
<td>pST30 (S. Tan) Sall/XhoI</td>
</tr>
<tr>
<td>p18-cs</td>
<td>18</td>
<td>cs 146</td>
<td>2</td>
<td>constructed from STO486 and STO487</td>
<td>pST30 (S. Tan) Sall/XhoI</td>
</tr>
<tr>
<td>p24-α3s (pUC19)</td>
<td>24</td>
<td>α3s 146</td>
<td>3</td>
<td>PCR with AMO9 and 10 from human alpha satellite DNA</td>
<td>pUC19 KpnI/BamHI</td>
</tr>
<tr>
<td>p24-α4s</td>
<td>24</td>
<td>α4s 146</td>
<td>3</td>
<td>PCR with AMO17 and 18 from α3s</td>
<td>pl-α3s BgII/BamHI</td>
</tr>
<tr>
<td>p24-α5s</td>
<td>24</td>
<td>α5s 146</td>
<td>3</td>
<td>PCR with AMO13 and 14 from α2s</td>
<td>pl-α3s BgII/BamHI</td>
</tr>
<tr>
<td>p16-α6s</td>
<td>16</td>
<td>α6s 146</td>
<td>3</td>
<td>constructed from AMO15 and 16</td>
<td>pl-α3s BgII/BamHI</td>
</tr>
<tr>
<td>p16-α7s</td>
<td>16</td>
<td>α7s 146</td>
<td>3</td>
<td>PCR with AMO19 and 20 from α6s</td>
<td>pl-α3s BgII/BamHI</td>
</tr>
</tbody>
</table>
| p24-α8s (pUC19) | 24 | α8s 147 | 4 | PCR with AMO21/22 and AMO23/22 for half sides a and b, respectively (template α3s)*** | a) pl-α3s BgII/BamHI b) pl-α8s-a BgII/BamHI (***)
| p24-α9s | 24 | α9s 147 | 4 | as above with AMO24/25 and AMO26/25 | a) pl-α3s BgII/BamHI b) pl-α9s-a BgII/BamHI (***)
| p24-α10s | 24 | α10s 145 | 4 | as above with AMO24/22 and AMO26/22 | a) pl-α3s BgII/BamHI b) pl-α10s-a BgII/BamHI (***)

(*** The two halves were cloned directly behind each other on the same plasmid.)
2.5 Protein methods

2.5.1 Chicken histone octamer purification

Chicken blood (Simon & Felsenfeld, 1979; procedure modified by T. Rechsteiner and L. Toldo) was collected directly into 15 mM sodium citrate and 15 mM NaCl to avoid clotting. The blood samples were put on ice immediately after collection. The blood was filtered through cotton gauze, diluted with 8 volumes of wash buffer (15 mM KCacodylate, 60 mM KCl, 15 mM NaCl, 0.15 mM spermine, 0.5 mM spermidine, 2 mM EDTA, 0.25 mM PMSF, 1 mM benzamidine, 0.5 M sucrose) and centrifuged in 150 ml COREX bottles (15 min, 4500 rpm, 4° C, HFA 12500 rotor, using 150 ml). The supernatant and the white blood cells were carefully removed and the red blood cells were washed again with 100 ml wash buffer. The cells were finally resuspended in 50 ml wash buffer, frozen in liquid nitrogen and stored at -20° C. Thawed blood cells were lysed by addition of 4 volumes of wash buffer with 0.5% NP-40 detergent and gentle swirling. The nuclei were washed by addition of 8 volumes of wash buffer supplemented with 40 mM β-mercaptoethanol and centrifugation as before. Washing the nuclei was repeated until the red haemoglobin color had disappeared completely. The nuclei were stored at -20° C after dilution with an equal volume of glycerol. After dialysis against digest buffer (wash buffer supplemented with 10 mM β-mercaptoethanol) to remove the glycerol, nuclei were washed by 3 cycles of centrifugation and re-suspension as before. The pH of the nuclei suspension was adjusted to 7.5 and the DNA concentration (measured by its absorbance at 260 nm in 100 mM NaOH) was set to 7.5 mg/ml. The suspension was equilibrated to 37° C, before CaCl₂ was added to a concentration of 1 mM. Micrococcal nuclease lyophilisate (Sigma) was resuspended in 25 mM TrisCl pH 7.5, 0.2 mM CaCl₂ to a concentration of 200 units/ml. Trial micrococcal nuclease digestions were carried out by addition of 2.5 μl of MNase to 1 ml samples incubated at 37° C. Several 50 μl aliquots were withdrawn.
during the first 10 min of digestion and diluted 100 fold with 0.1M NaOH. The samples were centrifuged (5 min, 13000 rpm, RT, benchtop centrifuge), and the absorbance of the supernatant was determined. Other 50 µl aliquots were extracted with phenol and analyzed on 0.8% agarose gels (see restriction digest analysis). Bulk digestion was then carried out for the time determined by the trial digest to give a plateau in DNA absorbance, which corresponds to 500 bp average fragment length. The reaction was stopped by addition of EDTA to 2 mM. The digestion mix was then centrifuged (10 min, 3000 rpm, RT, HFA12500 rotor) and the pellet re-suspended in one half volume of nuclysis buffer (10 mM TrisCl pH 7.5, 0.2 mM EDTA, 0.25 mM PMSF, 1 mM benzamidine, 10 mM β-mercaptoethanol), vortexed and stood on ice for 30 min. Long chromatin was isolated in the supernatant by centrifugation (10 min, 4500 rpm, 4° C, HFA 12500 rotor) and stored at -20° C after dilution with an equal volume of glycerol. Long chromatin, dialyzed against nuclysis buffer without benzamidine to remove glycerol, was loaded onto a freshly-packed, defined hydroxyapatite column (1g hydroxyapatite per 20 mg DNA to be loaded, 5 cm diameter Pharmacia TK50 column, flow rate 2 ml/min) equilibrated in loading buffer (10 mM TrisCl pH 7.5, 0.25 mM PMSF, 0.5 mM DTT) until the column was fully loaded (as measured by flow-through absorbance at 260 nm). The column was then washed well with loading buffer, and then by several column volumes of hydroxyapatite equilibration buffer (50 mM sodium phosphate pH 6.0, 65 mM NaCl, 0.25 mM PMSF, 0.5 mM DTT). H1/H5 was eluted using the same buffer containing 650 mM NaCl, and octamer by buffer containing 2.5M NaCl. For separation of dimer and tetramer, a 200 ml linear gradient from H1 to octamer eluting NaCl concentrations was used. Eluted fractions were identified by absorbance at 280 nm and using SDS polyacrylamide gels.
2.5.2 Recombinant histone octamer

Recombinant histone proteins

_E. coli_ BL21 cells containing the pET expression plasmids for _Xenopus laevis_ histones H2A, H2B, H3 and H4 (including mutants) were obtained from K. Luger. Cells were grown in 6-12 liter 2TY media (with 100 µg/ml ampicillin and 25 µg/ml chloramphenicol) at 37°C until the optical density measured at 600 nm was 0.4. The cells were induced with 0.4 mM IPTG and grown for another 2 to 2.5 hours. The cells were harvested by centrifugation (7 min, 7000 rpm, RT, HFA 12500 rotor), resuspended in 50 ml wash buffer (50 mM TrisCl pH 7.5, 100 mM NaCl, 1 mM benzamidine, 1 mM βME), flash frozen in liquid nitrogen and stored at -20°C. The cells lysed as they were thawing. 200 ml of lysis buffer were added and the cells were treated with a blender (Ultraturax) to reduce the viscosity by breaking down the chromosomal DNA. The suspension was spun for 20 min (12000rpm, 4°C, HFA 14290 rotor) and the pellet, containing the protein, was resuspended in 200 ml wash buffer containing 1% (v/v) Triton X-100 and allowed to stand for 10 min at RT. The insoluble material was pelleted by centrifugation (20 min, 12000rpm, RT, HFA 14290 rotor). The wash was repeated once with the same buffer and twice with wash buffer without Triton X-100. The last pellet was soaked in 1 ml DMSO for 30 min at RT. Then 50 ml unfolding buffer (6M guanidinium HCl, 20 mM NaAc pH 5.2, 1 mM DTT) was added and the suspension was stirred gently for 1 hour. Undissolved material was removed by centrifugation (RT, 12000rpm, 10 min, SS-34). The supernatant was applied to a S-200 gel filtration column (Pharmacia 5*100 cm) equilibrated with SAUDE 200 (7M urea, 20 mM NaAc pH 5.2, 200 mM NaCl, 1 mM βME and 1 mM Na/EDTA). The peak fractions as detected by the absorbance at 280 nm and 18% SDS gel electrophoresis were pooled and dialyzed against water in 8000 MWCO dialysis bags. The dialysate was lyophilized and stored at -20°C. Further purification of the protein was achieved by HPLC purification with a cation exchange column (TSK SP-5PW, Toyo Soda at 4
ml/min). The column was equilibrated in SAUDE 200. The protein was dissolved in the same buffer and approximately 15 mg were applied to the column/run. The histones were eluted with a gradient of 400 to 550 mM NaCl (40 min) and the peak fractions were pooled and dialyzed against water with 5 mM βME. Histone proteins obtained by this purification scheme were suitable for octamer refolding.

**Recombinant histone octamer**

Recombinant histone octamer was refolded from lyophilisates of individual histones using a method developed by K. Luger. Lyophilisates of the four histone proteins H2A, H2B, H3 and H4 were resuspended in 20 mM TrisCl pH 7.4, 7 M guanidinium chloride, 10 mM DTT and allowed to completely unfold for 30 min. The concentrations of the single histones were determined by absorbance at 276 nm (H2A, εM=4350 cm⁻¹M⁻¹; H2B, 7250 cm⁻¹M⁻¹; H3, 4640 cm⁻¹M⁻¹; H4, 5800 cm⁻¹M⁻¹; from K. Luger) and equimolar amounts of the histone proteins were combined (overall final concentration 1 mg/ml). This mixture was dialyzed in 6-8000 MWCO dialysis bags against three changes of refolding buffer (10 mM TrisCl pH 7.5, 2M NaCl, 1 mM EDTA, 5 mM β-mercaptoethanol) at 4° C. The refolded octamer was centrifuged in a Corex tube (20 min, 9000 rpm, 4° C, HFA 2250 rotor) and concentrated to 1 ml volume using a vacuum ultra thimble (Sartorius). The concentrated sample was purified on a Superdex 200 gel filtration column (1.6 cm x 60 cm; Pharmacia) equilibrated in refolding buffer. Octamer was monitored by absorbance at 280 nm and with SDS PAGE. Fractions containing octamer were pooled and the protein was concentrated to >2 mg/ml using a vacuum ultra thimble. Storage was at -20° C after addition of an equal volume of glycerol (86%).
2.5.3 Denaturing gel electrophoresis of histones

SDS polyacrylamide gel electrophoresis

Histone proteins were analyzed using a 18% (1:60 bisacrylamide:acrylamide), 0.75 M TrisCl pH 8.8, 0.1% SDS separating gel covered with a 5% (1:20 bisacrylamide: acrylamide), 120 mM BisTrisCl pH 6.8, 0.1% SDS stacking gel. The gel size was 10 cm x 10 cm x 0.5 mm (minigels). Samples were diluted with an equal volume of protein gel loading buffer (125 mM BisTrisCl pH 6.8, 20% glycerol, 4% SDS, 0.85M β-mercaptoethanol, with bromophenol blue dye as a marker) and boiled for 5 min before loading. Gels were run for 60 min at 10W using 50 mM TrisCl, 0.1M glycine, 0.1% SDS as running buffer (Laemmli, 1970).

Coomassie staining

After gel electrophoresis, protein bands were visualized by staining with Coomassie blue (stain: 45% ethanol, 9% acetic acid, 0.5% Coomassie blue R250) and destained by heating to 55°C in 7% ethanol, 5% acetic acid.

Triton-Urea-Acetic Acid gel electrophoresis

High resolution separation of histone proteins was achieved using Triton-Urea-Acetic Acid gels after a modification of the method of Zweidler (Lennox & Cohen, 1989; adapted by A. Flaus). The gel (60 cm x 20 cm x 0.4 mm) was prepared with a 11.7% (1:60 bisacrylamide: acrylamide), 7.3M urea, 0.38% Triton X-100, 5% acetic acid separating gel and a 5.8% (1:60 bisacrylamide:acrylamide), 7.3M urea, 0.38% Triton X-100, 5% acetic acid stacking gel. The gel was pre-run for 16 hours at 500V, then the lanes were pre-run twice with 2.5M cysteamine, 2.5M urea, 5% acetic acid and electrophoresed for 90 min at 750V, and once with 25 mg/ml protamine sulfate (Sigma), 2.5M urea, 5% acetic acid for 60 min at 500V. Samples (5-30 μM) were mixed with an equal volume of loading buffer (25 mg/ml protamine sulfate (Sigma), 5M urea, 10% acetic acid), loaded and electrophoresed for 14-18 hours at 600V. The buffer was
5% acetic acid in all cases, with the cathode at the bottom of the gel. Protein bands were visualized by Coomassie staining (see above).

2.5.4 HMG-17

The plasmid pET3a/HMG17F was constructed by Remo Amherd during his diploma work. It contained the human HMG-17 protein with a short N-terminal fusion sequence cloned into the NdeI/BamHI restriction sites of pET3a. The fusion peptide could be removed by digestion with NdeI and subsequent ligation of the plasmid.

_E. coli_ BL21 cells containing the plasmid pET3a-HMG-17 were grown in 2xTY media (100 µg/ml ampicillin and 25 µg/ml chloramphenicol) at 37° C until the optical density reached 0.5 at 600 nm. The cells were induced with 0.2 mM IPTG and grown for 2 additional hours. The cells were harvested by centrifugation (7 min, 7000 rpm, RT, HFA 12500 rotor), resuspended in 120 ml buffer (50 mM TrisCl pH 7.5, 100 mM NaCl), flash frozen in liquid nitrogen and stored at -20° C. The viscosity of the thawed cells was reduced in a blender (Ultraturax), and the cell debris was removed by centrifugation (30 min, 20000 rpm, 4° C, SS-34). The supernatant was incubated for 15 min at 70° C in a water bath. The suspension was spun as above to remove precipitated cellular proteins. To remove nucleic acids the supernatant was filtered through a filter funnel loaded with 20 ml SP-Sepharose (Pharmacia) equilibrated in the same buffer. The resin was washed with 30 ml buffer and the HMG-17 was eluted with 500 mM NaCl. The protein was diluted 5-fold and HPLC purified with a TSK SP-5PW (4 ml /min) cation exchange column equilibrated in 20 mM NaAc pH 5.2. HMG-17 eluted in a gradient between 250 and 630 mM NaCl (37 min). The protein was dialyzed against water and stored at -20° C. The concentration of the protein was measured by the absorbance at 242 nm: \[ \text{Abs}_{242} = 0.123 \] corresponds to 1 mg HMG-17/ml.
2.6 Fluorescence and chemical methods

2.6.1 DNA labeling

**Modification of the 5'-phosphate**

To attach fluorescence dyes to DNA a procedure developed to label oligonucleotides was modified (Heyduk & Lee, 1990; Heyduk & Lee, 1992). A typical reaction mix to couple the cystamine linker to the terminal 5'-phosphate of DNA contained 300 µl 100 mg/ml 1-Ethyl-3,3-dimethylaminopropylcarbodiimide (EDC, Fluka BioChemica) in H₂O, 150 µl DNA (max. 200 µg), 92.2 µl 1 M 1-Methylimidazole, pH 7 (Fluka puriss p. a.) and 460 µl cystamine dihydrochloride 1M, pH 7.1 (ox. form, Fluka purum). The reaction mixture was incubated at 50° C in a water bath. After 3 h the reaction was quenched by the addition of TE pH 8 buffer to a final volume of 10 ml. The DNA was then purified from the reactants by ion exchange chromatography using 1 ml Q Sepharose Fast Flow (Pharmacia). The DNA was ethanol precipitated twice and stored in TE buffer.

**Coupling the modified DNA to the fluorescence dye**

A solution containing a maximum of 150 µg DNA modified DNA in 100 µl TE was reduced with 5 µl 0.4 M dithiothreitol (DTT) for 1.5 h at RT. The DTT was removed by Sephadex G25 (Pharmacia medium grade, in water) spun column purification directly into 10 µl of 10 mM fluorescence dye in DMF.

The labeling reaction was carried out for 1 h at RT in the dark. 2 µl 0.4 M DTT were added to quench the reaction. The DNA was ethanol precipitated twice and stored in TE buffer (-20° C, dark). The labeled DNA was protected from light wherever possible. The labeling efficiency was determined by measuring the absorbance of the fluorescence dye attached to the DNA. The fluorescence dyes used were CPM.
(ε_{CPM/387nm}=30000 \text{ M}^{-1}\text{cm}^{-1}), \text{ FLU} \ (ε_{FLU/494nm}=70800 \text{ M}^{-1}\text{cm}^{-1}) \text{ and} \text{ RHO(ε_{RHO/542nm}= 90000 M^{-1}cm^{-1}) obtained from Molecular Probes, Eugene.}

**Labeling of symmetric 146mer DNA near the center**

Firstly, the 73mer DNA was digested with VentR® polymerase: 500 µg 73mer DNA were incubated in a volume of 1 ml containing 20 mM TrisCl pH 8.8, 10 mM KCl, 10 mM (NH₄)SO₄, 7 mM MgSO₄, 0.1% Triton X-100, 4 mM dTTP and 20 units VentR polymerase. After 240 min at 72° C, the digestion mix was extracted twice with phenol/CIA, once with CIA, and ethanol precipitated twice from a total volume of 10 ml. The DNA was stored in TE at -20° C.

Secondly, the DNA was further digested with T4 polymerase (New England Biolabs): 500 µg of the above material was digested in a volume of 0.5 ml containing 10 mM TrisCl pH 7.9, 50 mM NaCl, 10 mM MgCl₂, 1 mM DTT, 4 mM dTTP and 7.5 units T4 polymerase. After 50 min at 37° C, the digestion was stopped as after the VentR polymerase digest.

The progress of the digestion reactions was followed on denaturing gels (8%, 20:1 acrylamide:bisacrylamide, 7 M urea, 1xTBE) stained with Ethidium bromide.

**Preparation of the labeled oligonucleotide**

The phosphorothioate containing oligonucleotides 5'-CATC(ps)AAGCA-3' and 5'-CAT(ps)CAAGCA-3' were synthesized on an AB 380B DNA synthesizer according to standard protocols (Beaton *et al.*, 1991; Zon & Stec, 1991). The oligonucleotides were purified on RP HPLC on a Nucleosil C₈ column (Zon, 1990).

55 nmol of the oligonucleotide were kinased in 200 µl containing 70 mM TrisCl pH 7.6, 10 mM MgCl₂, 5 mM DTT, 1.5 mM ATP (Pharmacia) and 50 units T4 polynucleotide kinase (New England Biolabs). After 4 h at 37° C the kinasing mix was
lyophilized. The oligonucleotide was dissolved in 40 μl TE/1M NaCl and precipitated with 120 μl ethanol. The oligonucleotide was stored in water at -20° C.

The oligonucleotide was then labeled with 5-iodoacetamidofluorescein (5-IAF, Molecular Probes Eugene) according to a published method (Conway et al., 1991): 50 nmol oligonucleotide in 200 μl 30% DMF/25 mM TrisCl pH 8 and 15 mM 5-IAF were incubated at 50° C for 24 h in the dark. The solvent was removed in a Speedvac and the pellet was resuspended in 70 μl 1M NaCl and 210 μl ethanol. The oligonucleotide was pelleted by centrifugation for 20 min. at 4° C. The pellet was then resuspended in 400 μl TE. After another 5 min spin the supernatant with the labeled oligonucleotide was stored at -20° C in the dark.

Ligation of the full length DNA

The recessed 73mer DNA was ligated with the labeled phosphorothioate oligonucleotide to give the full length symmetric 146mer DNA fragment. The following ligation mix was used to ligate 800 μg DNA: 2 ml with 50 mM TrisCl pH 7.8, 10 mM MgCl₂, 10 mM DTT, 1 mM ATP, 800 μg recessed 73mer DNA, 25 nmol labeled oligonucleotide and 4000 units T4 ligase (New England Biolabs). The ligation was performed at 20° C. After 24 h, 6 nmol oligonucleotide and 1000 units T4 ligase were added, and the ligation was continued for another 8 h. The DNA was ethanol precipitated from a total volume of 14 ml. The DNA pellet was resuspended in 300 μl TE.

The centrally labeled symmetric DNA was purified with preparative PAGE (10 W, 20 ml of 8% 30:1 acrylamide:bisacrylamide, 0.5 TBE as running and elution buffer). The DNA in the peak fractions was ethanol precipitated and stored in TE buffer in the dark.
Chapter 2

Materials and Methods

2.6.2 Octamer labeling

Histone octamer containing histone mutants with accessible cysteine residues was labeled as NCP. 100 μl of NCP (containing a maximum of 200 μg NCP) were reduced with 5 μl 0.4 M DTT for 1h at 4° C. The DTT was removed on a spun column (Sephadex G25 medium in 20 mM TrisCl pH 7.5, 1 mM NaEDTA), and the NCP was directly spun into 10 μl of chromophore (FLU, RHO, 10 mM in 5% DMF/ 20 mM TrisCl pH 7.5, 1 mM NaEDTA). The labeling reaction was carried out for 1 h at 20° C in the dark. Free chromophore was removed by a second spun column and dialysis over night against NCP storage buffer. The labeling efficiency was determined by measuring the absorption of the chromophore attached to the NCP. Labeled NCP was analyzed with 18% SDS PAGE to check the labeling efficiency.

2.6.3 NCP acetylation

The core particle to be acetylated was dialyzed against 20 mM Cacodylate pH 6 and 1 mM EDTA. The acetylating reagent N-Acetyl sulfo succimide was prepared in the form of the sodium salt according to a procedure described earlier (Lewis et al., 1988; Staros, 1982). The degree of acetylation was mainly controlled by the pH and the molar ratio of reagent to NCP. A typical acetylation reaction contained 15 μl (32.4 μg) NCP, 5 μl 100 mM Na2HPO4 and 2.2 μl N-Acetyl sulfo succimide 10 mg/ml for the lower degree of acetylation (4.3 μl 20 mg/ml for the higher acetylation). After 1 h at 4 °C the reaction was quenched by the addition of 1 μl 0.5 M Tris-HCl pH 7.5. Larger scale acetylations were dialyzed against NCP storage buffer. The degree of acetylation was analyzed on a Triton-urea-acid gel (Zweidler, 1978).
2.7 Chromatin (-characterization) methods

2.7.1 Salt dialysis reconstitution

Nucleosome core particles were reconstituted by dialysis against a decreasing salt gradient (Richmond et al., 1988) starting from 400 ml 20 mM TrisCl pH 7.5, 1 mM K/EDTA, 1 mM DTT and 2M KCl. DNA and histone octamer were mixed in this buffer in a molar ratio of 1.1:1 at a DNA concentration of 0.5 mg/ml. Reconstitution volumes ranged from 25 μl in dialysis buttons to several ml in 14000 MWCO dialysis tubing. Under constant stirring buffer was continuously pumped out and replaced by adding 1600 ml 250 ml KCl (in the same buffer) at the same rate over a time of 36h. After centrifugation of the dialysate the NCP was purified either by anion exchange HPLC or preparative PAGE. For HPLC purification a TSK DEAE 5PW column at a flow rate of 4 ml/min was equilibrated in 10 mM TrisCl pH 7.5, 250 mM KCl, 0.1 mM K/EDTA and 1 mM DTT. The NCP was eluted with a gradient between 300 and 450 mM KCl (17 min) in the same buffer. The peak fractions containing the NCP were dialyzed against NCP storage buffer for 2 h and then concentrated to a concentration of 2-10 mg NCP/ml. Preparative PAGE gels (20 ml of a 5% 60:1 acrylamide:bisacrylamide) were run in the cold room at 10 W with 0.25 TBE as running- and NCP storage buffer (20 mM TrisCl pH 7.5, 1 mM EDTA, 1 mM DTT) as elution buffer. The fractions containing NCP were immediately concentrated with mini concentrators (Amicon, centicon 10) to a concentration of at least 2 mg NCP/ml. NCP for crystallization was dialyzed against 20 mM potassium/cacodylate pH 6, 1 mM K/EDTA. Reconstitution and purification was done at 4° C and unless otherwise stated. Purified core particles were incubated for 2h at 37° C before further use.
2.7.2 Native gel electrophoresis

Native (non-denaturing) polyacrylamide gels of nucleosomal complexes were run in 14 well 20 cm x 20 cm x 1 mm 5% (1:60 bisacrylamide:acrylamide) 0.2-0.25 x TBE gels using recirculated TBE running buffer at 4° C. The gels were temperature-equilibrated for at least 3 hours at 4° C and pre-run for 3 hours at 250V. Samples were loaded after addition of sucrose to 5%, and were electrophoresed for 3 hours. For visualization, gels were soaked in Ethidium bromide solution.

2.7.3 Fluorescence measurements

Equilibrium measurements were done on a SPEX FL112 fluorimeter. Unless stated otherwise, the experiment temperature was 20° C as measured in the sample holder. The emission signal was corrected automatically for lamp intensity fluctuations using an internal rhodamine reference. The excitation wavelength was set to 385 nm and 490 nm for the donor chromophores CPM or FLU/5-IAF, respectively. Quantitative Fluorescence Resonance Energy Transfer (FRET) measurements were done by measuring the donor emission at 483 nm for CPM (520 nm for FLU and 517 nm for 5-IAF) thus monitoring the quenching of the donor emission by the acceptor at different salt concentrations. Qualitative spectra were obtained with an emission scan. Starting sample volumes for these measurements were 1 or 2 ml with stirring during the entire measurement.

The conditions for the DNA end to end and DNA end to protein measurements (first transition) were as follows: the sample concentration at the start was 0.42 μM (85 μg NCP/ml) core particle ('cold' and labeled) in 10 mM buffer TrisCl pH 7.5, 50 mM NaCl, 0.1 mM NaEDTA (and 1 mM DTT for globular and recombinant NCP). The sample was allowed to equilibrate in the sample holder for an hour prior to the
measurements. NaCl was added from a 5 M stock solution in the appropriate buffer. The sample was incubated for one minute before starting the measurements.

For the dimer-tetramer dissociation measurements and the center DNA to H4 measurements, the starting NaCl concentrations were between 400 and 600 mM NaCl, 0.05% Triton-100 was added in some cases. NaCl was added from a 5 M stock, but the incubation time before data collection was between 1 and 2 hours. As a control (center to H4 measurements), the individual salt concentrations were prepared separately in a volume of 400 µl. The donor emission was measured after equilibration at 14 h at RT/dark. The NCP concentration was 0.42 µM (85 µg NCP/ml).

Data processing was done on a Macintosh computer. Microsoft Excel was used to correct the fluorescence emission intensity for the change in volume and for concentration adjustments.

The formula used to correct for donor emission changes in the absence of acceptor (eliminates all other effects such as salt and solvent effects as well) was as follows: 

\[ em_{corr} = em_{DA}[1 + (EMD-emD)/emD] \]

where \( em_{DA} \) is the emission of the donor-acceptor particle at a given NaCl concentration, \( emD \) the emission of the donor particle at the same NaCl concentration and \( EM_D \) the emission of the donor particle at 50 mM NaCl (any concentration could have been taken as reference).

## 2.8 Crystallization, X-ray work

### 2.8.1 General

Protein solutions were at concentrations of 2 to 8 mg of protein per ml. The material to be crystallized was spun (10 min, 4° C, 13000 rpm, benchtop) immediately before setting up crystals. If not stated otherwise all crystallization was done at constant 20° C. All solutions used for crystallization were prepared from chemicals of the
Chapter 2 Materials and Methods

highest obtainable quality using Milli-Q water. All solutions were filtered through 0.2 or 0.45 μm filters (Sartorius).

Dry cover slides were siliconized in 2% dimethyldichlorosilane in 1,1,1-Trichloroethane for 5 min. and subsequently dried in the oven at 60° C for 5 min. The cover slides were washed twice with ethanol and once with water.

Preparation of liquid propane: propane gas was condensed by blowing it into a brass tube which was closed at the bottom end and submerged in liquid nitrogen. The condensed liquid propane was transferred into 50 polypropylene tubes (with 2 holes in the cap), allowed to freeze in liquid nitrogen and then stored submerged in liquid nitrogen.

2.8.2 Crystallization of the Nucleosome Core Particle

The vapor-diffusion method was used for hanging- as well as for sitting drops. For hanging drops equal volumes (total 3-8 μl) of protein and precipitant were mixed on a siliconized cover slide. The reservoir solution was mixed in the wells of a Falcon 6x4 multiwell plate. The wells were sealed with silicon grease. Sitting drops were prepared as above except that the drop (14 μl) was sitting on a Micro Bridge.

Generally the method first described by (Rhodes et al., 1989) was used, though the exact conditions were dependent on the type of octamer, DNA, the purification batch and the protein concentration. NCP crystals were grown from solutions containing 3-4 mg NCP/ml, 65 -85 mM MnCl2, 50-80 mM KCl, 20 mM KCacodylate pH 6 equilibrated against a reservoir solution of half the salt concentration.
2.8.3 Crystal Handling

NCP crystals were manipulated and stored in glass dishes at 20° C. Crystals were routinely transferred into 24% MPD using the scheme summarized in table 2.6.

<table>
<thead>
<tr>
<th>Time</th>
<th>buffer</th>
<th>MPD conc. (% v/v)</th>
<th>Trehalose conc. (% w/v)</th>
<th>remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 min</td>
<td>HL-3.5</td>
<td>3.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 min</td>
<td>HL-7</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 min</td>
<td>HL-12.1</td>
<td>12.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 min</td>
<td>HL-18</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 min</td>
<td>HL-21</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 min</td>
<td>HL-24</td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥12h</td>
<td>HL-24/TL-2 24</td>
<td>2</td>
<td></td>
<td>for loop mounting</td>
</tr>
<tr>
<td>+ 10 min</td>
<td>HL-24/TL-5 24</td>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.6: Crystal soaking.

HL-0 is the crystal storage buffer (40 mM KCl, 37 mM MnCl₂, 10 mM KCacodylate pH 6.0; without MPD) used to harvest crystals. Soaking was done by exchanging the liquid surrounding the crystals with buffer of the new composition.

For data collection at 4° C and 20° C, crystals were mounted in a small buffer volume sticking to the side of a 0.7 mm glass capillary sealed with buffer on both sides of the crystal (1 mm away). For data collection at cryo-temperatures, either tapered capillaries (described elsewhere) or fiber loops were used. Fiber loops were made from a single fiber of an ordinary packing cord (diameter of 10 μm) glued into a 0.2 mm glass capillary (Cyanolite fast glue).

Frozen crystals were stored and transported in liquid nitrogen containers. Freezing and crystal storage was done using locally developed equipment (D. Sargent).
2.8.4 Data collection

Data collection was generally at cryo temperature. Characterization of crystals was either performed on a Huber camera (oscillation or precession) on film or using a MAResearch imaging plate. Data collection was exclusively done using imaging plates.

Exposure times ranged from 30 min to 6 h (depending on generator shape and crystal size) on the rotating anodes and from 6-90 sec using synchrotron radiation.

