Table of Contents

Figures VII
Tables XI
Abbreviations XII

1. Introduction 1

1.1 Chromatin 1

1.1.1 The Nucleosome 2

1.1.2 Nucleosomes Arrays to a 30 nm Fibre 4

1.1.3 30 nm Fibre to Chromonema 7

1.2 Histone Chaperones 9

1.2.1 Asf-1 10

1.2.2 CAF-1 and HIRA 11

1.2.3 Chz-1 11

1.2.4 Nucleoplasmin 11

1.3 Nucleosome Assembly Proteins 13

1.3.1 NAP1 15

1.3.2 NAP1 Structure 15

1.3.2.1 Unstructured N- and C-terminal regions 16

1.3.2.2 Novel protein fold 16

1.3.2.3 NAP1 oligomerisation 18

1.3.2.4 NES and accessory domain 18

1.3.3 Domain II 19

1.3.4 Localisation 19

1.3.5 Histone Binding 21

1.3.6 Non-Histone Binding Partners 23

1.4 Previous Studies on Xenopus laevis NAP. 24

1.5 Research Aims 25
2. **Materials and Methods**

2.1 Stock Solutions

2.2 SDS Polyacrylamide Gel Electrophoresis (PAGE)
   2.2.1 Gel Solutions
   2.2.2 Gel Compositions
   2.2.3 SDS PAGE Protocol

2.3 Acid Urea Triton (AUT) Gel Electrophoresis
   2.3.1 Gel Solutions
   2.3.2 AUT Gel Electrophoresis Protocol

2.4 Protein Expression
   2.4.1 Bacterial Cell Culture Solutions
   2.4.2 Competent Bacterial Cell Production
   2.4.3 Transformation of Plasmids into Competent Cells.
   2.4.4 Overnight Bacterial Culture
   2.4.5 Mini Preparations of Plasmid DNA
   2.4.6 Large Volume Expression of Proteins

2.5 Protein Purification
   2.5.1 xNAP1-C1 Purification Solutions
   2.5.2 Purification of xNAP1-C1
   2.5.3 Deuteration and Purification of xNAP1-C1
   2.5.4 Histone Purification Solutions
   2.5.5 Purification of Recombinant Histones

2.6 Histone Reconstitution
   2.6.1 Histone Reconstitution Solutions
   2.6.2 Reconstitution of Histones H2A and H2B
   2.6.3 Reconstitution of Histones H3 and H4

2.7 Formation of xNAP1-C1-Histone Complexes

2.8 Analytical Size Exclusion
2.9 Dynamic Light Scattering 39
2.10 Size Exclusion Chromatography Multi Angle Laser Light Scattering (SEC-MALLS) 40
2.11 Analytical Ultracentrifugation 40
   2.11.1 Sedimentation Velocity 40
   2.11.2 Sedimentation Equilibrium 41
2.12 Small Angle X-Ray Scattering 41
2.13 Small Angle Neutron Scattering 42
2.14 Crystallisation 43
   2.14.1 Limited Proteolysis 44
3. Purification of xNAP1-C1 and Histones 45
   3.1 Optimisation of Recombinant xNAP1-C1 Purification 46
   3.2 Optimisation of Recombinant Histone Purification 50
   3.3 Reconstitution of Recombinant Histone Dimers and Tetramers 57
      3.3.1 H2A/H2B Dimer 57
      3.3.2 H3/H4 Tetramer 59
   3.4 Discussion 61
4. Hydrodynamic Analysis of xNAP 63
   4.1 Introduction 63
   4.2 Size Exclusion Chromatography 64
      4.2.1 Theory 64
      4.2.2 Ionic Strength Dependence of xNAP Oligomerisation 64
      4.2.3 Concentration Dependence of xNAP Oligomerisation 67
   4.3 Light Scattering 68
      4.3.1 Theory 68
      4.3.2 Dynamic Light Scattering 69
      4.3.3 SEC-MALLS 70
   4.4 Analytical Ultracentrifugation 72
      4.4.1 Theory 72
4.4.1.1 Sedimentation Velocity 72
4.4.1.2 Sedimentation Equilibrium 75
4.4.2 Application and Results 76
4.4.2.1 Sedimentation Velocity 76
4.4.2.2 Sedimentation Equilibrium 78

4.5 Discussion 81

5. Hydrodynamic Analysis of xNAP in Complex with Histones 85
5.1 Size Exclusion Chromatography of xNAP-Histone Complexes 85
5.1.1 Stoichiometry of xNAP-Histone Complexes 87
5.2 Light Scattering 92
5.2.1 Dynamic Light Scattering 92
5.2.2 SEC MALLS 93
5.3 Analytical Ultracentrifugation 94
5.3.1 Sedimentation Velocity 94
5.3.2 Sedimentation Equilibrium 97
5.3.2.1 $xNAP_2\cdot H2A/H2B$ 98
5.3.2.1 $xNAP_2\cdot H3/H4$ 99
5.4 Discussion 101

6. Small Angle X-Ray and Neutron Scattering 103
6.1 Theory 103
6.2 Small Angle Scattering of xNAP 108
6.2.1 SAXS 108
6.3 Small Angle Scattering of $xNAP_2\cdot H2A/H2B$ 110
6.3.1 SAXS 110
6.3.2 SANS 112
6.3.3 Modelling 113
6.4 Small Angle Scattering of $xNAP_2\cdot H3/H4$ 114
6.4.1 SAXS 114
6.4.2 SANS 116
6.4.3 Modelling

6.5 Discussion

7. Future Projects and Preliminary Results

7.1 Recombinant versus Native Histones

7.1.1 Size Exclusion Chromatography

7.1.2 Sedimentation Equilibrium

7.1.3 xNAP₂·gH2A/H2B

7.1.4 xNAP₂·gH3/H4

7.1.5 Discussion

7.2 Crystallisation

7.2.1 Crystallisation trails

7.2.2 In situ Proteolysis

7.2.2.1 α-Chymotrypsin Digestion

7.2.2.2 Trypsin

7.2.2.3 Papain

7.2.2.4 Subtilisin

7.2.3 Discussion

7.3 Electron Microscopy

7.3.1 Discussion

8. Final Discussion

8.1 xNAP

8.2 xNAP-Histone Complexes

8.3 Hypothetical Model

8.4 Biological Relevance

8.5 Future Work

9. References

Appendix I - BenchMark™ Protein Ladder.

Appendix II – xNAP1-C1

   i) pET28 plasmid map
ii) pET28 plasmid sequence and restriction sites
i) xNAP1-C1 DNA and protein sequence

Appendix III - Histones
i) pET3 plasmid map
ii) pET3 plasmid sequence and restriction sites
iii) Histone H2A DNA and protein sequence
iv) Histone H2B DNA and protein sequence
v) Histone H3 DNA and protein sequence
vi) Histone H4 DNA and protein sequence

Appendix IV – Histone Alignment
i) Histone H2A
ii) Histone H2B
iii) Histone H3
iv) Histone H4

Appendix V – Beginners Guide to SAXS Analysis

Appendix VI – Beginners Guide to SANS Analysis

Appendix VII – Densitometry Trace
Figures

Figure 1.1 - Schematic representation of DNA packaging. 1
Figure 1.2 - A breakdown of the crystal structure of the nucleosome core particle (PDB code 1A0I, Luger et al., 1997a). 3
Figure 1.3 - Putative 30 nm fibre models and the crystal structure of a tetra-nucleosome. 7
Figure 1.4 - Theoretical models for DNA compaction within the chromosome. 8
Figure 1.5 - Crystal structure of the Asf-1-H3/H4 complex. 10
Figure 1.6 - Structure and models of the Nucleoplasmin-H2A/H2B complex. 12
Figure 1.7 - NAP family binding partners. 14
Figure 1.8 - Sequence alignment of NAP1 homologues. 17
Figure 1.9 - Crystal structures of pfNAPL and pfNAPS. 18
Figure 1.10 - Surface electrostatic potential maps of pfNAPS and yNAP1. 22
Figure 3.1 - Schematic domain alignments of xNAP1, xNAP1-C1 and the crystallised region of yNAP1. 45
Figure 3.2 - Nickel affinity chromatography elution profile of xNAP1-C1 and SDS PAGE of eluted fractions. 47
Figure 3.3 - Anion exchange chromatography elution profile of xNAP1-C1 and SDS PAGE of eluted fractions. 49
Figure 3.4 - UV Spectrum of xNAP1-C1. 50
Figure 3.5 - UV spectra of histone H2A, H2B, H3 and H4 before and after hydroxyapatite incubation. 52
Figure 3.6 - Cation exchange chromatography elution profile of histone H2A and SDS PAGE of eluted fractions. 53
Figure 3.7 - Cation exchange chromatography elution profile of histone H2B and SDS PAGE of eluted fractions. 54
Figure 3.8 - Cation exchange chromatography elution profile of histone H3 and SDS PAGE of eluted fractions. 55
Figure 3.9 - Cation exchange chromatography elution profiles of histone H4 and SDS PAGE of eluted fractions. 56
Figure 3.10 - Heparin affinity chromatography elution profiles of Histones H2A/H2B and SDS PAGE of eluted fractions.  
Figure 3.11 - Size exclusion chromatography elution profiles of Histones H3/H4 and SDS PAGE of eluted fractions.  
Figure 3.12 - SDS PAGE analysis of purified proteins used in further experiments.  
Figure 4.1 - Size exclusion chromatography elution profiles of xNAP in a range of ionic conditions.  
Figure 4.2 - Size exclusion chromatography elution profiles of xNAP at various protein concentrations.  
Figure 4.3 - Dynamic light scattering of xNAP oligomers.  
Figure 4.4 - SEC-MALLS of xNAP.  
Figure 4.5 - Schematic representation of a sedimentation velocity experiment.  
Figure 4.6 - Schematic representation of a sedimentation equilibrium experiment.  
Figure 4.7 - Sedimentation velocity of xNAP.  
Figure 4.8 - Sedimentation equilibrium of xNAP.  
Figure 4.9 - Plots showing the relationship between rotor speed, calculated molecular weight and percentage protein in solution.  
Figure 4.10 - xNAP sedimentation equilibrium analysis curves.  
Figure 4.11 - The NAP dimer interface.  
Figure 4.12 - Possible mechanism for NAP oligomerisation.  
Figure 4.13 - Oligomerisation states of xNAP.  
Figure 5.1 - Size exclusion chromatography monitoring xNAP-Histone complex formation.  
Figure 5.2 - Size exclusion chromatography of different stoichiometric mixtures of xNAP₂+H2A/H2B and xNAP₂+H3/H4.  
Figure 5.3 - Size exclusion chromatography of xNAP scavenging for histones.  
Figure 5.4 - Stoichiometry determination of the xNAP-Histone complexes by SDS PAGE and densitometry.  
Figure 5.5 - Size exclusion chromatography of the xNAP₂-H2A/H2B and xNAP₂-H3/H4 complexes at various concentrations.
Figure 5.6 - Dynamic Light scattering of the xNAP₂·H2A/H2B and xNAP₂·H3/H4 complexes.

Figure 5.7 - Size exclusion chromatography with MALLS of the xNAP·Histone complexes in 150 mM NaCl.

Figure 5.8 - Sedimentation velocity of xNAP₂·H2A/H2B.

Figure 5.9 - Sedimentation velocity of xNAP₂·H3/H4.

Figure 5.10 - Sedimentation equilibrium of xNAP₂·H2A/H2B.

Figure 5.11 - Sedimentation equilibrium global analysis of xNAP₂·H2A/H2B using a single species model.

Figure 5.12 - Sedimentation equilibrium absorbance scans of xNAP₂·H3/H4.

Figure 5.13 - Sedimentation equilibrium global analysis of xNAP₂·H3/H4 using a single species model.

Figure 5.14 - Hypothesis of xNAP·Histone oligomerisation.

Figure 6.1 - Schematic Diagram of a SAXS and a SANS beamline.

Figure 6.2 - Scattering lengths for SAXS and SANS.

Figure 6.3 - Theory of a SANS contrast match experiment.

Figure 6.4 - SAXS of BSA (ID14-3, ESRF).

Figure 6.5 - Kratky plot analyses.

Figure 6.6 - SAXS of xNAP (ID14-3, ESRF).

Figure 6.7 - Comparison of SAXS curves to check for radiation-induced damage.

Figure 6.8 - SAXS of xNAP₂·H2A/H2B (I22, Diamond).

Figure 6.9 - SANS of dxNAP₂·H2A/H2B (D22, ILL).

Figure 6.10 - Ab initio models of xNAP₂·H2A/H2B.

Figure 6.11 - SAXS of xNAP₂·H3/H4 (I22, Diamond).

Figure 6.12 - SAXS of xNAP₂·H3/H4 (ID2, ESRF).

Figure 6.13 - SANS of dxNAP₂·H3/H4 (D22, ILL).

Figure 6.14 - Ab initio models of xNAP₂·H3/H4.

Figure 6.15 - Hypothetical model of the xNAP₂·Histone complexes compared with a SAXS 3D model.
Figure 7.1 - Size exclusion chromatography of xNAP·xHistone and xNAP·gHistone complexes.

Figure 7.2 - Sedimentation equilibrium global analysis of xNAP₂·gH2A/H2B using a single species model.

Figure 7.3 - Sedimentation equilibrium global analysis of xNAP₂·gH3/H4 using a single species model.

Figure 7.4 - Limited proteolysis of the xNAP·Histone complexes with α-chymotrypsin.

Figure 7.5 - Limited proteolysis of the xNAP·Histone complexes with trypsin.

Figure 7.6 - Limited proteolysis of the xNAP·Histone complexes with papain.

Figure 7.7 - Limited proteolysis of the xNAP·Histone complexes with subtilisin.

Figure 7.8 - Negative stain EM images of the xNAP₂·H3/H4 complex.

Figure 7.9 - 2D Cryo-EM images of xNAP₂·H2A/H2B.
Tables

Table 1.1 - NAP family members .......................................................................................................... 14

Table 4.1 - Sedimentation velocity of xNAP ........................................................................................ 78

Table 4.2 - Sedimentation equilibrium analyses of xNAP using a single species model ............. 79

Table 5.1 - Sedimentation velocity of the xNAP-Histone complexes ................................................. 97

Table 5.2 - Sedimentation equilibrium analyses of xNAP₂·H2A/H2B using a single species model ........................................................................................................................................ 99

Table 5.3 - Sedimentation equilibrium analyses of xNAP₂·H3/H4 data using a single species model ....................................................................................................................................... 100

Table 6.1 - Summary of small angle scattering results. .................................................................... 119

Table 7.1 - Sedimentation equilibrium analyses of xNAP₂·gH2A/H2B using a single species model ....................................................................................................................................... 123

Table 7.2 - Sedimentation equilibrium analyses of xNAP₂·gH3/H4 using a single species model. .................................................................................................................................................. 125

Table 8.1 - Summary of Results. Experimental results are shown for xNAP, xNAP₂·H2A/H2B and xNAP₂·H3/H4. ............................................................................................................................. 134
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Å</td>
<td>Ångström(s)</td>
</tr>
<tr>
<td>APS</td>
<td>ammonium persulphate</td>
</tr>
<tr>
<td>AU</td>
<td>absorbance unit(s)</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine triphosphate</td>
</tr>
<tr>
<td>AUC</td>
<td>analytical ultracentrifugation</td>
</tr>
<tr>
<td>B₄C</td>
<td>Boron Carbide</td>
</tr>
<tr>
<td>bp</td>
<td>base pairs</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>Conc</td>
<td>concentration</td>
</tr>
<tr>
<td>C-terminal</td>
<td>Carboxyl-terminal</td>
</tr>
<tr>
<td>CV</td>
<td>column volumes</td>
</tr>
<tr>
<td>D</td>
<td>diffusion coefficient</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>dn/dc</td>
<td>refractive index increment</td>
</tr>
<tr>
<td>dH₂O</td>
<td>distilled water</td>
</tr>
<tr>
<td>DLS</td>
<td>dynamic light scattering</td>
</tr>
<tr>
<td>Dₘₐₓ</td>
<td>maximum dimension</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl-sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>E.coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EMSA</td>
<td>electrophoretic mobility shift assay</td>
</tr>
<tr>
<td>f</td>
<td>frictional coefficient</td>
</tr>
<tr>
<td>f/f₀</td>
<td>frictional ratio</td>
</tr>
<tr>
<td>g</td>
<td>gram(s)</td>
</tr>
<tr>
<td>H1</td>
<td>Histone H1</td>
</tr>
<tr>
<td>H2A</td>
<td>Histone H2A</td>
</tr>
<tr>
<td>H2B</td>
<td>Histone H2B</td>
</tr>
<tr>
<td>H3</td>
<td>Histone H3</td>
</tr>
<tr>
<td>H4</td>
<td>Histone H4</td>
</tr>
<tr>
<td>HAT</td>
<td>histone acetyltransferase</td>
</tr>
<tr>
<td>HCl</td>
<td>Hydrochloric acid</td>
</tr>
<tr>
<td>His-tag</td>
<td>6 histidine tag</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>HPLC</td>
<td>high pressure liquid chromatography</td>
</tr>
<tr>
<td>hr</td>
<td>hour(s)</td>
</tr>
<tr>
<td>HA</td>
<td>Hydroxyapatite</td>
</tr>
<tr>
<td>I</td>
<td>intensity</td>
</tr>
<tr>
<td>$I_0$</td>
<td>incident light</td>
</tr>
<tr>
<td>$I_s$</td>
<td>scattered light</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-β-D-Thiogalactopyranoside</td>
</tr>
<tr>
<td>k</td>
<td>Boltzmann constant</td>
</tr>
<tr>
<td>KCl</td>
<td>Potassium Chloride</td>
</tr>
<tr>
<td>kDa</td>
<td>kilo-Daltons</td>
</tr>
<tr>
<td>L</td>
<td>litre(s)</td>
</tr>
<tr>
<td>LB broth</td>
<td>Luria-Bertani broth</td>
</tr>
<tr>
<td>LS</td>
<td>light scattering</td>
</tr>
<tr>
<td>µg</td>
<td>microgram(s)</td>
</tr>
<tr>
<td>µL</td>
<td>microlitre(s)</td>
</tr>
<tr>
<td>µM</td>
<td>micromolar</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
<tr>
<td>mAU</td>
<td>milli absorbance unit(s)</td>
</tr>
<tr>
<td>MDa</td>
<td>Mega-Dalton(s)</td>
</tr>
<tr>
<td>mg</td>
<td>milligram(s)</td>
</tr>
<tr>
<td>min</td>
<td>minute(s)</td>
</tr>
<tr>
<td>mL</td>
<td>millilitre(s)</td>
</tr>
<tr>
<td>mM</td>
<td>millimolar</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-(N-morpholino)propanesulfonic acid</td>
</tr>
<tr>
<td>$M_r$</td>
<td>relative molecular mass</td>
</tr>
<tr>
<td>MWCO</td>
<td>molecular weight cut off</td>
</tr>
<tr>
<td>$N_a$</td>
<td>Avagadro’s number</td>
</tr>
<tr>
<td>NaAc</td>
<td>Sodium Acetate</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium Chloride</td>
</tr>
<tr>
<td>NAP</td>
<td>nucleosome assembly protein</td>
</tr>
<tr>
<td>NaPhos</td>
<td>Sodium Phosphate</td>
</tr>
<tr>
<td>NES</td>
<td>nuclear export signal</td>
</tr>
<tr>
<td>NLS</td>
<td>nuclear localisation signal</td>
</tr>
<tr>
<td>nm</td>
<td>nanometre(s)</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>N-terminal</td>
<td>Amino-terminal</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein Data Bank</td>
</tr>
<tr>
<td>pET</td>
<td>plasmid for expression by T7 RNA polymerase</td>
</tr>
<tr>
<td>pl</td>
<td>isoelectric point</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethanesulfonylfluoride</td>
</tr>
<tr>
<td>p(r)</td>
<td>distance distribution function</td>
</tr>
<tr>
<td>q</td>
<td>scattering vector</td>
</tr>
<tr>
<td>q&lt;sub&gt;max&lt;/sub&gt;</td>
<td>maximum scattering vector</td>
</tr>
<tr>
<td>R</td>
<td>gas constant</td>
</tr>
<tr>
<td>R&lt;sub&gt;A&lt;/sub&gt;</td>
<td>Rayleigh scattering</td>
</tr>
<tr>
<td>R&lt;sub&gt;G&lt;/sub&gt;</td>
<td>radius of gyration</td>
</tr>
<tr>
<td>R&lt;sub&gt;H&lt;/sub&gt;</td>
<td>radius of hydration</td>
</tr>
<tr>
<td>RI</td>
<td>refractive index</td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>sec</td>
<td>second(s)</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulphate</td>
</tr>
<tr>
<td>S</td>
<td>sedimentation coefficient</td>
</tr>
<tr>
<td>S&lt;sub&gt;S&lt;/sub&gt;</td>
<td>Svedberg</td>
</tr>
<tr>
<td>SANS</td>
<td>small angle neutron scattering</td>
</tr>
<tr>
<td>SAXS</td>
<td>small angle X-ray scattering</td>
</tr>
<tr>
<td>T</td>
<td>temperature in Kelvin</td>
</tr>
<tr>
<td>TEMED</td>
<td>N, N', N''-tetramethylethylene diamine</td>
</tr>
<tr>
<td>tris</td>
<td>tris (hydroxymethyl)-aminomethane</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
<tr>
<td>V</td>
<td>volts</td>
</tr>
<tr>
<td>V&lt;sub&gt;V&lt;/sub&gt;</td>
<td>volume</td>
</tr>
<tr>
<td>V&lt;sub&gt;<code>v</code>&lt;/sub&gt;</td>
<td>apparent volume</td>
</tr>
<tr>
<td>v&lt;sub&gt;av&lt;/sub&gt;</td>
<td>partial specific volume</td>
</tr>
<tr>
<td>v/v</td>
<td>volume per volume</td>
</tr>
<tr>
<td>w/v</td>
<td>weight per volume</td>
</tr>
<tr>
<td>w/w</td>
<td>weight per weight</td>
</tr>
<tr>
<td>η</td>
<td>viscosity</td>
</tr>
<tr>
<td>π</td>
<td>Pi</td>
</tr>
<tr>
<td>ρ</td>
<td>density</td>
</tr>
<tr>
<td>ρ&lt;sub&gt;v&lt;/sub&gt;</td>
<td>average scattering density of a macromolecule</td>
</tr>
</tbody>
</table>
During the writing of this thesis a new official nomenclature for naming proteins was adopted. The comparisons to the nomenclature used in this thesis and the new current nomenclature are below.