2.9 Computing

Analysis and manipulation of DNA sequences in this work was carried out using the University of Wisconsin Genetics Computer Group (GCG) package, versions 7.0, 8.0 (Devereux et al., 1984). Database searches were performed on Genbank and EMBL using the SRSWWW engine (http://www.ch.embnet.org/srs/srsc).

Apple Macintosh computers with common software were used for PC work. Data processing and model building were done on various DEC Alpha (VMS 6.1/Unix) and Silicon Graphics (Indigo2, Unix IRIX 5.3) machines, respectively. The software used for this purpose is listed in table 2.7. The use of the programs is described in the manuals, descriptions and tutorial files supplied with the programs.
Table 2.7: Summary of used software.

Protein, protein-DNA and DNA structures were retrieved as coordinate files from the Brookhaven protein data bank.

<table>
<thead>
<tr>
<th>Name</th>
<th>Function</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>XRRAW</td>
<td>data collection (labor.)</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>MAR-software</td>
<td>data collection (synchrotron)</td>
<td>MAReasearch</td>
</tr>
<tr>
<td>IPTORAW</td>
<td>oscillation photographs file conversion</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>RAW_SCALE</td>
<td>oscill. photograph scaling</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>HKL_HKLGEN</td>
<td>hkl generation</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>OSCAI_IP</td>
<td>data processing</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>OSCREF_IP</td>
<td>data processing</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>OSCKRUNCH</td>
<td>data processing</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>OSCMERGE</td>
<td>data processing</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>DIM</td>
<td>oscillation photographs display</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>XPLOR (3.1; 3.4)</td>
<td>structure refinement and analysis</td>
<td>(Brünger, 1992)</td>
</tr>
<tr>
<td>HKL_EDIT</td>
<td>data set editing</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>HKL_FFT</td>
<td>map calculation</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>HKL_LIST</td>
<td>data display</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>HKL_MERGE</td>
<td>data file merging</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>HA_LSCALE</td>
<td>local scaling of data</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>HKL_TO_XPL</td>
<td>data set editing</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>MAP_TO_OMAP</td>
<td>map conversion</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>HKL_SORT</td>
<td>data set editing</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>ATM_SF</td>
<td>structure factor calculation</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>ATM_RMHD</td>
<td>PDB file editing</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>PDBOTX</td>
<td>PDB file editing</td>
<td>T. J. Richmond (unpubl.)</td>
</tr>
<tr>
<td>Ono 6.0 (o)</td>
<td>map display, model building and graphics</td>
<td>(Jones et al., 1991)</td>
</tr>
<tr>
<td>Brix</td>
<td>map calculation</td>
<td>(Jones, 1992)</td>
</tr>
<tr>
<td>moleman2</td>
<td>utilities</td>
<td>(Kleyweg &amp; Jones, 1993-96)</td>
</tr>
<tr>
<td>LSQMAN</td>
<td>superposition</td>
<td>(Kleyweg &amp; Jones, 1994)</td>
</tr>
<tr>
<td>CURVES 5.1</td>
<td>DNA analysis</td>
<td>(Ravishankar et al., 1989)</td>
</tr>
<tr>
<td>MidasPlus</td>
<td>graphics</td>
<td>(Ferrin et al., 1988)</td>
</tr>
<tr>
<td>GRASP 1.2</td>
<td>surface display</td>
<td>(Nicholls et al., 1991)</td>
</tr>
</tbody>
</table>
3. Fluorescence energy transfer

3.1 Introduction

For a profound understanding of the function of large multisubunit complexes, characterization of the dynamic behavior of a macromolecular assembly is often crucial for understanding its structure-function relationship.

The assembly and disassembly of chromatin in general and of the nucleosome in particular is highly salt dependent. Therefore, most experiments have dealt with the effect of the ionic strength on nucleosome structure. Experiments were usually done in the salt range of 0-2 M of monovalent ions. At 2 M salt, the NCP is dissociated into histone octamer and free DNA.

Protein octamer obtained in this way was used to characterize its properties in experiments using gel filtration (Eickbush & Moudrianakis, 1978) and analytical ultracentrifugation (Stacks & Schumaker, 1979). The histone octamer was stable only above 2 M NaCl and was destabilized by increasing temperature, decreasing NaCl concentration, pH below 7 or above 10 and low concentrations of urea or guanidinium chloride. The standard free enthalpies for the cooperative dimer association with tetramer are -30.8 and -22.5 kcal respectively, as determined by calorimetric measurements (Benedict et al., 1984).

Most experiments with intact nucleosomes were done at lower ionic strength (0-1M NaCl). The ionic strength needed for dissociation of the DNA from the octamer was determined in several different ways. Sedimentation experiments showed increasing amounts of free DNA versus NCP already at 0.6-1M NaCl (Ausio et al., 1984; Stacks & Schumaker, 1979; Yager & van Holde, 1984). Circular dichroism (CD) and intrinsic tyrosine fluorescence resulted in slightly different results. An unidentified transition was found at 0.55 M NaCl (Oohara & Wada, 1987a; Oohara & Wada, 1987b) and the
dissociation of the dimers from the H3/H4 tetramer-DNA complex was assigned to a transition around 1M NaCl. Finally, the DNA dissociated from the tetramer at approximately 1.5 M NaCl.

In order to obtain more specific information, extrinsic chromophores were used to monitor the disassembly of the nucleosome. One such study measured the distance between fluorescence labels on Cys110 of H3 and the DNA termini in core particles using Fluorescence Resonance Energy Transfer (Eshaghpour et al., 1980). According to these measurements, the distance between Cys110 and the 3' ends of the DNA was identical at 0.32 mM and 0.6 M NaCl which would argue against dissociation of DNA at 0.6 M NaCl already. Another study showed a sharp transition between 0.4 and 0.6 M NaCl as observed with NCP labeled with NEM on the same Cys110 (Daban & Cantor, 1982). The transition was assigned to a structural change in the octamer rather than to a change in histone-DNA interactions. The interpretation of these results is somewhat questionable, as it is not clear whether octamer with a bulky label on H3 Cys110 forms native-like core particles. Another group used NMR spectroscopy to obtain information about the dissociation of the basic histone tails which were thought to be somehow linked to core DNA dissociation (Walker, 1984). The basic histone tails dissociated from core chromatin at 400 mM NaCl which could be indicative for DNA dissociation already in this salt range.

Acetylation or removal of the basic histone tails provided further insights into their function in DNA binding. Hyperacetylation with up to 17 acetylations per NCP (Ausio & Holde, 1986) or complete removal of the histone tails with trypsin (Ausio et al., 1989) did not alter NCP structure to a significant degree—the unfolding properties between 0.3 and 0.8 M NaCl measured by CD were invariant. However the thermal denaturation profile of the DNA of highly acetylated and trypsinized NCP was destabilized compared to native NCP. These findings were confirmed by hydroxyl radical footprints of trypsinized core particles (Hayes et al., 1991) which showed no signs for a change in DNA structure. In contrast, two recent studies (Lee et al., 1993;
Vettese-Dadey et al., 1994) showed that removal or acetylation of the basic histone tails resulted in facilitated binding of transcription factors such as GAL4-AH and TFIID to their cognate sequences organized in nucleosome core particles.

The methods described so far almost never gave information about defined parts of the nucleosome core particle, because the results from CD, intrinsic fluorescence measurements and other methods are generally difficult to link to a particular structural change in the core particle. In the use of extrinsic chromophores, care has to be taken not to disturb NCP structure during the labeling procedure. Labeling histone H3 Cys110 with bulky chromophores has been the only possibility due to the lack of other cysteine amino acids, except for rare variants. As this cysteine is buried in the tetramer interface and not accessible to solution in intact nucleosomes, labeled NCP may not be completely native (Bode & Wagner, 1980).

The idea, however, of monitoring the assembly processes of multisubunit complexes with the help of FRET is still very attractive. In our experiments, we were not interested in absolute distances between sites, those can be determined much more accurately using X-ray crystallographic methods. We focused on the relative change in distances as an indicator for dis/assembly steps. Fortunately, various cysteine mutants of histone proteins were available in our laboratory, produced mainly for the purpose of MIR phasing of the NCP. Using the specific labeling sites available on the histone octamer in combination with labeled DNA fragments allowed the design of completely new donor-acceptor arrangements for FRET measurements (see table 3.2 and appendices III/IV).

Whereas NCP characterization done in the past has focused mainly on the histone octamer part of the core particle, we were more interested in the dynamic behavior of the DNA. Therefore, we focused on measuring the movement of the DNA termini, as well as its center, relative to the histone octamer in increasing ionic strength. These measurements were done with different histone octamers, either isolated from chicken erythrocytes or using recombinant *Xenopus laevis* histones expressed in *E. coli*. 
Depletion or partial acetylation of the basic histone tails provided more information about their function in DNA binding. Combination of this data with dissociation measurements between tetramer and dimer completed the image of the NCP disassembly as a function for the ionic strength. The measurements provide a more detailed understanding of salt-dependent NCP disassembly than previously available.

3.2 The theory of FRET

Fluorophores are molecules that can emit light immediately after absorption of light of shorter wavelength. An absorbed photon excites the chromophore from its electronic ground state $S_0$ to a vibrational level of an excited state $S_1$ or higher. Vibrational energy is then lost in picoseconds, and the molecule drops to the lowest vibrational level of the excited state $S_1$ (see Meer et al., 1994 and Ref. therein). Emission of fluorescent light is paired with a drop from $S_1$ to any vibrational level of $S_0$ in the nanosecond time scale. This can be summarized in

$$D + h\nu_E \rightarrow D^* \xrightarrow{k_f} D + h\nu_F \xrightarrow{k_i} D$$

(3.1)

where $D$ is the chromophore in the ground state, $D^*$ the excited chromophore, $h$ Planck's constant, $\nu_E$ and $\nu_F$ the excitation and emission frequencies, $k_f$ the radiative decay constant and $k_i$ the summation over all nonradiative decay constants. The quantum yield $Q$ is the ratio between emitted photons and absorbed photons and is defined as

$$Q = \frac{k_f}{k_f + k_i}.$$  

(3.2)

The lifetime $\tau$ of the excited state is

$$\tau = \frac{1}{k_f + k_i}.$$  

(3.3)

The absorption and emission of photons by a chromophore is oriented along absorption and emission transition moments (dipole moments) and thus not uniform in space. As a
consequence when fluorophores are illuminated with polarized light, the emitted fluorescence will generally be polarized as well. The polarization $P$ (or anisotropy $r$) are then expressed as

$$P = \frac{I_V - I_H}{I_V + I_H}, \quad r = \frac{I_V - I_H}{I_V + 2I_H}$$

(3.4)

where $I_V$ is the fluorescence emission intensity with parallel excitation and emission polarizers (usually vertical) and $I_H$ with the emission polarizer perpendicular (usually horizontal). $P$ is a measure for the rotational freedom of the probe. Absorption and emission transition moments have a defined angle for a given molecule (at a given wavelength).

FRET is radiationless transfer by dipole-dipole interactions between two chromophores. There are three other possible transfer mechanisms - reabsorption, complex formation and collisional quenching. For FRET to occur the donor chromophore emission spectrum must overlap with the acceptor chromophore absorption spectrum and the distance between the chromophores must be in the 10-100 Å range. In addition the dipole moments of the two chromophores must be favorably oriented. In a system with both donor and acceptor chromophores, the efficiency of transfer $E$ is defined as

$$E = \frac{k_T}{k_T + k_D + k_{Di}}$$

(3.5)

where $k_T$ is the radiationless donor-acceptor transfer constant, $k_D$ is the emission decay constant of the donor, and $k_{Di}$ is the decay constant for the remaining nonradiative decay of the donor. $E$ can be obtained by measuring the quantum yield of the donor in presence ($Q_{DA}$) and absence ($Q_D$) of the acceptor chromophore:

$$E = 1 - \frac{Q_{DA}}{Q_D} \quad \text{and} \quad Q_{DA} = \frac{k_D}{k_T + k_D + k_{Di}}, \quad Q_D = \frac{k_D}{k_D + k_{Di}}.$$

(3.6)

The transfer rate constant $k_T$ is defined as follows (Förster):

$$k_T = (k_D + k_{Di}) \left( \frac{R_o}{R} \right)^6 = \frac{1}{\tau_D} \left( \frac{R_o}{R} \right)^6,$$

(3.7)
where $\tau_D$ is the donor excited lifetime in the absence of acceptor, $R_o$ is the Förster distance at which half the photons are transferred ($k_t = k_D + k_D$). Substituting (3.7) into (3.5) gives

$$E = \frac{R_o^6}{R_o^6 + R^6}.$$  

(3.8)

This function is shown in Fig. 3.1. It becomes immediately clear that $E$ changes rapidly only when the distance $R$ between the chromophores is approximately equal to $R_o$. $R_o$ has been defined by Förster as

$$R_o^6 = \frac{9000(ln 10)\kappa^2 Q_d J}{128\pi^2 n^4 N_A V}.$$  

(3.9)

With

- $Q_d$ = donor quantum yield in the absence of acceptor
- $n$ = index of refraction of the medium between the chromophores (approx. 1.4 for biological samples)
- $N_A V$ = Avogadro's number
- $\kappa^2$ = orientation factor
- $J$ = overlap integral between donor emission and acceptor absorption spectra.

$\kappa^2$ defines the orientation between the donor and acceptor dipoles:

$$\kappa^2 = (\cos \theta_T - 3 \cos \theta_D \cos \theta_A)$$  

(3.10)
with $\theta_T$ being the angle between the donor/emission transition moments and $\theta_D$, $\theta_A$ the angles between the donor-acceptor connection line and the respective transition moments. $\kappa^2$ can take values between 0-4 depending on the rotational freedom of the chromophores. It is often assumed to be $2/3$ for fast isotropically rotating donor and acceptor molecules.

For our purpose, the exact values of $Q_D$, $\kappa^2$ and $J$ are not of great importance as we were not interested in absolute distances. However, there are considerations for our experiments due to equation (3.8). Significant changes in transfer can only be expected when the sampled distance $R$ is in the range of 0.5 to 1.5 $R_0$. $R_0$ is roughly constant for a given donor acceptor pair. This has to be considered when selecting chromophore pairs and attachment sites on the macromolecular assembly.

### 3.3 Results

#### 3.3.1 Labeling of DNA and histone proteins

The chromophores

For our experiments, we have chosen cysteine reactive chromophores because the mutants prepared for MIR phasing were cysteine mutants. There are two reactive groups that are commonly used for labeling cysteines: maleimides and haloacetyl derivatives such as iodoacetamides. A summary of the chromophores used in this work is given in table 3.1. As the maleimides react somewhat more specifically, we mainly used those for labeling cysteines introduced into the histone proteins and 5' phosphates of DNA. A iodoacetamide derivative had to be used for labeling phosphorothioates, because maleimides are not reactive enough.
**Chapter 3**

**Fluorescence energy transfer**

<table>
<thead>
<tr>
<th>Chromophore</th>
<th>Reactive group</th>
<th>Use as</th>
<th>Excitation/emission wavelength</th>
<th>Labeled sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPM</td>
<td>Maleimide</td>
<td>donor</td>
<td>385/483 nm</td>
<td>DNA 5'</td>
</tr>
<tr>
<td>FLU</td>
<td>Maleimide</td>
<td>donor/acceptor</td>
<td>490/520 nm</td>
<td>DNA 5', protein Cys</td>
</tr>
<tr>
<td>RHO</td>
<td>Maleimide</td>
<td>acceptor</td>
<td>542/570 nm</td>
<td>DNA 5', protein Cys</td>
</tr>
<tr>
<td>5-IAF</td>
<td>Iodoacetamide</td>
<td>donor</td>
<td>490/517 nm</td>
<td>DNA phosphorothioate</td>
</tr>
</tbody>
</table>

Table 3.1: List of chromophores.

The FRET donor acceptor pairs used were CPM-FLU, FLU-RHO and 5-IAF-RHO with Ro around 50 Å (Meer et al., 1994).

Labeling the DNA termini

As a target site for end-labeling DNA, the 5' phosphate was favored over other possibilities such as the use of modified bases because it made the labeling independent of the DNA length restriction imposed by oligonucleotide synthesis. Any DNA sequence of any given length can be labeled by this method. Thereby, the chromophore was coupled via a cystamine linker to the phosphate which resulted in highly flexible attachment. An overview of the chemistry involved in the reaction is shown.

**Figure 3.2: Reaction scheme for 5' phosphate labeling.**

(I): Modification of the 5' DNA phosphate with a (oxidized) cystamine linker. The reaction is driven by the water abstraction through the corbodiimide derivative EDC. (II): After reduction of the disulfide with DTT, the free thiol group was reacted with the maleimide group of the chromophore.
Chapter 3  
Fluorescence energy transfer

I

\[
\begin{align*}
\text{DNA} & \quad + \quad \text{O-Methylimidazol, pH 7} \\
\text{50°C} & \\
\text{H}_2\text{N-S-S-} & \quad \text{NH}_2 \\
+ & \quad \text{EDC} \\
\end{align*}
\]

II

\[
\begin{align*}
\text{Chrom.} & \quad \text{DNA-SH} \\
\text{DNA} & \quad \text{Chrom.} \\
\text{H}^+ & \quad \text{DNA-S} \\
\end{align*}
\]
in figure 3.2. In the first step, one of the free phosphate oxygen was replaced by an amine in a reaction mediated by the carbodiimide derivative EDC (scheme I). The disulfide was then reduced with DTT, and the resulting thiol was reacted (scheme II) with the maleimide reactive group of the chromophore (Heyduk & Lee, 1990; Heyduk & Lee, 1992). The labeling was easy to perform with an efficiency higher than 95%.

When asymmetric DNA had to be labeled with different chromophores on each of the 5' phosphates, the strands were first separated on DEAE HPLC at high pH. The above labeling reaction was then done on single-stranded DNA. Reannealing gave double-stranded DNA with different chromophores on each end. Lanes 3 to 5 in figure 3.3 show single-stranded DNA after labeling. Formation of hairpins made them run higher than as146. Reannealing of the single-strands restored ds 146mer (lane 2) of similar electrophoretic mobility compared to asymmetric DNA labeled as a double-strand (lane 1).

![Figure 3.3: DNA gel of labeled asymmetric DNA (as146):](image)

Native 10% PAGE of labeled DNA fragments. The left side is stained with Ethidium bromide whereas on the right side the chromophores are visualized directly with UV light of 254 nm. The lanes are: lane 1: as146 labeled with RHO; lane 2: reannealed as146 with different labels on each end; lane 3: strand A, FLU labeled (top strand of the sequence, see Appendix 1); lane 4: strand B/RHO (bottom strand) and lane 5: strand B labeled with CPM. The DNA size marker has fragments of 540, 489, 404, 350, 242, 190, 147 and 110 bp length.
Labeling symmetric DNA near the dyad

For labeling the DNA center, an entirely different approach had to be used because of the lack of free phosphates within DNA. Again the attachment to the phosphate backbone was favored over the use of modified bases. In this way, the distortions of the DNA could be kept to a minimum. An overview of the labeling reaction for the symmetric DNA fragment (s146) is shown in figure 3.4. First a phosphorothioate containing oligonucleotide with one of the non-bridging phosphate-oxygen replaced by sulfur, was synthesized according to standard procedures (Beaton et al., 1991; Zon & Stec, 1991). As the synthesis is not stereoselective, the end product consists of two diastereomers. The chromophore (iodoacetamide derivative) was then attached to the phosphorothioate site on the DNA backbone in the oligonucleotide (Conway et al., 1991).

Symmetric fragments used were generally made through self-ligation of a DNA fragment with a sticky-end restriction site on one end (i.e. HindIII) and a dephosphorylated blunt end (EcoRV) at the other end. This has to be done because large palindromic sequences can not be grown in bacteria. To make room for the labeled oligonucleotide, the bottom strand of a 73mer half-side was selectively shortened using the 3’-5’ exonuclease activities of Vent and T4 DNA polymerase in the presence of the nucleotide triphosphate at which the polymerases were supposed to stop. Ligation of the labeled oligonucleotide with the recessed 73mer gave a 146 bp center-labeled symmetric DNA fragment.
Figure 3.4: Labeling of symmetric DNA (s146) near the dyad.

Figure 3.5 shows center-labeled, symmetric DNA. The DNA is visualized by direct fluorescence of the chromophores on the oligonucleotide. In the course of the ligation, labeled oligonucleotide was first incorporated into the recessed 73mer (A, lane 2). In a second step, labeled 73mer was ligated to the full length symmetric 146mer (A, lane 3). After purification of the labeled symmetric 146mer DNA on preparative PAGE, only a small fraction of unligated 73mer remained (figure 3.5 B, lane 1).
Chapter 3 Fluorescence energy transfer

**Figure 3.5: Center-labeled symmetric DNA.**

(A): Native 10% PAGE. End-labeled asymmetric DNA as marker in lane 1, ligation mix in lane 2 and purified labeled symmetric DNA in lane 3. (B) Denaturing 8% PAGE of purified center-labeled symmetric DNA in lane 1 and end-labeled asymmetric DNA as marker in lane 2. Visualization was by direct fluorescence of the chromophores.

**Labeling of histone proteins**

Cysteine amino acids in the histone proteins were labeled using the same approach and technique as for DNA end-labeling. The maleimide reactive group of the chromophores was directly reacted with the reduced cysteine in the protein. The labeling reaction was done on intact NCP for several reasons. Firstly, NCP is much easier to handle than histone octamer at 2 M NaCl. Secondly, it is very unlikely that NCP labeled in a native conformation suffers from structural changes due to the modification. Figure 3.6 shows a SDS gel of NCP after labeling with RHO on H4. The labeling reaction was very specific to the Cys47 of histone H4 and no significant labeling of the other core histones could be detected (lane 1).
Figure 3.6: SDS PAGE of labeled NCP.
Coomassie blue stain on the left and direct visualization with UV light on the right side. Lane 1 contains NCP with a Cys47 on H4 labeled with RHO. Lane 2 shows unlabeled NCP whereas NCP with labeled H2A Cys119 is shown in lane 3.

However, labeling of histones in the octamer at 2 M NaCl worked as well, but the high salt concentration made it difficult to remove unreacted fluorescence dyes due to solubility problems. Because of these problems, double-labeling of histone octamer with different chromophores for example on histone H2A and H4 involved the following steps: NCP with only one of the two mutant histone proteins was labeled with the appropriate chromophore. The two labeled core particles were then mixed in 2 M NaCl and reconstituted again. This ensured a high labeling efficiency at the correct site.

Unlike other authors who used the combined urea/salt method for reconstitution (Eshaghpour et al., 1980), we were not able to reconstitute histone octamer with a fluorescence dye attached to H3 Cys110 into core particles. In agreement with another study (Bode & Wagner, 1980), Cys110 was not accessible to fluorescence dyes in intact NCP at low salt.

Characterization of labeled NCP

Labeled DNA could readily be reconstituted into NCP in a way indistinguishable from the experiments with unlabeled DNA. The electrophoretic mobility of labeled core
particle (protein and DNA) was virtually identical compared to unmodified material.

Table 3.2 gives a summary of all labeled NCP used in this work. Fluorescence energy transfer was measured between the following sites in the core particle:
- DNA termini
- DNA termini and histone H4 (tetramer)
- DNA termini and histone H2A (dimer)
- DNA center and H4 (tetramer)
- H2A (dimer) and H4 (tetramer)

<table>
<thead>
<tr>
<th>Octamer</th>
<th>Cys sites</th>
<th>DNA</th>
<th>Chromophores and sites</th>
<th>Salt range of measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Xenopus</em> (= XNCP)</td>
<td>x2oct H4 C47</td>
<td>as146</td>
<td>CPM→FLU between DNA ends</td>
<td>0.05-1M NaCl</td>
</tr>
<tr>
<td><em>Xenopus</em> with an average of 18 acetylations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Xenopus</em> with an average of 30 acetylations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Xenopus</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Xenopus</em></td>
<td>x3oct H2A C119</td>
<td></td>
<td>FLU→RHO; FLU on H4 and RHO on the DNA ends</td>
<td>0.05-1M NaCl</td>
</tr>
<tr>
<td><em>Xenopus</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Xenopus</em></td>
<td>x2/x3oct H4 C47 H2A C119</td>
<td>as146</td>
<td>FLU→RHO; FLU on H4 and RHO on H2A</td>
<td>0.4-1.0 M NaCl</td>
</tr>
<tr>
<td><em>Chicken</em> (= CNCP)</td>
<td>H3 110 wt histones</td>
<td>as146</td>
<td>CPM→FLU and FLU→RHO between DNA ends</td>
<td>0.05-1M NaCl</td>
</tr>
<tr>
<td><em>Chicken</em> with averages of 11 or 18 additional acetylations</td>
<td></td>
<td></td>
<td>CPM→FLU between DNA ends</td>
<td></td>
</tr>
<tr>
<td><em>Xenopus</em> all globular</td>
<td>glob</td>
<td>as146</td>
<td>CPM→FLU between DNA ends</td>
<td>0.05-1M NaCl</td>
</tr>
</tbody>
</table>

*Table 3.2: Summary of labeled NCP.*
The asymmetric DNA fragment used in this study (as146) was a 146 bp long fragment from the 5S-RNA gene of *Lytechinus variegatus* (Richmond et al., 1988). The symmetric fragment (s146) was made from one half of as146 (Flaus et al., 1996). See Appendix I for complete DNA sequences and Appendices III/IV for protein sequences.

### 3.3.2 Characterization of the energy transfer system

When measuring FRET, donor emission measurements are generally preferred over acceptor emission measurements, because the acceptor spectrum is often contaminated with donor emission. For our set of chromophores, the donor emission maximum was free of any acceptor emission. However, for control reasons it was extremely useful to look at both the acceptor and donor emission. To quantify the amount of transfer, donor emission was measured in the presence and absence of an acceptor chromophore. The difference between donor-acceptor and donor-only measurements was a measure for the relative distance between the two sites. The sampling of this difference over a wide salt range allowed a determination of distance changes to be made. At first, energy transfer between the DNA ends is discussed as an example for all transfer systems presented later.

As146 DNA with chromophores attached to the 5' phosphates was prepared as discussed earlier. Two different donor-acceptor pairs were used to reduce artifacts eventually caused by the chromophores. The standard system used CPM and FLU as donor-acceptor pair with CPM as the donor chromophore. Replacement of CPM by RHO gave rise to the other pair. Here, FLU was the donor chromophore and thus the direction of transfer was inverted with respect to the labeled sites.

For the measurements, NCP with donor and acceptor chromophores as well as with the donor chromophore alone was prepared. To avoid dissociation of the NCP due
to low concentration (Stacks & Schumaker, 1979), the total particle concentration in the measurement mix was raised by the addition of identical but unlabeled NCP. Thus the starting concentration was 85 μg NCP/ml (see methods) of which 3.8 μg were labeled NCP. Both double-labeled core particles showed transfer between the DNA ends at 50 mM NaCl. Figure 3.7 shows the emission spectra of core particle and free DNA

![Emission spectra of CPM/FLU labeled DNA](image)

**Figure 3.7: Emission spectra of the CPM/FLU labeled DNA.**
The excitation wavelength was set to 385 nm for the donor chromophore CPM. Emission maxima as in table 3.1. Double-labeled NCP (---) and double-labeled free DNA (—) both at 50 mM NaCl, double-labeled core particle at 0.6 M NaCl (····).

for the CPM-FLU chromophore pair. When comparing the emission spectra of double-labeled NCP (dashed line) with the one of the corresponding free double-labeled DNA (same concentration and conditions; solid line), increased acceptor emission at 520 nm for the NCP is observed. This additional acceptor emission can only be explained by
energy transfer from one DNA end to the other due to close proximity of the DNA ends in the core particle. An increase of the NaCl concentration to 0.6 M NaCl resulted in a loss of transfer for the NCP (figure 3.7; dotted line).

To quantify this reduction of transfer, the donor emission was monitored by stepwise addition of 5 M NaCl. As a control, this was done for free DNA first. As can be seen in figure 3.8 there is no transfer between the DNA ends as the donor-emission

![Figure 3.8: Salt dependent donor fluorescence for double-labeled DNA. DNA with CPM (●), CPM and FLU (▲), with FLU alone (○) and FLU and RHO (△) is shown. Donor emission was measured at 483 nm (CPM-FLU pair) and at 520 nm for the FLU-RHO pair.](image-url)
is not altered by the presence of an acceptor chromophore attached to the other end of free DNA. That is true for both CPM-FLU and FLU-RHO donor-acceptor pairs. We then did the same for NCP reconstituted with this labeled DNA. Figure 3.9 shows the donor quenching curves for the four labeled NCP. The curves obtained

![Figure 3.9: Donor emission of chicken NCP reconstituted with end-labeled DNA.](image)

The donor emission of NCP with CPM only (●), with CPM and FLU (▲), NCP with FLU only (○) and with FLU and RHO (△) is shown. Corrected donor quenching curves for the CPM-FLU pair (●) and the FLU-RHO pair (△) are at the bottom of the graph.

for donor-acceptor NCP (CPM-FLU and FLU-RHO pair) were clearly different from the curves obtained for NCP with the donor alone and from labeled free DNA (figure 3.8). For both double-labeled (donor and acceptor) NCP an increase in donor emission (reduction of transfer rate) around 400 mM NaCl was observed. To eliminate donor
quenching effects other than transfer, the emission curve from donor only particles was subtracted using the following formula: \( \text{em}_{\text{corr}} = \text{em}_{\text{DA}}[1+(\text{EM}_{\text{D}}-\text{em}_{\text{D}})/\text{em}_{\text{D}}] \) where \( \text{em}_{\text{DA}} \) is the emission of the donor-acceptor particle at a given NaCl concentration, \( \text{em}_{\text{D}} \) the emission of the donor particle at the same NaCl concentration and \( \text{EM}_{\text{D}} \) the emission of the donor particle at 50 mM NaCl. Summarized in words this means that the salt induced change in donor emission in the absence of acceptor (including all salt and other effects) is scaled to the value of the donor-acceptor curve and then subtracted. The corrected donor quenching curves obtained by this method were practically identical for both chromophore pairs (figure 3.9, bottom).

No significant shift in the emission maximum of the donor chromophores could be observed during the measurements. Such effects are usually an indication for unspecific interactions of the chromophore with its environment (e.g. sticking to a hydrophobic surface) which can affect the position of the emission maximum. Disassembly of the particle would then lead to a considerable change in the position of the emission maximum. Our observations showed no evidence for such unspecific interactions which is in good agreement with the fact that rather long linkers were used for chromophore attachment. The presence of practically identical difference curves for the two different donor-acceptor pairs was another strong argument for the integrity of the approach.

### 3.3.3 Transfer between the DNA ends

As146 DNA labeled with CPM/FLU was used to measure dissociation of the peripheral DNA regions from various histone octamers in increasing ionic strength. Octamers included in the study were recombinant-(XNCP), chicken-(CNCP), globular and hyperacetylated NCP (see table 3.2). At 50 mM NaCl, XNCP and CNCP showed
about 16-17% transfer (donor quenching by the acceptor; relative to the corrected emission of the donor at 1 M NaCl, figure 3.10). Globular NCP (lacking all the tails) showed only little transfer (4%) under the same conditions. Equilibrium measurements in increasing salt produced a transition with a midpoint of approximately 400 mM for XNCP and CNCP (figures 3.10, 3.12). The binding of the DNA ends in XNCP seemed to be somewhat stronger than in CNCP. Donor emission of globular NCP increased in the same salt range, which was indicative for some remaining transfer even for this core particle.