<table>
<thead>
<tr>
<th>Thesis Nomenclature</th>
<th>Official Nomenclature</th>
</tr>
</thead>
<tbody>
<tr>
<td>hNAP</td>
<td>NAP</td>
</tr>
<tr>
<td>xNAP</td>
<td>Xenopus nap</td>
</tr>
<tr>
<td>dNAP</td>
<td>Drosophila nap</td>
</tr>
<tr>
<td>yNAP</td>
<td>Saccharomyces nap</td>
</tr>
<tr>
<td>pfNAP</td>
<td>Plasmodium nap</td>
</tr>
<tr>
<td>gH2A</td>
<td>Gallus h2a</td>
</tr>
</tbody>
</table>
1. Introduction

1.1 Chromatin

Eukaryotic DNA is packaged with a complex topology that is necessary to fit the entire genome into the relatively small nucleus requiring high levels of condensation. Exquisitely controlled, local de-compaction of the DNA is also essential for fundamental cellular processes. DNA replication and transcription rely on the fluidity of chromatin as large multi-subunit complexes, such as the replication fork and RNA polymerase, require access to specific regions of the DNA. Throughout the cell cycle the DNA is in a dynamic equilibrium between various levels of compaction (Figure 1.1).

![Figure 1.1 - Schematic representation of DNA packaging. This diagram taken from Horn and Peterson, 2002, and is a representation of the various stages of chromatin compaction, from nucleosomes to chromatids.]

Initial compaction occurs when the double stranded DNA wraps around a core of histone proteins, forming a nucleosome core particle. These particles are connected by regions of linker DNA, to which a linker histone can bind, producing a complex known as the nucleosome. Interaction between nucleosomes can promote further DNA compaction to form a 30 nm wide fibre, which can in turn self-associate to form the larger chromonema fibre with a width of 60-
80 nm. A final level of compaction produces the higher structural organisation of the well recognized metaphase X-shaped chromosomes (Woodcock and Ghosh, 2010).

An understanding of the biological and structural nature of these dynamic processes is essential in unravelling how cellular life-cycle is controlled. Multiple factors and signalling pathways have been shown to be involved in the regulation of DNA compaction yet the exact mechanisms remain unknown. This chapter provides a description of the current state of knowledge in field of chromatin organisation with a focus on the histone chaperone proteins.

1.1.1 The Nucleosome

Determining the structure and dynamics of the fundamental packaging unit, the nucleosome, is the primary step in trying to unravel the complex mechanism of DNA condensation. The 2.8 Å X-ray crystal structure of the nucleosome core particle solved by Luger et al. in 1997 was pivotal in demonstrating the detailed organisation of the histone-histone and histone-DNA interactions (Luger et al., 1997a) (Figure 1.2). Utilising human satellite DNA and Xenopus laevis histones, Luger et al. demonstrated that 146 base pairs (bp) of double stranded DNA is wrapped with 1.65 superhelical turns around an octamer of histone protein to form the nucleosome core particle (Luger et al., 1997a).

The histone octamer consists of two copies each of the four core histones: H2A, H2B, H3 and H4 (Figure 1.2 A-D). The core histones form hetero-dimers H2A/H2B and H3/H4, which have a characteristic hand-shake motif known as the histone fold (Figure 1.2 E+F). The H3/H4 dimer self-associates via a H3-H3 four helix bundle, into a (H3/H4)₂ tetramer. The (H3/H4)₂ tetramer interacts with the phosphodiester backbone of the DNA and thus organises ~70 bp, forming a complex known as the tetrasome. The nucleosome core particle is then formed by the addition of two flanking H2A/H2B dimers, which each bind to the tetrasome via a H2A-H4 four helix bundle. The addition of the H2A/H2B dimers is a two-step process; initially one dimer binds to form the hexasome followed by the incorporation of a second dimer to form the nucleosome core particle (Figure 1.2 G) (Kleinschmidt et al., 1990). The flanking H2A/H2B dimers interact with and organise about 40 bp of DNA each (Luger et al., 1997a). Linker histones, such as H1 bind to the linker DNA at the periphery of the nucleosome core particle, as the DNA enters and exits the nucleosome core particle (Thoma et al., 1979). The linker histone stabilises the nucleosome core particle by interacting with a further 20 bp to form the nucleosome (Parseghian et al., 2001).
In addition to the core histones there are also histone variants which vary in sequence from their canonical counterparts, usually in their N- and C-terminal regions. Histone variants are expressed throughout the cell cycle and are incorporated into nucleosomes in a replication-independent fashion (Tagami et al., 2004). The most researched variant is of H2A and
comparisons have shown conservation between diverse species such as *Gallus* (H2A.F), *Drosophila* (H2AvD), *Tetrahymena* (hv1) (Thatcher and Gorovsky, 1994) and *Saccharomyces* (H2A.F/Z) (Jackson et al., 1996). Knock-out experiments illustrated that the H2A variants were necessary for the viability of *Drosophila* (van Daal and Elgin, 1992) and *Tetrahymena* (Liu et al., 1996). The *Saccharomyces* H2A.Z homologue was found to be predominantly located in nucleosomes either side of the nucleosome-free promoter regions of genes, suggesting a role in transcriptional control (Raisner et al., 2005). Further studies using the *Saccharomyces* H2A.Z showed that the C-terminal tail was responsible for promoting gene transcription (Adam et al., 2001). A complementary study using the *Xenopus* H2A.Z variant showed that nucleosome arrays containing the variant did not form highly condensed chromatin above the 30 nm fibre (Fan et al., 2002). These data suggest that the incorporation of histone variants into nucleosomes has a destabilising effect which leads to the de-compaction of chromatin and nucleosome deformation, thus facilitating transcription.

Nucleosome assembly is an energetically favourable process and facilitated by the presence of a histone chaperone (e.g. NAP1), which ensures the correct histone-DNA interactions required for nucleosome formation (Andrews et al.). Nucleosomal DNA is, however, transcriptionally inactive due to the association of RNA polymerases being sterically hindered. Deformation or repositioning of the nucleosome is therefore necessary for DNA transcription and replication (Luger, 2006). The disassembly of the nucleosome occurs with the H2A/H2B dimers being removed first, followed by removal of the (H3/H4)_2 tetramer. Nucleosome disassembly is energetically unfavourable and was shown to require ATP-dependent DNA remodelling factors as well as histone chaperones (reviewed by Haushalter and Kadonaga, 2003). This is exemplified by the RSC (remodeler de structure de chromatin) chromatin remodelling complex, which can disassemble nucleosomes in the presence of the histone chaperone, NAP1 (Lorch et al., 2006). The incorporation of histone variants or the removal of the flanking H2A/H2B dimers destabilises the nucleosome and can lead to nucleosome sliding. This rather deceptive name refers to the histone octamer relocating along the DNA, hence altering the region of DNA which is inaccessible (Pennings et al., 1991). Nucleosome sliding is thought to be promoted by ATP-dependent chromatin remodelling machines, such as the SWI/SNF complex (Smith and Peterson, 2005) and histone chaperones, such as the nucleosome assembly protein (Park et al., 2005).

### 1.1.2 Nucleosomes Arrays to a 30 nm Fibre

Multiple nucleosomes are connected by regions of linker DNA ranging from 10-75 bp depending on the organism, which together form the nucleosome array. The nucleosome array
is sometimes referred to as the ‘beads on a string model’ due to its appearance in electron microscopy images (Gusse and Chevaillier, 1980). Chromatin is notoriously difficult to prepare in its native form, thus making characterisation of the higher order chromatin structure very difficult. A model system has therefore been devised for analysing DNA compaction, the 208-12 nucleosome array (Hansen, 2002). This consists of 12 repeats of DNA with a length of 208 bp each, assembled around purified histone octamers to form a chain of 12 nucleosomes. Characterisation of this simplified model has led to some major breakthroughs in our understanding of chromatin dynamics. AUC studies were performed with the 208-12 nucleosome array in the presence and absence of multivalent cations (e.g. Mg$^{2+}$) (Schwarz and Hansen, 1994). In the absence of these cations, the sedimentation coefficient of the array (29 S) suggested no compaction had occurred. However, in the presence of 1-2 mM divalent cations the sedimentation coefficient (40 S) suggested that the nucleosome arrays had condensed into a 30 nm wide fibre. As the cation concentration was increased further still, self-association of the fibres into larger complexes of 55 S were observed (Schwarz and Hansen, 1994).

The core histones have relatively long, positively charged, N-terminal tails which protrude from the nucleosome core. In the presence of divalent cations, nucleosomes can interact with each other via the N-terminal tails of the core histones leading to condensation of the chromatin (McGhee et al., 1983). Using sedimentation velocity and 2-4 mM MgCl$_2$, Tse and colleagues utilised a mixture of native and proteolysed histones without their N-terminal tails to illustrated that only the N-terminal tails of H3 and H4 were required for the formation of the 30 nm fibre (40 S), whereas, the N-terminal tails of all four core histones were necessary for the formation of the 60 nm fibre (55 S) (Tse and Hansen, 1997). Linker histones also have highly charged N- and C-terminal tails, which were shown to stabilise the higher order states of compaction through interaction with the linker DNA (Carruthers et al., 1998).

Histone acetyltransferases (HATs) can acetylate N-terminal histone tails, reducing chromatin compaction and increasing transcription by RNA polymerases (Tse et al., 1998). The activity of the HAT p300/CBP was highlighted in an *in vitro* assay by Georges *et al.* in 2002, in which they deciphered the minimal components needed for transcription of a specific gene in a 208-12 nucleosome array. The essential factors for transcription included the transcription factor CREB, the co-activator p300/CBP and an RNA polymerase. A mutant version of the 208-12 nucleosome array containing tail-less histone octamers was then examined. Transcription of genes in the mutant array did not require the presence of p300/CBP. These data suggest that
the histone tails are involved in nucleosome-nucleosome interactions which are disrupted by p300/CBP, presumably by acetylation (Georges et al., 2002).

Acetylation is one of several post translational modifications to occur upon the N-terminal histone tails. A combination of modifications including acetylation, methylation, ubiquitination and ADP-ribosylation, generates a complex code which acts as a labelling mechanism of DNA (Wolffe and Hayes, 1999). The modification sites are highly conserved between organisms, suggesting they play a key role within the cell (Strahl and Allis, 2000). Multiple modifications of one amino acid can occur, such as mono-, di- and tri- methylation of lysine 4 (K4) on histone H3, forming another level of complexity to the code (Eissenberg and Shilatifard). The three states of methylation of H3K4 have been associated with various degrees of transcriptional activity (Eissenberg and Shilatifard). This highly dynamic and complex code has been reviewed by Strahl and Allis, 2000, Murr, 2010 and Cosgrove and Wolberger, 2005.

The structure of the 30 nm fibre has been a contentious issue for over two decades, with three alternative hypotheses proposed on the basis of different techniques. The one-start solenoidal model (Widom and Klug, 1985), the two-start supercoiled model (Woodcock et al., 1984) and the two-start twisted model (Williams et al., 1986) (Figure 1.3 A-C) were predicted using X-ray diffraction, electron microscopy and small angle scattering respectively. A relatively recent crystal structure of four conjoined nucleosomes (a tetra-nucleosome) has helped to resolve this anomaly (Figure 1.3 D) (Schalch et al., 2005).

The relative positions of nucleosomes in the crystal structure of the tetra-nucleosome, were not consistent with the one-start fibre model (Figure 1.3 A). The two remaining putative fibre conformations were assessed using the tetra-nucleosome structure as a basis for nucleosome stacking using the contortional limitations of the linker DNA as a further constraint (Schalch et al., 2005). It had been identified previously that the exposed charged surfaces of the H4 and H2A are involved in nucleosome-nucleosome interactions (Dorigo et al., 2004). However, in the solenoid model, the H4 and H2A surfaces were too far away from each other to interact. A closer proximity was observed with the ‘zig-zag’ formation allowing for nucleosome-nucleosome interactions. The most likely conformation of the 30 nm fibre was therefore deemed to be the two-start twisted model (Schalch et al., 2005) (Figure 1.3 C).
1.1.3 30 nm Fibre to Chromonema

Despite clear low resolution images of the chromosomes, very little detailed information is known about the higher compaction states of chromatin, including the chromonema (60-80 nm fibre) and chromatids (100-300 nm fibre). In vitro experiments showed that DNA was predominantly present in a 30 nm fibre in low concentrations of divalent cations (1-2 mM), whereas, in higher ionic strengths the chromatin favoured the higher compaction states (Schwarz and Hansen, 1994). The divalent cation concentration of an interphase nucleus was determined by a 3D high resolution scanning ion microprobe to be 6-10 mM (Strick et al.,
suggesting that chromatin is likely to be highly compact. This is consistent with the
electron microscopy studies, in which 100 nm fibres were observed throughout mitosis
(Belmont et al., 1987), and 300 nm fibres were observed in interphase chromosomes (Belmont
et al., 1989). The quaternary arrangement of chromatin above the 30 nm fibre is also
controversial. The scientific community are split between several theoretical models which
have been reviewed by Woodcock and Ghosh and Ushiki et al., 2002 (Figure 1.4). Due to the
ambiguity of the current data, there are even some who believe no regular structure exists at
all (van Holde and Zlatanova, 1995).

Studies have been performed to determine if multiple compaction states of a DNA can occur at
any one time. A mixture of highly compacted chromatin has been observed in conjunction with
local de-condensation at certain loci in vitro (Tumbar et al., 1999) and in vivo (Muller and
Leutz, 2001). These findings suggest that local remodelling of nucleosomes can be
independent of the global compaction state. The control of DNA compaction is clearly a
dynamic process, requiring many factors to regulate fluidity. It is believed that multi-subunit
complexes including; ATP-dependent remodellers, histone chaperones, topoisomerases and

Figure 1.4 - Theoretical models for DNA compaction within the chromosome. This figure, which details four
putative models of chromatin compaction, was reproduced from Ushiki et al., 2002. A) The folded fibre model in
which the chromatin fibre is tangled in various directions. B) The successive helical coiling fibre model, the 30 nm
fibre coils into a 200 nm fibre which further coils into a sister chromatid. C) The scaffold and radial loop model,
chromatin fibres are folded into loops which are arranged into a chromatid in a radial conformation. D) The
helical coiling of radial loops model, the 30 nm fibres are folded into radial loops forming the 200 nm fibre which
are helically arranged within the chromatid.
transcription factors, are key players within this complex regulatory system (Krogan and Hughes, 2006).

### 1.2 Histone Chaperones

Histones are predominantly basic proteins which also contain hydrophobic and acidic patches. They therefore repel each other at physiological pH (Karantza et al., 1996), and form non-nucleosomal aggregates with DNA (Tyler, 2002). Histone chaperones are a group of acidic proteins which bind histones, neutralising their charge preventing them from forming non-specific interactions. Histone chaperones play a vital role in chromatin regulation including: shuttling of nascent histones from the cytoplasm into the nucleus, assembly and disassembly of nucleosomes, incorporation of histone variants into nucleosomes and histone storage (reviewed by Akey and Luger, 2003).

There is a diverse range of histone chaperones, which have a wide range of functions within the cell. Specialised chaperones, such as nucleoplasmin, can interact with the H2A/H2B dimer (Taneva et al., 2008), whereas other chaperones such as Asf-1 have a higher preference for H3/H4 (Natsume et al., 2007). Chaperones such as the nucleosome assembly protein (NAP), can bind all the core histones and the linker histone, although with varying affinities (Kepert et al., 2005; McBryant et al., 2003; Park et al., 2005). In contrast, the Chz-1 chaperone shows very high specificity, binding preferentially to the H2A.Z/H2B dimer over the canonical H2A/H2B dimer (Zhou et al., 2008).

Acidic patches are a common feature of histone chaperones such as NAP (Park and Luger, 2006b), nucleoplasmin (Dutta et al., 2001) and Asf-1 (Mousson et al., 2007). These acidic domains neutralise the charge on histones but are not involved in the affinities for specific histones. The acidic domains are therefore thought to stabilise the Chaperone-Histone interactions. Some histone chaperones share a common tertiary fold, involving an anti-parallel β-sheet. The Asf-1-H3/H4 crystal structure revealed that this β-sheet domain was interacting directly with the histones (Natsume et al., 2007) (Figure 1.5). Similar β-sheets can be seen in the NAP1 (Park and Luger, 2006b) and nucleoplasmin structures (Dutta et al., 2001) and are also predicted to be present in the histone binding domains of two other histones chaperones; HIRA (DeSilva et al., 1998) and CAF-1 (Verreault et al., 1996).

Many binding partners of histone chaperones have been identified, including ATP-dependent remodellers (Krogan and Hughes, 2006), protein kinases, nuclear import and export factors.
(Altman and Kellogg, 1997), transcription factors and also other histone chaperones (De Koning et al., 2007). A complex pathway of chaperone activity, dependent upon post translational modification signals, is becoming ever more apparent. Chaperones are thought to be involved in an energetically favourable pathway that includes folding, orientating, and oligomerising the histones ready for deposition onto the DNA (reviewed by Das et al. and Park and Luger, 2008). These pathways are thought to be controlled by signals, such as acetylation and phosphorylation. The diverse roles of histone chaperones are exemplified by a small subset of well characterised proteins including; Asf-1, CAF-1, HIRA, Chz-1, Nucleoplasmin and NAP1, which are described below.

1.2.1 Asf-1

The first crystal structure of a histone chaperone in complex with a histone fold dimer was that of Anti silencing factor (Asf-1) and H3/H4 (Natsume et al., 2007) (Figure 1.5). The concave hydrophobic region of Asf-1 is surrounded by charged surfaces and is responsible for binding the C-terminal regions of H3 and H4. An anti-parallel β-sheet is formed between H4 and the last β-strand of Asf-1 (Natsume et al., 2007). Asf-1 binds to H3/H4 dimer in an orientation which inhibits (H3/H4)₂ tetramer formation (English et al., 2005).

Asf-1 also binds to H3 variant complexes H3.1/H4 and H3.3/H4 presumably in a similar manner (De Koning et al., 2007). In vitro studies have shown an increased histone deposition in the presence of Asf-1 (Polo and Almouzni, 2006). However, in vivo analysis showed no signs of Asf-1 depositing the histones into the nucleosomes (Mello et al., 2002). This led to the conclusion that Asf-1 acts as a histone donor for other histone chaperones, which subsequently deposit
the histones for nucleosome assembly (Mello et al., 2002; Polo and Almouzni, 2006). Asf-1 is also known to interact with the HAT Rtt109, which consequently acetylates histone H3K56. This modification is thought to promote the donation of histones to other histone chaperones (Recht et al., 2006; Schneider et al., 2006).

1.2.2 CAF-1 and HIRA

The histone chaperones CAF-1 (Tyler et al., 2001) and HIRA (Daganzo et al., 2003) have been shown to bind both the H3/H4 dimer and Asf-1, making them putative histone acceptors. They have also been shown to subsequently deposit the histones for nucleosome formation (Tagami et al., 2004). The B-domain of HIRA interacts with Asf-1 via an anti-parallel β-hairpin which forms a β-sandwich with the Asf-1 N-terminal core domain (Tang et al., 2006). HIRA has been associated with histone deposition in a replication-independent manner (Ray-Gallet et al., 2002). Conversely, CAF-1 is thought to be the predominant chaperone in replication-dependent pathways (Quivy et al., 2001). It too binds the Asf-1 N-terminal domain via a p60 sub-domain, which is homologous to the B-domain of HIRA (Tyler et al., 2001). During replication, an RNA polymerase/DNA clamp recruits the CAF-1-Histone complex to the nascent DNA, suggesting a trigger for deposition (Shibahara and Stillman, 1999).

1.2.3 Chz-1

A Chz-1 trimer binds specifically to the variant dimer H2A.Z/H2B in preference to the canonical H2A/H2B (Luk et al., 2007). A conserved histone binding motif (CHZ domain) was identified and shown to be necessary and sufficient for histone binding (Luk et al., 2007). A structure of this CHZ domain bound to H2A.Z/H2B was later solved using NMR (Zhou et al., 2008). These NMR data suggests that the CHZ domain is unstructured in solution, and that it forms three α-helices upon histone binding (PDB code 2JSS, Zhou et al., 2008). Due to nucleosome disassembly being an energetically unfavourable process, the replacing of the canonical H2A/H2B dimer with the variant H2A.Z/H2B dimer requires an ATP-dependent remodelling complex. One such complex is the SWR1 remodelling complex which catalyses the histone exchange, which in turn promotes transcription (Luk et al., 2007).