![Figure 3.10: Donor quenching with different octamers.](image)

Acceptor induced donor quenching (corrected donor emission) for different recombinant NCP in increasing ionic strength. Corrected quenching curves of XNCP (■), XNCP with 18 acetylations (♦), 30 acetylations (◊) and globular NCP (▲) are plotted. Transfer rates were 17, 12, 9 and 4%. For clarity, only averaged curves are shown.
After addition of NaCl, the equilibrium was reached within seconds. The dissociation of the DNA termini was fully reversible up to 600 mM NaCl. NCP incubated at 600 mM NaCl/RT could be diluted to 50 mM NaCl without changing the electrophoretic properties on a band shift gel (no free DNA showing up) or the amount of transfer (data not shown). Our results suggest that the DNA ends can reversibly dissociate from the histone octamer between 200 and 600 mM NaCl.

This first transition was further characterized by the use of acetylated histone proteins. The histone tails were chemically acetylated with N-Acetylsulfosuccimide in both recombinant- and chicken NCP. This acetylation reaction is specific for lysines accessible on the surface of the core particle, mainly those in the basic histone tails (Lewis et al., 1988). Two different levels of acetylated recombinant NCP were prepared with averages of 18 and 30 acetylations per particle (2-4 acetylations per histone protein). Acetylation of chicken NCP under the same reaction conditions resulted in a similar degree of modification. The acetylation efficiency was determined on a Triton-Urea-Acid gel (figure 3.11 A, lanes 1-4). Recombinant NCP has no natural acetylations (see figure 3.11 A), whereas chicken NCP has some acetylations or other modifications (Zhang & Nelson, 1986; Zhang & Nelson, 1988a; Zhang & Nelson, 1988b).

Both donor-acceptor and donor only core particles were acetylated in order to have equivalent material for the control measurements. On a mobility shift gel (figure 3.11 B, lanes 2, 3, 5 and 6), acetylated NCP ran higher than unmodified NCP (lanes 1, 4 and 7), and the bands were not as well defined probably due to dissociated histone tails and an unequal charge distribution due to the random acetylation.
Chapter 3

Fluorescence energy transfer

Figure 3.11: Acetylation of NCP.

Triton-Urea-Acid gel (Coomassie blue stain, A) and mobility shift gel (Ethidium bromide, B) of acetylated XNCP. (A): M= unacetylated XNCP with the histones as named in the figure. The elution speed is mainly dependent on the number of positive charged Lys and Arg in the histones. Acetylated NCP are shown in lanes 1-4. Each band corresponds to a distinct number of Lys modified by acetylation (and thus neutralized). Both donor-acceptor NCP (lane 1+ 3) and NCP with donor alone (lane 2 + 4) were acetylated. Lanes 1+2 show NCP with approximately 18 acetylations whereas the higher acetylated NCP (30) are shown in lane 3 + 4. As can be seen from the pattern, the donor-acceptor and the donor-only particle are acetylated to the same degree. (B): M= DNA size marker. XNCP with CPM and FLU on the DNA (lane 1), the same NCP with 18 (lane 2) and 30 (lane 3) acetylations. Lanes 4-6 show the same for XNCP with the donor CPM alone. XNCP without any modification is shown in lane 7.

Acetylation of recombinant and chicken NCP decreased the amount of transfer at low salt. Transfer in recombinant NCP was reduced from 17% to 12% and 9% in core particles with 18 and 30 acetylations, respectively. The midpoint of the transitions was slightly shifted to about 360 mM NaCl (figures 3.10, 3.12). Thus all acetylated core particles (XNCP and CNCP) had the same transition midpoints, whereas the midpoints for the XNCP and the CNCP were slightly higher.
Figure 3.12: Overlay of donor quenching curves of recombinant- and chicken NCP.

The corrected donor quenching curves were all scaled to the same amplitude to compare the transition midpoints. From right to left recombinant NCP (■) and chicken NCP (●) are shown. Acetylated recombinant NCP (18 acetyl. (♦), 30 acetyl. (★) and corresponding acetylated chicken NCP ((▲) and (▲)) were scaled in the figure as well. For clarity, only averaged curves are shown.

3.3.4 Transfer between the DNA ends and H4 (H2A)

To complete data from transfer measurements between the DNA termini, transfer between the DNA ends and histone H4 in recombinant *Xenopus laevis* core particles was measured as well. For this purpose, NCP with FLU as donor on the mutant Cys47
of H4 and the acceptor RHO attached to both 5' phosphates of as146 DNA was prepared. As the two chromophores are rather close (see discussion below), the amount of transfer between the sites was much larger than between the DNA termini.

Transfer between H4 and the DNA termini was completely lost in a transition between 200 and 800 mM NaCl. Figure 3.13 shows the emission spectra of NCP at 50 mM (solid line) and at 1 M salt (dashed line). Salt disassembly as above produced a midpoint of the quenching curve of 500 mM (figure 3.15), which was clearly different from the one observed in the measurements between the DNA termini. The position of the donor emission maximum remained unchanged during disassembly of the particle. Again, this is a good indication that the donor label makes no unspecific interactions with its environment.

Where does the discrepancy in the transition midpoints in the transfer between the DNA termini and the transfer between H4 and the DNA termini come from. The preparation of the DNA was somewhat different in the sense that the DNA for the measurements between the DNA termini was reannealed DNA as compared to the DNA used in the system here. In order to compare the two asymmetric DNA fragments, core particle with three labels was prepared. NCP with the two chromophores CPM and FLU on the DNA (from the end to end transfer measurements) was additionally labeled with RHO on H4 Cys47. Transfer was then measured between FLU on one DNA end and RHO on H4. The transition midpoint was again at 500 mM NaCl which was identical to the H4-DNA termini measurements presented above (data not shown). Thus, the difference in the transitions is not an artifact caused by the different preparation of the DNA. As yet another control, transfer between the DNA ends and H2A was measured. Similar to the above systems, transfer was lost in the 400-500 mM NaCl range (data not shown).
3.3.5 Dissociation of the dimer from the tetramer-DNA complex

As reported in the literature (Oohara & Wada, 1987a; Oohara & Wada, 1987b), the dimer does dissociate from the tetramer-DNA complex around 950 mM NaCl and before the complete separation of DNA and tetramer. At 2 M NaCl, the core particle is dissociated into octamer and free DNA. Fluorescence transfer now allowed us to measure, whether the dimer dissociation was somehow concerted with the DNA.
release. To answer this question, core particle with donor and acceptor labels on H4 and
H2A, respectively were prepared as described above. Dimer dissociation was then
measured using different DNA fragments such as the as146, s146 and cs146 (see
Appendix I for complete sequences). cs146 is a sequence containing a repeat proposed
by Crothers (Shrader & Crothers, 1989) which is supposed to be an ideal NCP binding
sequence. A detailed discussion of this sequence is given in chapter 4. Figure 3.14

![Figure 3.14: Donor quenching curves for dimer -tetramer transfer.](image)

NCP labeled with FLU on H4 and RHO on H2A was reconstituted with s146 (○●) and cs146 DNA
(▲▲). The measurements were either done in the presence of 0.05% Triton (○,△) or without any
detergent added (●,▲). The curves shown are the result of multiple measurements.

shows the result of these measurements. Core particle made from cs146 was slightly
more stable than the s146 core particle. NCP with the sequence as146 had the same
unfolding characteristics as core particle with symmetric DNA (see figure 3.15). The dimer-tetramer interaction was significantly destabilized by the addition of the nonionic detergent Triton-100.

Control measurements at a NaCl concentration of 400 mM showed that the dimers readily exchange—after 10-12 hr incubation of labeled NCP with an excess of unlabeled core particle, transfer was reduced to the expected amount (all NCP had been incubated 10-12 hr previous to the measurements). At 2 M NaCl, tetramer-dimer transfer could only be restored when the NCP concentration was increased to 230 μg/ml (85 μg/ml in all other measurements) and at temperatures between 4-10° C. There was no transfer above 20° C which is again in good agreement with previous results (Eickbush & Moudrianakis, 1978).

3.3.6 Dissociation of the DNA center from H4

So far, we had only been measuring the peripheral regions of the DNA and the histone dimer dissociation. Of course, it was now of interest to measure the dissociation of the central region of the DNA as well. Therefore, core particle with octamer labeled on H4 and DNA labeled near the dyad was prepared as discussed in previous sections. As substitution of one of the non-bridging oxygen atoms on the phosphate with a sulfur does not cause significant changes in the DNA structure (Fidanza et al., 1992; Ozaki & McLaughlin, 1992), the labeled DNA should be very similar to the unmodified si46 DNA. The measurements were done using two different labeling positions on the DNA backbone. The chromophore was attached to the phosphate backbone 7 or 8 bp away from the DNA center in both halves. These sites were selected based on the 7 Å structure of the NCP (Richmond et al., 1984) and the known helical periodicity of nucleosomal DNA. As the DNA minor groove is oriented away from the histone
octamer at the DNA dyad, it could be expected that chromophores at the selected labeling sites would not interfere with the octamer.

The measurement of the complete release of the DNA from core particle leads to exchange of DNA between histone tetramers. In order to be able to measure this exchange, all particles in the measurement mix were labeled with acceptor. Therefore, NCP with labeled symmetric DNA and NCP with unlabeled DNA were mixed before

![Figure 3.15: Overlay of donor quenching curves.](image)

Dissociation of the DNA ends from the octamer, as measured between H4 and DNA ends is shown on the left (▲). Tetramer-dimer transfer with s146 DNA (■) and as146 DNA (□) in the presence of Triton is shown in the middle. The transition at the highest salt concentration shows the dissociation of the DNA center from the tetramer for DNA labeled 7 (●) or 8 (○) bp away from the DNA center.
labeling with the acceptor RHO on histone H4. By this method, all molecules were equally labeled with acceptor (in the measurements between the DNA termini only a small fraction of the core particle was labeled).

The DNA dissociation was measured in two ways: 1) by separate preparation and measurement of each salt concentration (incubation time 14 h), and 2) by stepwise addition of salt (incubation time 1 - 2 h). The results were independent of the method used. Figure 3.15 summarizes the results obtained. The tetramer-DNA center measurements (●○) are compared with the H4-DNA termini and dimer-tetramer results. All the transfer was lost in a transition with a midpoint around 880 mM NaCl (figure 3.15). Donor emission remained unchanged below 700 mM NaCl. Both DNA fragments, having the labels 7 or 8 base pairs from the DNA center, gave identical transition midpoints. Unlike the dissociation of the dimer from the tetramer, this transition was not sensitive to the addition of Triton, which was expected based on the nature of the DNA-protein interactions.

3.4 Discussion

3.4.1 General

With our new site-specific methodology, direct measurements of the dissociation of the DNA from the octamer/tetramer complex became possible. It could be shown that the release of the DNA is a multistep process taking place over a wide salt range.

Firstly, the peripheral regions of the DNA dissociate in a fast equilibrium at relatively low ionic strength (300-500 mM NaCl). The dimer-tetramer interactions remain entirely unaffected and so does the center of the DNA in general. The transition is not concerted with the dimer dissociation.
As expected, the dissociation of the dimer from the tetramer-DNA complex is sensitive to the presence of a nonionic detergent. With detergent added, the dimers dissociate around 750 mM NaCl, whereas the same transition seems to be concerted with the dissociation of the DNA center from the tetramer in the absence of Triton. The dimer appears to play a significant role in the stabilization of the complex. Without the dimer present, the tetramer/DNA interactions are destabilized. Surprisingly, the dimers do exchange already in 400 mM NaCl.

The last step in the disassembly of the NCP is the complete dissociation of DNA from the tetramer at 880 mM NaCl. Taken together with the data for the DNA termini, we postulate the following model for DNA dissociation. In a first transition, the terminal 30-40 base pairs at the DNA termini are released reversibly. In a second transition starting at 750 mM NaCl, the central 70-80 base pairs of the DNA dissociate from the tetramer. These two main steps are linked by the dimer dissociation which takes place in between. Our measurements, however, do not yet allow us to give the exact number of base pairs being released in each transition.

Since the two NCP crystal structures (chapter 5 and 6) are available now, the relevant results are included in the discussion here. Please note however, that this detailed information was not available at the time the fluorescence measurements were performed. Figure 3.16 shows two schematic drawings illustrating the positions of the labeled sites and the protein-DNA contacts. Only the DNA supercoil in the dyad region of the core particle is shown in figure 3.16A. The distances between the relevant sites are summarized in table 3.3. Figure 3.16B shows one DNA half with the points of interaction with the histone proteins. The numbering starts at the end of the DNA fragment and marks the most likely center of the interactions. At the first three contacts at the DNA center, the DNA is contacted by the H3/H4 tetramer. The next three contacts are made to the H2A/H2B dimers and the last contact at the DNA termini is made by histone H3 alone (discussed in more detail in chapter 5).
Figure 3.16: Positions of labels on the NCP surface.
A: NCP as seen when looking at the dyad axis position at the DNA center. The labeled sites are: the DNA 5' ends De-1/2, Cys47 on histone H4 (H4-1/2), Cys119 on H2A (H2A-1/2) and the phosphorothioates on the DNA center (Dc-1/2). B: DNA-histone contact sites. One DNA half is shown at the top with the fragment dyad axis as indicated. There are seven contact sites distributed over the 73 base pairs of DNA. The symmetry related 73 base pairs contact the symmetry related histones.

<table>
<thead>
<tr>
<th>Between</th>
<th>Sites</th>
<th>Distances</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA-termini</td>
<td>De-1/De-2</td>
<td>71 Å</td>
</tr>
<tr>
<td></td>
<td></td>
<td>83 Å (off-cent.*)</td>
</tr>
<tr>
<td>DNA termini- tetramer</td>
<td>De-1/H4-1</td>
<td>40 Å</td>
</tr>
<tr>
<td></td>
<td>De-1/H4-2</td>
<td>45 Å</td>
</tr>
<tr>
<td>dimer - tetramer</td>
<td>H2A-1/H4-1</td>
<td>25 Å</td>
</tr>
<tr>
<td></td>
<td>H2A-1/H4-2</td>
<td>30 Å</td>
</tr>
<tr>
<td>DNA termini- dimer</td>
<td>De-1/H2A1</td>
<td>17 Å</td>
</tr>
<tr>
<td>DNA center - tetramer</td>
<td>De-1/H4-1</td>
<td>23 Å</td>
</tr>
<tr>
<td></td>
<td>De-1/H4-2</td>
<td>45 Å</td>
</tr>
</tbody>
</table>

*: To simulate off-centered DNA, B-form DNA was modeled to one end of the nucleosomal DNA. The distance between the DNA ends of a 146mer DNA fragment off-centered by 8 or 9 base pairs was then determined from the model.

Table 3.3: Distances between labeled sites on the NCP.
The distances were derived from the α6s NCP crystal structure presented in chapter 6. Average values from the symmetry equivalent distances are shown.
Table 3.3 gives a summary of the distances between labeled sites on the NCP. The distances do not include the dimensions of the linker and the fluorescence chromophores. As end points for the distance measurements, the closest well defined atoms were taken. Based on the crystal structure, none of the chromophores is sterically restricted in any way.

3.4.2 Dissociation of the DNA from the octamer

The influence of the parameter Ro

The dissociation of the DNA ends was measured in two ways. Firstly, by transfer between the DNA termini, and secondly between histone H4 and the DNA termini. The transition midpoints were 400 mM for the end to end system and 500 mM NaCl for the transfer between H4 and the DNA ends. The source of this difference is considered. As shown earlier, the differences in DNA preparation do not influence the position of the transition midpoints. The DNA was labeled the same way in both experiments, and the chromophores on Cys47 of histone H4 are close to the central turn of the DNA not to the DNA termini (see figure 3.16). They should remain sterically unaffected by the dissociation of the DNA termini. In addition, donor emission in the other two transitions using the same H4 site (dimer-tetramer, center-H4) was absolutely unchanged below the actual transition.

The difference is more likely to be a consequence of the Ro value of the chromophores with respect to the measured distance. According to figure 3.1, significant changes in energy transfer can only be expected for distances in the range of 0.5 Ro - 1.5 Ro (Ro is approx. 50 Å for the chromophores used here). Whereas the distance between the DNA termini is approximately 70 Å, the other distance (H4 Cys47-DNA termini) is as short as 40 Å. The linkers could decrease the distance even further. The transfer between H4 and the DNA termini could therefore be insensitive
for the first small changes in distance, simply because the chromophores are too close
to each other.

However, our results clearly show a reversible release of DNA ends at NaCl
concentrations below 800 mM. The central turn of the DNA seems to remain
unaffected. This result is in opposition to previous experiments. Transfer measurements
between H3 Cys110 and the DNA ends (Eshaghpour et al., 1980) showed no change in
distance in NaCl concentrations up to 600 mM NaCl. However, the integrity of H3
Cys110 labeled NCP is at least questionable because of the site of labeling. Hydroxyl
radical footprinting experiments (Puhl & Behe, 1993) showed no change in the
cleavage pattern at 800 mM NaCl as compared to low salt. The problem of this
approach, however, is that information about the DNA ends is weak or even absent.

The data presented so far raise the question of the accuracy of high salt exchange
experiments (Drew, 1991). When the free energy of binding of the histone octamer to
DNA is measured in high ionic strength (≥ 800 mM), as has to be done in exchange
studies, one has to be aware of the absence of a considerable amount of DNA-protein
interactions in the flanking regions. Since the dimer is only weakly bound at 800 mM
NaCl, only the tetramer-DNA interactions are left.

Mechanism of the DNA release

The differences in the transition midpoints of the transfer between the DNA
termini and the transfer between H4 and the DNA termini allow us to make an
assumption about the mechanism in this first transition. Presumably the dissociation of
the DNA goes segment by segment, as defined by the distance between the individual
DNA-protein contact sites (see figure 3.16B). It would be of considerable interest to
know, how many segments actually dissociate and whether the process can be
approximated by a two state model. The latter would require a defined start point with
ends bound and transfer, and an endpoint with no transfer. In addition, no intermediate
states with transfer were allowed. These assumptions are probably fulfilled for a
dissociation reaction such as the dimer dissociation from the tetramer-DNA complex.
For a process such as the release of the DNA termini, it is not necessarily true. If a two
state model is assumed here, the same transition midpoints for both transfer systems
would be expected. Our data do not support this idea. Thus, at least the release of the
first base pairs (segments), which cause the change in the end to end transfer is not
approximated with a two state model. However, in case of very high transfer at the
beginning, one would not see intermediates with much less transfer. A transition would
then look like a two state process without being one. As discussed, it is thus of
importance to select the chromophores and sites to have $R = R_0$, the most sensitive
region for distance measurements.

In order to quantify the measured transition, an equilibrium for a two state process
can be formulated as follows: $NCP + n \text{NaCl} = NCP^*$, with $K = \frac{NCP^*}{NCP \times [\text{NaCl}]^n}$. Linearising this equation yields
\[ \ln(K^{-1}) - nx\ln([\text{NaCl}]) = \ln\left(\frac{NCP}{NCP^*-1}\right) \] with $NCP = NCP + NCP^*$. The number of Na$^+$ ions (slope) involved in the dissociation and
its free energy $\Delta G$ can then be obtained via a Hill plot. Applied to the H4-DNA termini
transfer data, approximately 5 Na$^+$ ions and a $\Delta G$ of 9 kJ/mol were calculated for the
DNA release. This number of Na$^+$ ions cannot be directly related to the number of
released phosphates because charge neutralization of DNA in solution is not by direct
binding of monovalent counter ions, but rather by formation of a charged hydration
shell (Clark & Kimura, 1990).

The exact determination of the number of dissociating segments would require
the labeling of the DNA in each individual segment so that the process could be
dissected into finer steps. However, it is probably save to assume that the movements
detected in the first transition are restricted to the outer three contact sites, because
otherwise the dimer binding would be effected.
Chapter 3

Fluorescence energy transfer

Effects of modification of the basic histone tails on the release of the DNA ends

It is known that the basic (N-terminal-) tails of the histone proteins bind the DNA probably on the outside of the NCP or in the linker region between nucleosomes, but it has been a matter of extensive discussions how these tails affect the binding of the 146mer DNA organized in the core particle. We therefore investigated the influence of histone acetylation on the disassembly process. Upon acetylation, the midpoint of the release of the DNA termini of modified recombinant- and chicken NCP was slightly moved to lower salt. All the hyperacetylated species had the same midpoints. Differences between these particles are observed only for the amount of transfer at low salt- the more acetylations , the less transfer.

Two different mechanisms could account for the release of the DNA ends from the histone octamer in increased ionic strength. Either the basic histone tails first dissociate from the DNA followed by a dissociation of the DNA from the octamer core, or interactions between the DNA and the octamer core are broken first. In this second mechanism the dissociating DNA would also free the histone tails. If the basic histone tails were indeed responsible for DNA dissociation, a significant shift of the transition midpoint after acetylation or removal of the tails would be expected. This was not observed. All species, including glob NCP and NCP lacking only one of the H2A, H2B, H3 or the H4 tails (data not shown) had their transition midpoints between 350 and 400 mM. From these results it is very unlikely that the basic histone tails make any of the important interactions.

Our results let us suggest that the dissociation of the DNA follows the second mechanism. The salt would thus disrupt the interactions between the DNA termini and the globular part of the histone octamer. These interactions remain mainly unchanged during acetylation and therefore give rise to roughly identical transition midpoints for all modified species. In terms of our model, the function of the tails would then be the stabilization of the DNA ends on the octamer core at low salt, presumably by binding to the outside of the DNA supercoil. The more modified the tails are (all globular NCP is
the extreme end on this scale), the lower is their stabilizing effect on the DNA ends at low salt and the smaller the population of tightly bound DNA at a given time point.

Recently, a theoretical model was proposed for the DNA dissociation from nucleosomes (Marky & Manning, 1995). The authors postulated a spontaneous release of 10 bp end segments from the octamer surface. Fragments between 10 and 60 bp were found to have intermediate states between bound and fully unwound (dissociated). In contrast 70 and 80 bp fragments do not seem to have stable intermediate positions and constitute a two state model. With respect to the results presented here, these findings could be interpreted as follows. The number (and acetylation state) of basic histone tails is determining for the equilibrium between closely bound and partially released DNA end segments in this spontaneous dissociation. The stronger the binding of the histone tails the larger is the fraction of closely bound end segments and thus the more transfer can be measured at low salt. Increasing salt concentration then weakens the binding forces between DNA and core octamer and thus reduces the stability of the DNA-bound state.

The conformation of NCP differs considerably between the hyperacetylated state and at 600 mM NaCl. Within hyperacetylated NCP, the basic histone tail interactions are weakened, but the DNA termini are still bound, allowing a significant amount of transfer. At 600 mM NaCl, the DNA termini are dissociated completely from the histone octamer. These findings support recent ideas that hyperacetylation does not alter the DNA cleavage pattern on the core particle DNA, but makes the DNA generally more accessible (Lee et al., 1993; Vettese-Dadey et al., 1994).

**Dimer dissociation and the complete release of the DNA**

A salt concentration of 900 mM NaCl is unexpectedly low for the dissociation of the DNA from the histone tetramer. Considering the results from the dimer dissociation, interesting observations can be made. Whereas in the presence of Triton,
the dimer dissociation is separated from the final DNA dissociation by approximately 100 mM, the transitions nearly superimpose in the absence of detergent. Obviously, the binding of the dimer does not directly affect the dissociation of the DNA center from the tetramer. On the other hand, the proper organization of the DNA on the tetramer seems to be lost in about the same salt range at which the dimer dissociates. Furthermore, the dimer dissociation is DNA dependent. This implies that the dimer-DNA interactions are at least as important for dimer binding as the direct dimer-tetramer contacts. cs146 is a strong nucleosome binding sequence which seems to stabilize the NCP (see also chapter 4). The shift of the dissociation curves between s146 and cs146 can be used to determine ΔΔG values for the dimer-tetramer interaction and the stabilizing effect by the cs146 sequence. Addition of Triton is worth a destabilization of the dimer binding of 4.2 and 4.9 kJ/mol for s146 and cs146, respectively. Replacing s146 by cs146 is worth a stabilization of the NCP of approximately 6.0 kJ/mol which is in good agreement with a value of 5.65 kJ/mol determined earlier for similar sequences (Shrader & Crothers, 1989). The data could be interpreted in a way that the DNA dissociation of the contact sites at position 37 (dimer) and 47 (tetramer) is tightly linked to the dimer dissociation and the dissociation of the entire complex.

In contrast to our results, a value of approximately 1.5 M NaCl following tyrosine fluorescence quenching by the bound DNA has been proposed for the DNA dissociation from the histone tetramer (Oohara & Wada, 1987a; Oohara & Wada, 1987b). Dimer dissociation was attributed to a transition around 1 M NaCl. However, tyrosine quenching is a rather unspecific way of measuring the DNA protein interactions. Unspecific (not positioned) association of tetramer/octamer to the DNA would probably produce the same quenching. In comparison, our system is very specific. If the center of the DNA is not properly positioned with respect to the histone dyad axis, transfer is lost. As a possible explanation for the different results, sliding of the DNA on the octamer surface could cause an earlier loss in transfer.
exchange experiments with an excess of unlabeled NCP show that there is DNA exchange already at 600 mM NaCl (data not shown). From this, we conclude that DNA dissociation is mainly responsible for the transition.

The label on the DNA backbone seems to affect the positioning of the DNA on the octamer. This is shown by two translational positions of the octamer on the labeled DNA after reconstitution compared to only one for unlabeled DNA (see chapter 4 for more details). However, NCP with the unlabeled symmetric DNA can occupy the same translational positions (data not shown). Although we can not rule out the possibility of a slight destabilization of the DNA tetramer binding, we do not believe this to be the main reason for the large difference to Oohara's results.

A common method to prepare core particles is the exchange method at 0.8 to 1M salt (Drew, 1991). Obviously, histone-DNA interactions must be significantly disrupted at these salt concentrations to allow for exchange. From that point of view, the salt range we measure for the loss of the specific interactions near the DNA center is compatible with the salt range observed for DNA exchange between octamers (Vettese-Dadey et al., 1994).
4. Crystallization of the NCP

4.1 Introduction

4.1.1 History of NCP crystallization

After the discovery of the highly repetitive packing of eukaryotic DNA, first attempts to crystallize the NCP were undertaken because X-ray diffraction studies on single crystals could reveal the fine details of DNA-histone interactions. At that time, one had also learned to isolate nucleosome core particles in preparative scale by limited micrococcal nuclease digestion of chromatin and subsequent purification (Noll & Kornberg, 1977). The treatment removed the linker DNA between nucleosomes, as well as the linker histone H1. The resulting core particles had a well defined structure containing DNA 146 ± 2 bp. Due to the source of the complex, the individual histone proteins as well as the DNA were not homogenous. In vivo, histone proteins are often acetylated or otherwise modified, and although the DNA length was rather uniform, it is of random sequence.

Crystals of NCP isolated from different sources such as rat liver, chicken erythrocytes, sea urchin sperm and many others gave the same diffraction pattern (Rhodes et al., 1989 and Ref. therein). For later studies, mainly core particles isolated from chicken erythrocytes were used because chicken blood is easily available and has only few proteases. NCP crystallizes in millimolar concentration of bivalent cations. The combination of MnCl$_2$ (40 mM), KCl (60 mM) and potassium cacodylate pH 6.0 (10 mM) has been most successful to produce diffraction quality crystals of up to 0.3x0.3x1.5 mm in size (Rhodes et al., 1989). Using these conditions applied to beef
kidney nucleosomes, the X-ray structure of the NCP was solved at 7 Å (Richmond et al., 1984).

To improve the homogeneity of the material and the crystals, NCP with an asymmetric, 146 bp defined sequence DNA and chicken octamer was prepared (Richmond et al., 1988). This DNA sequence was from the 5S RNA gene of *Lytechinus variegatus*, which had been shown to form stable nucleosomes in one translational position (Simpson & Stafford, 1983). These crystals were soaked in 22.5% 1,6-hexanediol and diffracted anisotropically to about 3-5 Å resolution at RT, which was an improvement of about 1 Å compared to the mixed sequence crystals. In order to improve crystal quality, recombinant histone proteins from *Xenopus laevis* were refolded into a histone octamer and used together with defined sequence DNA to make core particles (Richmond et al., 1993, K. Luger unpublished results). Besides the fact that these core particle preparations were homogeneous with respect to protein and DNA, the recombinant proteins also offered the possibility to make cysteine mutants, which could then be used for heavy atom labeling in MIR phasing. Crystals of this first all recombinant core particle grew in P2_12_12_1 with 1 molecule per asymmetric unit and diffracted anisotropically to 3-4 Å resolution.

In the meantime, the X-ray structures of some of the components of chromatin were solved. The structure of the histone octamer alone without DNA (Arents et al., 1991; Arents & Moudrianakis, 1993; Wang et al., 1994a) had been solved to 3.1 Å although the coordinates have never been released. The crystal structure of the globular domain of chicken linker histone H5 has been solved to 2.5 Å (Ramakrishnan et al., 1993).
4.1.2 Crystallization of all recombinant NCP

At the start of this thesis, the state of the art of NCP crystallization and X-ray structure determination was as follows: all recombinant NCP containing *X. laevis* histones (including several cysteine mutants) and as146 could be crystallized reproducibly. The crystallization was mainly done in hanging drops using initial salt concentrations of 60-80 mM MnCl₂, 50-80 mM KCl and 20 mM potassium cacodylate pH 6 against reservoir buffer of half the concentration at 20° C. To grow larger crystals, 50 µl sitting drops were allowed to equilibrate against water. Crystals were then stepwise transferred into 24% MPD/50 mM KCl, 40 mM MnCl₂ and 10 mM K-Cacodylate pH 6. MPD increased the maximum resolution of NCP crystals and also acted as a cryoprotectant in the cryotemperature experiments. For X-ray analysis at cryotemperatures, crystals were mounted in tapered capillaries. The following protocol had proved to be most successful for crystal cooling: 1) slow cooling from 4° C to -17° C in a cold air stream (5 min), 2) flash freezing into liquid propane at -90° C, and finally 3) another fast transfer to approximately -180° C as provided by a cryocooler. After this procedure, crystals could be stored for months in liquid nitrogen without losing their diffraction quality. The lifetime in the X-ray beam was only limited by the amount of primary radiation damage. This work and the associated techniques are the result of several years of work done in this laboratory by Dr. D. Sargent, Dr. K. Luger, L. Toldo, Dr. M. Struck, and M. Donatz.

4.1.3 Nucleosomal DNA and crystallization

When working with all recombinant core particles, the DNA has to be reconstituted with histone octamer to form the NCP. This is usually done by mixing the components at 2M NaCl, followed by dialysis against low ionic strength. It is known
that the octamer can take different translational positions during this process, whereas
the rotational phasing is generally conserved for the short fragments used here (Dong et al., 1990; Pennings et al., 1991). As a consequence, the translational position of the
octamer after reconstitution has to be checked carefully for each DNA fragment.

The crystallization with bivalent metal ions such as Mn$^{2+}$ strongly favors an
intermolecular cross-linking of the DNA strands as the main crystal contacts, thus
making the DNA most important for crystal formation. However, it was largely
unknown what DNA properties favor crystallization. In the structure determination of
smaller protein-DNA complexes, the DNA ends were often found to be useful for
crystallization. End to end stacking, sticky end interaction or short triple strand
formation can result in continuous alignment of the DNA fragments throughout the
crystal. The largest improvements for crystal quality are generally achieved when the 10
bp periodicity of the DNA can be conserved throughout the crystal. Based on the 7 Å
structure of the NCP, the importance of such contacts between the DNA ends was
suggested but remained unclear (Richmond et al., 1984). Base pair stacking in the NCP
crystals would presumably restrict the allowed DNA length to a narrow range.

Although the first, all recombinant NCP used in crystallization was made from
asymmetric DNA, the potential twofold symmetry of the histone octamer suggested the
use of a palindromic DNA fragment. The exact location of the dyad axis of the DNA on
the histone octamer would then be important for the design of the fragments. Assuming
that the space between two base pairs is positioned on the octamer dyad axis, an even
numbered palindromic DNA sequence could be symmetrically positioned. In case of a
base pair being positioned on the octamer dyad, an odd-numbered sequence of
palindromic DNA would be ideal for the formation of a entirely symmetric NCP.

The base sequence of nucleosomal DNA may also have an influence on the
crystallization of the complex, because certain sequences may prefer to adopt a
particular conformation (discussed earlier). However, at the moment it appears to be
difficult to link any of these parameters to the formation of high quality NCP crystals.
4.1.4 Objectives of the work

Even though the crystals available at the beginning of this thesis diffracted reasonably well, the work to improve crystal quality never ceased. One of the main concerns was asymmetry of the as146 DNA sequence. Assuming a two-fold symmetry for the histone octamer, the DNA asymmetry could disturb crystal packing. Work therefore focused on the DNA part of the NCP in the preparation of symmetric DNA fragments of different length and base sequence. Additionally, crystallization and the crystal handling were further optimized. In parallel, new histone protein cysteine mutants were prepared in order to find suitable sites for heavy atom labeling (K. Luger unpublished results). These sites were most important for phase determination by multiple isomorphous replacement (MIR).

4.2 Results

4.2.1 Translational position of the DNA in the NCP

Reconstitution of as146 with chicken or recombinant octamer generally resulted in three distinct bands (figure 4.1, lane 1), when analyzed on mobility shift gels (Richmond et al., 1993). The amount of NCP present in the different bands varied slightly from preparation to preparation. Such NCP could be readily crystallized giving crystals as discussed above. As described earlier for 5S rDNA (Pennings et al., 1991), a similar phenomenon had been observed already. In their experiments, the different bands were attributed to different translational positions of the octamer on the 207 bp DNA, and the distribution between the bands could be altered by incubation at 37° C.
When applied to the as146, all NCP was shifted into the fastest migrating band after incubation at 37° C (figure 4.1 lane 2, K. Luger unpublished results). Crystallization of such material was much more reproducible even though the quality of the crystals remained unchanged because only the fastest migrating NCP was incorporated into crystals anyway (K. Luger unpublished results). Based on these findings and on mapping results (Flaus et al., 1996), the assignment of bands was as follows. The fastest migrating NCP contained fully centered DNA (DNA positioned 0.5 bp from the dyad) whereas the DNA in the two slower migrating core particles was shifted by 9.5 bp to either side of the dyad. For palindromic DNA with two identical DNA termini, off-centering resulted in only one higher migrating band (see figure 4.3 lane 1+2).

Even though the 146 bp NCP is an artificial construct, it was interesting to further characterize these positioning effects because none of the DNA fragments used for crystallization was centered after reconstitution (see below). However, almost all fragments could be centered by incubation at elevated temperatures. This observation suggested that there must exist a very strong driving force for this process in the core
particle. The obvious conclusion was that once the core particle was formed, the octamer tried to maximize protein-DNA contacts by centering the DNA.

The fluorescence methods introduced in the previous chapter, offered a tool for further investigation of the shifting process. When transfer between DNA termini of freshly reconstituted material was compared with transfer of centered NCP, the latter clearly showed more transfer (data not shown). This was explained by a reduced distance between the DNA termini in centered NCP (see table 3.16 for the distances). Dissociation of DNA at 37° C as measured by transfer between H4 and the DNA termini was slightly destabilized (midpoint 40 mM lower), however, the amount of initial transfer was almost unchanged (data not shown). This is compelling evidence that the interactions between peripheral DNA and protein remain essentially unchanged at 37° C compared to 20° C. This has to be expected, if the driving force for the shifting to the centered position is indeed the maximization of histone DNA interactions. Given this information, increased ionic strength should inhibit shifting of the DNA towards the centered position because of the dissociation of the DNA ends from the histone octamer. As expected, incubation at 37° C and 600 mM NaCl did not alter the distribution between the positions (data not shown). In a related experiment starting with centered NCP, the octamer could be repositioned to the off-centered position by incubation at 37° C in the presence of more than 300 mM NaCl (figure 4.2).
Figure 4.2: Back shifting of NCP in increasing ionic strength

Centered NCP with s146 DNA is shown next to the DNA size marker (M). This NCP was incubated for 2 h at 37° C at the NaCl concentration indicated.

These shifting experiments are in absolute agreement with the results obtained from fluorescence measurements, where we show a dissociation of the DNA termini already at 300-400 mM NaCl.

So far the mechanism of the shifting process has not been discussed. There are three different ways, in which DNA could slide on the histone octamer: 1) the entire histone octamer could rotate in the DNA supercoil which would probably need a high activation energy because a large number of protein-DNA interactions has to be broken at the same time. 2) the DNA could be repositioned by a series of small torsional movements, for this to occur the DNA would have to be very flexible. Or, 3) DNA end segments could dissociate from the histone octamer and associate in a different translational (and rotational) position.
4.2.2 5S rDNA sequences

Starting from the asymmetric 5S rDNA sequence (as146), a symmetric 146 sequence (s146) was constructed using one half (Flaus et al., 1996, S. Tan). See appendix I for a sequence alignment. To test the influence of the DNA length on crystallization, shorter and longer DNA (138, 142, and 150 bp in length) was prepared. The dyad region with the HindIII site was identical for all these fragments, the differences in length were achieved by insertions/deletions at the DNA ends adjacent to the EcoRV site. The octamer position on s146 (and as146) had been determined using a tethered cutting reagent on the basis of hydroxyl radicals (Flaus et al., 1996).

![Figure 4.3: Mobility shift gel of NCP made from symmetric 5S sequences.](image)

NCP with the symmetric DNA fragments s146 (1,2), s142 (3,4) and s138 (5,6) were run on a mobility shift gel. After purification (1,3,5) and after incubation for 2 h at 37° C. M DNA size marker as described earlier.

Table 4.1 gives a summary of the data for NCP with these sequences. The octamer positions on the fragments are given as distances in bp from the dyad axis of the DNA fragment (between the central two base pairs for even-number, symmetric DNA). After reconstitution of octamer with s146 DNA, the DNA was predominately in the off-centered position. Figure 4.3 shows NCP with the three symmetric fragments
s146, s142 and s138 in the two translational positions. It is interesting to note that even a DNA fragment 8 bp shorter still behaved identical as s146. Shifting to the centered position was much slower for s146 than for as146: at 37° C an incubation time of 60 min was needed as compared to the 5-10 min for the as146 DNA.

<table>
<thead>
<tr>
<th>Seq. name</th>
<th>s146</th>
<th>s142</th>
<th>s138</th>
<th>s150</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position after reconstitution</td>
<td>8.5, 9.5</td>
<td>-10</td>
<td>-10, (0)</td>
<td>-10</td>
</tr>
<tr>
<td>Shifting to center position</td>
<td>&lt; 1h at 37° C</td>
<td>&lt; 2h at 37° C</td>
<td>&lt; 2h at 37° C</td>
<td>&lt; 2h at 37° C</td>
</tr>
<tr>
<td>Position after heating</td>
<td>0.5, 1.5</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
</tr>
<tr>
<td>Crystallization</td>
<td>a) 62-78 mM MnCl₂, 50-80 mM KCl</td>
<td>54-90 mM MnCl₂, 50-80 mM KCl</td>
<td>54-90 mM MnCl₂, 50-80 mM KCl</td>
<td>58-78 mM MnCl₂, 50-80 mM KCl</td>
</tr>
<tr>
<td>Crystals</td>
<td>a) plates (centered NCP)</td>
<td>no</td>
<td>needles</td>
<td>blocks</td>
</tr>
<tr>
<td>b) bipyramidal blocks (off-centered NCP)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diffraction limit</td>
<td>a) 4-6 Å</td>
<td>-</td>
<td>less than 30 Å</td>
<td>-</td>
</tr>
<tr>
<td>b) very poor diffraction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remarks</td>
<td>b) = Crystals with off-centered DNA</td>
<td>Crystals only for off-centered NCP</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(***: Starting salt concentration in the hanging drops. 20 mM KCacodylate pH 6 was always included. Vapor diffusion was against reservoir buffer of half the salt concentration. Initial NCP concentration was at 4 mg/ml.)

Table 4.1: Table of 5S rDNA crystals.
Characterization and crystallization of 5S rDNA containing NCP. Exact octamer positions before crystallization were determined for the s146 only (Flaus et al., 1996). The translational position of the nucleosomal DNA in the crystals was checked by running material from crystals dissolved with EDTA on a mobility shift gel.

Crystals could be obtained for both centered and off-centered s146 NCP. Starting from centered s146 NCP, as high a Mn²⁺ concentration as used in crystallization could induce the formation of off-centered NCP even at 20° C, which was rather surprising because it had been reported that the presence of bivalent cations inhibits the shifting process (Pennings et al., 1991). Crystals of centered NCP diffracted to 4-6 Å, which was no improvement compared to the as146 NCP crystals. Crystals of off-centered NCP
were of even poorer diffraction quality. Crystallization with bivalent cations of different ionic radius such as Mg$^{2+}$ (0.66 Å), Ca$^{2+}$ (0.99 Å), and Ba$^{2+}$ (1.34 Å), did not improve crystallization (Mn$^{2+}$ : 0.80 Å). In retrospect, the step from as146 to s146 did not extend the diffraction quality of NCP crystals. No diffraction quality crystals could be obtained for the other 5S sequences using Mg$^{2+}$, Mn$^{2+}$, Ca$^{2+}$ or Ba$^{2+}$ as precipitant.

4.2.3 Crothers sequence

Nucleosomes display a sequence preference for DNA binding, as suggested by the fact that DNA is strongly distorted in the core particle, and DNA bending is sequence dependent. It is obvious that factors affecting DNA bendability can be of great importance for the overall affinity of a DNA fragment to the histone octamer. Based on the sequence dependent DNA bendability (see above), sequences in which the distribution of AT and GC match the periodicity of the histone octamer surface had been designed (Shrader & Crothers, 1989; Shrader & Crothers, 1990). This DNA had the base pattern (A/T)$_3$NN(G/C)$_3$NN, with an overall repeat length of 10 bp. Presumably, the minor and major grooves of the DNA face the protein at AT and GC rich regions, respectively. Following Shrader & Crothers (1989), the repeat

(TCGGTGTAGAGCCTGTAAC)

was used to construct cs146 as follows (one half is shown with the EcoRI restriction site at the center):

AATTCGTAAAC - (TCGGTGTAGAGCCTGTAAC)$_3$ - TCGAT
GCAATT - (AGCCACAATCTCGGACATTG) -- AGCTA

EcoRI    EcoRV

This highly repetitive sequence had the first A/T stretch 5 bp away from the dyad axis where the minor groove of the DNA is supposed to be oriented towards the histone surface. The full length DNA was not intrinsically bent as compared to most of the symmetric DNA sequences used in this work (figure 4.5, lane 8 versus lanes 1-5).
Reconstitution of cs146 into NCP resulted in the same off-centered NCP as previously described for the symmetric 5S sequences. However, it was remarkably stable against shifting into the centered position. For complete shifting, the NCP had to be incubated for 2 h at 55° C. The translational positions before and after shifting were at 9.5 and 0.5 bp off the dyad (data not shown, together with A. Flaus). Similar to s146 NCP, core particle containing cs146 could be crystallized either in the centered or in the off-centered position. Crystallization of the centered NCP yielded rod-shaped single crystals with diffraction similar to the s146 crystals. Crystals containing NCP with off-centered DNA behaved very much the same as the analog material with the s146 DNA. No high resolution diffraction pattern could be recorded.

<table>
<thead>
<tr>
<th>Seq name</th>
<th>cs146</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position after reconstitution</td>
<td>9.5</td>
</tr>
<tr>
<td>Shifting to center position</td>
<td>2 h at 55° C</td>
</tr>
<tr>
<td>Position after heating</td>
<td>0.5</td>
</tr>
</tbody>
</table>
| Crystallization | a) 58-78 mM MnCl₂, 50-80 mM KCl  
                 b) 74-94 mM MnCl₂, 50-80 mM KCl |
| Crystals | a) rods (centered NCP)  
          b) bipyramidal blocks (off-centered NCP) |
| Diffraction limit | a) 4-6 Å  
                      b) very poor diffraction |

Table 4.2: Crothers sequence NCP crystals.
Characterization and crystallization of symmetric Crothers sequence NCP (cs146).
4.2.4 Human alpha satellite sequences

**General**

Alpha satellite sequences are stretches of tandemly repeated DNA that can span over thousands of base pairs in the genome. Such sequences are found in the constitutive heterochromatin region of centromeres of mitotic chromosomes of mammals. Alpha satellite sequences can account for as much as 5% of the total human genome. In humans, the higher order repeat is made from 2-14 (depending on the chromosome) 171 bp alpha satellite monomers arranged in imperfect, direct repeats. On the human x-chromosome, 12 monomers add up to a 2 kbp fragments which is then tandemly repeated over as much as 3 Mbp (Waye & Willard, 1985; Willard & Waye, 1987). See Appendix II for a complete sequence alignment. The monomers are 65-85% identical to each other. It has long been known that nucleosomes are rotationally phased on alpha satellite DNA. The translational position, however, was not clear. The 171 bp monomers have a 17 bp long binding site for CENP-B which is most likely involved in heterochromatin condensation (Masumoto et al., 1989). Mapping of a nucleosome reconstituted onto an asymmetric 151 bp fragment containing the sequence between the CENP-B sites on a consensus monomer showed a positioned nucleosome (A. Flaus, unpublished results).

**Repeat 8 derived sequences**

Sequences α3, 4, 8, 9 and 10s were derived from human alpha satellite DNA repeat 8 as shown in figure 4.4 (and Appendix II). The position of the histone octamer dyad as mapped on the asymmetric consensus fragment is indicated with an arrow (mapped by A. Flaus, unpublished results). Even-number, symmetric fragments were designed to have their center adjacent to this base, passing between base pairs in the
double-stranded DNA. For these palindromic sequences, an EcoRI restriction enzyme site with a 4 bp overhang was used to construct the DNA-center. HinfI, which creates a 3 bp overhang, was used for the construction of odd-number fragments having a base pair, the central base of the overhang, at their center. See Appendix I for an overview of the cloned sequences. Data from crystals containing these DNA sequences are summarized in table 4.3a.

Figure 4.4: Construction of human alpha satellite sequences.
A segment of the genomic sequence of repeat 8 is shown at the top. The exact positions of the halves of the symmetric fragments are shown underneath. On the left end of the DNA is the position of the dyad axis or center in the symmetric full length fragment. Note that for technical reasons, bases at the center had to be replaced by the recognition sequence for EcoRI (146mer) or HinfI (147/145mer). The last three bases at the right end were replaced by the EcoRV recognition sequence.

None of these sequences, when assembled into core particles, gave a centered position. However, all repeat 8 fragments could be shifted easily into that position. After shifting, the 146 bp fragments occupy the same positions as csl146, namely with the histone dyad axis going through the body of the base pair adjacent to the DNA dyad axis. As a consequence the resulting core particle is still not entirely symmetric. However with the 147/145 bp series of sequences, the octamer dyad axis was identical to the DNA dyad axis going through the central base pair and thus making a completely symmetric core particle (except for the central base pair).
As can be seen from table 4.3a, the conditions for crystal growth were very similar for all the fragments. The diffraction limits for NCP crystals with the sequences \( \alpha 3, 9 \) and \( 10s \) were practically identical as determined using synchrotron radiation. It was interesting to note that the crystals were not equally sensitive to the MPD soaking procedure. Whereas \( \alpha 8s \) never suffered from cracks, \( \alpha 3s, \alpha 10s, \alpha 9s \) and \( \alpha 4s \) had an increasing tendency to crack, with \( \alpha 9/4s \) crystals essentially disintegrating. However, the diffraction quality of \( \alpha 9s \) crystals was poor already before MPD soaking.

<table>
<thead>
<tr>
<th>Seq. name</th>
<th>( \alpha 3s146 )</th>
<th>( \alpha 4s146 )</th>
<th>( \alpha 8s147 )</th>
<th>( \alpha 9s147 )</th>
<th>( \alpha 10s145 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position after reconstitution.</td>
<td>9.5</td>
<td>9.5, 18.5,(0)</td>
<td>~10</td>
<td>9/10</td>
<td>9/10</td>
</tr>
<tr>
<td>Shifting to center position</td>
<td>&lt;1h at 37° C</td>
<td>&lt;1h at 37° C</td>
<td>&lt;1h at 37° C</td>
<td>&lt;1h at 37° C</td>
<td>&lt;1h at 37° C</td>
</tr>
<tr>
<td>Position after shifting</td>
<td>0.5</td>
<td>0.5</td>
<td>(0)?***</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Crystallization</td>
<td>58-78 mM MnCl(_2), 50-80 mM KCl</td>
<td>58-78 mM MnCl(_2), 50-80 mM KCl</td>
<td>62-82 mM MnCl(_2), 50-80 mM KCl</td>
<td>62-82 mM MnCl(_2), 50-80 mM KCl</td>
<td>62-82 mM MnCl(_2), 50-80 mM KCl</td>
</tr>
<tr>
<td>Crystals</td>
<td>rods</td>
<td>poor quality rods (cluster-like)</td>
<td>rods</td>
<td>blocks</td>
<td>rods/blocks</td>
</tr>
<tr>
<td>Diffraction limit</td>
<td>2 Å</td>
<td>-</td>
<td>2 Å</td>
<td>20 Å?</td>
<td>2 Å</td>
</tr>
<tr>
<td>Unit cell dimensions</td>
<td>( a = 105.0 ) Å</td>
<td>( b = 180.0 ) Å</td>
<td>( c = 109.3 ) Å</td>
<td>not exactly det. but similar to ( \alpha 3s ) NCP</td>
<td>not exactly det. but similar to ( \alpha 3s ) NCP</td>
</tr>
<tr>
<td>Remarks</td>
<td>***Very low reconst. yield</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3a: Human alpha satellite repeat 8 NCP.
Characterization and crystallization of human alpha satellite DNA core particles. Determination of the unit cell dimensions was based on precession photographs.

It is surprising that the movement of the DNA sequence dyad by one base pair (\( \alpha 3s146 \rightarrow \alpha 4s146 \)) can make a difference between diffraction to 2.0 Å and not useful. The crystal structures presented in chapter 5 and 6 will provide partial explanations for these observations.
Repeat 1 and consensus derived sequences

The sequence pairs α6s/α7s and α2s/α5s were made with the same starting positions as α3s/α4s except that repeat 1 and the satellite consensus sequence were used as templates, respectively (see Appendix I/II for the complete sequences). The repeat 1 sequences α6/7s146 were significantly bent as compared to the other sequences discussed earlier (compare lanes 6 and 7 in fig 4.5 with lanes 1-5). Another indication of this was the instability of the multicopy plasmids used for expression of these sequences at copy numbers greater than 12-16 (see material and methods section).

Nevertheless, all but α5s could be reconstituted into core particles. Again, mainly particles with off-centered DNA were obtained. The repeat 1 sequences seemed to be very stable against centering the DNA. Whereas α6s146 NCP had to be incubated for 1 h at 55° C to center the DNA on the particle, incubation for more than 5 h at the same temperature could not completely shift α7s146 into the centered position.
Chapter 4

Crystallization of the NCP

<table>
<thead>
<tr>
<th>Seq. name</th>
<th>αs146</th>
<th>α5s146</th>
<th>α6s146</th>
<th>α7s146</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position after reconstitution</td>
<td>0.5, 9.5, 18.5</td>
<td>-0, 10, unusual pattern</td>
<td>15.5 (0.5)</td>
<td>~15.5</td>
</tr>
<tr>
<td>Shifting to center position</td>
<td>&lt;2h at 37°C (appears to)</td>
<td>1h at 55°C</td>
<td>&gt;5h at 55°C</td>
<td></td>
</tr>
<tr>
<td>Position after heating</td>
<td>0.5</td>
<td>0.5?</td>
<td>0.5</td>
<td>?</td>
</tr>
<tr>
<td>Crystallization</td>
<td>58-78 mM MnCl₂, 50-80 mM KCl</td>
<td>no</td>
<td>54-74 mM MnCl₂, 50-80 mM KCl</td>
<td>no</td>
</tr>
<tr>
<td>Crystals</td>
<td>very poor quality rods (cluster like)</td>
<td>-</td>
<td>ugly rods</td>
<td>-</td>
</tr>
<tr>
<td>Diffraction limit</td>
<td>-</td>
<td>-</td>
<td>2.0 -2.3Å</td>
<td></td>
</tr>
<tr>
<td>Unit cell dimensions</td>
<td>-</td>
<td>a= 104.05 Å</td>
<td>b= 173.25 Å</td>
<td>c= 108.96 Å</td>
</tr>
<tr>
<td>Remarks</td>
<td>No sufficient amounts of NCP obtainable</td>
<td>-</td>
<td>No complete shift possible</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3b: Consensus and repeat 1 human alpha satellite NCP.

The intrinsic bending of these sequences seems to inhibit proper orientation on the histone octamer. Based on the crystal morphology, the high diffraction quality of the α6s146 crystals came somewhat unexpected. Even though they proved to be rather sensitive to the soaking procedure (the final concentration was 24% MPD as well), the diffraction limit was at 2.0 Å in the c axis and at 2.3 Å in the other two directions. The consensus sequence based fragments gave no usable crystals.

4.2.5 Optimization of crystal growth and handling

Optimization of crystal growth conditions and crystal handling was mainly done on crystals of α3s core particles. All steps such as crystallization, soaking and cooling/mounting were optimized during this process because the initial conditions were established for quite different material.

111
Chapter 4  
Crystallization of the NCP

Crystallization

Usually NCP was crystallized in the presence of Mn$^{2+}$ ions as the precipitant. Apart from that, other bivalent cations such as Ba$^{2+}$, Ca$^{2+}$ and Mg$^{2+}$ were tried. Whereas the first two cations did not yield crystals, small crystals could be obtained with Mg$^{2+}$ as the precipitant. However, no diffraction pattern could be recorded. The optimal temperature for crystal growth (4, 10, 20, 30° C were tried) was still 20° C.

For crystallization screening purposes, 4 μl hanging drop setups were used. However, crystals from these experiments were generally rather small. Therefore, a new sitting drop geometry was used to grow diffraction quality crystals. Drops of 14 μl containing NCP were immersed in micro-bridges containing 12 μl of DC200 110 Pa.s Silicon oil (Fluka). The oil had mainly two effects. It slowed down the equilibration time by reducing the exposed drop surface, and minimized the contact area between NCP drop and plastic. These experiments were performed in the same multiwell plates as used for hanging drop experiments. Together with a narrow gradient between drop (76 mM MnCl$_2$, 60 mM KCl, 20 mM KCacodylate pH 6) and the reservoir (300 μl of 43 mM MnCl$_2$, 48 mM KCl) one crystal per drop could be grown reproducibly within 2-4 weeks (as compared to a few days for hanging drops). The average size of the crystals with a rhomboidal cross section exceeded 0.3*0.25*1.0 mm, but often there was a hole in the center of the crystal, parallel to the long crystal axis. Micro-seeding in the supersaturation zone was also tried with success. However, crystals grown by this method showed no enhanced diffraction qualities.

Crystal soaking

As discussed earlier, the diffraction quality could be improved by stepwise soaking the crystals into a final concentration of 24% MPD. Fortunately, MPD also acted as a cryoprotectant which would be needed for low temperature data collection in any case. Because the optimization of the MPD concentration was originally done for
crystals of limited diffraction strength, other cryoprotectants such as glycerol and xylitol were evaluated in the 20-30% range. The results again underlined the unique role of MPD in increasing the diffraction strength of NCP crystals. Earlier experiments had shown that variation of the MPD concentration could alter the cell dimensions as well as the diffraction limit of the crystals. The concentration range of 22-30% was sampled for the new crystals. A concentration of 24% was still optimal, although the diffraction strength remained essentially unchanged between 22 and 26% MPD. The soaking procedure was changed from the addition of 40% MPD to the low percentage MPD solutions containing the crystals to a more reproducible soaking procedure using solutions of intermediate MPD concentrations and step times of 3-5 minutes (see methods section).

Crystal mounting and cryocooling

Previously, NCP crystals had always been mounted in tapered capillaries for data collection at cryotemperatures. The advantage of the method is the uniform amount of scattering material in the X-ray beam, easy handling, and good stability of the crystals. On the other hand, there are several disadvantages: 1) the volume to be cryocooled is large which slows down the cooling process and adds background to the diffraction pattern. 2) the orientation of the crystal is restricted because the morphologically long axis (c-axis) is always parallel to the capillary and to the rotation axis in a one circle geometry used with MAR image plate systems. Another problem was the presence of a strong and well located ice ring around 3.8 Å which restricted data processing in this region of the diffraction pattern. A slight change in the cooling protocol, namely the reduction of the liquid propane temperature from -90° C to -120° C, resulted in a flattening-out of this ice ring and improved data collection. Subsequently, cryocooling of capillary mounted crystals was done by slow cooling from 4° C to -16° C in a cold
air stream, flash freezing into liquid propane at -120° C, and a final quick transfer into the cold nitrogen gas stream at approximately -180° C. Using the optimized procedures discussed so far together with the α3s DNA, it was possible to collect data to 2.8 Å resolution (K. Luger unpublished results).

In order to overcome the disadvantages of the capillary mounting, we tried to mount the NCP crystals in fiber loops. This, however, needed protection of the crystals from drying out and skipping of the first cooling step with the cold air stream. The NCP crystals were not stable in 24% MPD alone. After transfer into the loop, they dried out within seconds which made the handling impossible. A solution to the problem was found by addition of 2-5% trehalose in addition to the 24% MPD (Colaço et al., 1992). Crystals were then stable for at least a week in the nitrogen gas stream. The 4° C to -16° C step of the original cooling protocol could be skipped without obvious consequences, although the mosaic spread of the crystals may have increased slightly. Due to the weak diffraction, however, it was generally difficult to analyze such effects using the laboratory rotating anode X-ray generator. The signal to noise ratio was improved further compared to the capillary mounting methods by shaping the loops to tightly accommodate the crystals, eliminating all surrounding buffer.

The loop technique also offered a solution to another problem possibly arising when data to 2.0 Å was to be collected. The maximum oscillation angle, before neighboring reflections overlap on successive exposures, is a function of both the resolution shell and the unit cell dimensions. With the long axis (b= 180 Å) perpendicular to the detector (ac plane in reflecting position) the maximum angle allowed reduces to 0.64° at 2.0 Å (see table 4.4). Subtracting the mosaic spread of the crystals, which was assumed to be between 0.3 and 0.8° (full width of the reflection), the actual oscillation angle would be restricted to very small values.
Chapter 4

Crystallization of the NCP

<table>
<thead>
<tr>
<th></th>
<th>ac in reflecting position (h0l)</th>
<th>bc in reflecting position (0kl)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.7 Å</td>
<td>2.0 Å</td>
</tr>
<tr>
<td>Maximum oscillation angle allowed before neighb. reflection overlap in the diffraction pattern.</td>
<td>0.86°</td>
<td>0.64°</td>
</tr>
</tbody>
</table>

**Table 4.4: Maximum oscillation angles.**

The rotation axis is parallel to the crystallographic c axis. Unit cell dimensions: a = 105.0 Å, b = 180.0 Å, a = 109.3 Å as determined from precession photographs.

Mounting crystals along the crystallographic b-axis would position the short reciprocal spacing along the rotation axis and thus allow the use of larger oscillation angles. Unfortunately, the crystallographic b-axis is parallel to the long diagonal of the rhomboidal cross section of the crystals. Using a kinked loop construction as shown in figure 4.6 allowed the mounting with the b-axis approximately parallel to the capillary. The approach was not straightforward because all crystals had to be trimmed to approximately equal dimensions in all directions which destroyed up the 50% of the crystals. The second occasion to destroy crystals was during the actual mounting procedure. The crystals had to be fished from the buffer solution while sitting perpendicular on top of the loop plane. However, for collection of the 2.0 Å data set, more than 70 crystals (from more than 200 crystals grown) could be mounted.

![Figure 4.6: Loops for mounting crystals along the crystallographic b-axis.](image)

Crystallographic axis assignment: a: short diagonal of cross section, b: long diagonal, c: long axis of the crystal. First the thread loop was glued into the open end of a 0.2 mm glass capillary. The loop was then shaped and kinked with a pair of tweezers under the microscope. During mounting, the loop plane generally filled with a thin buffer layer so that the frozen end product was mechanically stable.
For data collection, the cryocooling procedure for loop mounted crystals was flash cooling from RT to approximately -120° C (liquid propane) and quick transfer into the cold nitrogen gas stream at approximately -180° C. Crystals were soaked into a solution of 24% MPD, 2% trehalose, 40 mM MnCl₂, 37 mM KCl, and 10 mM K/Cacodylate pH 6 prior to cooling (see material and methods).

4.2.6 Crystallization in microgravity

The increasing number of space flights open for external experiments has lead to the possibility of performing crystallization experiments under microgravity conditions. Initiated by the microgravity program of the European Space Agency (ESA), the Advanced Protein Crystallization Facility (APCF) was designed and built by Dornier in Friedrichshafen (P. Lautenschlager, L. Potthast, R. Bosch). The facility provides room for total 48 reactors and is designed to be used on several different space crafts. Three different reactor types are available: hanging drop (HD) of 8 and 88 µl maximum volume, free interface diffusion (FID) and dialysis reactors.

Based on theoretical considerations, crystal growth under microgravity conditions should be optimized. Reasons for this are the reduction of density driven effects such as convection and sedimentation. Several reports from previous experiments indicated that the mosaic spread and the diffraction limit could indeed profit from crystallization in space (Borisova et al., 1996; Vaney et al., 1996; Weisgerber & Helliwell, 1995). However, the experimental set up seems to be very important as well. It has to be considered that, due to the predesigned reactors, the crystallization geometry cannot be adapted to the optimal laboratory setups. The same is true for the crystallization conditions. The gradient between protein and buffer volumes has to be chosen to allow crystallization in the time span available during the mission. Additionally, other
experimental factors such as vibrations or surface effects may be more important than the presence or absence of gravity.

We had the opportunity to use four 88 μl hanging drop reactors at 20° C during the spacetab flight on Space Shuttle mission STS-78 in June/July 1996. The hanging drop reactors were originally designed to accommodate a protein volume decreasing in the course of the experiment. However, our crystallization conditions start at a high salt concentration in the core particle volume which is then equilibrated against a lower ionic strength solution with the consequence of increasing the drop volume. To accomplish the experiments, the HD reactors had to be modified to allow the handling of larger volumes at the end of the experiment (see reactor manual for details).