1.2.4 Nucleoplasmin

Nucleoplasmin binds to the H2A/H2B dimer specifically, and is expressed in higher eukaryotes only, suggesting a specialised role. One such role of nucleoplasmin is to promote histone storage, observed in Xenopus oocytes (Mills et al., 1980). Each oocyte must store enough maternal histones to package approximately 10,000 somatic cells (Woodland and Adamson, 1977). The crystal structure of the nucleoplasmin core revealed a pentameric ring
conformation, with each monomer consisting of eight β-sheets forming a β-barrel (Dutta et al., 2001) (Figure 1.6 A). The unstructured C-terminal domains that were not visible in the crystal structure are thought to be involved in histone binding (Hierro et al., 2001).

Solution studies suggested that the nucleoplasmin pentamer self-associates into a decamer, facilitating the storage of five histone octamers (Dutta et al., 2001). Recent electron microscopy reconstructions have suggested that the pentameric ring can provide a storage environment for five H2A/H2B dimers (Ramos et al.). Ramos et al., taking into consideration the charged surfaces of the histones, have produced a model where the histones are bound to the nucleoplasmin in a novel pentameric ring conformation (Figure 1.6 B) (Ramos et al.). The 3D EM model consists of two circular discs, where the nucleoplasmin pentamer and the H2A/H2B pentamer are thought to be located. These discs are joined by five linker regions thought to be where the C-terminal region of the nucleoplasmin reside, forming the interactions with the histones (Figure 1.6 C).

Histone chaperones are key factors in chromatin remodelling and are often specialised for their roles in histone neutralisation, deposition, storage and nuclear import. The focus of this thesis is the histone chaperone NAP, which, rather than being specialised for one particular function, is thought to be involved in multiple cellular activities.
1.3 Nucleosome Assembly Proteins

The first protein in the nucleosome assembly protein (NAP) family was identified from HeLa cells in 1983 (Ishimi et al., 1983). NAP homologues have since been identified in all examined eukaryotes, suggesting a key role within the cell. NAP has been found in diverse organisms such as: *Saccharomyces cerevisiae* (Ishimi and Kikuchi, 1991), *Nicotiana tabacum* (tobacco), *Oryza sativa* (rice) (Dong et al., 2003), *Drosophila melanogaster* (Ito et al., 1996), *Xenopus laevis* (Steer et al., 2003), *Plasmodium falciparum* (Chandra et al., 2005) and *Homo sapiens* (Chang et al., 1997). NAP1 depletion studies led to abnormalities during development in *Xenopus laevis* (Abu-Daya et al., 2005) and were lethal in *Plasmodium falciparum* (Gill et al.) and *Drosophila melanogaster* (Lankenau et al., 2003).

In lower eukaryotes, such as *Saccharomyces*, only one NAP homologue is present - NAP1. In higher eukaryotes there are multiple NAP isoforms which can be divided into five subsets based on which system they were first identified: NAP1 (Kellogg et al., 1995), NAP1-Like (NAP1-L) (Simon et al., 1994), SE translocation (SET) (Matsumoto et al., 1995), CASK (calcium/calmodulin-dependent serine protein kinase)-interacting NAP (CINAP) (Wang et al., 2004) and testis specific protein Y-encoded (TSPY) (Ozbun et al., 2001). The family members were first identified within many different systems, leading to a nonsensical naming system. Some members are known by more than one name and there is no defining difference between proteins termed NAP1 and NAP1-L. The names and functions of the NAP family proteins are summarised in Table 1.1.
<table>
<thead>
<tr>
<th>Protein name</th>
<th>Other names</th>
<th>Organism</th>
<th>Remarks</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAP1</td>
<td></td>
<td>Yeast</td>
<td>Effect on expression of 10% of the genome</td>
<td>(Ohkuni et al., 2003)</td>
</tr>
<tr>
<td>xNAP1L</td>
<td></td>
<td>Frog</td>
<td>Required for signalling in primitive hematopoiesis.</td>
<td>(Steer et al., 2003)</td>
</tr>
<tr>
<td>hNAP1L</td>
<td></td>
<td>Human</td>
<td>Increased embryonic lethality</td>
<td>(Lankenau et al., 2003)</td>
</tr>
<tr>
<td>NAP1L1</td>
<td></td>
<td>Human</td>
<td>Human counterpart of yNAP1</td>
<td>(Ishimi et al., 1987)</td>
</tr>
<tr>
<td>NAP1L2</td>
<td></td>
<td>Mouse</td>
<td>Specifically expressed in brain tissue, embryonic lethality</td>
<td>(Rogner et al., 2000)</td>
</tr>
<tr>
<td>NAP1L3</td>
<td>MB20</td>
<td>Human</td>
<td>Expressed in brain tissue</td>
<td>(Shen et al., 2001)</td>
</tr>
<tr>
<td>NAP1L4</td>
<td>NAP2</td>
<td>Human</td>
<td>Gene regulation and histone shuttling</td>
<td>(Rodriguez et al., 1997)</td>
</tr>
<tr>
<td>NAP1L5</td>
<td>SET</td>
<td>Mouse</td>
<td>Stimulated the elongation of DNA replication</td>
<td>(Smith et al., 2003)</td>
</tr>
<tr>
<td></td>
<td>TAF-1β</td>
<td>Human/Mouse</td>
<td>Putative HLA-DR associated protein</td>
<td>(Matsumoto et al., 1993)</td>
</tr>
<tr>
<td></td>
<td>PHAPII</td>
<td>Putative HLA-DR associated protein</td>
<td>(Beresford and Boss, 2001)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>I22PP2A</td>
<td>Inhibits protein phosphatase A</td>
<td>(Li et al., 1995)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IGAAD</td>
<td>Inhibits granzyme-A-activated DNase</td>
<td>(Fan et al., 2003)</td>
<td></td>
</tr>
<tr>
<td>CINAP</td>
<td>Mouse</td>
<td>Widely expressed in all tissues but 3-fold more in the brain</td>
<td>(Lin et al., 2006)</td>
<td></td>
</tr>
<tr>
<td>TSPY</td>
<td>Human</td>
<td>Expressed in testis, encoded on the mammalian Y chromosome</td>
<td>(Schnieders et al., 1996)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.1 - NAP family members. This table was adapted from Park and Luger, 2006b. It summarises the five NAP family subsets including nomenclature, origin and main references.

![Figure 1.7 - NAP family binding partners.](image-url)

Figure 1.7 - NAP family binding partners. The figure is reproduced from Park and Luger, 2006b and shows the multiple binding partners of the NAP family. The colour coding scheme indicates the cellular activity that the binding partner is involved in when interacting with NAP.
Table 1.1 lists the NAP proteins which have been identified in various eukaryotes as well as summarising the main findings for each homologue. Given the central role of the NAP proteins, ubiquitous expression throughout all cell types is expected. Surprisingly, however, a degree of tissue specific expression has been identified for the NAP proteins in tissues including the brain, blood and testis. Figure 1.7 summarises the multiple binding partners of the NAP family and illustrates which cellular activity is prompted by each binding partners. This research is focussed on the highly conserved NAP1 isoform which is found throughout eukaryotes. The NAP1 binding partners and cellular activities are described in more detail in section 1.3.5.

1.3.1 NAP1

High conservation is seen throughout the NAP1 homologues, exemplified by the distant relatives Saccharomyces and Xenopus sharing 33% sequence homology (Figure 1.8). Information about the numerous biologically relevant binding partners and post translational modifications of NAP1 is becoming available, along with a more detailed understanding of cellular activities. Cellular functions include; histone shuttling (Mosammaparast et al., 2002), cell cycle control (Kellogg et al., 1995), transcriptional regulation (Walter et al., 1995) histone variant exchange (Okuwaki et al., 2005) and facilitating nucleosome sliding (McQuibban et al., 1998). These cellular functions are modulated by the interaction with other proteins described further in section 1.3.5. In addition to these cellular functions which are universal to all cells, the Xenopus laevis NAP1 protein was shown to perform a highly specific role in embryogenesis. Depletion of xNAP1 revealed that it is required for the progression of the hematopoiesis pathway within developing embryos (Abu-Daya et al., 2005).

The best described NAP1 is from Saccharomyces cerevisiae (yNAP1) which was shown to form a strong dimer by AUC (Ishimi et al., 1983) even in very high ionic strengths (Toth et al., 2005). In lower more physiological ionic strength, further oligomerisation of yNAP1 was observed using size exclusion chromatography (McBryant et al., 2003) and AUC (Toth et al., 2005). The first crystal structure of a NAP1 protein was also of the Saccharomyces homologue solved in 2006 by Park et al. Subsequent crystal structures of the Plasmodium falciparum NAPs (pfNAPS and pfNAPL) illustrated the structural similarity between the homologues and led to the determination of a putative histone binding surface.

1.3.2 NAP1 Structure

Currently there are crystal structures for only two proteins from the NAP1 sub-family; pfNAPL (Gill et al.) and yNAP1 (Park and Luger, 2006b). Here we focus on these NAP1 protein structures and the sequence homology with Homo sapiens and Xenopus laevis NAP1. Recently
another *Plasmodium falciparum* NAP was identified, pfNAPS, for which a crystallography structure is now available (Gill et al.). The pfNAPS protein is part of the SET sub-family but is also discussed for comparison (Figure 1.8). Figure 1.8 and 1.9 form the basis for sequence and structural comparisons for the following sub-sections.

1.3.2.1 Unstructured N- and C-terminal regions

The crystal structures for yNAP1, pfNAPL and pfNAPS contained ill defined density for their C- and N-terminal regions, suggesting that they were disordered (Gill et al.; Gill et al., 2009; Park and Luger, 2006b). Crystallisation of a truncated fragment of yNAP1, lacking the C- and N-terminal regions (residues 74-265) crystallised within the same space group as the wild type protein, confirming that the terminal regions were not involved in crystal packing contacts and that they were in fact unstructured (Park and Luger, 2006b). The N-terminal region, before the start of the first known or predicted α-helix, varies in length between the NAP1 homologues. *Saccharomyces* has the longest N-terminal region at 70 residues, *Xenopus* and *Homo* have 52 residues and *Plasmodium* (pfNAPL) has just 49 residues (Figure 1.8).

1.3.2.1 Novel protein fold

The yNAP1 structure had well defined electron density for the central residues 70-370 with data extending to a resolution of 3 Å (Park and Luger, 2006b). This central conserved domain has been shown to be necessary and sufficient for both histone binding and nucleosome assembly (Fujii-Nakata et al., 1992). The dimeric yNAP1 structure revealed a previously uncharacterised protein fold. Each monomer consists of two domains, I and II, responsible for dimerisation and protein folding respectively. Domain I consists of three α-helices which lead into domain II that comprises of an anti-parallel β-sheet, a protruding β-hairpin and three α-helices. This novel fold has subsequently been identified in the other NAP family members such as pfNAPL (Gill et al., 2009), pfNAPS (Gill et al.), and yVps75 (Tang et al., 2008). This fold can be split into 4 sub-domains which are denoted by four colours in Figure 1.8; the dimerisation helices (dark blue), the accessory helices (green), the amphipathic β-sheet (light blue) and the so called protective α-helices (red) which shield the hydrophobic residues in the β-sheet.
Figure 1.8 - Sequence alignment of NAP1 homologues. Two orientations of the crystal structure of yNAP1 (pdb 2Z2R) are shown (Park and Luger, 2006b) and the corresponding 3D structure is shown above the alignment below. The sequence alignment of xNAP1, xNAP1-C1 fragment (see section 1.4), hNAP1, yNAP1, pfNAPL and pfNAPS was created using the program Vector NTI (Invitrogen). Residues in blue and green are identical or conserved in relation to the xNAP1 sequence respectively.
1.3.2.2 NAP1 oligomerisation

The mechanism behind the strong NAP1 dimerisation was clarified by the yNAP1 crystal structure (Park and Luger, 2006b). Interactions between the long \( \alpha \)-helices of two NAP1 monomers form a novel anti-parallel non-coiled-coil motif which buries 20 hydrophobic residues leading to high stability. The dimerisation is further stabilised by two \( \alpha \)-helices and the connecting loops in domain I. The nature of the dimerisation leads to a sigmoidal curvature in one viewing plane and a dome shape in the other plane. This is not seen in the pfNAPL crystal structure due to crystallisation conditions rendering the protein a monomer (Gill et al., 2009). The hydrophobic residues involved in NAP1 dimerisation are well conserved among *Homo*, *Xenopus* and *Plasmodium* (Figure 1.8), suggesting that these dimerise in a similar fashion. The main dimerisation helix is 51 residues in both yNAP1 and pfNAPL with a similar length predicted for hNAP1 and xNAP1 (Gill et al.). The dimerisation helix for pfNAPS is notably shorter with only 41 residues (Gill et al.).

1.3.2.3 NES and accessory domain

A putative nuclear export sequence (NES) is located on the dimerisation helix which is conserved between yNAP1, hNAP1, xNAP1 and pfNAPL but not pfNAPS (Figure 1.8, blue box) (Park and Luger, 2006b). In the yNAP1 crystal structure, this putative NES is masked by an
accessory domain on the conjoining helix (Figure 1.8, green helix). The accessory domain is, however, absent in other NAP1 homologues such as pfNAPL and pfNAPS (Figure 1.8). This masking has been linked to the shuttling of histones into the nucleus and is discussed further in section 1.3.3.

1.3.2.4 Domain II

Domain II of the NAP1 family consists of a well conserved amphipathic β-sheet (Park and Luger, 2006b) that has also been identified in other histone chaperones such as Asf-1 (Figure 1.5) and nucleoplasmin (Figure 1.6). The Asf-1-H3/H4 structure reveals that this β-sheet is the site of histone binding, and is also involved in nucleosome assembly (Natsume et al., 2007).

One side of the yNAP1 β-sheet contains 53% hydrophilic residues, whereas, the other side is 87% hydrophobic. The hydrophobic side is buried by the protective α-helices in domain II increasing stability of the fold. Two regions with high conservation throughout the NAP1 species have been identified: sequence A and B (Figure 1.8, red boxes). Sequence A is involved in burying the hydrophobic β-sheet of the protein and thus is thought to be important for correct protein folding. Sequence B is located at a linker region connecting the amphipathic β-sheet (light blue) and the protective α-helices (red) (Park and Luger, 2006b).

The extended β-hairpin identified in the yNAP1 structure contains a putative nuclear localisation signal (NLS) (Dingwall et al., 1988; Park and Luger, 2006b). Sequence homology of the putative NLS is observed in hNAP1, xNAP1 and pfNAPL, but not however, for pfNAPS (Figure 1.8). Despite the sequence homology, the density of the β-hairpins was not well defined in the pfNAPL structure (Gill et al., 2009). This could be due to the flexibility of the hairpins when they are not involved in crystal contacts, as seen by Park et al., in 2006.

Mutations in the protruding β-hairpins blocked the oligomerisation of the yNAP1 to species larger than a dimer (Park et al., 2008). This suggests that oligomerisation occurs by the β-hairpins forming β-sheets with neighbouring NAP1 molecules (Park et al., 2008). Consistent with this hypothesis was the observation of β-sheet formation between three neighbouring β-hairpins from adjacent dimers within the crystal asymmetric unit (Park and Luger, 2006b).

1.3.3 Localisation

Drosophila (d) NAP1 (Ito et al., 1996) and hNAP1 (Rodriguez et al., 1997) are predominantly located in the cytoplasm from G2 through to G1 in the cell cycle. During S phase, however, they relocate to the nucleus, coinciding with DNA replication and the need for nascent histones for the packaging of the newly synthesised DNA (Ito et al., 1996) (Rodriguez et al., 1997). Immunohistochemistry studies of Xenopus laevis embryos illustrated that xNAP1 was
also located in both the cytoplasm and nucleus; suggesting that nucleocytoplasmic shuttling occurs (Steer et al., 2003). The yNAP1 homologue was thought to be predominantly cytoplasmic but disruption of the NES region resulted in the protein relocating to the nucleus (Miyaji-Yamaguchi et al., 2003), suggesting that the NES has a regulatory role in sub-cellular localisation of the NAP1 proteins.

The accessory helix in domain I was predicted by Park et al. in 2006 to modulate the NES activity. Complementary to this prediction is the fact that the pfNAPL protein contains a homologous NES and NLS sequence but no accessory domain, and is predominantly located in the cytoplasm with no suggestion of nuclear localisation (Gill et al., 2009). This suggests that if the putative NLS region is active, it requires other signalling factors for nuclear localisation. The pfNAPS homologue does not contain the NES, NLS or an accessory domain and is located in the nucleus (Gill et al.), suggesting that a NLS on the histones is responsible for its nuclear localisation.

The nuclear import factor karyopherin was identified as a binding partner for the yNAP1-Histone complex using both in vitro GST-tagging experiments and in vivo pull-down assays (Mosammaparast et al., 2002). The karyopherin recognises the NLS on the histone C-terminal tails, an interaction which has a higher affinity in the presence of yNAP1 (Mosammaparast et al., 2002). The karyopherin, when bound to the NAP-histone complex, can traverse the nuclear membrane by interacting with the nuclear pore complex (NPC). The NPC is a multi-subunit complex which spans the nuclear membrane, forming a hydrophilic core for the nuclear import and export of cellular components including proteins. Dephosphorylation of NAP1 at the G1/S boundary is thought to be the trigger for nuclear localisation (Rodriguez et al., 2000). This NAP import into the nucleus requires no NLS on the NAP itself, due to the one contained within the histone. This suggests a possible mechanism for pfNAPS, with no identified NLS, to cross the nuclear membrane.

The sub-cellular localisation of NAP is important as the local concentrations of NAP within the nucleus and the cytoplasm are significantly different. Immunodetection studies of proteins which have been expressed tagged with a high affinity epitope, allowed for more accurate determination of protein levels within the cell. The estimated population of yNAP monomers in a haploid cell was calculated to be approximately 8070 molecules (Ghaemmaghami et al., 2003). In G1 of the cell cycle the NAP is predominantly in the cytoplasm of the cell. Taking into consideration the average cell volume and the volume taken up by the sub-cellular organelles, the concentration of the yNAP dimer in the cytoplasm was determined to be approximately
0.5 µM (Toth et al., 2005). Relocation of NAP into the nucleus at S phase leads to a concentration increase due to the smaller accessible volume within the nucleus. The concentration of the yNAP dimer within the nucleus was calculated to be approximately 2 µM (Toth et al., 2005). This four fold change in local protein concentration has been shown to affect the proteins oligomerisation state in vitro, and may also have an in vivo relevance (Toth et al., 2005).

1.3.4 Histone Binding

NAP1 proteins are highly acidic consisting of almost 50% glutamate and aspartate residues (Figure 1.8). The central core of NAP was shown by Fujii and Nakata in 1992 to be necessary for nucleosome assembly. The concave area of this region is highly acidic and has therefore been highlighted as a putative histone binding region for yNAP (Figure 1.10 D-F) (Park and Luger, 2006b). Acidic concave regions were also observed for the pfNAPS and pfNAPL, suggesting a similar binding mechanism (Gill et al., 2009). Electrostatic surfaces for each of these proteins were rendered to provide a direct comparison of this unusual charge distribution (Figure 1.10 A-C). This mechanism for histone binding is also predicted for xNAP1 and hNAP1 due to the sequence homology between species in these regions and the conserved acidic properties.

Co-immunoprecipitation analysis indicated that dNAP1 binds to the H2A/H2B dimer in vivo (Ito et al., 1996). However, in vitro studies with hNAP1 and yNAP1 suggest that binding of all four core histones occurs (Ishimi et al., 1987; McBryant et al., 2003). The binding of hNAP and yNAP to histone variants was also identified in vitro (Okuwaki et al., 2005; Park et al., 2005). Further to these interactions, hNAP was also shown to remove the linker histone, H1, from chromatin extracted from HeLa cells (Kepert et al., 2005). A NAP1 preference was observed for the H2A/H2B dimer by Ito et al., 1996, Ishimi et al., 1987 and Chang et al., 1997. Conversely, in vitro, McBryant et al., 2003 observed a higher binding affinity to H3/H4 which was mediated by the N-terminal tails of the histones.
The general consensus is that the NAP1 dimer can bind a H3/H4 dimer, a H2A/H2B dimer or a H1 monomer (Kepert et al., 2005; McBryant et al., 2003; Toth et al., 2005; Zlatanova et al., 2007). However, an alternative stoichiometry has been suggested by Andrews et al., 2008 and McQuibban et al., 1998, where one NAP1 monomer binds one histone dimer. McQuibban et al. quantified the NAP-Histone binding stoichiometry via interactions between NAP1 and a cross-linked histone octamer. It is possible that NAP1 only interacts with the two flanking H2A/H2B dimers within the octamer, i.e. the H3/H4 tetramer is inaccessible, leading to the observed stoichiometry. Confusingly, the Andrews et al., 2008 and McBryant et al., 2003 papers which were produced from the same laboratory used two different techniques (fluorescence titrations and EMSAs respectively) and reported different NAP-Histone stoichiometries. Therefore, the question of the NAP1-Histone stoichiometry remains unresolved.

This histone-binding ability allows NAP1 to regulate the assembly and disassembly of nucleosomes by inserting or removing H2A/H2B dimers and by disrupting the stability of histone-DNA interactions (Levchenko and Jackson, 2004). The removal the H2A/H2B dimer,
facilitating nucleosome sliding, was also observed in vitro by Park et al., in 2005. In addition to histone binding and nucleosome assembly, NAP1 has been implicated in many more cellular functions, including histone shuttling (Mosammaparast et al., 2002), transcriptional regulation (Walter et al., 1995) and cell cycle control (Kellogg and Murray, 1995). This diverse range of cellular functions requires the involvement of multiple non-histone binding proteins.