<table>
<thead>
<tr>
<th>Reactor</th>
<th>Initial drop size</th>
<th>Reservoir buffer concentrations (MnCl₂/ KCl) 700 μl in two 350 μl absorbers</th>
<th>Crystal characterization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1S</td>
<td>50 μl</td>
<td>43/48</td>
<td>large number of small rods (100-200<em>100-200</em>500-1000 μm); often hexagonal; few holes</td>
</tr>
<tr>
<td>1G</td>
<td>&quot;</td>
<td>43/48</td>
<td>large number of small crystals</td>
</tr>
<tr>
<td>2S</td>
<td>&quot;</td>
<td>35/40</td>
<td>as 1S</td>
</tr>
<tr>
<td>2G</td>
<td>&quot;</td>
<td>35/40</td>
<td>large and small crystals</td>
</tr>
<tr>
<td>3S</td>
<td>&quot;</td>
<td>30/35</td>
<td>few big crystals of &gt;300<em>300</em>1000 μm; hollow rods as usual</td>
</tr>
<tr>
<td>3G</td>
<td>&quot;</td>
<td>30/35</td>
<td>large and small crystals</td>
</tr>
<tr>
<td>4S</td>
<td>&quot;</td>
<td>24/29</td>
<td>few large crystal with a regularly structured surface of the inner hole</td>
</tr>
<tr>
<td>4G</td>
<td>&quot;</td>
<td>24/29</td>
<td>large and small crystals</td>
</tr>
</tbody>
</table>

Table 4.5: APCF space experiments.
S stands for space, G for ground experiments. The starting protein concentration was at 8 mg/ml in 84 mM MnCl₂, 60 mM KCl and 20 mM potassium/cacodylate pH 6. For the space reactors, half of the initial drop volume was expelled from the syringe by using a 1 mm bushing (see reactor manual for details).

The hanging drop reactors were loaded according to the APCF reactor manual provided by Dornier. Appendix VI shows a drawing of the HD reactor. The crystallization conditions and results are shown in table 4.5. The morphology of the
space grown crystals did not differ from the ground control crystallization experiments. There were generally many nucleations in both systems with the consequence of limited crystal size. Most nucleations were observed on the Teflon piston of the syringe.

A first analysis of crystals at RT in HL-0 (without MPD, see materials and methods) on the laboratory rotating anode X-ray source did not show obvious differences in the diffraction pattern. Later, crystals grown in space as well as from the ground control crystallizations were analyzed at the SNBL at the E.S.R.F. in Grenoble (table 4.6). The results of the first analysis were confirmed. The crystals grown in space did not show enhanced diffraction strength. However, the large difference in diffraction resolution between crystals before and after cryocooling was surprising. The high resolution diffraction was lost due to radiation damage before the reflections could be measured. It would be interesting to see whether or not the high resolution reflections could be measured with a stronger X-ray beam. With respect to the mosaic spread of the crystals, no significant differences were seen. However, the MAR imaging plate used for data collection, was not ideal for such measurements.

<table>
<thead>
<tr>
<th>conditions</th>
<th>ground control (a,b,c) Å</th>
<th>space (a,b,c) Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>HL-0, RT</td>
<td>3.8/3.8/3.0</td>
<td>3.8/3.7/2.9</td>
</tr>
<tr>
<td>HL-24, RT</td>
<td>3.2/3.2/2.7</td>
<td>3.2/3.2/2.8</td>
</tr>
<tr>
<td>HL-24/TL-2, 100 K</td>
<td>2.2/2.2/2.0</td>
<td>2.2/2.2/2.0</td>
</tr>
</tbody>
</table>

Table 4.6: Diffraction of space grown crystals.

Diffraction limit for α3s NCP crystals grown in space and from the corresponding ground control experiment.
4.3 Discussion

The crystallization of the nucleosome core particle has been the subject of 20 years of intense research. Initially, core particles were only available as material isolated from higher organisms. Given the heterogeneity intrinsic to such preparations, it is still surprising that large crystals could be obtained at all. The diffraction strength, however, was clearly limited. With the development of cloning techniques and the availability of restriction enzymes in the last decade, the preparation of an entirely recombinant core particle became possible. However, the choice of suitable materials remained.

The task was to find a DNA sequence that binds in a unique position to the histone octamer, which is supposed to show only minor sequence preferences. In protein crystallography, it is frequently observed that single amino acid mutations can be decisive for crystal formation. In contrast, there is only little knowledge of the characteristics of large DNA surfaces in crystallization. Although some of the parameters of DNA-protein interactions in general and for core particle in particular are well established, their effect on crystallization is largely unknown. Still, the DNA primary sequence was virtually the only determining factor in NCP crystallization. Please note, that most of the sequences discussed in this thesis formed perfectly stable core particles with respect to standard analysis methods. The source of the histone octamer seemed to be less important in crystallization, as core particles made from chicken octamer and the \( \alpha3s \) sequence diffracted to similar, high resolution as well.

As part of this thesis, 4 DNA fragments were prepared that extended the diffraction limit beyond that seen in the laboratory before. Additionally, both crystallization and crystal handling were optimized and adapted to the new material. Based in part on the results presented here, the crystal structure of the nucleosome core particle was solved to a resolution of 2.8 Å (K. Luger). More recently, a high resolution
data set (2.0 Å) was collected for the α3s NCP, and the structure of another NCP made from the α6s DNA sequence was solved to 2.6 Å resolution.

The numerous DNA sequences also provided further insight into the nature of DNA binding by the histone octamer. It was unexpected that all of the fragments generally reconstituted in off-centered positions, leaving some peripheral histone-protein interactions unused. As the 146 bp core particle is an artificial entity compared to native chromatin, the salt dialysis method used for reconstitution is most likely not an exact replica of what happens in the cell. During the high salt reconstitution method, only the DNA-protein contacts at the DNA center are made because the DNA termini are dissociated above 400 mM NaCl (see chapter 3). Thus unless the DNA center is strongly positioning, there is no force for native positioning of the DNA with respect to the entire histone octamer surface. Nevertheless, the shifting behavior of the various sequences is worth consideration.

We assumed that the driving force in our shifting procedure had its origin in the maximization of histone-DNA interactions. In vivo, this is probably irrelevant because those interactions are made anyway. The enormous differences in shifting time, ranging from 10 min at 37° C (as146) to 2 h at 55° C (cs146) suggest that such sequences (or similar ones) must differ in their in vivo characteristics as well. For example, a nucleosome formed on a sequence like cs146 would not be easily displaced by sliding. With respect to the Crothers sequence, it is also interesting to note that the unusual high temperature for shifting is not paired with an intrinsic bend as seen in the α6s DNA sequence.
5. Refinement of the NCP structure

5.1 Introduction

Based on the crystals described in chapter 4, the X-ray structure of a nucleosome core particle containing the α3s DNA sequence was solved at 2.8 Å resolution using the method of multiple isomorphous replacement (K. Luger, A. Mäder, T.J. Richmond, submitted). The main reason for the somewhat arbitrarily chosen resolution limit for these data sets was the amount of synchrotron beam time available to solve the structure in the first place. The choice had to be made between collecting many heavy atom derivatives versus extending the high resolution limit. To assure successful MIR, data collection of as many heavy atom derivatives was favored. The current structure has good density for the histone globular core and for most of the DNA. In contrast, only about half the N-terminal histone tails are traced.

Fortunately, the diffraction of the crystals was not limited to 2.8 Å but extended to 2.0 Å in the best direction (along the crystallographic c-axis). Even though this data is weak, we were able to collect a high resolution data set in a recent synchrotron run. The new data will increase the accuracy of the atomic model substantially, should allow the localization of water molecules and metal ions, and possibly improve the electron density for the poorly defined regions in the 2.8 Å map.

The structural features presented in this chapter are the result of the crystallographic refinement of the α3s NCP structure to a resolution of 2.3 Å. Further refinement to 2.0 Å is currently in progress.
5.2 Results

5.2.1 Data collection

2.0 Å data was collected on insertion device 9 (BL-3) at the European Synchrotron Radiation Facility in Grenoble (see E.S.R.F. publications for details). The storage ring was running in hybrid mode at a maximum ring current of 140 mA. The half life of the electron beam was on the order of 30-45 h, with refills every 11-12 h (at 100 mA). The undulator beam was focused to a size of 150 μm horizontally and 300 μm vertically (size at half peak intensity) and the wavelength was set to λ = 0.8469 Å. Data collection was performed on a 30 cm MAR detector mounted on the original MAR base.

<table>
<thead>
<tr>
<th>Number of crystals</th>
<th>Mounted along; off-sets</th>
<th>Number of exposure/crystal</th>
<th>Δϕ</th>
<th>Exp. time (sec)</th>
<th>Resolution limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>b; ac approx. 10° off</td>
<td>20</td>
<td>0.4</td>
<td>90</td>
<td>2.0 Å</td>
</tr>
<tr>
<td>3</td>
<td>c; ab approx. 10° off</td>
<td>20</td>
<td>0.4</td>
<td>90</td>
<td>2.0 Å</td>
</tr>
<tr>
<td>1</td>
<td>b; ac approx. 10° off</td>
<td>115</td>
<td>0.8</td>
<td>10</td>
<td>3.5 Å (3.0 Å)</td>
</tr>
</tbody>
</table>

Table 5.1: Data collection summary.

The special loop construction discussed in the previous chapter was used to mount crystals along the b-axis. Mountings with the c-axis parallel to the rotation axis were achieved by using loops sized to the crystal shape.

Table 5.1 gives an overview of the crystals used for data collection. 90 degrees (-4° to 96° starting from the bc principal zone) were collected with 24 crystals increasing starting φ by 4° per crystal. The oscillation angle was at 0.4° and 8° were collected per crystal. Three crystals mounted along the crystallographic c-axis were used to collect the blind zone not covered by the crystals mounted along the b-axis (-1°
to 15° were collected starting from the bc principal zone). As a consequence of the long exposure times required, a significant portion of the low resolution data was missing due to saturation. Collection of a low resolution data set provided this missing data.

5.2.2 Data processing

The crystals were indexed to 2.0 Å resolution using locally written programs (T.J.R., unpublished). Diffraction data (high and low resolution) from the crystals mounted along the b-axis showed a linear decrease in the mean intensity of the reflections from 0kl to hk0. This was most likely a consequence of the mounting method used. Since the crystals were surrounded by virtually no liquid, and because the crystal dimensions were not equal in all three dimensions, the absorption conditions were anisotropic in the different reciprocal space directions. When available, local scaling of data should have been able to reduce this effect. For the moment, it has not been corrected. As a consequence, the three crystals mounted along the c-axis could not be merged satisfactorily with the rest of the data, even though they merged satisfactorily as a group. However, it turned out that the three crystals were not absolutely needed for a complete data set because of the large offset used with the other crystals.

The diffraction patterns showed a significant drop off in reflection intensity around 2.8 Å. This resulted in a rather poor signal to noise ratio for the high resolution data (which was the reason for the relatively long exposure times in the first place). Traces of ice gave strong diffraction as compared to the diffraction from the crystal. This was seen in form of an 'ice' ring around 2.25 Å. Several crystals suffered from this problem. The origin of the ice was probably traces of water accumulated in the liquid propane and/or nitrogen used to freeze and store the crystals.

Based on the Rmerge and in order to maximize the completeness of the data, a set of the 12 best high resolution crystals (mounted along the b-axis) and the low resolution
crystal were merged into one data set. Six more crystals could have been used, whereas
the remaining six crystals refused to merge satisfactorily. This was not so surprising as

<table>
<thead>
<tr>
<th>Resolution (Å)</th>
<th>20-2.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflections (observations/unique)</td>
<td>1.5*10^6/93572</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>98.8%</td>
</tr>
<tr>
<td>( R_{\text{merge}} ) (at 2.3 Å)</td>
<td>6.8% (18.7%)</td>
</tr>
<tr>
<td>Cell dimensions (a,b,c) in Å</td>
<td>106.03/182.35/109.49</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Resolution (Å)</th>
<th>8-2.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-factor/free R-factor (%)</td>
<td>23.7/30.5</td>
</tr>
<tr>
<td>Reflections, I &gt; 2σ (working/test)</td>
<td>87705/2245</td>
</tr>
<tr>
<td>Number of atoms</td>
<td>12415</td>
</tr>
<tr>
<td>mean B-factor (Å^2)</td>
<td>53.4/ 99.7</td>
</tr>
<tr>
<td>protein/DNA</td>
<td>12415</td>
</tr>
<tr>
<td>r.m.s.d. bond length</td>
<td>0.009/ 0.004</td>
</tr>
<tr>
<td>protein/DNA (Å)</td>
<td>1.31/ 0.821</td>
</tr>
<tr>
<td>r.m.s.d. bond angle</td>
<td>1.31/ 0.821</td>
</tr>
</tbody>
</table>

\[ \sum \sum \frac{t_{hkl} - \bar{t}_{hkl}}{t_{hkl}} \], where \( \bar{t}_{hkl} \) is the mean of measurements for a single \( hkl \).

Table 5.2: Statistics of the a3s NCP data set.

it was clear already during data collection that not all crystals were of the same quality;
some may have suffered during the rather complicated mounting procedure. In fact,
more than 70 crystals were mounted and examined in the beam in order to find the
crystals suitable for data collection.

A first high resolution data set to 2.3 Å resolution was derived from this reduced
set of crystals. Table 5.2 gives a summary of the data quality and the refinement
statistics. The atomic model of the 2.8 Å structure of the α3s NCP was used as starting
model. Crystallographic refinement was done exclusively with the program package X-
PLOR (Brünger, 1992) dividing the structure factors in a working set (R-factor) and a
2.5% test set (free R-factor). The refinement was started with a rigid body refinement
(20-4 Å) to allow the model to adapt to the slightly different cell dimensions of the new
data set (rigid body = entire model). This was followed by two cycles of
positional/grouped B-factor refinement in the 7-2.6 Å and 7-2.3 Å resolution shells. After refining an individual isotropic B-factor, the R-factor could be further improved by refining an overall anisotropic B-factor. Finally, statistics were optimized by applying a bulk solvent correction using a solvent radius of 1.0 Å.

This initial model was then refined by three cycles of model building and crystallographic refinement. Most changes were done to the DNA part of the model and to the water/ion positions, whereas the protein core remained almost unchanged. During the refinement cycles the resolution range was extended to 8-2.3 Å and the crystallographic pseudo-energy weight was set to \( w_a = 10^6 \). The upper limit for the maximum allowed B-factor was increased to 200 Å\(^2\). During positional refinement a harmonic restraint of 10 kcal mole\(^{-1}\)Å\(^6\) was applied to the protein main chain C-alpha atoms, the DNA phosphorus atoms, the water oxygen and the manganese ions. Simulated annealing in torsion angle space with the same harmonic restraints as described above could not improve the model significantly. Table 5.3 gives a summary of the residues included in the current model. The changes to the input model were minimal.

<table>
<thead>
<tr>
<th>Protein chain</th>
<th>Residues in the model</th>
<th>Well refined residues</th>
<th>Globular core</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3-1</td>
<td>38-135</td>
<td>39-134</td>
<td>27-129</td>
</tr>
<tr>
<td>H4-1</td>
<td>20-102</td>
<td>22-102</td>
<td>20-102</td>
</tr>
<tr>
<td>H2A-1</td>
<td>4-120</td>
<td>14-120</td>
<td>12-118</td>
</tr>
<tr>
<td>H2B-1</td>
<td>22-122</td>
<td>28-122</td>
<td>24-125</td>
</tr>
<tr>
<td>H3-2</td>
<td>20-135</td>
<td>37-135</td>
<td></td>
</tr>
<tr>
<td>H4-2</td>
<td>16-102</td>
<td>19-102</td>
<td></td>
</tr>
<tr>
<td>H2A-2</td>
<td>12-119</td>
<td>13-119</td>
<td></td>
</tr>
<tr>
<td>H2B-2</td>
<td>24-122</td>
<td>29-122</td>
<td></td>
</tr>
<tr>
<td>( \alpha_3 )s DNA</td>
<td>1-146</td>
<td>1-146</td>
<td></td>
</tr>
<tr>
<td>( \text{Mn}^{2+} )</td>
<td>10</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Water/ions</td>
<td>10</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.3: List of residues in the \( \alpha_3 \)s NCP 2.3 Å structure.
The last column lists the shortest stable peptide fragments after trypsin digestion of nucleosomes (Böhm & Crane-Robinson, 1984).
5.2.3 Water molecules and manganese ions

The crystallization of the NCP is unusual in the sense that bivalent cations were used as the only precipitant. The ion most successfully used in producing high quality crystals for the different octamers and DNA fragments was manganese. Therefore, it may have been expected to find Mn\(^{2+}\) ions in prominent positions between phosphate strands forming crystal packing interactions. Effectively, the Mn\(^{2+}\) ions found in the structure were predominately associated with edges of bases rather than with the DNA phosphates. This is in good agreement with observations that transition metals such as Mn\(^{2+}\) prefer binding to nucleotides rather than to phosphates alone (reviewed by Saenger, 1984). However, the Mn\(^{2+}\) ion defined by far the best is located at a protein-protein crystal contact between two acidic amino acids and is not associated with DNA at all (discussed in chapter 6).

<table>
<thead>
<tr>
<th>Number of Mn(^{2+})</th>
<th>Crystal contact</th>
<th>Contacted atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 yes</td>
<td>Yes</td>
<td>Phosphates of A67, A146 and B213</td>
</tr>
<tr>
<td>1 yes</td>
<td>Yes</td>
<td>Fills mg; A86(Thy) O2 and H2B-1 Ser120</td>
</tr>
<tr>
<td>1 yes</td>
<td>Yes</td>
<td>Phosphates of A140/A141 and B246(47) (Gua) N7</td>
</tr>
<tr>
<td>6 no</td>
<td>No</td>
<td>N7 of Gua A100, A121, A134, B217(76), B267(26), B280(13)</td>
</tr>
</tbody>
</table>

Table 5.4: Manganese ions in the α3s NCP 2.3 Å structure.

Numbers in brackets for bases starting with B-numbers indicate the base pair number in which the base occurs. The numbering of the DNA is as follows: one molecule of the DNA double strand is named A1 through A146 whereas the other molecule is named B147 through B292. In base pair 1, base A1 is therefore base pairing with base B292.

The identification of manganese ions associated with the DNA was based on the comparison of the B-factors obtained after crystallographic refinement of the positions as either water molecules or Mn\(^{2+}\) ions. As it is unlikely for a water molecule to refine with a B-factor lower than that of the macromolecule nearby, such water molecules were replaced with Mn\(^{2+}\) ions. Table 5.4 gives a list of the Mn\(^{2+}\) ions associated with
\(\alpha3s\) DNA. Particularly interesting are the 7 ions listed in the two bottom rows. These ions are close to the N7 atom of Gua residues in the major groove of the DNA. The distances to the nitrogen atoms are between 2.5 and 2.8 Å. The distances to the exocyclic O6 of the same Gua were between 2.9 and 3.9 Å. A third chelating atom was the non-bridging oxygen atom of the phosphate of the same or the preceding base in 5' direction. Those distances varied between 3.5 and 6.2 Å. Apart from these three partners, other close polar atoms were likely to participate in metal binding (see discussion in chapter 6).

### 5.3 Discussion

The crystallographic refinement of the 2.8 Å atomic model of the \(\alpha3s\) NCP to 2.3 Å resolution did not reveal many new features with respect to the protein and DNA components of the model. The histone tails poorly defined in the starting model remained so in the 2.3 Å structure. The same is true for the DNA segments with weak electron density and associated high B-factors. This was not so surprising as the work was based on the same core particle (and crystals) in both cases.

Because the actual determination of the 2.8 Å structure of the \(\alpha3s\) NCP was not part of this thesis (see Luger et. al., submitted), only the general features of the complex are presented here. The main focus will be on the DNA and its interactions with the histone octamer, since this is the main topic of this thesis. Detailed analysis of the \(\alpha3s\) DNA will be presented along with data from the \(\alpha6s\) DNA in chapter 6.
5.3.1 General view of the NCP

Figure 5.1: The model of the α3s NCP.

Ribbon diagrams of the NCP using the program MidasPlus (Ferrin et al., 1988). The histone proteins are colored blue (H3), green (H4), yellow (H2A) and red (H2B). The positions of the N-termini are as indicated. The DNA strand A is shown in magenta whereas strand B is colored white. The picture at the top shows one copy of the histone proteins and the first 73 bp of the DNA. The DNA center is near the first contact site between H3/H4 and the DNA. The contact sites between DNA and histone protein are numbered starting from the dyad.
Figure 5.1 shows a model of the $\alpha$3s NCP using the ribbon representation for the histones. On the top, only one half of the particle is shown whereas the entire model is shown underneath. As expected, the histones are mainly alpha-helical proteins. The histone-fold motive between H3/H4 and H2A/H2B, respectively, is formed by the close packing of helix $\alpha$1 (short), $\alpha$2 (long) and $\alpha$3 (short) of each histone. Histone H3 has a fourth helix at the N-terminus (named $\alpha$N) and so does H2B which has an additional helix at the C-terminus (named $\alpha$C). In addition to the 3 helices forming the histone-fold, H2A has 3 additional shorter helical stretches.

Contacts between the two H3/H4 'dimers' are exclusively made by histone H3 at the C-terminus of the long helix $\alpha$2 and helix $\alpha$3, which form a four-helix bundle at the central region of the DNA. Thereby, amino acids H3 Ile130, Leu126, Cys110 and His113 form a hydrophobic core. In contrast, the contact between dimers around H2A Asn38 seems to be of minor importance. However, the dimers make extensive contacts to both H4 and H3 by forming a bridge between the two H3/H4 'dimers'. The most extensive contacts are formed as a small hydrophobic core around H4 Tyr88 (C-term of $\alpha$2) and H2B Tyr80 (C-term of $\alpha$2) which stack on top of each other. This dimer-tetramer interface is near the DNA between contact site 3 and 4. The same copy of H2B contacts the other copy of H4 in the region of H4 Tyr98 (helix $\alpha$3). It is in this region as well where H2A makes extensive contacts with H3 and H4. These contacts are mainly mediated by hydrogen bonds.

The trypsin sensitive, N-terminal histone tails as defined in table 5.3 (Böhm & Crane-Robinson, 1984), remain largely untraced or show weak electron density. Given the knowledge from the crystal structure, specific binding to (nucleosomal) DNA seems unlikely, the interactions are most likely electrostatic. Some tails (N-term of H2A-1, H3-2) were traced, although the density was weak. The protein chain seems to follow the DNA grooves of symmetry related particles without forming secondary structure elements.
The histone octamer forms a left-handed positively charged ramp for the DNA supercoil. The surface to which the DNA is bound is surprisingly smooth and regularly bent. The histone surface has the form of a rim rather than a groove (see figure 5.2). This reduces the contact sites between DNA and octamer to a minimum. Consequently, the accessibility of the DNA to the solvent is much higher than what might have been expected for such a complex. The positively charged surface seems to function like sticky tape. The DNA is attracted by the large number of positive charges on the ring-like surface and held in place by phosphate-protein main chain interactions. At the top and bottom faces of the histone octamer, where no DNA is bound, the local charge is predominately negative (figure 5.2).

Figure 5.2: Surface potential representation of the histone octamer.

The surface potential of the histone octamer was calculated with the program GRASP (Nicholls et al., 1991). The left image shows the NCP from the top, so that the 'flat' surface of the octamer (one histone dimer) can be seen. The center of the DNA is oriented towards the bottom. The image on the right is a view down the dyad axis of the particle. The molecular surface shown is the boundary between solvent (water) and protein atoms, both as hard spheres. Negative potential is shown in red, positive potential in blue (between $-17k_B T$ and $+38.9k_B T$, where $k_B$ is the Boltzman constant and $T$ is the absolute temperature). The c66s NCP was used for the generation of the images.
The small contact interface between the histone dimers may account for the cooperativity observed in dimer-tetramer binding (Benedict et al., 1984). However, binding of one dimer could also lock the tetramer in the conformation seen in the two crystal structures and therefore enhance binding of the second dimer. A structural change in the tetramer upon dimer binding may provide an explanation for the results of the fluorescence measurements presented in chapter 3. Those results obtained in elevated ionic strength suggested that the tetramer-DNA complex falls apart shortly after (or simultaneously with) the dissociation of the histone dimers.

5.3.2 The DNA in the α3s NCP

The helical path of the DNA on the histone surface seems relatively uniform (see figure 5.1 and 5.2). There are no significant distortions other than those needed for the curvature around the histone octamer surface. However, DNA bending appears to be slightly stronger 1.5 and 4.5 helical turns away from the DNA center. From biochemical mapping experiments (Flaus et al., 1996) it was suggested that the body of a DNA base pair rather than the space between two bases is positioned on the protein dyad. The X-ray structure revealed the same orientation of the DNA resulting in a partially asymmetric particle. In the crystal structure, base pair 73 (of 146) is positioned on the protein dyad. To compensate for the missing base pair in the first DNA half as compared to the second half, the DNA is overwound (stretched) by one base pair between base pairs 50 and 60. This puts both DNA termini in identical positions, so that the NCP is largely symmetric again. The DNA ends of neighboring molecules in the unit cell are stacked on each other, thus forming an almost continuous DNA helix (see next chapter for details).

Generally, the DNA contacts the protein only via one or two phosphates per strand, at locations where the minor groove faces the octamer. The interactions are
predominately sequence unspecific and are often made to the protein main chain amide atoms. The main chain atoms generally have a well defined position, which restricts the positional freedom of the phosphates. However, they only have a partial positive charge, so that the presence of positive charged amino acids such as lysine and arginine is needed for charge neutralization. A combination of both interactions, where the phosphates are held down in position by binding to the protein main chain amide nitrogen, combined with a general charge neutralization by clustering positively charged residues around these binding sites, may represent an ideal compromise between site specific (but DNA sequence unspecific) binding on the octamer and the need for high accessibility.

There are seven points of interaction between histones and one DNA half, one approximately every ten base pairs (see figure 5.1). Each contact site has an arginine side chain inserted in the DNA minor groove. The first three contacts are made through H3/H4 such that at a given contact, each DNA strand is contacted by either H3 or H4. The next three contacts are made similarly by H2A/H2B, with H2B in the place of H3 and H2A replacing H4. Contact sites 1, 3, 4 and 6 are formed by the loops between helix α1 and α2 of one histone in combination with the loop between helix α2 and α3 from the other histone (e.g. H4 and H3 at contact 1). At contact site 1, a H3 Pro, from the same N-terminal region of αN which contacts the DNA end, is inserted in the minor groove. DNA-octamer contact sites 2 and 5 are mediated by the C-terminal regions of the α1 helices of the histones. The last contact at the DNA end is made exclusively by the N-terminus of αN (it is the other copy of H3). At this site, contacts are made predominately to one of the DNA strands, and the amino acids tyrosine and histidine are inserted into the minor groove.

Surprisingly, the crystallographic B-factors are very high for the DNA segments which are not contacted by the histone octamer. This was also documented in the high overall B-factor for the DNA. This is an indication for relatively high positional freedom of nucleosomal DNA.
6. Structure of the α6s 146 NCP

6.1 Introduction

The cloning and characterization of more than a dozen DNA sequences was described in detail in chapter 4. From this set of DNA fragments, four resulted in NCP crystals diffracting to high resolution. The structure of the α3s NCP was solved, and the general features were presented in chapter 5. In order to proof read the crystallographic information obtained by this first structure, the X-ray structure of a NCP made from a different DNA sequence was solved. The α6s DNA sequence was chosen for this, because it also diffracts to high resolution, and because the DNA primary sequence was significantly different from the α3s DNA. Figure 6.1 shows the α6s DNA with the sequence changes as compared to α3s DNA highlighted. AT to GC base pair changes are shown in magenta whereas purine to pyrimidine changes (AT to TA, GC to CG) are shown in green. The two sequences differ considerably especially in the peripheral 40 bp whereas the central region is rather homologous. Both sequences lack a regular

Figure 6.1: The α6s DNA sequence.

One half side of the α6s 146 DNA sequence is shown in the middle flanked by the α3s 146 sequence. AT to GC base pair changes are high lighted in magenta. Purine to pyrimidine changes within the same base pair are colored in green.
pattern that would position AT base pairs in regions where the minor groove contacts the histones (5, 15, ... base pairs away from the DNA center). In the α6s DNA, GC base pairs are rather equally distributed throughout the DNA fragment.

From the biochemical experiments, it is also known that the α6s DNA must have some intrinsic curvature as this fragment runs more slowly in non denaturing PAGE. Given the observation, that complete centering of this DNA on the histone octamer needed a shifting time of 1 h at 55° C, it was of interest to know what the path of the DNA on the histone octamer surface looked like. From the characterization of the crystals it was also known that the cell dimension in the direction of the end to end stacking (b-axis) of the DNA strands was approximately 6 Å shorter as compared to the α3s crystals.

### 6.2 Results

#### 6.2.1 Data collection

Data collection was performed on E.S.R.F. ID-9/BL-3 in the same run as the 2.0 Å data (see chapter 5) was collected. Therefore, the machine parameters were

<table>
<thead>
<tr>
<th>Number of crystals</th>
<th>φ range (starting from bc main zone)</th>
<th>Number of exposure/crystal</th>
<th>( \Delta \phi )</th>
<th>Exp. time (sec)</th>
<th>Resolution limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-2 - 37.5°</td>
<td>56</td>
<td>0.7</td>
<td>45</td>
<td>2.5 Å</td>
</tr>
<tr>
<td>1</td>
<td>32.5 - 61</td>
<td>47</td>
<td>0.6</td>
<td>55</td>
<td>2.5 Å</td>
</tr>
<tr>
<td>1</td>
<td>56 - 76.5</td>
<td>41</td>
<td>0.5</td>
<td>63</td>
<td>2.5 Å</td>
</tr>
<tr>
<td>1</td>
<td>65 - 69 / 75.5 - 92</td>
<td>42</td>
<td>0.5</td>
<td>63</td>
<td>2.5 Å</td>
</tr>
<tr>
<td>1</td>
<td>-1 - 92.8</td>
<td>67</td>
<td>1.4</td>
<td>6</td>
<td>4 Å</td>
</tr>
</tbody>
</table>

*Table 6.1: Data collection with α6s NCP.*

All crystals were mounted with the c-axis parallel to the rotation axis and offsets of approximately 5 degrees around the a and/or b-axis.
Chapter 6

Identical. Table 6.1 shows a summary of the data collection strategy. Four crystals were used to collect high resolution data and a low resolution data set was collected from an additional crystal. The oscillation angle was decreased when approaching the ac principle zone in order to reduce the number of overlapping reflections (see chapter 4).

### 6.2.2 Data processing and structure determination

#### Data Processing

Data was scaled and indexed using locally written programs (T.J.R. unpublished). The high resolution crystals were indexed from 20-2.6 Å whereas for the low resolution crystal, data between 20 and 4.5 Å was used. Data from all five crystals was merged to give a 98.7% complete data set ranging from 20-2.6 Å. The final statistics of this data set are summarized in table 6.3. 4% of the reflections were used as a test set for the calculation of R_free and were thus not used during molecular replacement and refinement (but used for map calculation).