### 1.3.5 Non-Histone Binding Partners

Specific non-histone proteins such as the nuclear import factor karyopherin (Mosammaparast et al., 2002) and the chromatin assembly factor CAF-1 (McQuibban et al., 1998) bind to NAP1 and influence histone binding and chromatin remodelling. Karyopherin, described in section 1.3.3, relocates the NAP-Histone complex from the cytoplasm into the nucleus. Whilst in the nucleus, NAP1 can facilitate nucleosome sliding unaided, but CAF-1 and ATP are also required for the nucleosome spacing to be of physiological length (McQuibban et al., 1998). It was also shown that the large multi-subunit RSC chromatin remodelling complex requires NAP1 and ATP in order to disassemble the nucleosome structure (Lorch et al., 2006), suggesting another putative NAP1 binding partner.

Further diverse roles of NAP1 have been identified, such as transcriptional and cell cycle control which are also mediated by the interactions with non-histone binding partners. In vitro studies with γNAP1 and HeLa nuclear extracts demonstrated that NAP1 assisted the binding of transcription factors to DNA. This level of transcriptional control is facilitated by the removal of the H2A/H2B histones from the nucleosome by NAP1 (Walter et al., 1995). NAP1 from Saccharomyces and Xenopus have been shown to play a role in cell cycle control due to their interaction with a B-type cyclin-activated kinase complex which mediates the progression into mitosis (Kellogg et al., 1995). In vivo studies in Saccharomyces showed that the activity of the kinase complex was suppressed in cells depleted of NAP1 (Kellogg and Murray, 1995). These results suggest that the presence of NAP1 is important for cell cycle control and specifically the progression into mitosis.

In addition to NAP1 interacting with multiple cellular proteins, it has also been demonstrated to be targeted by retroviral proteins. The retrovirus life cycle requires the diversion of cellular resources for viral integration and ultimately the production of viral progeny. This is achieved by synthesising specific virally encoded proteins, such as reverse transcriptase, and hijacking available cellular proteins and modifying their function. It is not surprising, therefore, that NAP is targeted directly by several retroviral proteins. The hNAP1 protein has been implicated in
the regulation of HIV via interactions with Tat-1 (Vardabasso et al., 2008). Tat-1 mediates the expression of viral genes which have been incorporated into the host genome. The domains within each protein responsible for the Tat-1-NAP1 interaction were determined using proteomic analysis. A small basic domain of Tat-1 was shown to mimic the basic nature of the histones and bind to the acidic domain II of hNAP1, (Vardabasso et al., 2008). Vardabasso et al. also carried out in vivo studies which revealed that upon over expression of hNAP1, Tat-1-mediated gene expression increased along with HIV infection. Using a VSG-luciferase HIV construct, hNAP1 knockout cells were shown to have a decrease in Tat-1 mediated gene expression along with a decrease in HIV infection (Vardabasso et al., 2008).

Furthermore, hNAP1 has been shown to interact with the HIV Rev protein (Cochrane et al., 2009). Rev is essential for viral replication and is involved in binding to viral RNAs and exporting them into the cytoplasm (Cullen, 1998). Proteomic assays revealed that hNAP1 binds to Rev via the central region of the protein which contains the NLS and oligomerisation domain. In vitro, binding of hNAP1 to Rev prevented Rev from forming large inactive aggregates in the nucleus (Cochrane et al., 2009). Binding of hNAP1 did not, however, affect the function of the Rev protein, suggesting that the NAP1 binding domain is distinct from the RNA-binding domain. The Rev-hNAP1 interaction increases the available Rev within the nucleus and increases RNA exportation (Cochrane et al., 2009).

1.4 Previous Studies on Xenopus laevis NAP.

Previous work in this laboratory focused upon the properties of NAP1 from Xenopus laevis (x), an ideal model system due to the 92% sequence homology with hNAP1 (Figure 1.8). In vivo studies were carried out looking at the expression patterns in embryos following xNAP1 over-expression (Steer et al., 2003). An up-regulation of specific genes was identified, consistent with the findings of Walter et al., which suggested a role for NAP in transcriptional regulation (Walter et al., 1995). Subsequently, developing embryos were examined from which xNAP1 had been depleted. The results indicated a specific role of xNAP1 in the progression of the hematopoiesis pathway (Abu-Daya et al., 2005).

Also in this laboratory, limited proteolysis on the xNAP1 protein was performed to analyse the tertiary structure of the protein (Friedeberg et al., 2006). xNAP1 was digested with either trypsin or α-chymotrypsin which produced two main fragments termed T2 and C1 respectively. Using a combination of N-terminal sequencing, mass spectroscopy and predictions of protease cleavage points, the two fragments were mapped onto the full length protein. The C1 fragment lacked 49 residues from the N-terminal region and the T2 fragment lacked 36
residues from the N-terminal region and 109 residues from the C-terminal region. The two fragments were individually sub-cloned into the pET-28b vector for bacterial expression of the fragments with C-terminal His-tags (the xNAP1-C1 expression product sequence is shown in Figure 1.8). Friedeberg et al. also performed hydrodynamic analysis of the T2 fragment in 50 mM KCl conditions using dynamic light scattering and AUC. The resulting hydrodynamic radius and Mr were estimated as 9 nm and 420 kDa respectively. Given the T2 monomer was 31 kDa including the His-tag, it was clear that the fragment was oligomerising. Full-length, C1 or T2 fragments were over-expressed in developing embryos to determine in vivo affects of the truncations. Similar phenotypes were observed when both the wild type and the proteolytic fragments were injected (Friedeberg et al., 2006).

The subsequent crystal structure of the yNAP1 protein confirmed that the xNAP1-C1 fragment contains all of the structured domains of the protein. The full length xNAP1 was highly insoluble and therefore difficult to characterise in solution. The xNAP1-C1 fragment was significantly more soluble than the full length protein and was therefore used for further biochemical and biophysical characterisation of xNAP1.

1.5 Research Aims

The main aims of this study were as follows:

1) To perform biophysical analyses of the xNAP1-C1 fragment, characterising the ionic strength and concentration dependence of its oligomeric states.

2) To characterise the xNAP1-C1-Histone complexes including the binding stoichiometries and oligomeric status.

3) To obtain the first structural information on the xNAP1-C1-Histone complexes.
2. Materials and Methods

2.1 Stock Solutions

All solutions, unless otherwise stated, were made with deionised water and filtered through a 0.2 µm membrane using a vacuum pump.

- 0.2 M EDTA, pH 5.2 Stored at 4°C
- 1 M Tris HCl, pH 7.5 at 16°C Stored at 16°C
- 200 mM NaAc, 20 mM Tris, pH 5.2 Stored at room temperature
- 10% w/v APS Stored at -20°C
- 1 M IPTG Stored at -20°C
- 50 mg/mL kanamycin Stored at -20°C
- 10 mg/mL ampicillin Stored at -20°C
- 25 mg/mL chloramphenicol Dissolved in ethanol and stored at -20°C
- 1 M PMSF Dissolved in ethanol and stored at -20°C
- 1 M DTT Stored at -20°C
- Water-saturated butanol Stored at room temperature

2.2 SDS Polyacrylamide Gel Electrophoresis (PAGE)

2.2.1 Gel Solutions

<table>
<thead>
<tr>
<th>ProtoGel (National Diagnostics)</th>
<th>Acrylamide</th>
<th>30% w/v</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bis-acrylamide</td>
<td>0.8% w/v</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ProtoGel stacking buffer (National Diagnostics)</th>
<th>Tris HCl, pH 6.8</th>
<th>0.5 M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SDS</td>
<td>0.4% w/v</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ProtoGel resolving buffer (National Diagnostics)</th>
<th>Tris HCl, pH 8.8</th>
<th>1.5 M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SDS</td>
<td>0.4% w/v</td>
</tr>
</tbody>
</table>
10 × SDS running buffer
- Tris 0.25 M
- Glycine 1.92 M
- SDS 1% w/v

3 × SDS Loading Dye
- Tris HCl, pH 6.8 0.36 M
- SDS 9%
- Glycerol 15%
- β-mercaptoethanol 15%
- Bromophenol blue 0.2%

2.2.2 Gel Compositions

Resolving Buffer with 12.5% acrylamide
- ProtoGel (National Diagnostics) 3.35 mL
- Resolving Buffer (ProtoGel National Diagnostics) 2 mL
- dH₂O 2.60 mL
- 10% w/v APS 45 µL
- TEMED 5 µL

Stacking Buffer with 4% acrylamide
- ProtoGel Stacking Buffer (National Diagnostics) 3.35 mL
- ProtoGel (National Diagnostics) 2 mL
- dH₂O 2.60 mL
- 10% w/v APS 45 µL
- TEMED 5 µL
2.2.3 SDS PAGE Protocol

An ATTO mini system was used; the glass plates (15 cm x 15 cm) were cleaned with dH₂O and then degreased with 70% ethanol. The plates were clamped together, spaced with a 1 mm gasket. Resolving buffer (6 mL) was poured between the plates and 200 µL of water-saturated butanol was layered on top, to form a flat interface. After 30 min, the butanol was removed and the gel surface washed with dH₂O. The stacking gel was poured on top and the comb was placed into the gel immediately. When the gel had set, the comb and spacer were removed. SDS running buffer was added to the tank along with the gels, when all air bubbles had been removed from underneath the gels, they were clamped into place and the central reservoir was filled to the top with SDS running buffer.

Samples were prepared by adding 3 x SDS loading dye to the sample in a 1:2 ratio respectively. The samples were incubated for 10 min to allow the protein to denature, 2-30 µL was then loaded into a well. Each gel was calibrated by running 7 µL of Benchmark™ protein ladder (Appendix I) in the first well as a standard (Laemmli, 1970). The gels were run for 3 hr at 100 v towards the positive anode. The Gel was put into SimplyBlue™ Safestain (Invitrogen) for an hour and then destained in water.

2.3 Acid Urea Triton (AUT) Gel Electrophoresis

2.3.1 Gel Solutions

<table>
<thead>
<tr>
<th>AUT Stacking Buffer</th>
<th>Solid Urea</th>
<th>4.8 g</th>
</tr>
</thead>
<tbody>
<tr>
<td>60% Acrylamide</td>
<td>1 mL</td>
<td></td>
</tr>
<tr>
<td>Acetic Acid</td>
<td>0.6 mL</td>
<td></td>
</tr>
<tr>
<td>2.5% 2-Bis-acrylamide</td>
<td>0.65 mL</td>
<td></td>
</tr>
<tr>
<td>NH₄OH</td>
<td>30 µL</td>
<td></td>
</tr>
<tr>
<td>4% Riboflavin</td>
<td>0.65 mL</td>
<td></td>
</tr>
<tr>
<td>TEMED</td>
<td>50 µL</td>
<td></td>
</tr>
<tr>
<td>dH₂O</td>
<td>10 mL</td>
<td></td>
</tr>
</tbody>
</table>

| AUT Resolving Buffer | Solid Urea | 9.6 g |
2.3.2 AUT Gel Electrophoresis Protocol

Make stacking and resolving solutions without TEMED and triton, heat until Urea has dissolved. Let cool, add the rest of ingredients and mix. Using an ATTO mini system (16 cm x 16 cm), the glass plates were cleaned with dH₂O, then degreased with 70% EtOH. The plates were clamped with a 1 mm gasket. 8 mL of resolving buffer was poured between the plates, 400 µL of water-saturated butanol was layered on top to aid setting. The gel was sandwiched between silver foil and a lamp to speed up setting. After 2 hrs the butanol was removed and the gel surface washed with dH₂O. The stacking gel was poured in, placing the comb into the gel immediately. When the gel had set the comb and spacer were removed and the gel was clamped in place in the gel tank.

Samples were prepared by adding AUT Loading Dye to the sample in a 1:5 ratio respectively. The samples were left for 10 mins to denature and then loaded into the wells. The gels were run for 2hrs at 350 volts towards a negative node.

2.4 Protein Expression

2.4.1 Bacterial Cell Culture Solutions

<table>
<thead>
<tr>
<th>FB1</th>
<th>Rubidium Chloride</th>
<th>100 mM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MnCl₂.4H₂O</td>
<td>50 mM</td>
</tr>
<tr>
<td></td>
<td>Potassium acetate</td>
<td>30 mM</td>
</tr>
<tr>
<td></td>
<td>CaCl₂.2H₂O</td>
<td>10 mM</td>
</tr>
<tr>
<td>Component</td>
<td>Concentration</td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>Glycerol</td>
<td>15%</td>
<td></td>
</tr>
<tr>
<td>Adjusted to pH 5.8 with acetic acid</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FB2**

- MOPS: 10 mM
- Rubidium Chloride: 10 mM
- CaCl$_2$.2H$_2$O: 75 mM
- Glycerol: 15%

Adjusted to pH 6.8 with NaOH

**LB Broth**

- Tryptone: 10 g
- Yeast extract: 5 g
- NaCl: 5 g
- dH$_2$O: Made up to 1 L

Adjusted to pH 7.2 with NaOH, autoclaved same day.

**LB Plates - ~30 plates**

- Tryptone: 10 g
- Yeast extract: 5 g
- NaCl: 5 g
- Agar: 12.5 g
- H$_2$O: Made up to 1 L

Adjusted to pH 7.2 with NaOH, then autoclaved. 50 µg/mL of kanamycin or 100 µg/mL of ampicillin and 25 µg/mL chloramphenicol were added upon cooling, the plates were poured and then stored at 4°C.

**2 × YT Broth**

- Tryptone: 16 g
- Yeast extract: 10 g
2.4.2 Competent Bacterial Cell Production

Two strains of competent cells were utilised; *E. coli* DH5α for plasmid production, and *E. coli* BL21 (DE3) pLysS for protein expression. The cells were plated onto fresh agar containing no antibiotic, and grown at 37°C overnight. A single colony was used to inoculate 10 mL of 2 × YT broth and this was again incubated overnight at 37°C. The overnight culture was used to inoculate 500 mL of 2 × YT (pre-warmed to 37°C), then incubated at 37°C with shaking at 220 rpm. When an OD₆₀₀ of 0.6 was achieved, the cells were collected by centrifugation in a swing-out rotor at 4,000g and 4°C for 20 min. The cells were resuspended in 166 mL FB1 and left on ice for 15 min and the centrifugation step was repeated. The cells were resuspended in 40 mL FB2 and kept on ice for a further 15 min. Aliquots of 1 mL were transferred to Eppendorf tubes on ice, snap frozen with liquid N₂, and stored at -70°C until required.

### 2.4.3 Transformation of Plasmids into Competent Cells.

The xNAP1-C1 gene was cloned into the pET28b (+) expression plasmid using Ncol and Xhol restriction sites, (Appendix II) by C.E.Friedeberg (Friedeberg et al., 2006). Expression plasmids containing the histone genes were provided by the Richmond lab (Luger et al., 1997b) (Appendix III), originating from gene clusters characterised by Old et al. (Old et al., 1982). An aliquot containing the appropriate strain of competent cells was thawed on ice, 120 µL was transferred to a chilled Eppendorf and 2 µL of plasmid was added. The cells were heat shocked for 75 sec at 42°C and returned to ice for 5 min. The cells containing the histone plasmids were

<table>
<thead>
<tr>
<th>Solution</th>
<th>Component</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>xNAP1-C1 wash buffer</td>
<td>Tris HCl, pH 7.5</td>
<td>20 mM</td>
</tr>
<tr>
<td></td>
<td>NaCl</td>
<td>100 mM</td>
</tr>
<tr>
<td>Histone wash buffer</td>
<td>Tris HCl, pH 7.5</td>
<td>50 mM</td>
</tr>
<tr>
<td></td>
<td>NaCl</td>
<td>100 mM</td>
</tr>
<tr>
<td></td>
<td>Benzamidine</td>
<td>1 mM</td>
</tr>
<tr>
<td></td>
<td>β-mercaptoethanol</td>
<td>1 mM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NaCl</th>
<th>5 g</th>
<th>dH₂O</th>
<th>Made up to 1 L</th>
</tr>
</thead>
</table>
|      |     |      | Adjusted to pH to 7.2 with NaOH, autoclaved same day.
then transferred to grow on an ampicilin plate. The cells containing the xNAP1-C1 plasmid required an extra growth step to allow them to develop kanamycin resistance. The cells were therefore added to 900 µL of chilled 2 × YT broth and incubated at 37°C at 220 rpm for 1 hr. The culture was centrifuged at 16,100g for 1 min then 800 µL of the supernatant was discarded and the pellet was resuspended in remaining broth.

An aliquot of 100 µL of the prepared recombinant cells, were spread onto agar plates containing the appropriate antibiotic(s): 100 µg/mL of ampicilin and 25 µg/mL chloramphenicol for histone expression and 50 µg/mL of kanamycin for the xNAP1-C1 expression. Ampicilin and kanamycin were used to select for maintenance of the expression plasmids, whilst chloramphenicol was used to select for the pLysS plasmid. The plates were incubated overnight at 37°C.

2.4.4 Overnight Bacterial Culture

The appropriate antibiotic(s) was added to 10 mL of LB broth: 100 µg/mL of ampicilin and 25 µg/mL chloramphenicol for recombinant histone cultures and 50 µg/mL of kanamycin for the recombinant xNAP1-C1 culture. The broth was inoculated with a single colony from the agar plates. The cultures were incubated at 37°C at 220 rpm for 16 hr.

2.4.5 Mini Preparations of Plasmid DNA

For plasmid production E.coli BL21 pLysS strains were used. Overnight bacterial cultures were centrifuged at 6,800g for 5 min, 25°C. The supernatant was removed and a Qiagen miniprep protocol was followed, this method used alkaline lysis to break open the cells, the bacterial lysate was then removed and the plasmid DNA absorbed onto a silica membrane. The DNA was eluted using water and the UV spectrum measured with a Labtech International ND-1000 Nanodrop spectrophotometer. All of the plasmid preparations were in the region of 100 ng/mL, 25 µL of each was sent to Cogenics Lark for sequencing. T7 promoter and terminator primers were used in the sequencing reaction.

2.4.6 Large Volume Expression of Proteins

For expression of recombinant proteins, plasmids were transformed into E.coli (DE3) pLysS cells, grown on agar plates and then a colony was transferred into a 10 mL overnight culture as previously described (section 2.3.3 and 2.3.4). Overnight cultures were used to inoculate 1 L of pre-warmed 2 × YT broth. The broth contained either 100 µg/mL of ampicilin and 25 µg/mL chloramphenicol for histones preparations or 50 µg/mL of kanamycin for xNAP1-C1 preparations. Protocols for the expression of the four histone proteins (Luger et al., 1997b) and xNAP1-C1 (Friedeberg et al., 2006) have been previously published. Optimisation of the protocols has been achieved as described below.
To prepare histone pellets, the cells were incubated at 37°C shaking at 220 rpm until an OD$_{600}$ of 0.8 was reached. The cultures were induced with 0.4 mM IPTG and returned to incubate for 2.5 hr (1.5 hr for H4). The bacterial cells were harvested at 23,000g for 30 min at 4°C. The pellets were resuspended in histone wash buffer, decanted into Falcon tubes and centrifuged in a swing out rotor at 5,500g for 10 min at 4°C. The supernatant was removed and the pellets were stored at -20°C.

To prepare xNAP1-C1 pellets, the cells were incubated at 37°C shaking at 220 rpm until an OD$_{600}$ of 0.6 was reached. The cultures were induced with 1 mM IPTG and returned to incubate for 3 hr. The bacterial cells were harvested at 23,000g for 30 min at 4°C. The pellets were resuspended in xNAP1-C1 wash buffer, decanted into Falcon tubes and centrifuged in a swing out rotor at 5,500g for 10 min at 4°C. The supernatant was removed and the pellets were stored at -20°C.

### 2.5 Protein Purification

All chromatography columns were purchased from GE Healthcare and the columns were connected to, and controlled by, an ÄKTA basic or ÄKTA purifier system.

#### 2.5.1 xNAP1-C1 Purification Solutions

<table>
<thead>
<tr>
<th>Nickel buffer A</th>
<th>Tris HCl, pH 7.5 at 16°C</th>
<th>20 mM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NaCl</td>
<td>500 mM</td>
</tr>
<tr>
<td></td>
<td>Imidazole</td>
<td>10 mM</td>
</tr>
<tr>
<td></td>
<td>PMSF</td>
<td>0.1 mM</td>
</tr>
<tr>
<td></td>
<td>DTT</td>
<td>1 mM</td>
</tr>
<tr>
<td></td>
<td>Glycerol</td>
<td>10% w/v</td>
</tr>
<tr>
<td></td>
<td>MgCl$_2$</td>
<td>3 mM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nickel buffer B</th>
<th>Tris HCl, pH 7.5 at 16°C</th>
<th>20 mM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NaCl</td>
<td>500 mM</td>
</tr>
<tr>
<td></td>
<td>Imidazole</td>
<td>1 M</td>
</tr>
</tbody>
</table>
2.5.2 Purification of xNAP1-C1

The purification protocol described is for use with one pellet (~1.6 g) from the 1L, large volume expressions (section 2.3.6). The pellet was thawed on ice, and resuspended in 10 mL of nickel buffer A. The suspension was sonicated for a total of 4.5 min (5.5 sec pulse with 9.9 sec rest) with amplitude of 25% at 9°C. The sample was centrifuged for 30 min with 39,200g at 4°C. The supernatant was kept for purification and the pellet was discarded.

The supernatant was run onto a 1 mL HisTrap HP column in nickel buffer A and eluted from the column using nickel buffer B with a gradient of 0-30% over 100 CV. The eluted fractions

<table>
<thead>
<tr>
<th>Buffers</th>
<th>Constituents</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>MonoQ buffer A</td>
<td>Tris HCl, pH 7.5 at 16°C</td>
<td>20 mM</td>
</tr>
<tr>
<td></td>
<td>NaCl</td>
<td>50 mM</td>
</tr>
<tr>
<td></td>
<td>Glycerol</td>
<td>10% w/v</td>
</tr>
<tr>
<td></td>
<td>MgCl₂</td>
<td>3 mM</td>
</tr>
<tr>
<td>MonoQ buffer B</td>
<td>Tris HCl, pH 7.5 at 16°C</td>
<td>20 mM</td>
</tr>
<tr>
<td></td>
<td>NaCl</td>
<td>1 M</td>
</tr>
<tr>
<td></td>
<td>Glycerol</td>
<td>10% w/v</td>
</tr>
<tr>
<td></td>
<td>MgCl₂</td>
<td>3 mM</td>
</tr>
<tr>
<td>Size exclusion buffer</td>
<td>Tris HCl, pH 7.5 at 16°C</td>
<td>20 mM</td>
</tr>
<tr>
<td></td>
<td>Glycerol</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>MgCl₂</td>
<td>3 mM</td>
</tr>
<tr>
<td></td>
<td>NaCl</td>
<td>Varying from 20 mM-1 M</td>
</tr>
</tbody>
</table>
containing xNAP1-C1 were pooled and desalted into MonoQ buffer A using a 50 mL HiPrep 26/10 Desalting column. Finally the sample was run onto a 1 mL MonoQ HR 5/5 and eluted using MonoQ buffer B with a gradient of 0-100% over 100 CV. Collected xNAP1-C1 fractions were either used directly or stored at –20°C.

2.5.3 Deuteration and Purification of xNAP1-C1

Protocol 2.3.3 for transforming the xNAP1-C1 containing plasmid into *E.coli* BL21 (DE3) pLysS competent cells was followed. However, instead of spreading the recombinant cells on an agar plate, the cells were stabbed into a Universal tube containing 10 mL of set agar and 50 µg/mL of kanamycin. The culture was sent to Dr Michael Haertlein at the D-lab, Grenoble, France, for the production of deuterated protein. Enfors minimal media containing 85% D$_2$O, with hydrogenated glycerol as the carbon source, was used to grow the bacteria. This gave rise to a 75% deuteration level, such that the expressed dxNAP1-C1 protein had a theoretical contrast match point equivalent to 100% D$_2$O. A bacterial slab of 50 g produced by this method was transported back to the University of Portsmouth on ice. The same purification method, which was used for the protonated xNAP1-C1 above was used with 5 g of the deuterated bacterial lysate with the following modification: A 5 mL HisTrap HP column was used with a 0-30% nickel buffer B gradient over 20 CV instead of the 1 mL HisTrap HP column.

2.5.4 Histone Purification Solutions

<table>
<thead>
<tr>
<th>Wash buffer H</th>
<th>Tris HCl, pH 7.5 at 16°C</th>
<th>50 mM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NaCl</td>
<td>100 mM</td>
</tr>
<tr>
<td></td>
<td>Benzamidine</td>
<td>1 mM</td>
</tr>
<tr>
<td></td>
<td>β-mercaptoethanol</td>
<td>1 mM</td>
</tr>
<tr>
<td></td>
<td>PMSF</td>
<td>1 mM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Wash buffer Ht</th>
<th>Tris HCl, pH 7.5 at 16°C</th>
<th>50 mM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NaCl</td>
<td>100 mM</td>
</tr>
<tr>
<td></td>
<td>Benzamidine</td>
<td>1 mM</td>
</tr>
<tr>
<td></td>
<td>β-mercaptoethanol</td>
<td>1 mM</td>
</tr>
<tr>
<td></td>
<td>Triton</td>
<td>1% v/v</td>
</tr>
<tr>
<td></td>
<td>PMSF</td>
<td>1 mM</td>
</tr>
</tbody>
</table>
### Purification of Recombinant Histones

The purification protocol described is for use with one pellet from the 1L, large volume expressions (section 2.3.6). The pellet was thawed on ice, and resuspended in 25 mL of Wash buffer H. The suspension was sonicated for a total of 1 min (5 sec pulse with 9.9 sec rest) with amplitude of 25% at 10°C. The sample was centrifuged for 15 min at 23,000g at 4°C (these spin conditions were used for all of the washes). The resulting pellet was resuspended in 25 mL wash buffer Ht and centrifuged. To remove Triton from the sample, the pellet was then

<table>
<thead>
<tr>
<th>Solution/Buffer</th>
<th>Components</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denaturation</td>
<td>Guanidinium HCl</td>
<td>6 M</td>
</tr>
<tr>
<td></td>
<td>Sodium acetate, pH 5.2</td>
<td>20 mM</td>
</tr>
<tr>
<td></td>
<td>DTT</td>
<td>1 mM</td>
</tr>
<tr>
<td>Ion exchange A</td>
<td>Urea, deionised</td>
<td>4 M</td>
</tr>
<tr>
<td>Buffer A</td>
<td>Sodium acetate, pH 5.2</td>
<td>20 mM</td>
</tr>
<tr>
<td></td>
<td>Tris</td>
<td>2 mM</td>
</tr>
<tr>
<td></td>
<td>NaCl</td>
<td>0.2 M</td>
</tr>
<tr>
<td></td>
<td>β-mercaptoethanol</td>
<td>2 mM</td>
</tr>
<tr>
<td></td>
<td>EDTA</td>
<td>0.1 mM</td>
</tr>
<tr>
<td>Ion exchange B</td>
<td>Urea, deionised</td>
<td>4 M</td>
</tr>
<tr>
<td>Buffer B</td>
<td>Sodium acetate, pH 5.2</td>
<td>20 mM</td>
</tr>
<tr>
<td></td>
<td>Tris</td>
<td>2 mM</td>
</tr>
<tr>
<td></td>
<td>NaCl</td>
<td>1 M</td>
</tr>
<tr>
<td></td>
<td>β-mercaptoethanol</td>
<td>2 mM</td>
</tr>
<tr>
<td></td>
<td>EDTA</td>
<td>0.1 mM</td>
</tr>
<tr>
<td>Desalting</td>
<td>Formic acid</td>
<td>0.5% v/v</td>
</tr>
</tbody>
</table>
resuspended in 25 mL of wash buffer H and centrifuged, this was repeated once. The pellet was carefully mixed in 0.5 mL of DMSO and left to solubilise for 30 min. To denature the proteins, 4.5 mL of denaturing solution was added and the sample was left mixing for 1 hr. The sample was then spun at 23,000g for 30 mins at 25°C.

The supernatant was mixed with 1 g of hydroxypatite resin for 10 min, and then passed through a fine filter to remove the resin. A further 15 mL denaturing solution was added to the filter column to wash residual unbound protein through. The sample was transferred into Ion Exchange buffer A using a 50 mL HiPrep 26/10 Desalting column (10 mL / run). The sample was pooled and run onto a 5 mL HiTrap SP HP ion exchange column in ion exchange buffer A. The sample was eluted off the column with ion exchange buffer B, using a gradient of 0-100% over 10 CV. Eluted fractions containing the histone proteins were pooled, buffer exchanged into 0.5% v/v formic acid, lyophilised and stored at -70°C.

### 2.6 Histone Reconstitution

#### 2.6.1 Histone Reconstitution Solutions

<table>
<thead>
<tr>
<th>Buffer</th>
<th>Composition</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Heparin buffer A</strong></td>
<td>Tris HCl, pH 7.5</td>
<td>20 mM</td>
</tr>
<tr>
<td></td>
<td>NaCl</td>
<td>50 mM</td>
</tr>
<tr>
<td></td>
<td>EDTA</td>
<td>1 mM</td>
</tr>
<tr>
<td></td>
<td>Glycerol</td>
<td>10%</td>
</tr>
<tr>
<td><strong>Heparin buffer B</strong></td>
<td>Tris HCl, pH 7.5</td>
<td>20 mM</td>
</tr>
<tr>
<td></td>
<td>NaCl</td>
<td>2 M</td>
</tr>
<tr>
<td></td>
<td>EDTA</td>
<td>1 mM</td>
</tr>
<tr>
<td></td>
<td>Glycerol</td>
<td>10%</td>
</tr>
<tr>
<td><strong>Denaturing buffer</strong></td>
<td>Guanidinium</td>
<td>8 M</td>
</tr>
<tr>
<td></td>
<td>Sodium acetate, pH 5.2</td>
<td>20 mM</td>
</tr>
<tr>
<td></td>
<td>DTT</td>
<td>10 mM</td>
</tr>
</tbody>
</table>
2.6.2 Reconstitution of Histones H2A and H2B

The individually lyophilised histones H2A and H2B were resuspended in 0.5 mL 10 mM HCl each, mixed in equimolar amounts and then refolded by slowly adding 9 mL heparin buffer A. The mixture was left at 4°C for one hour before being run onto a 1 mL heparin column in heparin buffer A, the sample was eluted off the column using a 0-100% gradient into heparin buffer B over 100 CV. Eluted fractions were analysed by acid urea triton gel electrophoresis (section 2.3) by which H2A and H2B have a different electrophoretic mobility (Waterborg and Harrington, 1987). Fractions containing the H2A/H2B heterodimers were pooled and used immediately or stored at -20°C for up to one month.

2.6.3 Reconstitution of Histones H3 and H4

The individually lyophilised histones H3 and H4 were resuspended in 0.5 mL denaturing buffer and mixed in equimolar amounts and incubated at room temperature for 1 hr. The sample was dialysed into heparin buffer B at 4°C in a 3,000 Da MWCO Slide-A-Lyzer (Thermo Fisher Scientific). The sample was recovered via the injection ports and run on a Superdex 75 size-exclusion column in heparin buffer B. Eluted fractions were run on SDS PAGE to identify the histone H3/H4 hetero-dimers. Fractions containing the H3/H4 heterodimers were pooled and used immediately.

2.7 Formation of xNAP1-C1-Histone Complexes

Reconstituted histones were prepared as above; native histones were a gift from Dr John Baldwin et al., and were prepared using a variation on a previously published protocol described by (Lambert et al., 1999). These were added to an equimolar amount of xNAP1-C1 in a 1 histone dimer : 1 xNAP1-C1 dimer ratio. This was dialysed overnight at 4°C, using 10,000 Da MWCO Slide-A-Lyzer into size exclusion buffer containing 150 mM NaCl. The recovered sample was centrifuged for 30 min at 4°C at 16,100g and run on a Superose 6 column equilibrated in the size exclusion buffer, again with 150 mM NaCl. The sample homogeneity in eluted fractions was verified using dynamic light scattering.
2.8 Analytical Size Exclusion

To use size exclusion as a comparative tool, the following protocol was performed, unless otherwise stated. A Superose 6 column was equilibrated at 0.5 mL/min in 2 CV of SE buffer containing the same NaCl concentration as the sample to be analysed. The sample was loaded onto the column via a 200 µL loop full of protein; 1 mL of buffer was used to inject the entire contents of the loop onto the column. Buffer was flowed through the column for 1.5 CV to elute all the proteins prior to the next injection. Fractions of 0.5 mL were collected throughout the elution. The absorbance of the elution was measured at 254 and 276 nm to monitor the presence of DNA and protein respectively.

2.9 Dynamic Light Scattering

Dynamic Light Scattering was performed on protein samples post size exclusion. Measurements were performed at 16°C using a Protein Solutions DynaPro, with temperature-controlled micro sampler. The instrument was equipped with a flow kit connected to a low volume quartz flow cell. The flow system was washed with filtered buffer prior to use. The protein samples were centrifuged at 16,100g for 30 min at 16°C to remove dust particles and aggregation then 300 µL of the supernatant was injected into the flow cell via the inlet port.

Samples were measured using an incident laser beam at 100% power of wavelength 830 nm. Multiple readings over ten minutes were measured to produce an averaged correlation curve and analysed using DYNAMICS (Wyatt Technology Corp.). The diffusion coefficient (D) of the samples was obtained using the relationship to the scattering vector (K) and the inverse of the exponential decay time (Γ = 1/τ);

$$D = \frac{Γ}{K^2}$$

The diffusion coefficient was converted to hydrodynamic radius ($R_H$) via the Stokes-Einstein equation;

$$R_H = \frac{kT}{6\pi\eta D_0}$$

where k is the Boltzmann constant, T is the temperature in Kelvin and η is the viscosity of the solvent. The percentage of the sample corresponding to each $R_H$ was plotted in the form of a histogram. The polydispersity of the sample was also estimated on this basis.
2.10 Size Exclusion Chromatography Multi Angle Laser Light Scattering (SEC-MALLS)

xNAP1-C1 and xNAP1-C1-Histone complexes were prepared as described in sections 2.4.2 and 2.6 respectively. The samples were dialysed into 20 mM Tris HCl, pH 7.5, 150 mM NaCl and 3 mM MgCl₂ (refractive index of 1.331), and then injected onto a pre equilibrated Superose 6 column and eluted for 2 CV at 0.3 mL/min. Light scattering and refractive index were sequentially measured of the eluate (both from Wyatt Technologies, Inc). Light scattering (LS) was measured using a DAWN HELEOS machine with laser wavelength of 658 nm at 11 different angles. The refractive index (RI) was subsequently monitored with an Optilab rEX machine. Profile analysis was performed using the software ASTRA 5.3.4 (Wyatt Technology). Elution peaks were defined manually and the LS baselines were also adjusted. The Mᵣ across the peaks was determined using the Rayleigh-Gans-Debye approximation (described in section 4.3.1) and a dō/dc of 0.186.

2.11 Analytical Ultracentrifugation

Analytical ultracentrifugation experiments were carried out using a Beckman Optima XL-A analytical ultracentrifuge (Beckman-Coulter, Palo Alto, CA, USA), equipped with an An50-Ti rotor, equilibrated to 10°C. Cells with charcoal filled Epon centrepieces, 12 mm optical path lengths and quartz windows were used for velocity and equilibrium experiments, with 2 or 6 sectors respectively. Prior to sample loading, the rotor was cooled to 10°C and a radial calibration was performed at 3,000 rpm. The samples were then loaded and the temperature and vacuum were allowed to equilibrate whilst the rotor was stationary. The first and last scans of the experiment were overlaid to check for cell leakage by divergence of the menisci. In equilibrium studies, the four scans at each speed were also overlaid to verify that equilibrium had been reached.

2.11.1 Sedimentation Velocity

For sedimentation velocity experiments, 400 µL of sample and 425 µL of buffer were loaded into adjacent sectors in a cell. The rotor was accelerated to 15,000 rpm and scans commenced immediately. Absorbance scans at 276 nm were obtained every ten minutes in step mode with a radial step size of 0.005 cm. The partial specific volume of the protein and solvent density values were calculated from the protein and buffer compositions using the program SEDNTERP (Philo, 2006). These parameters were input into SEDFIT for data analysis (Schuck, 2000). Every
scan was included for analysis of xNAP1-C1 and scans 1-50 were used for the xNAP1-C1-Histone complexes.

The Lamm equation was fitted using a continuous c(s) distribution model. The menisci and baseline were manually estimated and along with the values calculated in SEDNTERP used in analysis following the standard protocols (Brown et al., 2009). The distribution of sedimentation coefficients was shown in the form of a c(s) plot. The distribution of M, was also obtained by using a continuous c(M) model.

### 2.11.2 Sedimentation Equilibrium

For sedimentation equilibrium studies, 90 µL of sample at three concentrations were loaded into three sectors with the other three sectors containing 100 µL of the respective buffer. Three spin speeds were measured for each experiment; these are specified individually in the text. Absorbance scans at 276 nm were obtained in step mode with a radial step size of 0.001 cm. At each spin speed, samples were left for 18 hr before the first scan was taken. Three subsequent scans were taken with 3 hr intervals to verify equilibrium had been reached. The rotor was accelerated to 40,000 rpm until meniscus depletion. The average baseline absorbance values were calculated from the scans taken at 40,000 rpm for each sector using EXCEL (Microsoft Office).

The partial specific volume of the protein and solvent density values were calculated from the protein and buffer compositions using the program SEDNTERP (Philo, 2006). Analysis was performed with these parameters as well as the solvent baselines using the program Origin 6.0 (Microcal Software, Inc.) with a specially designed AUC package developed by Beckman Coulter. Absorbance scans were cropped into high, medium and low concentration sectors. A weighted fit was employed to reduce the influence of outliers. A single species model was used to analyse each equilibrium curve individually. The scans from three concentrations at a given speed were then globally fitted. The curves from three speeds and a single concentration were also globally fitted. Finally a global fit using all nine data curves was performed.

### 2.12 Small Angle X-Ray Scattering

xNAP1-C1 and xNAP1-C1-Histone complexes were measured directly after size exclusion (as described in section 2.7) in 150 mM NaCl, 3 mM MgCl₂, 20 mM Tris HCl, pH 7.5. The eluate containing protein was concentrated to 1-2 mg/mL using vivaspin 30 MWCO spin columns at 4,000g and 4°C. The running buffer for the size exclusion was used as a buffer blank. Data were obtained from three high brilliance beamlines; ID14-3 and ID2 at the European Synchrotron
Radiation Facility (ESRF) and I22 at the Diamond Light Source (Diamond). Size exclusion was performed at Portsmouth prior to the collection at Diamond, Oxford. Size exclusion was performed in the EMBL laboratories, Grenoble, prior to use of the ESRF beamlines. The following table contains individual beamline characteristics:

<table>
<thead>
<tr>
<th>Beamline</th>
<th>ID14-3, ESRF</th>
<th>ID02, ESRF</th>
<th>I22, Diamond</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray wavelength (Å)</td>
<td>0.90</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>Sample-detector distance (m)</td>
<td>2.5</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>q range (Å⁻¹)</td>
<td>0.01 to 0.48</td>
<td>0.005 – 0.332</td>
<td>0.005 - 0.548</td>
</tr>
<tr>
<td>Sample environment</td>
<td>In-vacuum capillary sample holder</td>
<td>Sample changer with quartz cuvettes</td>
<td>Quartz windows clamped in cassette.</td>
</tr>
<tr>
<td>Detector</td>
<td>2D Pilatus 1M Dectris detector</td>
<td>FReLoN Kodak CCD</td>
<td>RAPID gas wire detector</td>
</tr>
<tr>
<td>Length of measurement</td>
<td>10 x 1 sec</td>
<td>3 x 0.1 sec</td>
<td>180 x 1 sec</td>
</tr>
<tr>
<td>Radial averaging software</td>
<td>BsxCuBE (P Pernot1, 2010)</td>
<td>BsxCuBE (P Pernot1, 2010)</td>
<td>DREAM (Mark Malfois, Diamond)</td>
</tr>
</tbody>
</table>

Buffer measurements were taken before and after the sample measurement and overlaid to check for fluctuations in the beam intensity. Scattering data were radially averaged using the radial averaging software setup for the individual beamlines. Scattering curves were overlaid in PRIMUS to check for radiation-induced damage and averaged. The buffer curves were subtracted from the sample curves. Guinier plots were analysed also using PRIMUS. Data were input into GNOM and the Dmax was optimised to minimise the discrepancy between the p(r) back transformation curve and the experimental curve (for more details see Appendix V).

GNOM output files were input into the SAXS MoW program, to calculate an average M, by integration under the Kratky curve. GNOM files were also input into DAMMIF, ten models were produced with P1, P5 and P6 symmetry imposed. The DAMMIF ab initio models were individually checked for outliers and then input into DAMAVER for filtering and averaging.

### 2.13 Small Angle Neutron Scattering

xNAP1-C1-Histone complexes were measured directly after size exclusion (as described in section 2.7) in 150 mM NaCl, 3 mM MgCl₂, 20 mM Tris HCl, pH 7.5 (H₂O solvent). This size exclusion was carried out at the EMBL laboratories, Grenoble. The eluate containing protein
was concentrated to 1-2 mg/mL using vivaspin 30 MWCO spin columns at 4,000g and 4°C. The samples were then dialysed using 30 MWCO Slide-A-Lyzers into buffers containing 150 mM NaCl, 3 mM MgCl₂, 20 mM Tris HCl, pH 7.5 with either 0%, 40% or 100% D₂O in the solvent. The respective dialysates were used as the buffer blanks.

Data were obtained from D22 beamline at the Institut Laue Langevin (ILL). The samples were loaded into a temperature controlled sample changer which held 20 quartz cuvettes. Measurements were collected at 10°C for one hour for each sample and buffer. Background radiation was also measured by placing a cadmium block in one of the cuvette holders. The neutron wavelength used was 6 Å with a sample to detector distance of 17 m resulting in a q range of 0.003 – 0.35 Å⁻¹. Scattered neutrons were measured by a 96 x 96 cm multidetector with pixel size of 7.5 x 7.5 mm. Data were radially averaged using GRASP, the buffer and background scattering were subtracted and the measured transmission for each sample (obtained from the attenuated beam measurement) were input for beamline calibration (for more details see Appendix VI). Guinier plots were analysed in PRIMUS.