#### Molecular replacement

The structure of the α6s NCP was solved by molecular replacement (MR) using a truncated α3s NCP structure as the search model. Table 6.2 summarizes the input parameters for the MR procedure. The actual MR was done using the series of programs provided by the X-PLOR package (Brünger, 1992). The orientation of the α6s NCP in the asymmetric unit was determined by the cross-rotation function in the following way. Patterson vector sets were calculated for the search model as well as for the native data, and the 3000 strongest Patterson vectors were used in the rotation function. The Patterson vectors from the search model were then rotated around Eulerian angles using a constant angular grid of 2.5°. The 6000 highest rotation-function grid points, resulting
from the product function of the two Patterson vector sets, were selected. This resulted in a set of 71 best rotation solutions. PC refinement (Brünger, 1990) was carried out to optimize each of the 71 possible solutions using the complete search model as a rigid body. This yielded two orientations with a scores of 0.152 and 0.150, respectively, whereas the rest of the possible solutions gave PC scores of 0.02. Individual translation-function searches were then performed on a 2 Å grid, searching 1/8 of the unit cell. The initial R/Rfree for the best translational solutions for the two rotational positions were 50.7/51.4% and 48.6/53.1% (8-4 Å), respectively. The initial positions of the two models were then refined in a three step rigid body refinement as follows. Firstly, the entire model was taken as a rigid group, secondly, the model was divided into the histone octamer and 10 bp segments of the DNA, and in the final cycle, the DNA was refined in groups of 5 bp. The R-factors after the refinement procedure were at 38.1/39.4% and 37.8/40.0%, respectively. Solution 1 was used for further refinement. Compared to the crystal packing in the structure of the search model, the new position corresponded to a symmetry related particle but rotated by 180° around the octamer dyad axis. Solution 2 was in a position equivalent to the search model.

<table>
<thead>
<tr>
<th>Search model</th>
<th>α6s NCP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P212121</td>
</tr>
<tr>
<td>Unit cell dimensions</td>
<td>106.04/181.78/110.2 Å</td>
</tr>
<tr>
<td>Residues included in the model</td>
<td>H2A-1 14-117, H2A-2 16-117</td>
</tr>
<tr>
<td></td>
<td>H2B-1 31-122, H2B-2 29-122</td>
</tr>
<tr>
<td></td>
<td>H3-1 38-135, H3-2 38-135</td>
</tr>
<tr>
<td></td>
<td>H4-1 24-102, H3-2 20-102</td>
</tr>
<tr>
<td></td>
<td>α6s 146: bp 3-144 (of 146 bp)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Target</th>
<th>native α6s 146 DNA data set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P212121</td>
</tr>
<tr>
<td>Unit cell dimensions</td>
<td>105.30/175.69/109.53 Å</td>
</tr>
<tr>
<td>Resolution</td>
<td>15 - 4 Å</td>
</tr>
</tbody>
</table>

Table 6.2: Summary of MR parameters.

The initial model was refined using alternating cycles of crystallographic refinement in X-PLOR (Brünger, 1992) and model building in O (Jones et al., 1991). In
Chapter 6

\( \alpha 6s \) NCP structure

the first cycle, the DNA sequence was changed to the \( \alpha 6s \) sequence without changing the position of the overwound (stretched) DNA region between base pairs 50 and 60, as seen in the structure of the \( \alpha 3s \) NCP. Assuming that the \( \alpha 6s \) made similar stacking interactions at the DNA termini, a stretched region would have been expected, although not necessarily at the same location. From the starting model it was then tried to find the correct DNA conformation and orientation using several cycles of positional refinement, combined with grouped B-factor refinement, simulated annealing (in x, y, z as well as in torsion angle space) and 2Fo-Fc omit maps for the entire DNA region (in 20 bp segments). As the distribution of pyrimidine and purine bases around the DNA center was consistent between electron density and model, this was taken as an indication for the correct orientation of the DNA. Base pairs 63 through 146 appeared to be modeled correctly even though the density was poor in several regions of the model. The DNA ends were symmetrically positioned as in the \( \alpha 3s \) NCP structure which implicated the existence of a similar stretched region as seen in the \( \alpha 3s \) NCP structure. Therefore, base pairs 1-62 went through several cycles of major rebuilding because the fit between model and electron density was not satisfactory. This finally positioned the stretched DNA region between base pair 20 and 30.

A total of 10 cycles of crystallographic refinement and model building was needed to achieve reasonable statistics which are summarized in table 6.2. In the later cycles, statistics could be improved by applying a bulk solvent correction using a solvent radius of 1.0 Å. The crystallographic pseudo-energy weight was set to \( wa=10^6 \). The upper limit for the maximum allowed B-factor was increased to 200 Å\(^2\). During positional refinement a harmonic restraint of 10 kcal mole\(^{-1}\)Å\(^6\) was applied either to the protein main chain C-alpha atoms, the DNA phosphorus atoms, the water oxygen and manganese ions, or to all atoms. Simulated annealing (SA) in torsion angle space with the same harmonic restraints as described above was especially helpful in positioning
Table 6.3: Statistics of α6s data set.

Table 6.4: Summary of residues in the α6s NCP structure.
Chapter 6  \( \alpha 6s \) NCP structure

phosphate free to move) proved useful. Table 6.4 gives a summary of the residues included in the current model of the \( \alpha 6s \) NCP. The number of protein residues included in the model is almost identical to the structure of the \( \alpha 3s \) NCP described previously.

6.2.3 General view of the \( \alpha 6s \) NCP

At a first glance, the two crystal structures of the \( \alpha 6s \) and the \( \alpha 3s \) NCP are strikingly similar (figure 6.3). The histone octamer and the DNA-octamer contact sites are particularly highly conserved. These interactions are shown schematically in figure 6.2. The gray bars indicate the location of the minor groove between the contacted phosphates. Base pair 73 is positioned on the octamer dyad axis in a way similar to the

\[
\begin{align*}
\text{a6s DNA:} & \quad \text{ATCTCAGATATCGCTGACTCGTAGAAAGAGGTGCTCAAACTCGGGCTAAAGGGAAACTTCAACTGAA} \\
\text{a3s DNA:} & \quad \text{TTGTTATAGGCAACGTCCTTAGCTCGATAGTTTTCTTTGAAAGTTGACTT}
\end{align*}
\]

Figure 6.2: Schematic drawing of the interactions between histones and DNA.

Comparison of the \( \alpha 6s/\alpha 3s \) sequences with respect to the stretched regions and the interaction sites between histones and DNA. The DNA bases are colored according to the histone protein that contacts the phosphate. The gray bars indicated the location of the minor groove between the contacted phosphates.

\( \alpha 3s \) crystal structure. The \( \alpha 6s \) DNA is stretched between base pairs 20 and 30 as compared to base pairs 50-60 in the \( \alpha 3s \) sequence by overwinding the DNA strand not contacted by the histone octamer. For the \( \alpha 6s \) NCP, this leaves only the 20 bp at the
DNA ends to be symmetrically related to the level of sequence. The conformational flexibility of the DNA is restricted to the regions which do not contact the octamer directly (figure 6.3).

Similar to the α3s NCP structure, the basic histone tails are only poorly defined. In some cases (H2B-1, H3-2; see table 6.4), some density could be traced but the fit was generally poor and the B-factors high.

**Figure 6.3: Superposition of the α6s and α3s NCP structure.**
The histones are colored in blue for H3, green for H4, yellow for H2A and red for H2B. The DNA strands are shown in magenta and white. Backbone traces are shown using the C-alpha atoms for the protein and the C1’ atoms for the DNA. The molecules were superimposed in the orientation as they occur in the unit cell. The left side shows one set of the histone proteins and the first 73 base pairs of DNA, whereas the second half is shown on the right side. Numbering is according to the α6s structure. The DNA-octamer contact sites are numbered from 1-7, starting at the DNA center. The images were generated with MidasPlus (Ferrin et al., 1988).
6.2.4 The crystal packing interactions

To understand the crystallization characteristics of the different DNA sequences, it was important to investigate the interactions responsible for crystal packing. With regard to the DNA, such interactions could induce major distortions which then are detected in the form of unusual DNA helix parameters. There are two dominant crystal packing interactions which are observed in both crystal structures: a H4-H2A contact and stacking interactions between the DNA ends. In the structure of the α6s NCP, H4-1 Arg23 is bound in a cavity formed by H2A-2 Glu56, H2B-2 Glu110 and H2B-2 Gln44 from a symmetry related particle. Nearby, there is a manganese ion bound between H3-1 Asp77, H4-1 Leu22 (main chain O) and H2A-2 Glu64/ H2B-2 Val45 (main chain O) of the neighboring particle. The identification of this ion as manganese is well proven, as the B-factor is practically identical to the direct neighbors and the occupancy is 1.3. This Mn$^{2+}$ ion trapped at the protein-protein interface was a somewhat unexpected observation but could eventually explain the absolute requirement of Mn$^{2+}$ in the growth of high quality crystals. These two crystal contacts are made by the other copies of the histone proteins in the α3s NCP.

The second site of contact is the stacking interaction between the DNA termini. In the α3s structure, the twist from base pair 146 to base pair 1 of the next particle is approximately 130° (4x normal B-form DNA twist). The two DNA coils are displaced by approximately 5 Å along the long axis of base pair 146 and away from base A146. The stacking of the termini is inclined with a mean rise of 3.9 Å. For the α6s sequence the twist is only 90° (3x normal B-form DNA twist). The helix axis is not displaced and thus forms a DNA coil continuous in appearance with a rise of 3.7 Å between molecules. The H4-H2A contact and the stacking interaction can presumably be formed by every core particle with the right DNA length, regardless of the DNA primary sequence. This has been observed already, because that was how NCP isolated from
higher organisms was originally crystallized. Unfortunately, these crystal contacts alone do not guarantee highly ordered crystal packing.

The remaining crystal contacts appear to be of minor importance and differ between the two X-ray structures. In fact, superposition of the crystal packing of the two structures shows a slightly different orientation of the particles against each other based on a variation in the stacking interactions, resulting in a shorter cell dimension along the b-axis (the altered DNA stacking may be the result rather than the origin). Figure 6.4 summarizes the crystal contacts (and close approaches) involving the DNA.

Three of the minor contact sites involving DNA as well as Mn$^{2+}$ ions are discussed on the basis of the α6s NCP structure. The DNA from base pairs 3-7 forms an extended contact site with DNA from a symmetry related particle between base pairs 95 and 100. Between the strands, 4 Mn$^{2+}$ ions are trapped, mediating contacts between phosphates or between phosphates and edges of bases (see discussion below). In the α3s structure, these contacts extend over a longer stretch of the DNA, and the contact seems to be closer. However, only one Mn$^{2+}$ ion can be located between the neighboring strands in this case. The contact site between base pairs 65/66 and 77 also involves the binding of a Mn$^{2+}$ ion between phosphates and the N7 of A78(Gua). In the α3s NCP (bp 80-90) this contact site seems to be even more pronounced resulting in an slight distortion of the DNA backbone (see discussion below). Another Mn$^{2+}$ ion is bound in the minor groove around base pair 62 where it is held in place by a serine of symmetry related H2B contacting the DNA.

**Figure 6.4: Crystal contacts in the NCP structures.**
Comparison of the two sequences α6s/α3s with respect to crystal contact sites. The overwound segments of the DNA are colored in magenta. Crystal contacts considered as important are colored in red with the interaction site marked on the side. Minor contact sites are colored blue. Numbers in parenthesis give the shortest distance for a given contact site. Mn$^{2+}$ ions involved in crystal contacts are indicated as stars (*).
Chapter 6

α6s DNA

Stacking on A1/B292 (3.6)

H2B-1 116-120 (5.2)

H2B-1 110-113 (5.4)

H2B-1 102-105 (4.0)  B203

A23/A24

A3-A6 (5.9)

B232 (6.1)

A65/A66 (6.1)

B214/A77 (6.1)

H2B-2 113-120 (2.6)

A75/A76 (6.1)

H2B-2 113-120 (3.5)

H2B-2 119/120 (4.9)

H2B-1 119/120 (5.7)

A136/A137

A25/A26 (4.4)

A117/B173 (9.1)

A66/A67, A146 (5.2)

A106/A107 (8.1)

H3-2 80/81 (2.9)

B198/B199 (3.5)

H3-1 76-80 (2.7)

A96/B204 (9.0)

B246

A44/A45 (4.7) in MG

Stacking on A146 (3.6)

α3s DNA

Stacking on A146 (3.8)

H2B-2 116-120 (5.8)

H2B-2 109-113 (3.8)

A139-A141, B148/B149

A140/A141

B156 (4.7)

H2A-2 71-74 (9.0)

B213/B214 (5.2)

A106/A107 (8.1)

A44/A45 (4.7) in MG

Stacking on A1/B292 (3.8)
6.2.5 General DNA bend and twist

Helical DNA parameters used in the following analysis were calculated from the DNA PDB files with the program CURVES (Lavery & Sklenar, 1988). The overall helix curvature in the nucleosomal DNA is mainly a result of non-zero roll angles and to a smaller extent of non-zero tilt parameters. The contribution of negative and positive angles is about equal. Figure 6.5 plots the DNA axis curvature together with the twist values for both fragments. The data points shown are mean values over three base pair steps in order to reduce the noise in the data. The average values for 10 bp segments are shown as dots of the same color. For the DNA helix axis curvature, there is no obvious periodic pattern. However, the DNA bend is stronger 1.5 and 4.5 helical turns (near base pairs 28, 58, 88 and 108) from the DNA center, which is the DNA region between the first and the second contact made by H3/H4 or H2A/H2B (see also figure 6.3). The last turn at the DNA termini appears to be relatively straight. The DNA bending in the dimer contact regions (bp 15-45 and 105-135) is generally stronger than on the tetramer surface.

The overall average twist values for the α6s and α3s sequences are 34.97° and 34.78°, respectively. These twist values are calculated in a "laboratory" frame (as defined earlier, see discussion below) and measure the average rotation from one base pair to the next (Klug & Travers, 1989). The corresponding helical periodicities are 10.30 and 10.35 bp/turn, respectively. The DNA in the two X-ray structures is therefore overwound as compared to free B-form DNA with a periodicity of 10.5 bp/turn. There is, however, a compelling shorter periodicity in the helical twist. The sites of DNA-protein interaction and narrow minor groove are characterized by significant overwinding. From a mechanical point of view, DNA overwinding helps to form a narrow minor groove and increases the angle between the minor groove and the global helical axis of the DNA supercoil.
Figure 6.5: Twist and global axis curvature.
The global twist and global axis angles as calculated with the program CURVES (Lavery & Sklenar, 1988) are shown as a comparison of the two sequences (α3s blue, α6s red). Noise in the data was reduced by plotting the average over three neighboring bp steps. At the top, the helical twist parameter is shown for both sequences. Overlaid are the average values of the twist angle for 10 bp segments. At the bottom of the graph, the same was done for the global helix axis curvature. In the center (green circles), the periodicity in bp/turn is shown for the individual 10 bp segments of the DNA. Average values over both sequences are shown here. Additionally, the center of the DNA-histone interaction sites is indicated by black bars. Mn²⁺ ions associated with DNA are shown as diamonds of the same color as the DNA they bind to (α3s blue, α6s red). Mn²⁺ ions bound to the N7 atoms of Gua are shown as closed diamonds, whereas ions bound to the DNA minor groove are shown as open diamonds.

The average twist value calculated for 10 bp long DNA segments was surprisingly uniform over the entire length of the nucleosomal DNA. There is no indication of an underwinding of the DNA center as proposed earlier (Hayes et al., 1990; Wolffe, 1995). The average periodicity of both sequences (figure 6.5; green circles) was even more uniform. The increased twist around base pair 30 and 50 is a result of the stretching (overwinding) of α6s or α3s DNA in these regions, respectively. The irregularity around base pair 130 for the α6s sequence comes from a distortion induced by a crystal contact around bp 130 (see discussion below).

6.2.6 Manganese ions

The problems associated with the identification of Mn²⁺ ions bound to the DNA were very similar to the situation as encountered in the α3s NCP structure. The decision in assignment between water molecules and manganese ions was based on the same criterion as discussed in the previous chapter. Table 6.5 summarizes the Mn²⁺ ions participating in DNA binding. In contrast to the α3s NCP structure, there are a total of 6 Mn²⁺ ions involved in DNA-DNA crystal contacts and 10 Mn²⁺ ions are located near the N7 atoms of Gua at distances between 2.5 and 3.5 Å. Comparison of the exact
location of the Mn\(^{2+}\) ions (diamonds in figure 6.5) with the center of the DNA-protein contacts (black bars in figure 6.5) showed a compelling correlation. The Mn\(^{2+}\) ions are predominately bound at sites of large DNA twist where the minor groove faces the protein. These sites also tend to have a local maximum in the overall DNA helix axis curvature. There is no preference for binding of manganese ions to pyrimidine-purine base pair steps over purine-purine steps.

<table>
<thead>
<tr>
<th>Number of Mn(^{2+})</th>
<th>Crystal contact</th>
<th>Contacted atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>yes</td>
<td>Phosphates of A100, B285(8), B286(7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phosphate of A98 and B288 (5) (Gua) N7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phosphates of A3, A99 and A100 (Gua) N7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phosphate of B287(6) and A97 (Ade) N3</td>
</tr>
<tr>
<td>1</td>
<td>yes</td>
<td>Fills mg; A62(Ade)/B234 and H2B-2 Ser120</td>
</tr>
<tr>
<td>1</td>
<td>yes</td>
<td>Phosphates of A77, A65 and A78 (Gua) N7</td>
</tr>
<tr>
<td>7</td>
<td>no</td>
<td>N7 of Gua A20, A59, A142, B174 (119), B217 (76), B227 (66), B280 (13),</td>
</tr>
<tr>
<td>7</td>
<td>no</td>
<td>In mg: A32 (Ade) N3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>In mg: B254 (41) (Ade) N3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A80 (Thy) O4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A82 (Ade) N7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B213 (Ade) N6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B179 (114) (Ade) N7 and B180 (Gua) O6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A126 (Cyt), B166 (Gua) O6, B167 (Gua) O6</td>
</tr>
</tbody>
</table>

Table 6.5: Manganese ions in the α6s NCP 2.6 Å structure.

Numbers in brackets for the B-strand give the base pair number (number). The numbering of the DNA is as follows: one molecule of the DNA double strand is named A1 through A146 whereas the other molecule is named B147 through B292. In base pair 1, base A1 is therefore base pairing with base B292.

Two main types of manganese binding sites could be identified and are shown in figure 6.6. Firstly, there are sites where the Mn\(^{2+}\) ion is bound by the atoms of one base only, represented here by the site at Gua B217 in α6s DNA (figure 6.6B). Such sites generally are correlated with a large twist (>45°) with respect to the previous base pair in 5' direction. Secondly, Mn\(^{2+}\) ions bound by the phosphate of the previous base in the 5' direction generally reduce the twist between the two base pairs (figure 6.6A, Gua A20). For the site shown in figure 6.6A, the DNA backbone is significantly distorted, virtually making a bend around the manganese ion (6.6C). This distortion could be
crucial for positioning the phosphates correctly on the octamer surface. It is noteworthy that the phosphates of both bases A19 and A20 are contacted by the histones.

**Figure 6.6: Manganese binding sites.**

Two Mn$^{2+}$ binding sites of the α6s DNA are shown. The Mn$^{2+}$ ion at Gua A20 is shown in A, whereas the Mn$^{2+}$ binding site at Gua B217 is shown in the drawing on the top right (B). C: Minor groove-octamer contact site 6 as seen from the octamer (superhelical axis of DNA supercoil aligned vertically). The Mn$^{2+}$ binding site at Gua A20 is right adjacent to the contact site. The distances between the Mn$^{2+}$ ion and the N7, O6 and PO atoms are 2.47, 3.18 and 5.0 Å for A20 and 2.56, 3.76 and 4.65 Å for B217.
The two binding sites shown are examples for the stabilization of DNA distortions by manganese ions. However, each binding site is slightly different and other factors such as the local base sequence or DNA parameters influence the exact binding mode. The DNA distortions stabilized by the manganese ions may be quite significant for the overall stability of the octamer-DNA complex.

6.2.7 Positional flexibility of the DNA

In order to quantify the positional flexibility of the nucleosomal DNA, the deviation between equivalent phosphate positions in the two structures was determined after superposition of the histone tetramers. This was done in two ways: Firstly, the particles were compared in the same orientation as they occur in the unit cell (figure 6.7B, top), and secondly, the alignment was repeated after rotation of the α3s particle by 180° around the octamer dyad axis (6.7B bottom). In both cases, only the tetramers were used for superposition of the two structures. The r.m.s.d. of the C-alpha atoms calculated for the tetramer and for dimers 1 and 2 were 0.38, 0.67 and 0.37 Å, respectively (using the numbering of the α6s NCP; this is the superposition shown in figure 6.3 and in 6.9). For the second superposition, the values were 0.24, 0.41 and 0.54 Å. The average distance deviation of the phosphates was 1.87 Å in the first alignment and 1.64 Å for the second superposition. As a control, the DNA halves of individual core particles were superimposed based on the tetramer as well (figure 6.7A).

The large deviations near base pairs 51-61 (α3s) and 18-27 (α6s) were caused by the stretching of one DNA strand in this region (figure 6.7A). These distortions are also clearly visible in the cross comparisons between the particles. In contrast, the distortions near base pairs 120-146 (α6s figure 6.7B, 1-26 in figure 6.7A bottom) were caused by a close contact between base pairs 130-135 and histone H2B of a symmetry related particle in the α6s NCP crystals (see figure 6.4). Thereby, the DNA is displaced.
(and distorted) by several angstroms along the superhelical axis towards the central DNA turn of the particle (figure 6.9). This distortion between the histone-DNA contact sites at base pairs 130 and 140 induces, as a compensation, curvature in the opposite direction in the adjacent 10 bp segment starting at base pair 130. There are no crystal contacts in that region, so that the DNA is free to move. This comparison clearly demonstrates the flexibility of the DNA on the histone surface. The distortion in this region is presumably the reason for the unusual twist, base pair displacement and inclination parameters observed between base pairs 120 and 146 for the α6s DNA sequence (discussed below). Between base pairs 80 and 90, the α3s DNA phosphates show an unusually high distortion as well (figure 6.7B bottom), caused by a rather close crystal contact.

In the remaining regions, the phosphate deviations of the individual DNA halves in the same particle (figure 6.7A) are uniform. The average deviation of the phosphate positions of the α3s DNA halves is 1.19 Å as compared to 1.60 Å for the α6s DNA. Excluding the regions discussed earlier, the superposition of the α3s DNA halves seems to be better than that of the α6s DNA. This may be a consequence of the DNA being symmetrically positioned, except for the central 40 base pairs. The deviations around the NCP dyad are generally lower than in the flanking regions of the particle.

The B-factor distribution for the DNA backbone atoms (phosphate and deoxyribose) exhibits a similar periodicity as seen in the deviation plots. Phosphates in close contact to the histone octamer generally have low positional freedom and low B-factors whereas the outer strand has rather large positional freedom (figure 6.8). The DNA around the dyad has reasonably low B-factors for both strands.
Figure 6.7: Deviation in the phosphate backbone positions

Plots of the distances between equivalent DNA phosphates of the DNA backbone. NCP were superimposed based on the globular parts of the H3/H4 tetramer. The distance between phosphates in equivalent positions was then plotted for each base and both strands. (A): Direct comparison of the two DNA halves in the α3s NCP (top) and the α6s NCP (bottom). The numbering starts at the DNA ends, the DNA center is at base pair 73. (B): Comparison of α6s DNA with α3s DNA superimposed in the same orientation as they occur in the crystal (top) and after rotation of the α3s NCP by 180° around the octamer dyad axis (bottom). DNA strand A (A1-A146) is color coded in blue whereas strand B (B147-B292) is colored in red. The bars indicate the sites of interaction between phosphates and histone octamer. Superposition and distance calculations were done using the programs O (Jones et al., 1991) and LSQMAN (Kleywegt & Jones, 1994).
**Figure 6.8: B-factors of the DNA backbone.**
Plots (A) and (B) show the DNA backbone (phosphate and deoxyribose) B-factors for the two sequences. DNA strand A (A1-A146) is color coded in blue whereas strand B (B147-B292) is colored in red. The bars indicate the sites of interaction between phosphates and histone octamer.

**Figure 6.9: DNA superposition.**
Two segments of nucleosomal DNA are shown after the tetramers have been superimposed using the same alignment as in the crystal. The superhelical axis of the DNA supercoil is horizontally aligned. The α6s DNA is shown in green, whereas the α3s DNA is colored red. The numbering is according to the α6s DNA sequence. The DNA dyad is at base pair 73.
6.2.8 The DNA groove dimensions and helical parameters

The DNA groove dimensions as calculated by the program CURVES (Lavery & Sklenar, 1988; Stofer & Lavery, 1994) show a remarkably regular periodicity throughout the DNA. The minor groove width changes from 3 Å at the protein-DNA contact sites to 9 Å, where the major groove faces the protein. The depth of the minor groove varies between 7 Å and 3 Å at the same time. The major groove width alternates between 9 and 13 Å being narrower when facing the octamer. The groove is very shallow opposite the DNA-protein contact sites (1-3 Å) and deeper in the remaining segments of the DNA (maximum 7-8 Å).

Figure 6.10 shows a plot of the minor groove width aligned with the DNA sequence, the base pair displacement in the direction of the major groove (Xdisp) and the opening angle (positive for opening into the major groove). The regular pattern of narrow and wide minor groove is remarkable. The minimal width varies only between narrow boundaries, and there is no obvious dependence on the underlying DNA sequence. Even stretches of 3 GC base pairs do not inhibit the formation of a narrow minor groove (e.g. around bp 88 in α3s, bp 48, 58, 98 in α6s). The average minor groove widths calculated separately for AT and GC base pairs are around 6.1 Å with a standard deviation of 1.7 Å, and are independent of the base pair. There is no apparent correlation between minor groove width and DNA sequence for these DNA fragments. The presence of GC base pairs at locations with narrow minor groove is mainly compensated by displacement of the base pair into the major groove and an increased opening angle (towards the major groove as well, figure 6.10). However, large opening angles especially with GC base pairs can only be achieved with a large concomitant propeller twist at the same time, as seen in the NCP structures. In the α3s sequence for example, the peripheral four contact DNA-octamer contacts are mainly AT-rich and the corresponding Xdisp and opening parameters show no periodicity. The three histone-DNA contact sites at the center are GC-rich, with the consequences as described. The
propeller twist values show the same general pattern with a clustering of large non-zero values around the DNA center (data not shown). In the α6s DNA, the pattern is not so clear as there are GC base pairs at every site with narrow minor groove except for the DNA-H3 interaction near the DNA ends. Likewise, the Xdisp, opening, and propeller twist values are high over the entire length of the fragment. The Xdisp values show maxima almost at every position with narrow minor groove. The unusual low Xdisp values for the α6s DNA between base pair 120 and 140 is a consequence of the distortion induced by the crystal contact in this region.

A comparison of the AT base pair distribution in positions of narrow minor groove for both sequences revealed that the last octamer-DNA contact, where the DNA bending is relatively low, is the only site where both sequences have exclusively AT. For the α6s sequence the narrowest part of the minor groove effectively moved by two base pairs from a region with GC base pairs into an AT stretch. This could be an indication that straight B-form DNA indeed prefers AT base pairs at sites of narrow minor groove. However, the actual insertion of the H3 Tyr/His into the minor groove tolerates AT base pairs (α3s) as well as GC base pairs (α6s).

The nucleosomal DNA was further analyzed by examination of some additional helical parameters. The DNA is very close to B-form DNA, which is consistently shown by the average rise between base pairs of 3.36 Å and the average propeller twist of approximately -9°. The mean distance between phosphates in the same strand is 6.63 Å which is close to the 7.0 Å of straight B-form DNA (A-form DNA: 5.9 Å).

Apart from the periodicity induced by the octamer-DNA contacts discussed earlier, twist values also show the typical periodicity originally predicted for B-form DNA and since then experimentally proven by several oligonucleotide X-ray structures (Calladine, 1982; Klug et al., 1979; Nelson et al., 1987; Yoon et al., 1988). Purine(pu)-pyrimidine(py) base steps are significantly underwound, whereas the py-pu steps are overwound. Pu-pu base pair steps show average periodicity. The same pattern of alternating twist angles is observed for all possible pu/py mixed base pair steps. For
example, d(AT) steps have an average twist of approximately 30° as compared to 38° for the d(TA) steps. It should be noted, that the average intra-strand phosphate distances for the same base pair steps do not show any periodicity.

A closer look at the propeller twist values revealed some interesting features. The absolute values for AT base pairs were not significantly different from the GC base pairs. It has been speculated that propeller twist values are mainly dependent on the neighboring base pairs, as large propeller twist in purine-purine steps would cause fewer clashes than in purine-pyrimidine steps (Calladine, 1982; Saenger, 1984). To test this hypothesis, average (absolute) propeller twist values were calculated for the central base pair in pu-pu-pu, pu-pu-py and py-pu-py triplets. The first group showed the largest average propeller twist (14.3/13.9° for the a3s/a6s sequences), whereas the values for the other two triplets were smaller (10.8/10.9 and 10.7/13.0°); the standard deviation was on the order of 7°. The roll angles were thought to be linked to the propeller twist values in a way that positive roll (opening in the minor groove) can reduce clashes in the minor groove of py-pu base pairs and vice versa with negative roll angles for pu-py base pairs (Calladine, 1982; Saenger, 1984). However, analysis of the available data from the two X-ray structures did not give conclusive results which could proof this hypothesis. Other parameters such as rise or Ydisp (displacement along the long base pair axis) which could also reduce the mechanical stress induced by large propeller twist values in mixed base pair steps were not found to show a consistent pattern either.

Figure 6.10: DNA minor groove dimensions.
The DNA parameters, as calculated with the program CURVES (Lavery & Sklenar, 1988), are shown for the two sequences α6s (A) and α3s (B). The minor groove width (blue) is aligned with the DNA sequence, the base pair displacement into the major groove (Xdisp, red) and the opening angle (light red). The black bars indicate the center of the histone-DNA interaction sites.
6.3 Discussion

6.3.1 DNA sequence dependent crystallization

The ability to form the stacking interactions between the DNA termini as observed, appears to be obligatory for crystallization. The allowed DNA length varies between 145 and 147 base pairs, but is also dependent on the sequence. Based on this, it is clear why the sequences si38, s142 and s150 form crystals of poor quality or no crystals at all.

For the α3s DNA, the stretching of the sequence is very close to the DNA center. This raises the question of whether this is an intrinsic feature of this DNA sequence or whether this could happen during crystallization. The method of the site directed cutting developed in our laboratory (Flaus et al., 1996), applied to the outer turn of the DNA, could eventually resolve this issue. Because of the crystallization behavior of the other fragments of the α3s family, the possibility of the stretch being intrinsic to the sequence is favored. However, there is no obvious sequence pattern that would favor one position over the other. At least, the predominate occurrence of GC base pairs in regions with the minor groove facing the protein in the central 70 base pairs is likely not to inhibit repositioning.

Insertion (α8s 147) or removal (α10s 145) of one base pair at the dyad of the α3s sequence presumably creates particles with two normal branches or two stretched branches, respectively. Both NCP crystallize readily and the crystals diffract to 2.0 Å resolution. However, their crystal structures remain to be solved. Addition of one base pair at the end of α10s 145 (α9s 147) creates another 147mer fragment which forms beautiful looking crystals of a slightly different morphology which do not diffract to more than 20 Å. This was explained by the presence of one additional base pair at each DNA end with both DNA halves being stretched as in the corresponding α10s
sequence, and thus distortion of the packing in the crystal. α4s 146 which was constructed from α3s by removal of two base pairs at the dyad and addition at the ends, probably suffered from the same problem as the α9s 147 sequence, only more so.