### 2.14 Crystallisation

Crystallisation of xNAP1-C1, was performed with 500 and 50 mM NaCl concentrations in an attempt to crystallise the dimeric and larger oligomeric species. Crystallisation attempts of the xNAP1-C1-Histone complexes contained 150 mM NaCl. The following molecular dimensions crystallisation screens were used: Proplex, Macrosol, Morpheus, JCSG-plus, PACT-premier, Clear Strategy-Screens I and II, Structure screen I and II. Sitting well, vapour diffusion drops were set up in a 96 condition MRC-style two well plate using the Cartesian dispensing system (Genomic Solutions) run with Honey Bee software. The robot dispenses 100 nL of sample and 100 nL of precipitant into each well, and 60 µL mother liquor was manually put into the wells. Trays were stored either at 4 or 16°C.

If crystalline material developed in these drops, optimisation trials were set up manually. Hanging drops were set up in 24 well plates with 1:1, 1:2 and 2:1 µL protein to buffer ratio tested 500 µL of mother liquor was put into the wells for vapour diffusion. X-ray data were collected from any crystals that developed in the larger drops using an Xcalibur Nova X-ray diffractometer (Oxford Diffraction).
2.14.1 Limited Proteolysis

Limited proteolysis was performed using a JBS Floppy Choppy kit (Jena Bioscience), which contained four proteases; α-chymotrypsin, trypsin, subtilisin and papain. The four protease stock solutions (1 mg/mL) were diluted to 1/100 and 1/1000 and incubated with either the xNAP₂·H2A/H2B or the xNAP₂·H3/H4 complexes which were at 1.8 mg/mL in a 1:1 ratio (final volume 40 µL). Aliquots of 10 µL were taken at time 0, 0.5, 3 and 18 hr and mixed with 5 µL SDS Loading dye immediately. Fractions were examined by SDS PAGE.

Crystallisation trials of the xNAP-Histone complexes were performed at 0.45 and 1.8 mg/mL with in situ protease and the PACT premier screen. Directly prior to the robot uptake of the sample a 1/1000 dilution of α-chymotrypsin or a 1/10,000 dilution of papain were added to the sample. The trays were incubated at 4°C.
3. Purification of xNAP1-C1 and Histones

Analysis of hydrodynamic and structural characteristics of macromolecular assemblies requires a high yield of purified protein subunits. Previous research on the full length xNAP1 was hindered due to the low solubility of the protein. This prompted the search for more soluble domains using limited proteolysis, and the xNAP1-C1 fragment was obtained as a result of partial chymotryptic digestion (Friedeberg et al., 2006). The xNAP1-C1 fragment has 49 amino acids absent from the N-terminus and was found by Friedeberg et al., to be significantly more soluble than the full length protein.

At the same time as this work, the first crystal structure of the NAP1 family was obtained by Park et al. in the US. The *Saccharomyces cerevisiae* [yeast (y)] NAP1 X-ray diffraction data revealed well defined density for the central region of the protein, domains I and II; however, density was lacking for the N and C termini (Park and Luger, 2006a). Park et al. 2006 also crystallised a yNAP1 fragment lacking the N- and C- terminal domains which yielded an identical structure in the same space group, thus indicating that the termini were not involved in crystal packing and were in fact disordered. This central domain of the NAP1 protein contains the conserved nuclear localisation sequence (NLS) and the nuclear export sequence (NES). The NAP1 domain has been shown to be necessary and sufficient for nucleosome assembly and histone binding (Fujii-Nakata et al., 1992).

The xNAP1-C1 fragment and the full length yNAP1 share 33% identity. By aligning the xNAP1-C1 sequence with the yNAP crystal structure, xNAP1-C1 is shown to contain the entire functionally relevant and structured domains I and II (Figure 1.8). A simplified domain alignment is shown in Figure 3.1. Interestingly the crystal structure for the malarial NAP1 has recently been solved (pfNAPL) and shares a similar fold to that of yNAP1 (Gill et al., 2009). Dr C. Friedeberg in our laboratory sub-cloned the xNAP1-C1 fragment into the kanamycin resistant,
pET28b(+) vector (Appendix II), with a C-terminal His-tag, to allow for selective purification of fully translated proteins.

Expression clones of the four core Xenopus laevis histones: H2A, H2B, H3 and H4 were obtained from the Richmond lab (Institut für Molekularbiologie und Biophysik, Zürich). The genes had been cloned into ampicillin resistant pET3a vectors with one exception, H3, which was cloned into pET3d (Luger et al., 1997b) (Appendix III). The published protocol by Luger et al. 1997 for purifying these histones was labour intensive and resulted in a relatively low yield of histone (~2-5 mg/L culture). Optimisation of the purification protocols for both xNAP1-C1 and the histones was necessary to reduce purification time and improve protein yield. This chapter describes the optimised protocols for recombinant proteins, designed for use with 1L of over-expressed bacterial culture.

3.1 Optimisation of Recombinant xNAP1-C1 Purification

The published protocol for xNAP1 production (Friedeberg et al., 2006) was modified to optimise the purity and yield of the xNAP-C1 protein (section 2.5.2). Firstly the buffering agent was changed from 20 mM sodium phosphate to 20 mM Tris HCl, with the pH maintained at 7.5. Sodium phosphate stock solutions can crystallise out of solution over time, modifying the pH, and Tris HCl was therefore used to maintain a stable pH over time. Secondly solubility assays were performed to monitor protein aggregation under various conditions. The addition of 10% glycerol and 3 mM MgCl₂ to the buffers was found to increase the yield of the protein during purification and its solubility during storage.

Immobilised metal affinity chromatography (IMAC) was used as the first stage of protein purification, employing the xNAP1-C1 C-terminal His-tag. The IMAC column consisted of chelating ligands charged with Ni²⁺, that were immobilised on a 34 µm Sepharose matrix that created a high affinity surface for histidine residues. The soluble fraction of the cell lysate was injected onto a 1 mL HisTrap HP column (GE Healthcare) to bind the His-tagged xNAP1-C1 to the resin. The imidazole gradient was optimised to allow separation of the His-tagged protein from the histidine-containing contaminants, whilst keeping the sample volume as small as possible. A small sample volume was sought to avoid the requirement for subsequent concentration and to reduce the overall purification time and therefore minimise sample degradation. Eluted fractions from the nickel column were examined by SDS PAGE (section 2.2). xNAP1-C1 migrates on a 12.5% SDS polyacrylamide gel as if it were 55 KDa, the sequence mass is however only 41 KDa (Friedeberg, 2005).
Figure 3.2 - Nickel affinity chromatography elution profile of xNAP1-C1 and SDS PAGE of eluted fractions. A) An elution profile from the HisTrap column is shown with the imidazole gradient in green and UV absorbance at 276 nm (blue) and 254 nm (red). Fractions of interest, marked by the numbers in light blue, were examined by SDS PAGE (section 2.2). B) A region of the elution profile at the point in which xNAP1-C1 elutes from the column. C) SDS PAGE is shown of the selected fractions with a standard marker in the left hand lane. xNAP1-C1 has a sequence mass of 41 KDa and can be visualised as the dominant band on the gel highlighted by the arrow running anonymously to approximately 55 KDa.
The UV trace corresponding to eluting proteins from the HisTrap column shows that the majority of the bacterial lysate did not bind to the column and eluted in the flow through (0-15 mL) (Figure 3.2). SDS PAGE analysis shows that a number of bound contaminating proteins were separated from the xNAP1-C1 by the optimised imidazole gradient. The fractions 5-10 which contained predominantly xNAP1-C1 were pooled, for further purification using anion exchange chromatography.

To prepare the sample for anion exchange chromatography, the protein needed to be contained in a solvent with no imidazole and a lower NaCl concentration. This buffer change was achieved using a 50 mL HiPrep 26/10 Desalting column; this column works on the principal of size exclusion described in section 4.2.1. The proteins were separated from the buffer constituents due to their differences in size. The proteins therefore eluted in the buffer that the column was first equilibrated in, in this case the MonoQ buffer A.

Fractions that eluted from the Desalting column were examined by SDS PAGE. Fractions containing protein were pooled and injected onto a MonoQ 5/5 column. This anion exchange column contains a quaternary ammonium matrix, and binds molecules with a net negative surface charge. xNAP1-C1 has an unusual isoelectric point (pl) of 4.58, influenced by the negative surface on the NAP domain, thought to be involved in histone binding (Park and Luger, 2006a). At a pH of 7.5, xNAP1-C1 carries a net negative charge and therefore should bind to the MonoQ column. Positively charged contaminants will flow through the column, acting as the first level of contaminant separation. An increasing NaCl gradient of 50 -1000 mM NaCl was then applied to the column to selectively elute the xNAP1-C1 away from the remaining negatively charged contaminants.
**xNAP1-C1 – Anion Exchange Chromatography**

![Graph showing elution profile](image)

**Figure 3.3** - Anion exchange chromatography elution profile of xNAP1-C1 and SDS PAGE of eluted fractions. 

A) An elution profile from a MonoQ column is shown with the NaCl gradient in green and UV absorbance at 276 nm (blue) and 254 nm (red).  
B) A region of the elution profile at the point in which xNAP1-C1 elutes from the column. Fractions of interest, marked by the numbers in light blue were examined by SDS PAGE.  
C) SDS PAGE of the selected fractions is shown with a standard marker in the left hand lane. xNAP1-C1 has a sequence mass of 41 KDa and can be visualised as the dominant band on the gel highlighted by the arrow running anonymously to approximately 55 KDa.
The elution of xNAP1-C1 from the MonoQ column resulted in a sharp UV absorbance peak (Figure 3.3 A-B). A small proportion of contaminants were removed by this step, as seen by comparing the SDS PAGE of post HisTrap NAP (Figure 3.2C) and post mono Q NAP (Figure 3.3C), as well as concentrating the xNAP1-C1 sample for further use. Fractions 2-6 containing predominantly xNAP1-C1, as determined by SDS PAGE, were pooled and used in further research. Although some contaminants remained, the optimised protocol led to an excellent yield of xNAP1-C1 (~6 mg/L culture) with approximately 85% purity. A UV spectrum of the resultant protein sample was measured to determine the level of DNA contamination and assess the condition of the protein present.

![UV Spectrum of xNAP1-C1](Figure 3.4 - UV Spectrum of xNAP1-C1. A UV spectrum is shown of the xNAP1-C1 post MonoQ purification (Fraction 4). The Absorbance at 280 nm is shown as well as the ration between the absorbance at 260 and 280 nm.)

The UV spectrum taken of the xNAP1-C1 sample contained no visible light scattering between 300-320 nm, indicating the protein was not aggregated (Figure 3.4). The ratio between absorbance at 260 nm and 280 nm \((A_{260}/A_{280})\) was 0.61, this suggested that no DNA contamination was present.

### 3.2 Optimisation of Recombinant Histone Purification

Optimisation of the published protocol for the four core histones was performed to increase protein yield and reduce purification time (section 2.5.5). The previous method washed the cell pellets rigorously, and then denatured the sample. The supernatant was then applied to a S200 size exclusion column followed by HPLC and then lyophilisation. This chapter describes the optimised method, purifying each of the four core histones in the same manner. The histones were expressed as insoluble inclusion bodies from E.coli. The cell lysate pellet was first washed to remove bacterial cell debris, however less so than in the published protocol to reduce histone loss (section 2.4.5). The proteins were then solubilised and denatured to allow for further purification.
As histones are naturally occurring DNA binding proteins, it was first necessary to remove contaminating bacterial DNA from the sample whilst the proteins were unfolded in guanidinium-containing buffer. This was achieved using hydroxyapatite (HA), a mineral with a strong DNA binding capacity. HA contains positively charged calcium residues which can bind to DNA strongly via calcium coordination with the phosphoryl groups on the DNA. These positive calcium ions in the HA will also bind to some of the negatively charged contaminating proteins, further aiding histone purification (Hjerten et al., 1956). HA also contains negative phosphate groups, and under certain conditions these can act as a weak cation exchangers, binding small positively charged proteins. In the conditions used, there was no significant binding to the histone proteins seen. To monitor the amount of contaminating DNA in the protein sample, UV spectra were taken before and after HA incubation and the $A_{260/280}$ was examined.

The UV spectra of the histone samples before and after HA incubation (Figure 3.5), showed a significant decrease in the $A_{260/280}$ ratio for all four histones to approximately 0.7. The total UV absorbance at 280 nm for all of the samples decreased by approximately two fold for H2A, H2B and H3 and three fold for H4. For further purification of the histones, cation exchange chromatography was employed. The histones had to be exchanged into a buffer containing no guanidinium as this charged molecule would affect protein-column binding. Buffer exchange was carried out using a 50 mL HiPrep 26/10 Desalting column (as described in section 3.1) placing the proteins into ion exchange buffer A. Conditions were however, still denaturing as the ion exchange buffers contained 4 M of the non-charged denaturant, Urea.

Eluate from the desalting column was examined by SDS PAGE and the fractions containing protein were pooled and injected onto a HiTrap Sepharose (SP) high performance (HP) column. This cation exchange chromatography column contained negatively charged side chains ($\text{CH}_2\text{CH}_2\text{CH}_2\text{SO}_4^-$), chemically linked to a polypropylene matrix, which provided a binding surface for proteins with a net positive charge. The four core histones have relatively high pl values of 10.87, 10.17, 10.82 and 11.02 for H2A, H2B, H3 and H4 respectively. In pH 5.2, at which the cation exchange column was run, all of the histones carried a very strong net positive charge and therefore should bind to the column along with any positively charged contaminants. Contaminating proteins with a net negative charge at this pH should flow through the column, to provide further purification. An increasing NaCl gradient of 0.2 – 1 M, was applied to the column to selectively elute the histones from the other bound contaminating proteins.
Figure 3.5 - UV spectra of histone H2A, H2B, H3 and H4 before and after hydroxyapatite incubation. The UV spectra of the H2A (A), H2B (B), H3 (C) and H4 (D) lysate before and after HA incubation are shown. The A_{260/280} ratio is shown alongside the absorbance at 280 nm in the top right of each spectrum.
Histone H2A – Ion Exchange Chromatography

Figure 3.6 - Cation exchange chromatography elution profile of histone H2A and SDS PAGE of eluted fractions. A) An elution profile is shown with the NaCl gradient in green and UV absorbance at 276 nm (blue) and 254 nm (red). Fractions of interest, marked by the numbers in light blue were examined by SDS PAGE (section 2.2). B) A region of the elution profile at the point in which H2A elutes from the column. C) SDS PAGE of the selected fractions is shown with a standard marker in the left hand lane. Histone H2A has a sequence mass of 14.6 KDa.
Histone H2B – Ion Exchange Chromatography

Figure 3.7 - Cation exchange chromatography elution profile of histone H2B and SDS PAGE of eluted fractions. A) An elution profile is shown with the NaCl gradient in green and UV absorbance at 276 nm (blue) and 254 nm (red). Fractions of interest, marked by the numbers in light blue were examined by SDS PAGE (section 2.2). B) A region of the elution profile at the point in which H2B elutes from the column. C) SDS PAGE of the selected fractions is shown with a standard marker in the left hand lane. Histone H2B has a sequence mass of 14.3 KDa.
Figure 3.8 - Cation exchange chromatography elution profile of histone H3 and SDS PAGE of eluted fractions. A) An elution profile is shown with the NaCl gradient in green and UV absorbance at 276 nm (blue) and 254 nm (red). Fractions of interest, marked by the numbers in light blue were examined by SDS PAGE (section 2.2). B) A region of the elution profile at the point in which H3 elutes from the column. C) SDS PAGE of the selected fractions is shown with a standard marker in the left hand lane. Histone H3 has a sequence mass of 16.4 KDa.
Histone H4 – Ion Exchange Chromatography

Figure 3.9 - Cation exchange chromatography elution profiles of histone H4 and SDS PAGE of eluted fractions. A) An elution profile is shown with the NaCl gradient in green and UV absorbance at 276 nm (blue) and 254 nm (red). Fractions of interest, marked by the numbers in light blue were run on an SDS PAGE (section 2.2). B) A region of the elution profile at the point in which H4 elutes from the column. C) SDS PAGE of the selected fractions is shown with a standard marker in the left hand lane. Histone H4 has a sequence mass of 12.1 KDa.
Elution profiles from the cation exchange column show the remaining contaminants eluting in the flow-through volume (0-15 mL) (Figures 3.6-3.9). SDS PAGE of selected fractions indicated that the contaminating proteins eluted first from the column, followed by the histone proteins, which eluted in NaCl concentrations of 1-2 M. Fractions containing predominantly histones; 8-10 for H2A (Figures 3.6), 7-9 for H2B (Figure 3.7), 8-11 for H3 (Figure 3.8) and 8-11 for H4 (Figure 3.9) were pooled. The samples were buffer exchanged into 0.5% formic acid, again using the 50 mL HiPrep 26/10 desalting column. The eluted protein from the desalting column was flash-frozen, lyophilised and stored at –70°C until required. From SDS PAGE it is evident that some contaminants and/or degradation products were still present in the samples at this stage.

The total yield of protein from each 1 L bacterial pellet varied between histones: approximately 12 mg/L for H2A, 18 mg/L for H2B, 5 mg/L for H3 and 6 mg/L for H4. The yield obtained also varied by up to 2 mg/L between preparations from pellets from different cultures. The optimised purification protocol took ~11 hr from frozen bacterial lysate pellet to the start of lyophilisation, therefore reducing the time by two thirds from the published protocol. Some contaminants and/or degradation products of the histones were still present at this stage however the purity reached was still 90-99% depending on the histone.

3.3 Reconstitution of Recombinant Histone Dimers and Tetramers

3.3.1 H2A/H2B Dimer

Refolding of the histones H2A and H2B, followed by purification of the heterodimer was achieved by following a previously published method (Schmiedeke et al., 1989). Given that the protocol was a fast and efficient way of isolating H2A/H2B heterodimers, no optimisation was required and therefore the method is only briefly described (section 2.6.2).

Lyophilised aliquots of H2A and H2B were resuspended in 10 mM HCl and mixed with a 1:1 molar stoichiometry. Heparin buffer A was then added slowly to refold the proteins. The sample was injected onto a 1 mL HiTrap Heparin HP column. The column matrix has a structure similar to that of DNA, hence providing affinity for the histone dimer, a native DNA-binding protein complex. An increasing NaCl gradient of 50 mM – 2 M was applied to elute the heterodimers separately from the homodimers, monomers and any remaining contaminants.
Histone H2A/H2B – Heparin Affinity Chromatography

Figure 3.10 - Heparin affinity chromatography elution profiles of Histones H2A/H2B and SDS PAGE of eluted fractions. A) An elution profile is shown with the NaCl gradient in green and UV absorbance at 276 nm (blue) and 254 nm (red). B) A region of the elution profile at the point in which the heterodimer eluted from the column. Fractions of interest, marked by the numbers in light blue were examined by SDS PAGE (section 2.2) and AUT gel electrophoresis (section 2.3). C) SDS PAGE of the selected fractions is shown with a standard marker in the left hand lane. H2A/H2B dimer can be visualised as the dominant band on the gel and is highlighted by the arrow. D) AUT electrophoresis with individual histones H2B and H2A as markers in the first two lanes and eluate fraction 4 in the final lane.
The elution profile from the heparin column contained one predominant absorbance peak (Figure 3.10). The histone H2A and H2B have the same electrophoretic ability in SDS PAGE so AUT gel electrophoresis was used to examine the eluate from the heparin. The predominant peak contained approximately equal amounts of H2A and H2B suggesting that it corresponds to the elution of the H2A/H2B heterodimer (Figure 3.10 D). A small peak was seen at an elution volume of 18 mL, representing histone monomers. A third elution peak was seen at the elution volume of 27 mL corresponding to contaminating protein (seen in fractions 8 and 9 of SDS PAGE). No misfolded homodimers were seen eluting from the column as they have an extremely strong affinity to heparin and require 6 M guanidinium to dislodge them (Gloss and Placek, 2002).

Typically, 1 L growth culture of H2A and 1 L of H2B produced a total yield of purified H2A/H2B dimer of approximately 12 mg. Eluted fractions 2-5, containing pure H2A/H2B dimers were pooled, thus removing any visible contamination identified by SDS PAGE.

### 3.3.2 H3/H4 Tetramer

Refolding the H3/H4 tetramer was based upon a published protocol (Luger et al., 1999) and optimised to reduce the length of purification (section 2.6.3). The lyophilised histones H3 and H4 were resuspended in denaturing buffer, mixed together in equimolar amounts and incubated for 1 hr. The sample was dialysed into heparin buffer B to allow the proteins to refold in a high salt environment so as to minimise aggregation. The soluble protein was recovered by centrifugation and subjected to size exclusion chromatography, also in high salt conditions, to separate the H3/H4 tetramers from individual subunits and DNA contaminants as previously described by Luger et al.
**Histone H3/H4 – Size Exclusion Chromatography**

**Figure 3.11** - Size exclusion chromatography elution profiles of Histones H3/H4 and SDS PAGE of eluted fractions.

A) An elution profile is shown with UV absorbance at 276 nm (blue) and 254 nm (red). B) A region of the elution profile at the point in which the H3/H4 tetramer eluted from the column. Fractions of interest, marked by the numbers in light blue were examined by SDS PAGE (section 2.2). C) SDS PAGE of the selected fractions is shown with a standard marker in the left hand lane. The H3/H4 proteins can be visualised as the dominant bands on the gel and are highlighted by the arrows.
Three main eluting species were seen on the Superdex elution profile (Figure 3.11); the main species eluted at ~9 mL and most likely corresponded to the (H3/H4)4 tetramer due to the high NaCl content of the buffer (Banks and Gloss, 2003). The species that eluted at ~10.5 mL could correspond to partially dissociated histone dimers as seen by SDS PAGE of fraction 8, however, the presence of histone monomers could not be ruled out. The species which eluted at ~7.5 mL, did not contain protein, as determined by SDS PAGE of fractions 1 and 2, the high A260/280 ratio of the absorbance peak suggests that it was remaining DNA contaminants.