For the other sequences, the interpretation of the crystallization behavior is more speculative than for the α3s family of sequences. The s146 probably suffers from two problems. Firstly, the fragment is not properly positioned (see chapter 4), and secondly, long (8bp) AT stretches 25 and 60 bp away from the dyad could be inhibitory, as such sequences are known to be rather stiff (see Appendix 1 for the sequence). In addition this sequence is 146 base pairs long as well, so high quality diffraction crystals are unlikely to be grown without the stacking interactions between DNA termini being made. The cs146 sequence, constructed as an ideal nucleosome binding sequence with an almost perfect positioning of AT, may have no driving force for stretching in order to make the stacking interactions. If this assumption is correct, a 147mer Crothers sequence should result in crystals that diffract to high resolution. The α7s sequence (equivalent of α4s for α6s) could not be centered properly either due to the high intrinsic bend of this DNA (see chapter 4), or because centering would have moved a GC doublet into a possibly unfavored position 6 or 7 base pairs from the DNA ends. This region is AT rich in all sequences that have given high quality crystals and is located immediately adjacent to the H3 Tyr/His insertion site. For the human alpha-satellite consensus sequence (α2s, α5s), it is not clear why diffraction quality crystals could not be obtained. The high AT content (48/73 versus 42/73 in α6s) is the only outstanding feature of this sequence. It is likely that either the stacking interactions cannot form because the sequence is not stretched, or that some special DNA conformation elsewhere in the particle abolishes crystal formation.

An issue not yet discussed is the observation that all core particles made from even-numbered symmetric DNA fragments are asymmetric due to the positioning of the body of a base on the histone dyad. It is not clear a priori, whether or not the two orientations of the particle can be distinguished during crystallization. In the two
structures, crystal contacts are located near the symmetry related sites of DNA stretching. That is presumably how the two orientations are distinguished in the present crystal lattices. DNA stretching at any other site may remain undetectable and the resulting crystals may suffer degeneracy in particle packing.

6.3.2 General view of the NCP

In this thesis, the crystal structures of two nucleosome core particles made from different DNA sequences were presented. The structure of the histone octamer and the interactions between DNA and protein are conserved impressively. In spite of this, there remains plenty of positional freedom for the DNA, but it is restricted to the segments which are not contacted by the histone proteins. This is well documented by the high B-factor of these DNA segments as well as by the observed deviation in the phosphate positions, when the DNA of the two structures is superimposed. Neither of the two X-ray structures provides secure information about the function and possible binding properties of the majority of the N-terminal basic histone tails.

6.3.3 DNA binding in the NCP

The question of how the DNA is bound in the nucleosome core particle has been a matter of extensive discussion in the past. Major issues of concern were the sequence preferences of DNA binding by the histone octamer and the helical periodicity of the nucleosomal DNA.

Even though it is not entirely clear what the in vivo situation is, it was assumed that AT sequences should preferentially bind in positions where the minor groove of the DNA faces the histone octamer (Satchwell et al., 1986; Shrader & Crothers, 1989; Shrader & Crothers, 1990; Wolff, 1995). This was shown to be true in numerous in
vitro mapping and binding experiments. In one of the original experiments (Satchwell et al., 1986), the DNA of 177 core nucleosome sequences isolated from chicken erythrocytes was analyzed for the dependence of the AT content on the rotational position on the histone octamer. A periodicity for AT sequences facing the octamer at sites where the minor groove interacts with the octamer was found throughout the particle, except for the central two DNA turns, where it was inverted. Binding studies using sequences with a 10 bp AT periodicity explicitly showed the formation of more stable nucleosomes (Shrader & Crothers, 1989; Shrader & Crothers, 1990, chapter 3). Longer polyA (more than 4 bp) stretches, which are known to be relatively stiff, were localized predominately near the DNA ends where the DNA bending is reduced.

The two crystal structures clearly show that there is no absolute requirement to have exclusively AT base pairs at sites where the minor groove is facing the protein. There is no correlation between minor groove width and DNA sequence for the two fragments. Since the DNA is bound only at a few points, sufficient flexibility remains to react with conformational changes to the stress eventually induced by GC bp at locations with narrow minor groove. With respect to minor groove formation, the presence of GC bp at such sites is compensated by increased Xdisp, opening and Propeller twist values. Local overwinding of the DNA in the protein contact regions is also a very important contribution to the formation of a narrow minor groove. As a side effect, overwinding also increases the angle between the minor groove path and the overall helix axis. This is likely to improve the fit of the phosphates on the amide main chain atoms of the histone octamer. Of course, all these distortions can only be achieved at the expense of a certain energy input. However, in a 146 mer core particle the maximization of the number of DNA-octamer contacts is probably more favorable than most of the sequence dependent positioning effects. For longer fragments, this means that AT-rich sites are most likely preferred in the DNA-octamer contact regions, because DNA sequence dependent distortions can be reduced. However, it should again be emphasized that the data available from the two crystal structures clearly argues
against an absolute requirement for AT base pairs at sites of narrow minor groove. This is in agreement with molecular dynamic simulations of straight oligonucleotides, which showed narrower minor grooves for GC sequences than for AT sequences (Boutonnet et al., 1993), even though experimental results generally showed the opposite (see above). It was also shown, that narrowing of the minor groove causes increase in DNA twist and displacement of the base pairs into the major groove.

The overall helical periodicity of nucleosomal DNA was shown experimentally to be approximately 10.15-10.35 bp/turn (Drew & Travers, 1985; Hayes et al., 1990; Hayes et al., 1991a; Wolffe, 1995). However, some studies showed that the three central turns were underwound to 10.7 bp/turn whereas the outer DNA regions were overwound to 10.0 bp/turn (Hayes et al., 1990; Wolffe, 1995). These helical periodicities were obtained mainly using hydroxyl radical or DNasel footprinting techniques. The cutting efficiency of such methods, however, depends not only on the helical twist of the DNA, but is also a function of the accessibility of the DNA on the octamer surface. The results may therefore be biased in a way difficult to predict. Based on the two crystal structures, the average helical periodicity has been calculated for 10 bp segments all around the DNA for both sequences. The calculated average for these 28 10 bp segments is 10.33 bp/turn with a standard deviation of 0.27. Since the values calculated with the program CURVES (Lavery & Sklenar, 1988) are in the "laboratory" frame as compared to the footprinting results which give helical periods in the "local" frame of references, the twist values have to be corrected for the superhelical pitch (Klug & Travers, 1989). Based on the results of the two crystal structures, the DNA is wrapped around the NCP in -1.66 superhelical turns with a pitch of approximately 4.5-5°. Applied to the average helical twists calculated with CURVES, the helical periodicities change to 10.20 and 10.25 for the two sequences α6s and α3s, respectively (as measured in the "local" frame of references), a periodicity which now can be directly compared to the results obtained from footprinting experiments. There is no evidence for DNA underwinding in the central three turns of the DNA, which is
consistent with earlier results (Drew & Travers, 1985). The large helical periodicity of the DNA appears to be fixed by the position of the DNA-protein interactions and can only be altered within narrow borders. However, local distortions induced by binding of the DNA to the histone surface are allowed. The periodic overwinding to form narrow minor grooves at sites of histone-DNA interaction is a good example. Such local distortions, additional protein-DNA interactions, or sequence specific effects may affect the local cutting efficiency for external reagents. For example, the αN helix of H3, which contacts the DNA at the first histone-DNA contact site next to the DNA center (insertion of a proline into the minor groove), is likely to affect the position of maximum cutting and thus change the apparent helical periodicities obtained by this footprinting method. The two stretched DNA regions provide an interesting example of how the DNA can be locally overwound without changing the histone protein interactions. Underwinding by the same mechanism appears to be more difficult. Whether such effects are relevant for nucleosomes in vivo, can not be decided at the moment.

Analysis of the helical parameters such as twist and rise showed that the nucleosomal DNA has B-form conformation. Even the alternating twist pattern of pu-py versus py-pu steps, typically seen in B-form DNA, is present in all possible mixed base steps (Klug et al., 1979; Yoon et al., 1988). However, some characteristics of nucleosomal DNA, as seen in the two X-ray structures, deviate significantly from the expected behavior. Minor groove width and propeller twist values are largely independent of the base sequence. The same applies for the predicted compensation of large propeller twist by roll between base pairs (Calladine, 1982; Saenger, 1984). The helical parameters most affected by DNA binding to the octamer are the roll (and as a consequence propeller twist) and tilt angles, as they are the determinants of overall DNA bending. In order to allow DNA binding to the histones, the octamer is able to override a certain amount of the DNA sequence intrinsic parameters (Hayes et al., 1991a). However, there is a limit to what can be done, and so long stretches of either
homo dA or dG are preferentially excluded from nucleosome formation. The large propeller twist of such sequences is thought to inhibit the formation of base step roll, which is needed for binding to the histone octamer (Nelson et al., 1987; Saenger, 1984). In contrast, the overall helical periodicity and base-step rise need not be changed and therefore retain the dimensions typical of B-form DNA.

6.3.4 Stability of different histone-DNA complexes

The two sequences used for crystallization were significantly different from each other. Whereas the base sequence of the central 70 base pairs is highly similar, the flanking 80 base pairs differ considerably as half of the base pairs have AT exchanged for GC or vice versa. Neither of the two sequences, however, showed an AT pattern which would make it an especially good nucleosome binding sequence. The biochemistry was slightly different in the way that the α6s DNA appeared to be intrinsically bent as compared to all other DNA sequences crystallized. Centering the DNA on the particle needed an incubation time of 1h at 55° C as compared to the normal 37° C. It would be worth knowing if any structural parameters can be found that can account for this effect.

Comparison of the overall helical path of the DNA in the two structures showed striking similarities. The α3s DNA appeared to have a somewhat stronger bend 1.5 helical turns away from the DNA center, but this is in a region of high sequence similarity between the two DNA fragments. In contrast, the overall bend in the region contacted by the dimers appeared to be stronger for the α6s DNA resulting in a more regular axis curvature over the entire length of the particle. The α6s NCP gives the impression of being slightly more compact than the α3s NCP which could be a result of the increased average helix axis curvature. The significance of this for the stability of the entire complex is not clear.
The explanation for the increased shifting stability of the α6s NCP is probably also dependent on the mechanism assumed to be involved. Based on the smooth histone octamer-DNA binding surface and the regular bend of the DNA, a simple rotation of the histone octamer cylinder in the DNA supercoil by approximately 10 bp seems to be an attractive shifting mechanism. DNA fragments which can intrinsically accommodate a lower radius of curvature may slow down such a process significantly. The slightly higher average bend of the α6s DNA as compared to the α3s DNA may account for the observed slower shifting rate. The cs146 which was even more stable towards shifting, may do so, because it accommodates the histone surface in an ideal way due to the periodic AT distribution. There is, however, one problem with the theory. The idea works fine for shifts which retain the rotational position of the DNA on the octamer, which is true for most fragments discussed here. The α6s DNA is the exception in the sense that the DNA was offset by 15 bp after reconstitution (there are several other examples, A. Flaus unpublished). Consequently, the positional as well as the rotational settings had to change during centering of the DNA. For this to happen, the DNA must either have a large torsional flexibility or segments of the DNA can temporarily dissociate and rebind in a different rotational position. Experimental evidence for dissociation comes from recent experiments studying the accessibility of buried restriction sites in the NCP and was supported by theoretical considerations (Marky & Manning, 1995; Polach & Widom, 1995). However, in order to have a driving force (maximization of the DNA-protein interactions) for the shifting in 146mer core particles, the DNA ends must be bound predominately to the histone surface. The dissociation could thus be only a fast oscillation. With respect to the torsional flexibility, the biological significance of the observed DNA stretching in the crystal structures would be of great interest. Torsional movements by one base pair seem to be possible. For a 15 bp shift, a mechanism without dissociation seems unlikely.

Based on the experimental data available from the two crystal structures, there is no obvious reason for the different biochemical properties of the two sequences. It is
possible, that a number of small sequence differences add up to the full effect, which would be difficult to detect.

6.3.5 DNA flexibility as a function of additional factors

The crystallization/storage buffer contained two components, namely cations and the cryo-protectant MPD, both of which can affect the structure of DNA. Bivalent cations are known to bind nucleic acids in different ways, depending on the type of ion (reviewed by Saenger, 1984). Whereas Mg$^{2+}$ increases the melting temperature of DNA, the presence of Cu$^{2+}$ destabilizes the DNA which is correlated with the way the cations interact with DNA. There are three possible metal binding sites in nucleic acids: the phosphates, the sugar hydroxyl groups and the base keto oxygen and ring nitrogen atoms. Cations such a Na$^+$, K$^+$, Mg$^{2+}$ and Ca$^{2+}$ can bind at all sites but prefer the phosphates and sugar hydroxyls. The transition metals such as Mn$^{2+}$ are preferentially bound by the ring atoms of the bases and the phosphates. One of the best metal binding site in nucleotides is between the ring nitrogen N7 and the keto oxygen O6 on the major groove side of guanine bases.

This is where several Mn$^{2+}$ ions were found in the nucleosomal DNA. The binding sites seem to cluster in regions of high twist (overwinding) which also correlate with the sites of DNA-histone contact formation. Crosslinking DNA bases to the phosphate backbone, as seen in X-ray structures, will certainly stabilize a particular DNA conformation and could therefore provide an explanation for the stabilizing effect of low concentrations of bivalent cations with respect to shifting (Pennings et al., 1991). The retardation of DNA fragments in PAGE experiments due to increased bending may also be caused by such a Mg$^{2+}$ induced bend at GC base pairs (Dlakic & Harrington, 1996). Recent X-ray structures of RNA and DNA molecules confirmed the outstanding properties of Gua N7 in metal binding. For example in the structures of the
hammerhead ribozymes, the catalytic, bivalent metal ion is bound between the Gua N7, O6 and the proR oxygen of the preceding phosphate similar to the Mn$^{2+}$ at base A20 in the $\alpha$6s DNA (Pley et al., 1994; Scott et al., 1995). Binding of Mg$^{2+}$, Ba$^{2+}$, Co$^{2+}$ and Cu$^{2+}$ ions to Z-DNA revealed a preference for the N7 position of guanine bases as well (Gao et al., 1993; Jean et al., 1993). Guanine N7 sites are also preferentially attacked by the anticancer drug cisplatin (cis-Pt(NH$_3$)$_2$Cl$_2$), which has been shown to crosslink GG base pair steps (van Boom et al., 1996). The exact mode of metal coordination in theses examples depends on the metal ion as well as on the ligands. Since most of the biologically relevant bivalent ions have an octahedral first hydration shell, the ion is bound by 6 ligands. The positions in the hydration shell are either occupied by a water oxygen, a heteroatom such as oxygen or nitrogen or a water mediating the contact to a heteroatom. Distances between the metal ion and a heteroatom in the range of 2.0-3.0 Å are typical for direct interactions, whereas distances between 3.5 and 6 Å are mediated by a water molecule of the first hydration shell. However, we do not see water molecules bound to the Mn$^{2+}$ ions in the two crystals structures, even though the measured distances would suggest their presence in most positions of the hydration shell. Considering the millimolar concentration of bivalent ions such as Mg$^{2+}$ and Ca$^{2+}$ in the cell, binding to DNA is likely to be biologically relevant. Binding of bivalent ions to GC base pairs positioned where the DNA contacts the octamer may be a way to reduce the energy differences between AT and GC base pairs in such locations.

Pentanediol (MPD) which was used at a concentration of 24% as a cryo-protectant is known to influence the DNA structure by dehydration of the DNA surface. In the core particle crystals, MPD also increased the diffraction resolution and decreased the spacing along the crystallographic b-axis (Richmond et al., 1984; Struck et al., 1992). The b-axis is the direction in which the DNA ends stack on each other. Slight conformational changes in the last DNA turn can therefore induce significant changes in packing. The rest of the nucleosomal DNA does not seem to undergo significant conformational changes as its overall position is defined by the protein-DNA contacts.
Recently it had been suggested that MPD could reduce DNA curvature by making DNA A-tracts stiffer (Dlakic et al., 1996). As a consequence of the MPD and the crystal packing interactions, the final turn of DNA could adopt a conformation different from NCP in solution.

### 6.3.6 Accessibility of nucleosomal DNA

The DNA stretches of the nucleosomal DNA, that do not directly contact the histone octamer, exhibit a large amount of positional freedom, whereas the position of the phosphates contacting the octamer is generally conserved. The displacement of the DNA helix induced by the crystal contact site in the α6s NCP around base pair 130 shows the degree of flexibility of the DNA. In this region, the entire DNA supercoil is distorted and displaced by as much as 7 Å without significantly altering the DNA-protein contacts. This surprising level of flexibility is a direct consequence of the limited size of the histone-DNA contact sites.

The nucleosomal DNA is highly accessible and initial binding of a transcription factor to its cognate site in the nucleosome appears to be possible. Only if the recognition site was hidden in the part of the minor groove directly contacted by the histone octamer surface, would the obstacle most likely be to large to overcome. After initial binding and in combination with partial dissociation of the DNA from the octamer (Polach & Widom, 1995), a factor could then gain full access to its recognition site. During the process of binding, the NCP may be significantly disrupted or the translational and rotational settings of the DNA may be altered. However, the importance of partial DNA dissociation for factor binding in chromatin, where the DNA termini are not free to move, is not clear. In fact, binding of transcription factors to their cognate sites in the core particle had been shown in several in vitro experiments (Côté et al., 1994; Vettese-Dadey et al., 1996; Vettese-Dadey et al., 1994). In these cases, the
typical periodic footprinting pattern of nucleosomal DNA was maintained even after binding of the additional protein.
7. HMG-14/17 and NCP

7.1 Introduction

7.1.1 The HMG-14/17 family of proteins

Proteins of the HMG-14/17 family are of the most abundant non-histone proteins in the nucleus (approximately 1 copy per 10 nucleosomes) and are found in a wide variety of species (Goodwin & Mathew, 1982). Human HMG-17 is a 9.2 kd (89 aa, HMG-14: 99 aa) chromosomal protein with an unusual amino acid composition: there are no aromatic amino acids but 11 Pro, 22 Lys, 4 Arg and 14 Asp/Glu resulting in a net charge of 12 (see figure 7.1).

The HMG-14/17 protein family exhibits a modular structure in which the conserved residues are clustered into well defined regions. Figure 7.1 shows a sequence alignment of the two human proteins with the 4 conserved domains (Trieschmann et al., 1995b). The nucleosomal binding domain alone is sufficient for NCP binding (Crippa et al., 1992). The N-terminal 4 amino acids (domain A) are conserved in all members of the family, however, removal of this N-terminus does not influence cooperative NCP binding or transcription activation in vitro. The C-terminus (beginning with domain C) has been shown to be crucial for transcription activation in vitro and has no NCP binding activity. Whereas the N-terminus has a positive net charge, the C-terminus is negatively charged.

The HMG-14/17 group of proteins is one of the very few protein classes known to stably bind to nucleosomes. In vitro, the nucleosome binds two copies of HMG-14/17 in a cooperative manner under nearly physiological conditions at the entry and exit points.
of the DNA (Goodwin & Mathew, 1982; Postnikov et al., 1994). Bound HMG-14/17 could be crosslinked to H3, underlining a direct mode of binding to the NCP (Brawley & Martinson, 1992). Core particles are not destabilized upon HMG-14/17 binding. The binding of HMG-14 to NCP at 140 mM NaCl (close to physiological conditions) has been shown to be cooperative even though overall binding is weaker than in the absence of salt (Postnikov et al., 1994). Point mutations in the nucleosome binding domain (Ala 21 to Pro, Lys 26 to Cys) decreased cooperative binding by a factor of 3-7; non-cooperative binding at low salt was not affected. Under the same cooperative conditions from mixtures with HMG-14/17 and shorter peptides, only complexes between NCP and homodimers are formed (Postnikov et al., 1995). The effect seems to be based on structural changes in the NCP as direct interactions between HMG molecules could almost certainly be ruled out. The binding of the full length protein as well as of a 30 amino acid peptide containing the NCP binding domain are lost upon removal of the core basic histone tails with trypsin.

![Sequence alignment of human HMG-14/17](image)

**Figure 7.1:** Sequence alignment of human HMG-14/17.

The N-terminal domain is conserved in all HMG-14/17 proteins characterized so far. The NCP binding domain has been defined as indicated (adapted from Trieschmann et al., 1995b).
Chapter 7

It had been shown that incorporation of HMG-14/17 into minichromosomes reconstituted in *Xenopus laevis* egg extracts can enhance transcription (Trieschmann et al., 1995a). Furthermore HMG-14/17 can increase spacing of NCP on a DNA template (Tremethick & Drew, 1993; Tremethick & Hyman, 1996). The C-terminal acidic domain was found to be indispensable for this effect. The presence of HMG-14/17 only enhanced transcription from chromatin templates while transcription from free DNA templates remained unaffected.

Based on the current knowledge the following image of HMG-14/17 function is emerging. The HMG-14/17 group of proteins seems to have a similar domain structure as many of the transcription activators. The NCP binding domain (N-terminus) is responsible for the site specific attachment whereas the activation domain (C-terminus) mediates the enhancer effect. The acidic activation domain of HMG-14/17 could be part of a 'cow catcher' mechanism for the basic histone tails. The resulting destabilization of the higher order structure could be an explanation for general transcription enhancement in the presence of HMG-14/17. Again, the exact mode of action as well as the targeting mechanisms remain obscure.

7.1.2 Objectives of the work

To understand the function of the HMG-14/17 proteins structural data would be of great interest. Therefore co-crystallization of HMG-17 together with the nucleosome core particle was attempted. Given the knowledge from biochemical experiments, it seemed reasonable to crystallize the full-length protein rather than shorter peptides. Firstly, the use of full length protein guarantied more stable binding especially at low salt concentration. In theory, that should be advantageous for crystallization because it is less restrictive to the crystallization conditions. Secondly, the functional domains seem to be distributed over the entire length of the protein.
7.2 Results

7.2.1 Purification of human HMG-17

Full length human HMG-17 was expressed in *E. coli* bacteria in high yields. After lysis of the cells, the supernatant with the soluble protein was incubated at 70° C for 15 minutes. This precipitated most of the cellular proteins leaving highly enriched HMG-17 in the supernatant. After two steps of cation exchange chromatography the protein was over 95% pure as probed by SDS gel electrophoresis (see figure 7.2A) and mass spectroscopy. The retention time of HMG-17 on SDS PAGE is unusually high which is probably caused by the special amino acid composition. There was a low percentage of a faster migrating contamination in the final preparation of the HMG-17 protein. This is unlikely to be due to proteolysis of the full length protein because there were no indications of proteolysis during purification. A similar contamination has been observed in earlier *E. coli* preparations of the same protein (Bustin et al., 1991). The reason could be a translational mistake, in which a base triplet is skipped (Kane et al., 1992).

7.2.2 Binding of HMG-17 to NCP

The integrity of purified HMG-17 was checked by its binding to nucleosome core particles. Binding of HMG-14/17 induces a supershift of the NCP when analyzed by native gel electrophoresis (Bustin et al., 1991). NCP species with 1 or 2 HMG-14/17 molecules bound run as distinct bands with a longer retention time than free core particle. Figure 7.2B shows an experiment in which a constant amount of α3s146 NCP was titrated with an increasing amount of HMG-17. As expected a total of two molecules of HMG-17 could be bound to the NCP. Higher HMG-17/NCP ratios did not
result in more molecules of HMG-17 bound to NCP. Similar supershifts could be obtained using NCP made from other DNA sequences or longer DNA fragments (147, 150, 154 bp). A ternary complex with 180mer NCP, H1 (material from T. Rechsteiner) and two copies of HMG-17 could be obtained in a similar way (data not shown).

![Figure 7.2: HMG-17.](image)

(A) 18% SDS PAGE of purified human HMG-17 stained with Coomassie blue. Recombinant Xenopus octamer was used as size marker (M; 15.3, 13.9, 13.8, 11.2 kD). (B) 5% Mobility shift gel of core particle with HMG-17. M is the DNA size marker. HMG-17/NCP ratio as stated above the lanes (0, 1, 1.25, 1.5, 1.75, 2, 2.25, 2.5, 2.75, 3, 4, 5, 6, 7, 8). Samples were prepared in 20 mM Tris/Cl pH 7.5 and loaded in 5% sucrose. The bands were visualized by staining with Ethidium bromide.

### 7.2.3 Attempts to crystallize a NCP/HMG-17 complex

The solubility characteristics of NCP with HMG-17 bound were very similar to free NCP. Therefore, the restrictions with respect to crystallization conditions were very much the same as encountered in NCP crystallization. The complex precipitated in the presence of millimolar concentrations of bivalent (multivalent) cations and was not stable against high salt concentrations (ammonium sulfate, citrate) often successfully used in protein crystallization. Therefore, the search for crystallization conditions was restricted to the use of low concentrations of multivalent cations and of different
alcohols often used in nucleic acid crystallization. Table 7.1 gives a summary of the different compounds tried. Trials have been set up with chicken as well as recombinant octamer using the DNA sequences α3s146, α8s147, s150 and s154. A ternary complex with s180NCP/H1/HMG-17 (together with T. Rechsteiner) was included in the trials.

<table>
<thead>
<tr>
<th>Multivalent ions</th>
<th>Monovalent ions</th>
<th>Alcohols/precipitants</th>
<th>Detergents</th>
<th>Additives</th>
</tr>
</thead>
<tbody>
<tr>
<td>MnCl₂, CaCl₂, MgCl₂, MgSO₄, MnSO₄, Spermine, Spermidine, Citrate, Jeffamine 230/400, Diaminopentane, -hexane, -heptane, -octane, Histamine</td>
<td>KCl, LiCl, NaCl</td>
<td>MPD, isopropanol, PEG 600, 1000, 6000</td>
<td>Triton X100, NP-40, ODG, N,N-Dimethyl-dodecylamin-N-oxide</td>
<td>Glycerol, Trehalose, DMF, urea, DMSO</td>
</tr>
</tbody>
</table>

Table 7.1: Sampled crystallization conditions for NCP/HMG-17 complexes.

Generally, a multivalent ion (1-10 mM) was used together with a monovalent ion (0-100 mM) and/or a precipitant (0-60%). Additives were added to the promising looking trials. The precipitant or crystalline like material was analyzed on mobility shift gels to quantitatively check the composition. However, no quality crystals could be grown so far.

In the following, I will focus on a subset of the crystallization experiments to show the problems encountered with the NCP/HMG-17 system. Crystallization under the same conditions as applied for the core particle alone yielded NCP crystals with HMG-17 remaining in solution. The NCP/HMG-17 complex often precipitated by forming an oil. 1-4 mM concentrations of Mn²⁺, Mg²⁺, spermidine, spermine or jeffamine 230 as precipitant yielded small crystal embedded in oil for the different NCP/HMG-17 complexes (α3s146/HMG-17: 5% MPD, 2 mM Spermidine, 10 mM K/Cacodylate pH6). However, due to the limited size of these crystals, a detailed analysis of their content was not possible. It is likely that the crystalline material only
contained NCP without HMG-17 (even though the oil clearly contained the correct stoichiometry of both NCP and HMG-17). The pH was generally 6 (K/Cacodylate). Lowering the pH by the use of acetate buffer did not improve crystallization. Alteration of the HMG-17 to NCP ratio (1:1 - 4:1) or addition of nonionic detergents and other additives (see table 7.1) was tried without success. Epitaxial growth experiments using positively and negatively charged lipid layers (1-palmitoyl-2-oleoyl-phosphatidyl choline/ dodecylamin, cardiolipine) or positively charged coverslips were not successful either (Darst & Edwards, 1995; Edwards et al., 1994; Okamoto & Yamada, 1982).

7.3 Discussion

It is likely that crystallization of a NCP/HMG-17 complex is on the borderline of what can be done by X-ray crystallography. Specific association between NCP and HMG-17 does not necessarily mean that the structural features of the complex are well defined. From the two NCP crystal structures presented in this thesis, we learned that the tail regions of the histones are not very well defined as no or only poor electron density could be found for them. The mode of interaction with their target (DNA) seems to be electrostatic in nature without the formation of well defined contacts. Based on the primary sequence of the HMG-14/17 proteins with their lack of secondary structure elements and a high number of charged amino acids, it is likely that at least the acidic domains of the protein show the same mode of action as the tail regions of the histones. X-ray crystallographic studies would then be very difficult to perform.

Another factor complicating crystallization is the dominant influence of the DNA on the solubility parameters and crystal packing. The high B-factors of the DNA outer strand probably inhibit the formation of proper crystal contacts between the DNA supercoils of neighboring NCP. In fact, both crystal structures only show two well defined crystal contact sites per NCP molecule: one stacking interaction on the DNA superhelix.
ends and one protein-protein contact involving H4/H3 and H2A/H2B from symmetry related particles. This crystal packing does not allow the binding of additional proteins (near the DNA ends), and a different crystal packing is unlikely to occur because of the high B-factors of the outermost phosphates.

Additionally the core particle as well as the NCP/HMG-17 complex precipitate at millimolar concentrations of multivalent cations without any significant supersaturation zone. As discussed earlier, the only way of getting large crystals of the NCP on its own was by going from intermediate to a low ionic strength. As discussed in the previous section, this method was not successful for NCP/HMG-17 or NCP/H1 (T. Rechsteiner, unpublished results) complexes.
8. Past goals and future perspectives

8.1 General

This thesis is divided into three different groups of experiments, all with the nucleosome core particle as the object of interest. Firstly, a new biochemical characterization method for the dissociation of the nucleosomal DNA was established using Fluorescence Resonance Energy Transfer (FRET). Secondly, the diffraction quality of the NCP crystals was extended to a resolution of 2.0 Å, which provided the basis for the X-ray structure determination of the NCP with two different DNA sequences. Thirdly, co-crystallization of HMG-17 in a complex with the nucleosome core particle was attempted. So far, no quality crystals could be grown of this complex.

8.2 Fluorescence studies on the NCP

We introduced a new method for monitoring the disassembly of the nucleosome core particle. Fluorescence Resonance Energy Transfer in combination with recombinant proteins and defined sequence DNA fragments offers the possibility to measure the disassembly of large, multisubunit protein-DNA complexes such as the NCP at specific sites of interaction. Fluorescence methods are generally not restricted by measurement conditions.

With this method, we could for the first time obtain data about regions of the core particle inaccessible by other methods such as intrinsic fluorescence or hydroxyl radical footprints. In this present work, we measured directly the start- and the endpoint of the DNA dissociation from the histone proteins. Combined with the dissociation of the
dimer from the tetramer-DNA complex, individual disassembly steps could be identified.

The dissociation of individual 10 bp DNA segments as defined by the histone-DNA contacts can be measured by the change in FRET between labeled sites in the 10 bp segments and new sites in the histone octamer. Table 8.1 summarizes potential labeling sites on the DNA and the protein (see also figure 5.1). The DNA sites were chosen so that phosphates pointing away from the protein can be labeled. In order to cover the ideal distance range for the FRET measurements, the donor-acceptor distances should be similar to the Ro of the chromophore pairs (approximately 50 Å for most common chromophore pairs). The protein amino acids that would have to be mutated into cysteine residues, were chosen to fulfill the distance criterion. Apart from this, the selected amino acids are located in alpha helical segments of the histones, are accessible from solution, and mutation to cysteine does not change the overall charge of

<table>
<thead>
<tr>
<th>DNA segment between DNA-protein contact</th>
<th>DNA near DNA phosphates</th>
<th>Protein sites</th>
<th>Approximate distances (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 and 3</td>
<td>H3/H4</td>
<td>A54/B241</td>
<td>H4 Gly56 (H3 Gln76)</td>
</tr>
<tr>
<td>3 and 4</td>
<td>H3/H4 and H2A/H2B</td>
<td>A44/B251</td>
<td>H2B Ser109 (H2A Asn68)</td>
</tr>
<tr>
<td>4 and 5</td>
<td>H2A/H2B</td>
<td>A34/B261</td>
<td></td>
</tr>
<tr>
<td>5 and 6</td>
<td>H2A/H2B</td>
<td>A24/B273</td>
<td></td>
</tr>
<tr>
<td>6 and 7</td>
<td>H2A/H2B and H3</td>
<td>A13/B284</td>
<td></td>
</tr>
</tbody>
</table>

Table 8.1: Potential labeling sites for FRET measurements.

H4 Gly56 is at the center of helix α2, H3 Gln76 is at the C-terminus of helix α1, H2B Ser109 is at the N-terminus of helix αC and H2A Asn68 is located at the C-terminus of helix α2. The distances were determined for the first protein mutant in the list. Use of the amino acid site given in brackets would reduce the donor-acceptor distance.
the histone octamer. With this set of donor-acceptor sites, it should be possible to measure the dissociation of individual 10 bp DNA segments.

The FRET methodology can also be used to measure the effect of different histone mutants on DNA dissociation. For example, the transfer between the DNA termini could be used to investigate the DNA binding potential of H3 variants with mutations in the αN-helix that contacts the DNA termini.

8.3 X-ray crystallography of the NCP

With the two crystal structures of the nucleosome core particle, an enormous amount of new information became available, especially for the DNA part of the complex. This resulted in entirely new insights into the nature of histone-DNA interactions in the nucleosome. For the first time, there is now structural data available for DNA in one of the most important in vivo conformations. There are, however, still unresolved questions. For example, the DNA sequence dependent stability of the different NCPs could not yet be explained satisfactorily.

In the immediate future, the crystal structure of the α3s NCP will have to be refined to 2.0 Å resolution which should further improve the accuracy of the atomic model. Together with the 2.6 Å structure of the α6s NCP, we then should have a good starting point for future experiments. A high resolution crystal structure of one of the odd numbered, symmetric NCP (α8s or α10s DNA sequence) should be solved as well, because these core particles are presumably totally symmetric.

The alpha-satellite sequences used for structure determination are devoid of an AT periodicity as assumed to be optimal for binding to the histone octamer. With a crystal structure of a NCP with DNA of correctly oriented 10 bp AT periodicity, some of the hypothesis resulting from the two crystal structures could be verified. A DNA fragment such as the Crothers sequence with a regular pattern of AT and GC base pairs
would therefore be ideal for such purposes. Unfortunately the cs146 did not form high quality crystals, but hopefully a cs147 would. If this or a similar sequence cannot be successfully crystallized, further characterization of nucleosomal DNA from crystal structures would be complicated, because if it is impossible to rationalize the crystallization in a way that sequences of interest can be crystallized, the results will always be biased in one way or another. In general, a crystallization method which is independent of crystal contacts formed by the DNA part of the NCP would allow a more unbiased interpretation of the DNA helical parameters.

The stacking interactions observed at the DNA termini may have consequences for other crystallization experiments. Assuming no overwound region in the nucleosomal DNA and conservation of the other crystal contacts, DNA 3 or 4 base pairs longer (150 or 151 bp) could form a continuous DNA helix with a correct orientation of the grooves. Variation of this DNA length in steps of 5 base pairs, so that the overall helical periodicity is conserved, could eventually help in the co-crystallization of a NCP/H1 complex. However, new crystal packing interactions are generally difficult to predict, and sequences longer than 146 base pairs often have poor translational positioning. Given the conformational flexibility of the last 5 base pairs at the DNA termini in the two crystal structures, the approach may not be straightforward. One has also to consider the consequences of the high crystallographic B-factors of the nucleosomal DNA, which is an indication for the relatively high positional flexibility of the DNA. Any crystal packing that relies on the formation of packing interactions between the DNA supercoils, which cover most of the NCP surface and exhibit a dominant influence on the solubility, is likely not to form well ordered crystals.

Reduction of the influence of the DNA on crystallization could eventually facilitate crystal formation, which could be useful in the crystallization of NCP as well as of the NCP/HMG-17 complex. There are several possibilities to reach this goal: 1) co-crystallization with an antibody fragment, 2) ternary complex formation or 3), construction of a fusion protein between one of the components and an additional
protein of known structure. Co-crystallization of protein-antibody fragments has been used successfully in the past and should not cause fundamental problems when applied to the NCP (reviewed by Kovari et al., 1995). The second approach is related to the first idea. Here, a complex between NCP and a ternary protein specifically binding to the octamer or DNA could be formed and crystallized. The third approach is based on the construction of a fusion protein between one of the NCP peptides and an additional protein with a known X-ray structure. Preliminary results from crystallization of a NCP/HMG-17-MBP fusion complex (MBP= 40.7 kD soluble monomeric maltose binding protein from E. coli with a known X-ray structure to 2.3 Å; plasmid available as part of a protein fusion system from NEB) showed significantly increased solubility of the complex in the presence of low concentrations of multivalent cations. First crystallization trials gave little rod-shaped crystals without the formation of oily precipitate. However, the HMG-17 fusion was probably not present in these crystals (as analyzed by native PAGE of dissolved crystals).
9. References


References


References


References


References


191


References


References


References


References


Appendix

Appendix I

DNA fragments (138-150bp)

as146

\[
\begin{align*}
\text{atccgggattataagacgagctattaaactctttaatttaatgc} \\
\text{tagggcccttaatttccgtcatctgactattttaggatgagttc} \\
\text{gactatatttt} & \text{agcttgcttgatgaaagttaagctatttaaaggctagggatgttatgacgtcatcggcttataaatccttgagat} \\
\text{acgaactactttcaattcgataaatattttaggacacta} \\
\end{align*}
\]

s146

\[
\begin{align*}
\text{agcttgcttgatgaaagttaagctatttaaaggctagggatgttatgacgtcatcggcttataaatccttgagat} \\
\text{acgaactactttcaattcgataaatattttaggacacta} \\
\end{align*}
\]

s150

\[
\begin{align*}
\text{atccgggattataaga} & \text{agcttgcttgatgaaagttaagctatttaaaggctagggatgttatgacgtcatcggcttataaatccttgagat} \\
\text{acgaactactttcaattcgataaatattttaggacacta} \\
\end{align*}
\]

s142

\[
\begin{align*}
\text{agcttgcttgatgaaagttaagctatttaaaggctagggatgttatgacgtcatcggcttataaatccttgagat} \\
\text{acgaactactttcaattcgataaatattttaggacacta} \\
\end{align*}
\]

s138

\[
\begin{align*}
\text{agcttgcttgatgaaagttaagctatttaaaggctagggatgttatgacgtcatcggcttataaatccttgagat} \\
\text{acgaactactttcaattcgataaatattttaggacacta} \\
\end{align*}
\]

cs146 (Crothers sequence)

\[
\begin{align*}
\text{aattcagctgacataccctgggcagctgtagttttaatcatctttaattttaggatgagttc} \\
\text{gactatattttt} & \text{agcttgcttgatgaaagttaagctatttaaaggctagggatgttatgacgtcatcggcttataaatccttgagat} \\
\text{acgaactactttcaattcgataaatattttaggacacta} \\
\end{align*}
\]

Human alpha satellite repeat 8

\[
\begin{align*}
\text{atccgggattataaga} & \text{agcttgcttgatgaaagttaagctatttaaaggctagggatgttatgacgtcatcggcttataaatccttgagat} \\
\text{agcttgcttgatgaaagttaagctatttaaaggctagggatgttatgacgtcatcggcttataaatccttgagat} \\
\text{acgaactactttcaattcgataaatattttaggacacta} \\
\end{align*}
\]

\[
\begin{align*}
\text{agcttgcttgatgaaagttaagctatttaaaggctagggatgttatgacgtcatcggcttataaatccttgagat} \\
\text{agcttgcttgatgaaagttaagctatttaaaggctagggatgttatgacgtcatcggcttataaatccttgagat} \\
\text{acgaactactttcaattcgataaatattttaggacacta} \\
\end{align*}
\]
**Appendix**

<table>
<thead>
<tr>
<th>α4s 146</th>
<th>α8s 147</th>
<th>α9s 147</th>
<th>α10s 145</th>
</tr>
</thead>
<tbody>
<tr>
<td>aattccagctgacacgcttttcgatggagcagtttccaaatacacttttgtagaatctgcaggtggatatttgatgcggacttgtacggaaaactacctcgtcaaaaggttttatgtgaaaaaaccatcatagacgtccacctataacta</td>
<td>10 20 30 40 50 60 70</td>
<td>10 20 30 40 50 60 70</td>
<td>10 20 30 40 50 60 70</td>
</tr>
<tr>
<td>aattccagctgacacgcttttcgatggagcagtttccaaatacacttttgtagaatctgcaggtggatatttgatgcggacttgtacggaaaactacctcgtcaaaaggttttatgtgaaaaaaccatcatagacgtccacctataacta</td>
<td>aaccccagctgacacgcttttcgatggagcagtttccaaatacacttttgtagaatctgcaggtggatatttgatgcggacttgtacggaaaactacctcgtcaaaaggttttatgtgaaaaaaccatcatagacgtccacctataacta</td>
<td>aaccccagctgacacgcttttcgatggagcagtttccaaatacacttttgtagaatctgcaggtggatatttgatgcggacttgtacggaaaactacctcgtcaaaaggttttatgtgaaaaaaccatcatagacgtccacctataacta</td>
<td>aaccccagctgacacgcttttcgatggagcagtttccaaatacacttttgtagaatctgcaggtggatatttgatgcggacttgtacggaaaactacctcgtcaaaaggttttatgtgaaaaaaccatcatagacgtccacctataacta</td>
</tr>
</tbody>
</table>

**Human alpha satellite repeat 1**

<table>
<thead>
<tr>
<th>α6s 146</th>
<th>α7s 146</th>
<th>α2s 146</th>
<th>α5s 146</th>
</tr>
</thead>
<tbody>
<tr>
<td>aaccccagctgacacgcttttcgatggagcagtttccaaatacacttttgtagaatctgcaggtggatatttgatgcggacttgtacggaaaactacctcgtcaaaaggttttatgtgaaaaaaccatcatagacgtccacctataacta</td>
<td>aaccccagctgacacgcttttcgatggagcagtttccaaatacacttttgtagaatctgcaggtggatatttgatgcggacttgtacggaaaactacctcgtcaaaaggttttatgtgaaaaaaccatcatagacgtccacctataacta</td>
<td>aaccccagctgacacgcttttcgatggagcagtttccaaatacacttttgtagaatctgcaggtggatatttgatgcggacttgtacggaaaactacctcgtcaaaaggttttatgtgaaaaaaccatcatagacgtccacctataacta</td>
<td>aaccccagctgacacgcttttcgatggagcagtttccaaatacacttttgtagaatctgcaggtggatatttgatgcggacttgtacggaaaactacctcgtcaaaaggttttatgtgaaaaaaccatcatagacgtccacctataacta</td>
</tr>
</tbody>
</table>

**Human alpha satellite consensus**

Only one half side is shown for the symmetric DNA sequences(s). Hereby, the sticky end restriction site which forms the dyad of the fragment after ligation is located at the left end. Note that for the odd number fragments self ligation was with a DNA fragment differing at the central base of the overhang. Sequence s150 was cloned by T. Rechsteiner.
Appendix

Appendix II

Sequence alignment of human alpha satellite higher order repeat sequences of a 2055 bp BamHI fragment from the human x-chromosome. The segment consists of 12 monomers with an average length of 171 bp. They are numbered from X1 through X12 corresponding to the order in the genome. The consensus sequence (Con; shown at the bottom) is the result of a comparison of at least 130 repeats from more than 14 human chromosomes. The numbering within the repeats is somewhat arbitrary as it depends on the restriction enzyme used to cut out the higher order repeat and thus defining the start of the first monomer (Waye & Willard, 1985; Willard & Waye, 1987).

Sequences used for fragments α3s (repeat 8/9), α6s (repeat 1) and α2s (consensus sequence) are in magenta. The CENP-B box is shown in blue.
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>GATCCGCAAG</td>
<td>.GGATATTTG</td>
<td>GAC.CCTTT</td>
<td>GAAGGTTTTC</td>
</tr>
<tr>
<td>X2</td>
<td>AATGTCGCAAG</td>
<td>TGGACCTTTA</td>
<td>GCG.GGCTTG</td>
<td>GAGGACTGTTG</td>
</tr>
<tr>
<td>X3</td>
<td>AATCTCGCAAG</td>
<td>TGGAGATTTCG</td>
<td>GAC.CCGTTT</td>
<td>GAGGCCCTGTTG</td>
</tr>
<tr>
<td>X4</td>
<td>AATCTGCGGAG</td>
<td>GAC.TCTCTT</td>
<td>GACTCTTCCTG</td>
<td>GAGATTTCCTG</td>
</tr>
<tr>
<td>X5</td>
<td>AATCTGCAAG</td>
<td>TGTATATTTT</td>
<td>GAC.CACTTT</td>
<td>GATGCTCTTG</td>
</tr>
<tr>
<td>X6</td>
<td>AATCTGCAAG</td>
<td>TGGACCTTTG</td>
<td>GAC.CTTTCT</td>
<td>GAGCATTTCA</td>
</tr>
<tr>
<td>X7</td>
<td>AATCTGCAAG</td>
<td>TGGACGTTTG</td>
<td>GAG.GGCTTT</td>
<td>GAG.CTGTTG</td>
</tr>
<tr>
<td>X8</td>
<td>AATCTGCAAG</td>
<td>GAC.TGCTTT</td>
<td>GACCTCAGTG</td>
<td>GAGTGCACAG</td>
</tr>
<tr>
<td>X9</td>
<td>AATCTGCAAG</td>
<td>TGGAGATTTG</td>
<td>GAC.CACTTT</td>
<td>GATGCTCTTG</td>
</tr>
<tr>
<td>X10</td>
<td>AATCTGCAAG</td>
<td>TGGATATTTG</td>
<td>GAC.CACTTT</td>
<td>GATGCTCTTG</td>
</tr>
<tr>
<td>X11</td>
<td>AATCTGTCAGG</td>
<td>TGGACCTTTG</td>
<td>GAC.CCTTTT</td>
<td>GAGCATTTCA</td>
</tr>
<tr>
<td>Con</td>
<td>AATCTGCAAG</td>
<td>TGGATATTTG</td>
<td>GAC.CACTTT</td>
<td>GATGCTCTTG</td>
</tr>
</tbody>
</table>

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>...</td>
<td>GGATATA</td>
<td>TCTTCCACTTA</td>
<td>AAAGCTAAAC</td>
</tr>
<tr>
<td>X2</td>
<td>...</td>
<td>GAAATA</td>
<td>TCTTCCACTTA</td>
<td>AAATCCAGAC</td>
</tr>
<tr>
<td>X3</td>
<td>...</td>
<td>GAAAGA</td>
<td>ACTTTCATATA</td>
<td>AAAACGACAT</td>
</tr>
<tr>
<td>X4</td>
<td>...</td>
<td>GGATCA</td>
<td>CCTTCCCATATA</td>
<td>ACTG..AACC</td>
</tr>
<tr>
<td>X5</td>
<td>...</td>
<td>TCTTATA</td>
<td>TCTTCCACTACA</td>
<td>AAACCTGACT</td>
</tr>
<tr>
<td>X6</td>
<td>...</td>
<td>GGATCA</td>
<td>TCTTCCCATATA</td>
<td>AAAACTAACAA</td>
</tr>
<tr>
<td>X7</td>
<td>...</td>
<td>GAAATA</td>
<td>TCTTCCACTACA</td>
<td>AAACCTGACT</td>
</tr>
<tr>
<td>X8</td>
<td>...</td>
<td>GAAGGA</td>
<td>ACTTTCATATA</td>
<td>AAAAGCCTAA</td>
</tr>
<tr>
<td>X9</td>
<td>...</td>
<td>GGAATA</td>
<td>TCTTCCCTATA</td>
<td>ACTA..AACAA</td>
</tr>
<tr>
<td>X10</td>
<td>...</td>
<td>GCTATA</td>
<td>GCTTCCAGCTTA</td>
<td>AAATCTAAAG</td>
</tr>
<tr>
<td>X11</td>
<td>...</td>
<td>GGAATA</td>
<td>TTTCCACAGAG</td>
<td>AAATCTAAAG</td>
</tr>
<tr>
<td>X12</td>
<td>...</td>
<td>TGGCCTTTTA</td>
<td>TCTTCCACTACA</td>
<td>AAAGCCGAGG</td>
</tr>
<tr>
<td>Con</td>
<td>...</td>
<td>GNAATA</td>
<td>TCTTCCCANATA</td>
<td>AAAACTGACG</td>
</tr>
</tbody>
</table>

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>TTTGCATTCA</td>
<td>CCTCACAGAG</td>
<td>TCTTGGGATG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X2</td>
<td>AATGATCCTAG</td>
<td>TGGACCATTTG</td>
<td>CTTTGGGATG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X3</td>
<td>AATGGTGTTTA</td>
<td>ACTCAGGGTAC</td>
<td>CTTTGGGATG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X4</td>
<td>TTTGTATTCA</td>
<td>ACTACAGTATG</td>
<td>CTTTGGGATG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X5</td>
<td>TTTGTATTCA</td>
<td>ACTACAGTATG</td>
<td>CTTTGGGATG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X6</td>
<td>TTTGTATTCA</td>
<td>ACTACAGTATG</td>
<td>CTTTGGGATG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X7</td>
<td>TTTGTATTCA</td>
<td>ACTACAGTATG</td>
<td>CTTTGGGATG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X8</td>
<td>TTTGTATTCA</td>
<td>ACTACAGTATG</td>
<td>CTTTGGGATG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X9</td>
<td>TTTGTATTCA</td>
<td>ACTACAGTATG</td>
<td>CTTTGGGATG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X10</td>
<td>TTTGTATTCA</td>
<td>ACTACAGTATG</td>
<td>CTTTGGGATG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X11</td>
<td>TTTGTATTCA</td>
<td>ACTACAGTATG</td>
<td>CTTTGGGATG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>Con</td>
<td>TTTGTATTCA</td>
<td>ACTACAGTATG</td>
<td>CTTTGGGATG</td>
<td>TTGGAAACG</td>
</tr>
</tbody>
</table>

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>TCTAC...</td>
<td>TTGGAAACG</td>
<td>TTGGAAAAG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X2</td>
<td>TTTTGTAG</td>
<td>TTGGAAACG</td>
<td>TTGGAAAAG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X3</td>
<td>TTTTGTAG</td>
<td>TTGGAAACG</td>
<td>TTGGAAAAG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X4</td>
<td>TTTTGTAG</td>
<td>TTGGAAACG</td>
<td>TTGGAAAAG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X5</td>
<td>TTTTGTAG</td>
<td>TTGGAAACG</td>
<td>TTGGAAAAG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X6</td>
<td>TTTTGTAG</td>
<td>TTGGAAACG</td>
<td>TTGGAAAAG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X7</td>
<td>TTTTGTAG</td>
<td>TTGGAAACG</td>
<td>TTGGAAAAG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X8</td>
<td>TTTTGTAG</td>
<td>TTGGAAACG</td>
<td>TTGGAAAAG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X9</td>
<td>TTTTGTAG</td>
<td>TTGGAAACG</td>
<td>TTGGAAAAG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X10</td>
<td>TTTTGTAG</td>
<td>TTGGAAACG</td>
<td>TTGGAAAAG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X11</td>
<td>TTTTGTAG</td>
<td>TTGGAAACG</td>
<td>TTGGAAAAG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>X12</td>
<td>TTTTGTAG</td>
<td>TTGGAAACG</td>
<td>TTGGAAAAG</td>
<td>TTGGAAACG</td>
</tr>
<tr>
<td>Con</td>
<td>TTTTGTAG</td>
<td>TTGGAAACG</td>
<td>TTGGAAAAG</td>
<td>TTGGAAACG</td>
</tr>
</tbody>
</table>
### Appendix III

**Xenopus laevis histone octamers**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>x7oct</td>
<td>wt</td>
<td>A7P</td>
<td>C28S</td>
<td>wt</td>
<td>H3  C110</td>
<td>Xenopus wt</td>
</tr>
<tr>
<td>x2oct</td>
<td>wt</td>
<td>A7P</td>
<td>C28S, C110A</td>
<td>S47C</td>
<td>H4  C47</td>
<td></td>
</tr>
<tr>
<td>xoct</td>
<td>wt</td>
<td>A7P</td>
<td>C28S</td>
<td>S47C</td>
<td>H3  C110 H4 C47</td>
<td></td>
</tr>
<tr>
<td>x6oct</td>
<td>wt</td>
<td>A7P</td>
<td>C28S/ C110A</td>
<td>wt</td>
<td>none</td>
<td></td>
</tr>
<tr>
<td>x3oct</td>
<td>K119C</td>
<td>A7P</td>
<td>C28S/ C110A</td>
<td>wt</td>
<td>H2A C119</td>
<td></td>
</tr>
<tr>
<td>x4oct</td>
<td>wt</td>
<td>87 C</td>
<td>C28S</td>
<td>wt</td>
<td>H2B C87 H3 C110</td>
<td></td>
</tr>
<tr>
<td>octG2</td>
<td>glob (13-119)</td>
<td>A7P</td>
<td>C28S</td>
<td>wt</td>
<td>H3  C110</td>
<td>H2A glob</td>
</tr>
<tr>
<td>octG1</td>
<td>wt</td>
<td>glob (24-122)</td>
<td>C28S</td>
<td>wt</td>
<td>H3  C110</td>
<td>H2B glob</td>
</tr>
<tr>
<td>x8oct</td>
<td>wt</td>
<td>glob</td>
<td>C28S/ C110A</td>
<td>wt</td>
<td>none</td>
<td>H2B glob</td>
</tr>
<tr>
<td>octG3</td>
<td>wt</td>
<td>A7P</td>
<td>glob (27-135)</td>
<td>wt</td>
<td>H3  C28/ C110</td>
<td>H3 glob</td>
</tr>
<tr>
<td>x5oct</td>
<td>wt</td>
<td>A7P</td>
<td>C28S</td>
<td>glob (20-102)</td>
<td>H3  C110</td>
<td>H4 glob</td>
</tr>
<tr>
<td>glob</td>
<td>glob</td>
<td>glob</td>
<td>glob</td>
<td>glob</td>
<td>H3  C28/ C110</td>
<td></td>
</tr>
</tbody>
</table>

All cloning done by K. Luger and T. Rechsteiner. Some of the globular octamers have been prepared by Jeannette Hayek during her diploma work.
Appendix IV

Sequence alignment of chicken (G. gallus) and X. laevis core histone proteins

The chicken histone sequence is shown at the top whereas the X. laevis sequence is shown underneath. The numbering is aligned for the X. laevis sequence starting with the first residue after the Met.

Point mutations between chicken and X. laevis are in green, other mutation sites introduced are in blue. Sequences shown in magenta correspond to the globular part of the protein (as cloned by K. Luger).

H2A

<table>
<thead>
<tr>
<th>1</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>SGRKQGGKAARKA</td>
<td>RSSRAGLQFPVGRHSLRKGNYAEVGAGAVYLAAYGYYLYTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSGRKQGGKAARKA</td>
<td>RSSRAGLQFPVGRHSLRKGNYAEVGAGAVYLAAYGYYLYTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

H2B

<table>
<thead>
<tr>
<th>1</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>PEFAKSAAPKGSKKAVTQKGGKRRKSRKESYSVYYYYKQVHDPGTKISSKAMIMN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFAKSAAPKGSKKAVTQKGGKRRKSRKESYAIYYKQVHDPGTKISSKAMIMN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

203
Appendix

**H3**

<table>
<thead>
<tr>
<th>1</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

ARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKSTEL
MARTKQTARKSTGGKAPRKQLATKAARKCAPATGGVKKPHRYRPGTVALREIRRYQKSTEL

<table>
<thead>
<tr>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
<th>110</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

LIRKLPQQLVREIAQDFKDLRFQSSAVMALQEASEAYLVGLPEDETNLCAIHAKRVTIM
LIRKLPQQLVREIAQDFKDLRFQSSAVMALQEASEAYLVALPEDETNLCAIHAKRVTIM

<table>
<thead>
<tr>
<th>130</th>
<th>140</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
</tr>
</tbody>
</table>

PKDIQLARRIRGERA

**H4**

<table>
<thead>
<tr>
<th>1</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

SGRGKGGKLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVGKRISGLYEETRGVLKV
MSGRGKGGKLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVGKRISGLYEETRGVLKV

<table>
<thead>
<tr>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

FLENVIRDAVTYEHAKRKVTAMDVYYALKRQGRTLYGFGG
FLENVIRDAVTYEHAKRKVTAMDVYYALKRQGRTLYGFGG

204
## Appendix V

### Oligonucleotides

<table>
<thead>
<tr>
<th>Name</th>
<th>Length</th>
<th>Sequence</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM03</td>
<td>9</td>
<td>5'CATC(PS)AAGCA</td>
<td>Phosphorothioate reverse PCR primer for s142</td>
</tr>
<tr>
<td>AM04</td>
<td>34</td>
<td>3'CAGTAGCCGAATTTAGGACTATAGGCTTACGAGTCGAGGTGTAAGCCCTAGGGCC</td>
<td>reverse PCR primer for s138</td>
</tr>
<tr>
<td>AM05</td>
<td>33</td>
<td>3'GCGAGTCAGCAATATTTAGCTAGGCCCTAGGGCC</td>
<td>Crothers top strand</td>
</tr>
<tr>
<td>AM06</td>
<td>9</td>
<td>5'CATC(PS)CAAGCA</td>
<td>Crothers bottom strand</td>
</tr>
<tr>
<td>AM07</td>
<td>84</td>
<td>5'TCGACGGAATTCGTGAACCTGGTCGTGGTGGTGAACCTGGTGCTAAGCCCTAGGGCC</td>
<td>PCR primer for α3s</td>
</tr>
<tr>
<td>AM08</td>
<td>84</td>
<td>3'GCTTAAGCATGAGCCACAAATCTCAGACCTGAGTGTCGTGGTGAACCTGGTGCTAAGCCCTAGGGCC</td>
<td>reverse PCR primer for α3s</td>
</tr>
<tr>
<td>AM09</td>
<td>47</td>
<td>5'GGGTCACGCAATCAGATCTGATTAGAGCTGTCGAGGTGTAAGCCCTAGGGCC</td>
<td>PCR primer for α5s</td>
</tr>
<tr>
<td>AM10</td>
<td>33</td>
<td>3'CITAGACGCTCCTAACTAAGCCTAGGGCC</td>
<td>reverse PCR primer for α5s</td>
</tr>
<tr>
<td>AM13</td>
<td>35</td>
<td>5'GAAGATCTGAAATCGCTGAACCTGGTGCTAAGCCCTAGGGCC</td>
<td>reverse PCR primer for α5s</td>
</tr>
<tr>
<td>AM14</td>
<td>37</td>
<td>3'CATTTAGACGTTCACTAAGCCCTAGGGCC</td>
<td>α6s top strand</td>
</tr>
<tr>
<td>AM15</td>
<td>85</td>
<td>5'GATCTGAAATCGCTGAACCTGGTGCTAAGCCCTAGGGCC</td>
<td>α6s bottom strand</td>
</tr>
<tr>
<td>AM16</td>
<td>85</td>
<td>3'ACTTAAGTCGAACCTAAGCCCTAGGGCC</td>
<td>α6s bottom strand</td>
</tr>
<tr>
<td>AM17</td>
<td>34</td>
<td>5'GAAGATCTGAAATCGCTGAACCTGGTGCTAAGCCCTAGGGCC</td>
<td>PCR primer for α4s</td>
</tr>
<tr>
<td>AM18</td>
<td>37</td>
<td>3'GATTTAGCGTCACTAAGCCCTAGGGCC</td>
<td>reverse PCR primer for α4s</td>
</tr>
<tr>
<td>AM19</td>
<td>35</td>
<td>5'GAAGATCTGAAATCGCTGAACCTGGTGCTAAGCCCTAGGGCC</td>
<td>PCR primer for α7s</td>
</tr>
<tr>
<td>AM20</td>
<td>33</td>
<td>3'GGGTTCACGCAATCAGATCTGATTAGAGCTGTCGAGGTGTAAGCCCTAGGGCC</td>
<td>reverse PCR primer for α7s</td>
</tr>
<tr>
<td>AM21</td>
<td>32</td>
<td>5'GAAGATCTGAAATCCAGCTGAACCTGGTGCTAAGCCCTAGGGCC</td>
<td>PCR primer for α8s halfside a</td>
</tr>
</tbody>
</table>

205
| AMO22  | 44 | 3'GTGAAAAACCATCATAGACGTCC ACCTATAACTATAGCCTAGGGC | reverse PCR primer for α8s |
| AMO23  | 32 | 5'GAAGATCTCGATTCCAGCTGAAC ATGCCCTTTTG | PCR primer for α8s halfside b |
| AMO24  | 31 | 5'GAAGATCTGATTCAGCTGAACA TGCCCTTTTG | PCR primer for α9s halfside a |
| AMO25  | 34 | 3'CATTAGACGTCCACCTATAAAC ATAGCCTAGGGC | reverse PCR primer for α9s |
| AMO26  | 31 | 5'GAAGATCTGAATCGCTGAAC ATGCCCTTTTG | PCR primer for α9s halfside b |
| AMO27  | 25 | 5'GGGGTACCGAAACGTAAGGCT GAAG | PCR primer for MBP-HMG17 fusion construct 1 |
| AMO28  | 37 | 5'GGGGTACCGCTAATTCGCCGA AACGTAAGGCTAAG | PCR primer for MBP-HMG17 fusion construct 2 |
| TR51 (*) | 42 | 5'GGAATTTCGATATCAGATCTAAG CTGCTTGATGAAAGTTGAAG | α2s top strand |
| STO486 (** | 84 | 5'TCGAGATATCAATATCCATG GAGCTCAATCAAAGAAGGTTTT CAAAACCTGCTCTATCAAAAGGA AGTTCAACTGAATTG | α2s bottom strand |
| STO487 (** | 84 | 3'CTATAGTTATAGGTGGAACGTCT AAGATGTTTCTCAAAAGGTT TGACGAGATAGTTTTCTCCAA GTGACTTTAAGCAGCT | (*) from T. Rechsteiner, (**) from S. Tan
Appendix VI

Appendix VI shows a drawing of the APCF hanging drop reactor as built by Dornier. The protein hanging drop volume is colored in red. Before and after activation, the protein sample is enclosed in the syringe. During the crystallization experiment, the drop is allowed to equilibrate against the reservoir solution stored in two 350 µl absorber blocks (not visible in the drawing).

![APCF Hanging Drop reactor](image)

On the next page crystals of NCP with *Xenopus* octamer and different DNA sequences are shown. The sequence used is indicated with each crystal. Please note that only crystals of NCP with the $\alpha_3s_{146}$, $\alpha_8s_{147}$, $\alpha_{10}s_{145}$ and $\alpha_6s_{146}$ DNA sequences diffracted to high resolution.