Typically 1 L of bacterial culture of H3 and 1L of H4, produced a total yield of purified H3/H4 tetramer of 2 mg. Fractions containing the H3/H4 tetramer (4-6) were pooled and utilised in further research, and at this stage no contaminants were apparent by SDS PAGE.

### 3.4 Discussion

**An optimised purification protocol for xNAP1-C1 was developed.** Optimisation of purification buffers, addition of an anion exchange chromatography step and the reduced purification time have resulted in a much improved protein yield and purity. Some of the remaining contaminants were also observed by Dr C. Friedeberg and were shown to be degradation products of xNAP1-C1 by mass spectroscopy and sequencing (Friedeberg, 2005). It is possible that as well as degradation products, some other contaminating proteins may be present, as it is highly unlikely that these proteins have the same nickel binding affinity and pI as xNAP1-C1, it is probable that they are binding to the xNAP1-C1 protein. xNAP1-C1 is highly charged and contains a very acidic domain, therefore making it very susceptible to non-specific binding with other proteins via electrostatic interactions (Park and Luger, 2006a).

**An optimised purification protocol for the histones has been developed.** Optimisation of histone purification protocols successfully provided higher yields in approximately one third of the time of the previously published method. Incubation with hydroxyapatite (HA) significantly reduced the A260/280 ratio of the sample, indicating a reduction in DNA contamination which greatly improved the efficiency of downstream cation exchange chromatography. Cation exchange chromatography proved to be a quick and efficient way of purifying the histone proteins away from other contaminants and the resulting elution peaks corresponding to the histones H2A and H2B were symmetrical and sharp indicating monodisperse samples. The
histones H3 and H4 eluted over a larger volume than H2A and H2B, possibly indicating greater DNA contamination. Purified histones were then combined and further purified as heterodimers.

Production of high quality proteins (Figure 3.12) was an essential prerequisite for the biophysical characterisation of xNAP1-C1 alone and in complex with histones. Natively produced Gallus gallus (g) H2A/H2B and H3/H4 complexes are also shown in Figure 3.12, these were a kind gift from John Baldwin of the John Moores University (for details see section 7.1). The purified xNAP1-C1 fragment was used in all following chapters and will be referred to simply as xNAP.
4. Hydrodynamic Analysis of xNAP

4.1 Introduction

The hydrodynamic behaviour of a molecule is based upon its size and shape as well as the viscosity and density of the solvent in which it moves. Hydrodynamic techniques can be executed on a system at rest to analyse a particle’s Brownian motion, or to a system under a known force to examine the resulting translational / rotational movement of a particle. Hydrodynamics can be utilised to obtain low resolution information about the molecule in question, usually by comparing it to a theoretical sphere of equivalent hydrodynamic properties, the radius of which is called the Stokes radius or hydrodynamic radius, \( R_H \).

Previous sedimentation velocity studies of yNAP1 in conditions containing 10 mM KCl, have described the formation of a dimer with a sedimentation coefficient of 4.4 S (Toth et al., 2005). This is consistent with the yNAP1 and pfNAPS crystal structures (Figure 1.8 and 1.9), which demonstrate that the repeating biological unit is a dimer (Gill et al.; Park and Luger, 2006a). As the KCl concentration was increased to 100 mM, the yNAP1 dimer (yNAP₂) was seen to oligomerise. The model found to best fit the data was a yNAP dimer-octamer-hexadecamer equilibrium, a clear systematic error can, however, still be seen (Toth et al., 2005). Sedimentation equilibrium analyses also proved problematic where the models could not fully account for the complex equilibria seen in the analysis (Toth et al., 2005). Preliminary sedimentation equilibrium studies were also carried out with the Xenopus NAP T2 fragment (section 1.4), which also suggested xNAP1-T2 oligomerisation (Friedeberg et al., 2006). The protein concentration was, however, too low and the data could not be analysed with any degree of accuracy.

Following the production of high yields of pure proteins (as described in Chapter 3); it was possible to carry out analytical hydrodynamic experiments on NAP and histones from Xenopus laevis. This chapter describes the use of analytical size exclusion, dynamic light scattering and analytical ultracentrifugation to characterise the oligomerisation states of xNAP at various protein concentrations and ionic conditions.
4.2 Size Exclusion Chromatography

4.2.1 Theory

Size exclusion chromatography is commonly used to separate molecules according to their molecular mass and shape (Porath and Flodin, 1959). The chromatography columns are made up of a continuous liquid phase contained within a solid gel matrix, available in a range of pore sizes. Molecules are retarded within the pores as they flow down the column, smaller molecules are retarded the most and elute last, while larger molecules are retarded less due to steric hindrance and therefore elute faster. Some very large molecules are excluded from the pores altogether and elute in the void volume of the column.

This technique can be used to purify proteins away from contaminants, as illustrated in Chapter 3 (section 3.3.2). Molecular weight (Mr) estimations can also be determined using size exclusion, based on a standard curve. However, because elution volume depends on the hydrodynamic radii (R_H) of the molecules, it is only accurate for spherical, and in practise fairly globular proteins (Andrews, 1970).

Size exclusion chromatography is a useful analytical tool for comparative analysis of macromolecules in various conditions. It not only provides an estimate of the oligomerisation state of a molecule, but it can also reveal conditions which affect their equilibrium. This can be valuable information for planning further biophysical studies which require a monodisperse species, such as small angle scattering and X-ray crystallography. This section describes the use of size exclusion to monitor the effects of NaCl and protein concentration on the oligomerisation state of xNAP.

4.2.2 Ionic Strength Dependence of xNAP Oligomerisation

Proteins can selfassociate via electrostatic interactions in a salt dependant manner. To investigate the effect of ionic strength upon xNAP oligomerisation, size exclusion chromatography was performed in buffers ranging from 50 mM to 1 M NaCl. Included within the range tested was 150 mM NaCl, to obtain information at an ionic condition similar to that of the cellular environment. xNAP was dialysed overnight into the appropriate SE buffer prior to injection onto a Superose 6 column (described in section 2.7).
Comparison of the six size exclusion chromatography profiles in various ionic conditions (Figure 4.1) illustrated that xNAP is capable of forming multiple oligomeric species. Four predominant oligomeric species highlighted by the red lines were identified; I, II III and IV, with a peak of elution at 14.8, 13.6, 11.7 and 10.3 mL respectively. In 1 M NaCl, xNAP appeared to be present mainly as species I, with some further oligomerisation into species II. Conversely, at low ionic conditions, such as 50 mM NaCl, xNAP was present mostly as species IV. At intermediate NaCl concentrations, a complex equilibrium was seen involving the four oligomeric species. This equilibrium, between species I, II, III and IV, was also seen at 150 mM NaCl, approaching physiologically relevant ionic conditions.

None of the conditions tested resulted in a single elution peak, thus representing heterogeneous solutions. It was notable that in the experimental series, species I in 1 M NaCl had the sharpest, most highly resolved peak. The other peaks were comparatively broader suggesting that a more complex equilibrium was present.
Figure 4.1 - Size exclusion chromatography elution profiles of xNAP in a range of ionic conditions. The elution profiles from a Superose 6 column are shown with an input of 200 µL of NAP at 1.3 mg/mL. The six runs were performed in SE buffer (section 2.5.1) containing the NaCl concentration indicated in the top right of each profile. The red lines highlight the four main eluting species; I, II, III and IV.
4.2.3 Concentration Dependence of xNAP Oligomerisation

A change in protein concentration can alter the equilibrium between oligomeric species in solution. As discussed in Chapter 1, the concentration of NAP *in vivo* varies between the cytoplasm and the nucleus. Many physiological systems utilise concentration dependent oligomerisation to alter the function of a protein. To analyse the effect of protein concentration upon xNAP oligomerisation, size exclusion chromatography was performed in 150 mM NaCl at various protein concentrations. The sample was first concentrated to 3 mg/mL and then diluted to the appropriate concentrations to size exclusion.

![Size exclusion chromatography elution profiles of xNAP at various protein concentrations](image)

Figure 4.2 - Size exclusion chromatography elution profiles of xNAP at various protein concentrations. Elution profiles of xNAP from a Superose 6 column, in SE buffer containing 150 mM NaCl (section 2.5.1), are shown overlaid. The samples were loaded at either 0.5 (light blue), 1 (green), 2 (red) or 3 mg/mL (dark blue). The red lines are shown indicating the four putative species observed in Figure 4.1.

At relatively low protein concentrations, 0.5 mg/mL, xNAP appears to be in two main states of oligomerisation (Figure 4.2). The two elution peaks are similar in intensity and roughly correspond to species II and III described in the NaCl dependence assay in section 4.2.2. When the protein concentration is increased, the equilibrium shifts, so that more of the protein is present as species III than species II, thus indicating that xNAP oligomerisation is concentration dependent. At all four protein concentrations, small elution peaks corresponding to species I, were also present. Elution peaks were also seen at 7.5 mL in all of the profiles, corresponding to aggregated material eluting in the void volume. This aggregation increased as the concentration of protein loaded onto the column increased.
4.3  Light Scattering

4.3.1  Theory

Dynamic light scattering (DLS) is a technique that measures a particle’s diffusion characteristics by analysis of scattered light. Diffusion of a particle via Brownian motion is dependent upon particle size and shape, and is related to the particle’s hydrodynamic radius ($R_H$). DLS can be used to calculate the diffusion coefficient of particles. This is achieved by analysing fluctuations in the scattered light from the sample, following irradiation with a monochromatic source, in this case an 830 nm laser (Serdyuk I.N., 2007). These fluctuations arise due to particle diffusion, altering their relative positions to the detector. The intensities are measured over short time intervals and the correlation between these intensities at different times is analysed. If the particles are exhibiting Brownian diffusion, the correlation function is of exponential decay (Bloomfield, 1981). The decay constant of the correlation curve is proportional to the diffusion coefficient of the molecule which, in turn is related to the $R_H$ using the Stokes-Einstein equation:

$$R_H = \frac{kT}{6\pi\eta D_0}$$

where $k$ is the Boltzmann constant, $\eta$ is the viscosity of the solvent, $T$ is the temperature in Kelvin and $D_0$ is the diffusion coefficient of the molecule (Bloomfield, 1981).

Another light scattering method using size exclusion chromatography in conjunction with multi angle laser light scattering (SEC MALLS) can be used to determine absolute $M_r$ of samples. As the sample elutes from the size exclusion chromatography column, the refractive index (RI) and light scattering (LS) at multiple angles are measured continuously (Serdyuk I.N., 2007). The weight averaged molar mass ($M$) of the eluate is derived from the Rayleigh-Gans-Debye approximation which in its simplest form is:

$$\frac{Kc}{R_\theta} = \frac{1}{M}$$

where $K$ is a constant, $c$ is the concentration of the solute and $R_\theta$ is the excess Rayleigh scattering. $K$ is an optical constant which includes parameters for the refractive index of the solvent, the solute $d\tilde{n}/dc$ (RI increment), Avogadro’s number and the wavelength of the incident light. The concentration of the solute ($c$) is derived from converting the RI voltages to g/mL by multiplying by a RI calibration constant and dividing by the $d\tilde{n}/dc$ of the solute (Knobloch and Shaklee, 1997). The Rayleigh ratio ($R_\theta$) is the ratio of scattered light ($I_s$) to incident light ($I_o$), measured at a fixed distance ($r$) from the sample at a given angle ($\theta$):
\[ R_\theta = \frac{I_s}{I_0} \left( \frac{r^2}{1 + \cos^2 \theta} \right) \]

A Debye plot of \( Kc/R_\theta \) against \( \sin^2(\theta/2) \) can be plotted for all the LS angles and concentrations measured, a line of best fit is drawn and where it intercepts the y-axis equates to \( 1/M \) (Knobloch and Shaklee, 1997).

### 4.3.2 Dynamic Light Scattering

DLS can be used to qualitatively determine the heterogeneity of a sample, estimate the level of aggregation and, if the sample is monodisperse, derive an \( R_n \). This section of the chapter describes the use of dynamic light scattering as a comparative analysis tool for three of the main oligomeric species of xNAP, following size exclusion chromatography (section 2.9). The NaCl dependence assay shown in section 4.2.2, indicated that four main species I, II, III and IV were formed. In 50 mM NaCl conditions, species IV was well defined and the corresponding fraction which eluted at 10.3 mL was examined by DLS. The elution profile for xNAP in 1 M NaCl contained well defined peaks for species I and II; the corresponding fractions which eluted at 13.6 and 14.8 mL were therefore, also analysed by DLS. The polydispersity was noted to determine if the species were present in an equilibrium.

![Figure 4.3 - Dynamic light scattering of xNAP oligomers. The elution profiles of xNAP from a superose 6 column are shown, run in SE buffer (section 2.5.1) with (A) 50 mM NaCl and (B) 1 M NaCl. DLS measurements of the eluting NAP species III, II and I are shown in plots C, D and E respectively. The polydispersity and resulting \( R_n \) for each DLS measurement are presented in the grey boxes.](image-url)
The resulting hydrodynamic radii for the three NAP species I, II and IV, were 46, 63 and 89 Å respectively (Figure 4.3). The different R_h values obtained for each eluted species confirmed that they were in fact three distinct oligomeric states or conformations. The relatively low polydispersity value obtained for species I (9%), suggested that this was the most homogeneous species, and did not readily associate into the larger species. The polydispersity values for species II and III were slightly larger at 13% and 16% respectively, suggesting that these larger species might be involved in a more complex equilibrium.

### 4.3.3 SEC-MALLS

SEC-MALLS was performed at the N.I.M.R centre in Mill Hill, London, to assign molecular weights to the different oligomeric species formed by xNAP (section 2.10). Due to time restrictions, all four species were analysed in one run using SE buffer containing 150 mM NaCl (section 2.5.1). The xNAP sample was injected onto the Superose 6 column with a concentration of 0.3 mg/mL and eluted for 1.5 CV. During the size exclusion chromatography run the refractive index (RI) of the eluting fractions was measured in conjunction with the light scattering (LS). LS was measured at 11 different angles to allow for determination of the M_r of the eluting species. The baselines for the individual LS and RI curves were manually adjusted and using the ASTRA 5.3.4 software (Wyatt technology). Determination of the M_r spanning across eluate peaks was performed using all 11 scattering angles.
Figure 4.4 - SEC-MALLS of xNAP. A) An elution profile from the Superose 6 column is shown with the refractive index (RI) in red and the light scattering (LS) in blue. The molecular weight estimates for corresponding peaks are shown in black on the y-axis. B) A region of the elution profile where species III and IV are eluted; arrows to the minimum and maximum M_r are shown for clarity. The four xNAP oligomeric species I, II, III and IV are labelled in red as well as a fifth region of elution (V).

The four main oligomeric species of xNAP (I, II, III and IV) were seen on the resulting RI profile (Figure 4.4A, red curve). The predominant species present were III and IV as determined by the relative heights of the peaks in the RI trace. The M_r estimates are shown in black; if this curve has a plateau across the elution peak it indicates that the species is homogeneous. Such plateaus were seen for species I and II and estimates of 96 and 165 kDa were calculated respectively. Despite the low signal, the values for species I, corresponded to a xNAP dimer (xNAP_2) of 81.4 kDa, within the experimental error. The M_r determined for species II matched very closely to a dimer of dimers (2·xNAP_2) of 162.8 kDa.
The Mr determination curve showed no such plateau across the peak corresponding to the xNAP species III. The peaks observed in Figure 4.4 A+B are offset indicating that the sample is polydisperse. This occurs because the RI measures the total concentration of the eluting species, whereas, LS is weighted for molecular mass. Across peak III / IV a range of Mr was seen between 400 – 720 kDa (10-18mer). The average Mr over the entire peak was 578 kDa, corresponding to 7·xNAP₂. In the trailing edge of the peak, labelled as species V, an extremely sloped Mr curve was observed corresponding to non specific oligomerisation.

4.4 Analytical Ultracentrifugation

4.4.1 Theory

Analytical ultracentrifugation (AUC) is a hydrodynamic technique that is based upon Newton’s second law; force is equal to mass multiplied by acceleration. Placing a molecule under a known force results in a measurable acceleration that facilitates the determination of the mass. On small biological molecules the effects of opposing forces such as diffusion and thermal currents are significant; therefore to measure their acceleration, a relatively large force must be applied. Strong centrifugal forces can be obtained using an ultracentrifuge, allowing analysis of very small molecules. The Beckman XL-A analytical optical centrifuge has an in-built spectrophotometer which can measure the absorbance of the samples as they spin. By following these profiles during centrifugation, the movement of solute particles can be measured in real time (Ralston).

There are two complementary AUC methodologies; Sedimentation velocity and Sedimentation equilibrium. In sedimentation velocity experiments, a high centrifugal force is applied to samples in a cell with a relatively long chamber length. This is optimal for obtaining information on particle size, shape and inter-molecular interactions. Sedimentation equilibrium employs a lower centrifugal force and uses a smaller cell length, optimal for gaining shape independent molecular weights.

4.4.1.1 Sedimentation Velocity

The high centrifugal forces used in sedimentation velocity cause a translational motion of the particle, greater than that of Brownian diffusion, and thus the molecules sediment to the bottom of the cell. The rate of sedimentation is dependent upon the speed of the rotor (centrifugal force) and the particle size and shape (buoyancy and frictional forces). A UV
spectrum of the cell is measured during the experiment; as the particles sediment the scan changes accordingly, following the movement of solutes into the outer regions of the cell. A solute boundary is formed that moves across the cell as sedimentation progresses. The speed at which this boundary moves is dependent upon the sedimentation coefficient (s) of the solute and the slope of the boundary is related to the diffusion coefficient (D) (Figure 4.5).

Figure 4.5 - Schematic representation of a sedimentation velocity experiment. A typical cell (grey circle) contains two sectors, one for buffer and one for sample. More buffer (25 µL) is used than sample, so that the buffer and sample menisci are distinguishable in the scan. The rotor is accelerated to the chosen speed. At the beginning of the experiment (t=0), a scan is taken prior to particle sedimentation and the positions of the two menisci and the sample absorbance are visualised. At t=1 and t=2 the particles begin to sediment, creating an absorbance boundary. The progression of the midpoint can be followed across the cell throughout the experiment. The gradient of the boundary will decrease as the back diffusion of the particles becomes greater.

The time resolved scans are collated and ln \( r_b(t)/r_b(t_0) \) vs. (t-t₀) can be plotted, where \( r_b \) is the radius of the boundary midpoint at time t relative to the boundary midpoint at the start (t₀) (McRorie D. K.). The slope of this plot, divided by the angular velocity squared (\( ω^2 \)), yields the sedimentation coefficient (s). When using the sedimentation coefficient to calculate the molar mass (M) of the molecule, the buoyancy term (1-\( ν \rho_0 \)) and the frictional coefficient (f) must be considered. These terms together with Avogadro’s number (\( N_A \)) are related by the Svedberg equation:

\[
s = \frac{M(1-\nu \rho_0)}{N_A f}
\]

To analyse sedimentation velocity data, experimental parameters including the shape and diameter of the sector cell and the radius of the cell position must be considered. The Lamm
equation relates the sedimentation and diffusion coefficients with the experimental conditions (Serdyuk I.N., 2007):

\[
\left( \frac{dC}{dt} \right)_r = -\frac{1}{r} \left\{ \frac{d}{dr} \left[ \omega^2 r^2 sC - D r \left( \frac{dC}{dr} \right)_t \right] \right\}_t
\]

The Lamm equation has two unknowns, the sedimentation (s) and diffusion (D) coefficients, therefore an exact solution can only be obtained when one of these unknowns is equal to 0. SEDFIT (Schuck, 2000) is an AUC analysis program which uses numerical solutions to solve the Lamm equation with the best fit to the experimental data resulting in a root mean square deviation (RMSD), illustrating the goodness of fit.

The ratio of the experimentally gained s and D values gives the molecular weight of the solute using another form of the Svedberg equation:

\[
M = \frac{sRT}{D(1 - \bar{\varrho} \rho_0)}
\]

where R is the gas constant, T is the temperature in Kelvin and \((1 - \bar{\varrho} \rho_0)\) is the buoyancy term. The frictional coefficient (f) of the solute is related to the diffusion coefficient by \(D = RT/N_A f\) and relays information about the size and shape of the molecule. The frictional ratio \((f/f_0)\) of a solute relates the f to that of a theoretical spherical molecule with the same M, and specific volume \((f_0)\); thus if the solute was spherical it would have an \(f/f_0\) of 1, divergence from spherical will increase the value. The calculation of M, is dependent upon the s and \(\bar{\varrho}\), thus making it shape dependent.

Using a continuous c(s) model to solve the Lamm equation in SEDFIT provides a distribution plot of sedimentation coefficients within the sample. The c(s) plots are a good estimation of the numbers of species present and their size distribution. A continuous c(M) model can also be used to provide a distribution of molecular weights within the sample. The c(s) and c(M) models use the approximation that all species in the model have the same frictional ratio. The effects of this approximation, when multiple species are present with various \(f/f_0\) are negligible in the c(s) plot but can lead to distortions in the c(M) plot.
4.4.1.2 Sedimentation Equilibrium

Particles spun at low centrifugal speeds, migrate to a position in the cell where the centrifugal and diffusional forces are equal, an exponential concentration gradient increasing towards the bottom of the cell is formed (Figure 4.6). The concentration gradient can be visualised after equilibrium has been reached via the UV scan. The samples are spun at a range of speeds to obtain the multiple concentration gradients for analysis. The time taken to reach equilibrium is dependent on particle shape however the radial position of equilibrium is shape independent (Ralston).

![Figure 4.6 - Schematic representation of a sedimentation equilibrium experiment. A typical cell contains 6 sectors (grey squares), and three concentrations of sample are measured together with three corresponding buffers. More buffer (10 µL) is used than sample, so that the buffer and sample menisci are distinguishable in the absorbance scan. A scan is taken at the beginning of the experiment to measure overall protein absorbance. The rotor is accelerated to the chosen speed and after approximately 18 hr (t=1), when the sample has reached equilibrium a second scan is taken. Further scans at 3 hr intervals are performed to confirm that equilibrium has been achieved. An exponential absorbance curve is produced.](image)

For a monodisperse sample, the distribution of molecules in equilibrium can be used to obtain a shape independent molecular mass with the second Svedberg equation:

\[
C(r) = C(a) \exp \left[ \frac{-\omega^2 M (1 - \bar{\rho}) (r^2 - a^2)}{2RT} \right]
\]
where \( C(r) \) and \( C(a) \) are the concentration of solute at the radial position \( r \) and the meniscus (a) respectively (Serdyuk I.N., 2007). The solvent density (\( \rho \)), gas constant (\( R \)) and the temperature in Kelvin (\( T \)) are accounted for. Weighted fits are usually performed to reduce the influence of outliers in the analysis. The variance for each \( M_i \) is calculated from the sum of squares of the residuals divided by the number of degrees of freedom (McRorie D. K.).

### 4.4.2 Application and Results

#### 4.4.2.1 Sedimentation Velocity

To further investigate the xNAP oligomerisation states, sedimentation velocity experiments were performed at two ionic conditions: 20 mM and 1 M NaCl with a xNAP\(_2\) concentration of 3.5 \( \mu \)M (section 2.11.1). These conditions were chosen due to the previous size exclusion data suggesting that xNAP sample was the most homogenous at these ionic conditions. Sedimentation coefficients and frictional ratios of the species I and III were sought to elucidate their size and shape and for comparison to those published of the yNAP homologue.
Figure 4.7 - Sedimentation velocity of xNAP. Experiments were carried out with xNAP at a concentration of 0.5 mg/mL at 10°C and 40,000 rpm. The experiments were performed in SE buffer with either 20 mM or 1 M NaCl (section 2.5.1). A-B) Experimental data curves with fit lines, residual bitmaps and residual plots from the analysis in 1 M and 20 mM salt conditions respectively. C) Overlay of C(s) plots of xNAP in high (blue) and low (red) salt, the corresponding sedimentation coefficients (s) are shown above the relevant peaks.
Table 4.1 - Sedimentation velocity of xNAP. The frictional ratio ($f/f_0$), sedimentation coefficients ($s$) and estimated Mr for xNAP are shown at two concentrations.

<table>
<thead>
<tr>
<th>Sample</th>
<th>$f/f_0$</th>
<th>$s$</th>
<th>$M_r$ (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>xNAP - 1 M NaCl</td>
<td>1.65</td>
<td>4.2</td>
<td>82</td>
</tr>
<tr>
<td>xNAP - 20 mM NaCl</td>
<td>2.01</td>
<td>8-18</td>
<td>330-1240</td>
</tr>
</tbody>
</table>

Modelling sedimentation velocity data to the Lamm equation using SEDFIT resulted in excellent fits to the experimental data with low RMSD values (Figure 4.7). Analysis of the data at 1 M NaCl, indicated that under these conditions xNAP was present as a monodisperse species with a sedimentation coefficient of 4.2 S. The calculated $M_r$ was 82 kDa, corresponding to a xNAP dimer with a theoretical $M_r = 81.4$ kDa. In low NaCl conditions (20 mM), xNAP was seen to be heterogeneous with the absence of a well-defined peak in the $c(s)$ plot. The sample contained species with sedimentation coefficients that ranged from approximately 8 to 18 S. As SEDFIT is limited to solving the Lamm equation using a single frictional coefficient, accurate $M_r$ estimations could not be determined for this broad range of species. The large range of species present in low ionic conditions was complementary to the size exclusion chromatography results performed in low ionic conditions (Figure 4.1).

4.4.2.2 Sedimentation Equilibrium

Sedimentation equilibrium studies were carried out to obtain a shape-independent, average molecular weights of xNAP oligomers (section 2.11.2). Equilibrium data were collected of xNAP at three protein concentrations; 0.5, 0.25 and 0.125 mg/mL. Five replicate scans were obtained for each measurement to produce one weighted average curve. Measurements were taken at 3,000, 10,000, 20,000 and 30,000 rpm. Initially the rotor was accelerated to 3,000 rpm and scans were obtained at 18, 21, 24 and 27 hr to verify equilibrium had been reached. The rotor was then accelerated to the next speed and the process was repeated. Sedimentation equilibrium experiments were carried out on xNAP at high (1 M NaCl) and low (20 mM NaCl) ionic conditions, similar to those examined by sedimentation velocity and size exclusion for comparative analysis. Individual analysis of the equilibrium curves was performed with a single species model using ORIGIN software (Beckman Coulter).
Figure 4.8 - Sedimentation equilibrium of xNAP. Overlays of the final equilibrium scans are shown of xNAP in A) 1 M NaCl and B) 20 mM NaCl. Analysis was carried out at three protein concentrations; 0.5, 0.25 and 0.125 mg/mL, and these sectors are seen in the plots from left to right respectively. Scans at rotor speeds; 3,000 (dark blue), 10,000 (red), 20,000 (orange) and 30,000 (light blue) rpm are shown.

Table 4.2 - Sedimentation equilibrium analyses of xNAP using a single species model. The M_r obtained from individual analysis of equilibrium curves of xNAP in A) 1 M NaCl and B) 20 mM NaCl are shown with kDa units (blue boxes). The variance of the fits to the experimental data is shown in brackets. Global fits of the curves at each concentration are shown in the bottom row (purple boxes). Global fits of the curves at each speed are shown in the right hand column (purple boxes). The bottom right-hand box shows the global analysis of all nine curves.

The overlaying menisci in the final scans at each speed indicate that none of the experimental sectors had leaked (Figure 4.8); this meant that the concentration within the cells remained constant throughout the analysis. Comparison of the scans taken at 18, 21, 24 and 27 hrs overlaid with one another suggesting that the curves had reached equilibrium (data not shown). The calculated M_r from data analysis ranged from 63.9 – 1334 kDa for xNAP in 1 M NaCl and 33.8 – 1446 kDa for xNAP in 20 mM NaCl (Table 4.2). This enormous range of M_r observed as well as the clear dependence of M_r upon rotor speed suggested that multiple species were present in the sample and led to further investigations.
A large difference in absorbance was seen in the scans of xNAP in 20 mM NaCl between the 3,000 rpm and 10,000 rpm data (Figure 4.8B). This large absorbance change was not however seen in the analysis of NAP in 1 M NaCl (Figure 4.8A). To examine this different sedimentation of the samples, the equilibrium curves were integrated between the radial points 5.92 – 6.11 cm for each rotor speed. The percentage of protein left in solution at each rotor speed was quantified by normalising to the 3,000 rpm value, prior to any visible sedimentation.

![Figure 4.9 - Plots showing the relationship between rotor speed, calculated molecular weight and percentage protein in solution. The plots shown were derived from sedimentation equilibrium data of xNAP at 0.5 mg/mL. A) The global fits from sedimentation equilibrium data of xNAP in high and low NaCl concentrations have been plotted against the rotor speed at which the experiment was carried out. A trend line has been added to help visualise the data points. B) The estimated percentage of protein remaining in solution is shown for xNAP in high and low ionic conditions.](image)

In high and low ionic conditions the M_r gained from equilibrium analysis was heavily dependent upon the rotor speed at which the data was obtained (Figure 4.9A). The resultant M_r obtained for the 1 M and 20 mM NaCl conditions, both followed the same trend of reducing with increasing rotor speed. However, integration of the equilibrium curves revealed that the amount of protein contributing to each analysis varied significantly between the two NaCl conditions.

In 20 mM NaCl, 70% of the protein had sedimented before the 10,000 rpm scan was taken, whereas for the 1 M NaCl data only 6% had been removed from solution (Figure 4.9B). Therefore, to obtain a representative M_r for xNAP in low ionic conditions only the 3,000 rpm curves could be considered. A further 40% of the protein in high ionic conditions had sedimented by 20,000 rpm and another 40% by 30,000 rpm, thus indicating a heterogeneous sample dominated by the species seen at 10,000 and 20,000 rpm. As the variance values suggested a good fit for the global analysis of xNAP in 1 M NaCl at 10,000 and 20,000 rpm, the M_r gained from these were also considered.
Figure 4.10 - xNAP sedimentation equilibrium analysis curves. Sedimentation equilibrium curves and fits to a single species model are shown with absorbance versus radial position, at high, medium and low protein concentrations. A) The curves from a global fit analysis are shown of xNAP in 1 M NaCl conditions with a rotor speed of 20,000 rpm. B) The individual fits of NAP in 20 mM NaCl with a rotor speed of 3,000 rpm are shown. Above each plot is the residual error between the fitted and experimental values as well as the apparent Mw from each analysis with the residual variance shown in red.

The average Mr estimate obtained from data collected for xNAP in 1 M NaCl conditions at 20,000 rpm was 83 kDa, which corresponds closely to a xNAP dimer (NAP₂) (Figure 4.10A). A larger average Mr was obtained at 10,000 rpm of 145 kDa, indicating that larger species were also present at the lower speeds. These Mr values could therefore not be relied upon as they are the product of a single species fit analyses when clearly multiple species were present. The three individual analyses of xNAP in low ionic strength at 3,000 rpm produced three significantly different Mr estimates and variance values (Figure 4.10B). These results indicated that multiple species are present, once again rendering a single species fit unsuitable despite the deceptively good fit to the experimental data.

4.5 Discussion

xNAP forms a stable dimer. Size exclusion chromatography indicated that xNAP oligomerisation occurred in a salt dependant manner with the smallest species (I) dominating in high salt conditions (0.5 – 1 M NaCl). Species I was determined to be a dimer using SEC-
MALLS, sedimentation velocity and DLS with a sedimentation coefficient, \( f/f_0 \) and \( R_u \) of 4.2 S, 1.65 and 46 Å respectively. The observation of a stable xNAP dimer is consistent with the data for the yNAP1 and pfNAPS homologues, which crystallised in a dimeric form (Gill et al.; Park and Luger, 2006a). Solution studies have also shown that even at high ionic strengths the yNAP1 and pfNAPS form stable dimers and do not disassociate into monomers (Gill et al.; Toth et al., 2005). Phylogenetic analyses illustrate that the NAPs from *Plasmodium* and *Saccharomyces* are more divergent than those from *Saccharomyces* and *Xenopus* (Gill et al., 2009). Therefore, given that pfNAPS and yNAP1 have similar protein folds and dimerise via the same \( \alpha \)-helix, it seems likely that xNAP1 also has this conserved protein fold and dimerises in the same fashion. The hydrophobic residues along the dimerisation helix which form the basis of the strong dimerisation are remarkably conserved between *Saccharomyces*, *Plasmodium* and *Xenopus*. These two interacting dimerisation helices can be observed for the yNAP1 structure in orange in Figure 4.11.

\[ \text{xNAP}_2 \text{ further oligomerises in an ionic strength and concentration dependent manner.} \]

Size exclusion chromatography performed in conditions approaching the physiological ionic strength (i.e. 150mM NaCl) revealed a complex self-associating equilibrium of xNAP2. Consistent with solution studies of yNAP1 (Toth et al., 2005), an ionic strength and concentration dependence of the xNAP oligomerisation was observed. Mutation studies revealed that the \( \beta \)-hairpins, seen protruding from the main body of yNAP, are involved in high-order oligomerisation (Park et al., 2008). The sequence within these \( \beta \)-hairpins is conserved between *Saccharomyces* and *Xenopus*. It is therefore possible that xNAP also oligomerises via the \( \beta \)-hairpins, forming a \( \beta \)-sheet with adjacent xNAP dimers (Figure 4.12). The \( \beta \)-sheets formed between adjacent dimers in the asymmetric unit of the yNAP1 crystal structure were used as a basis for modelling this interaction in xNAP.
**xNAP₂ can associate into multiple high-order species.** In physiological ionic strengths multiple species were consistently observed via SEC-MALLS, sedimentation velocity and sedimentation equilibrium. Previously, research groups have attempted to determine Mₐ values for the yNAP1 oligomeric species using sedimentation velocity and equilibrium and have encountered difficulties (Friedeberg et al., 2006; Toth et al., 2005). A systematic analyses of the AUC data was performed, which revealed the limitations of the techniques and highlighted that the deconvolution of the data, for such a complex system, was not possible. Although distinct Mₐ values of individual species could not be obtained, the average Mₐ from equilibrium AUC confirmed the presence of multiple high-order species. SEC-MALLS, which separates the various species prior to light scattering analysis, was successfully employed to analyse the range of NAP oligomers in physiological conditions. (Figure 4.13).
The smallest species observed was again NAP₂ (species I) and species II was calculated to have a Mₘ of 165 kDa consistent with a dimer of dimers (2NAP₂). The larger species (III and IV) were not resolved on the column suggesting a fast equilibrium between the two, perhaps accentuated by the separation bias of the technique. An estimated Mₘ was calculated for species III and IV of 400 and 750 kDa respectively, most probably corresponding to putative complexes of 5-NAP₂ and 10-NAP₂ (Figure 4.13) with theoretical Mₘ values of 405 and 810 kDa respectively.
5. Hydrodynamic Analysis of xNAP in Complex with Histones

Following the successful characterisation of xNAP alone, similar hydrodynamic techniques were used to examine the interactions with its biologically relevant binding partners, the histones. Despite multiple laboratories working on these interactions, much controversy surrounds the binding stoichiometry and affinities for the NAP-Histone complex (as described in section 1.3.4). Previous studies have focused on *Saccharomyces* (y) and *Plasmodium* (pf) NAP interacting with either *Xenopus* (x) or *Saccharomyces* (y) histones. Previous AUC experiments have suggested that one dimer of yNAP associates with one xhistone dimer, forming a species with a sedimentation coefficient of 5.5 S (Toth et al., 2005). In 100 mM KCl conditions Toth *et al.* observed these complexes associating into higher order assemblies estimated to be \((\text{NAP}_2 \cdot \text{H}_2\text{A/H}_2\text{B})_4\) and \((\text{NAP}_2 \cdot \text{H}_3/\text{H}_4)_4\), both with sedimentation coefficients of 16.6 S, although the data contained low signal to noise. No research to date has focused on the interactions between NAP and histones from the higher eukaryotic *Xenopus laevis*.

This chapter describes the interaction between xNAP and the four core histones; H2A, H2B, H3 and H4. It is commonly accepted that H2A/H2B forms a dimer, whereas H3/H4 is thought of as a tetramer \((\text{H}_3/\text{H}_4)_2\). Although the tetramer is observed in high ionic conditions, it is in fact unstable at physiological ionic conditions leading to dissociation into dimers (Karantza *et al.*, 1996). The studies described in this chapter investigate the ionic strength and protein concentration dependent oligomerisation of the xNAP₂ dimer with the xH2A/H2B or the xH3/H4 dimer.

5.1 Size Exclusion Chromatography of xNAP-Histone Complexes

As illustrated in chapter 4, size exclusion chromatography can be used successfully as a comparative analysis tool. Primarily, size exclusion chromatography was employed to determine if the recombinantly produced xNAP and histones formed complexes *in vitro*. The effects of ionic condition and protein concentration upon binding were then examined along with the binding stoichiometry. Mixtures of xNAP₂+H2A/H2B and xNAP₂+H3/H4 were compared with the individual components by analytical size exclusion chromatography on a Superose 6 column. Two ionic conditions were tested; 150 mM NaCl to mimic physiological
ionic strength, and 500 mM NaCl to increase the solubility of the histones. The proteins were mixed at a 1:1 molar ratio of xNAP dimer to histone dimer and dialysed overnight into SE buffer containing either 150 or 500 mM NaCl (section 2.5.1). The individual components were also individually dialysed into these two buffers and run as standards.

The profiles from size exclusion chromatography of the xNAP₂+H2A/H2B and xNAP₂+H3/H4 mixtures in 500 mM NaCl allude to some association between the two components (Figure 5.1 A+B). The change in elution between the xNAP alone and xNAP in complex with the histones was, however, minor. Conversely, near to physiological ionic conditions (150 mM NaCl), the xNAP₂+H2A/H2B and xNAP₂+H3/H4 mixtures resulted in an elution peak at 11.5 mL, distinct from those of the individual components (Figure 5.1 C+D). These novel elution peaks provide evidence that complex formation occurs between xNAP₂ and the histones dimers, resulting in species with relative large Rₚₚ values. It was also noted that the main xNAP₂+H3/H4 peak eluted slightly before the main xNAP₂+H2A/H2B peak at 150mM NaCl conditions, indicating that this, the former, had a slightly larger Rₚₚ.

A clear salt dependence was also observed for the individual components. The elution profiles for xNAP alone, in high and low ionic conditions, were consistent with the findings in section
4.2.2. H2A/H2B had a slightly later elution in the lower ionic strength buffer, possibly due to some interaction with the column matrix or filter. A larger shift of the elution peak was seen for H3/H4 - from 17 to 18 mL in high and low salt respectively (Figure 5.1B+D). In high ionic strength buffer the (H3/H4)$_2$ tetramer is stable. In contrast, at lower ionic strength the tetramer dissociates into a H3/H4 heterodimer and therefore elutes later from the column. Some interaction of the unstable H3/H4 dimer with the column or filter is also possible.

5.1.1 Stoichiometry of xNAP-Histone Complexes

To determine the binding stoichiometries of the xNAP-Histone complexes, three ratios were analysed by size exclusion chromatography, including all those published in the literature. The xNAP$_2$ dimer was added to the H2A/H2B dimer or the H3/H4 dimer at molar ratios of 2:1, 1:1 or 1:2, samples were then dialysed into 150 mM NaCl buffer overnight and run individually on a Superose 6 column.

![Figure 5.2 - Size exclusion chromatography of different stoichiometric mixtures of xNAP$_2$+H2A/H2B and xNAP$_2$+H3/H4.](image)

Figure 5.2 - Size exclusion chromatography of different stoichiometric mixtures of xNAP$_2$+H2A/H2B and xNAP$_2$+H3/H4. Elution profiles from a Superose 6 column are shown of three ratios of xNAP$_2$+H2A/H2B and xNAP$_2$+H3/H4 complexes; 2:1 (A+D), 1:1 (B+E) and 1:2 (C+F). The xNAP$_2$ input was at a concentration of 7.5 μM for each sample and the H2A/H2B and H3/H4 concentrations varied accordingly. The red stars correspond to fractions examined by SDS PAGE in Figure 5.4.
For all three stoichiometries examined for xNAP₂⁺H₂A/H₂B, the characteristic elution peak at 11.5 mL was observed in the elution profiles (Figure 5.2 A-C). A shoulder to the right of the main peak was observed in the 2:1 ratio profile (Figure 5.2 A), this was shown to be unbound xNAP by SDS PAGE analysis of the corresponding fraction. This shoulder was replaced by a more defined peak in the 1:1 and 1:2 ratios (Figure 5.2 B+C), which by SDS PAGE examination was shown to be another xNAP-Histone complex.

Similarly for xNAP₂⁺H₃/H₄, a main peak with 11.5 mL elution was observed for all of the ratios examined (Figure 5.2 D-E). Unbound xNAP was again observed in the 2:1 ratio profile as a shoulder on the right-hand side of the main peak (Figure 5.2 D). No elution peaks were observed corresponding to unbound histones in any of the elution profiles for the xNAP₂⁺H₂A/H₂B or xNAP₂⁺H₃/H₄ mixtures.

To determine why excess histones were not eluting from the column, assays were performed to identify if the histones were sticking to the filter at the top of the Superose 6 column, but the results indicated that this was not the case (data not shown). Another hypothesis was that the histones were sticking to the column matrix itself, retarding their elution. To investigate this theory, xNAP was injected onto the superose 6 column directly after each of the 1:1 and 1:2 ratio xNAP-Histone mixtures (described in Figure 5.2), with an aim to disrupt any interaction between the histones and column matrix, the eluates were then examined by SDS PAGE.
Figure 5.3 - Size exclusion chromatography of xNAP scavenging for histones. Elution profiles of xNAP from a superose 6 column are shown, when run directly after (A) 1 xNAP₂ : 2 H2A/H2B and (B) 1 xNAP₂ : 2 H3/H4. Selected fractions denoted by light blue numbers were examined by SDS PAGE, and a standard marker (M) is shown in the left hand lane. The positions at which the NAP and the histones migrate are highlighted by red arrows.

When the xNAP-Histone mixture of ratio of 1:2, was run down the Superose 6 column, the excess histones interacted with the column matrix and did not elute. When xNAP alone was subsequently loaded onto the column, it successfully disrupted the interactions between the histone and column, resulting in the histones elution (Figure 5.3). In contrast, no histones remained on the column after the 1:1 stoichiometric ratios were applied. Unbound histones are likely to be prone to non-specific interactions due to their highly charged and hydrophobic surfaces.

The binding stoichiometry for both xNAP-Histone complexes was likely to be a 1:1 ratio since no excess xNAP was observed in the elution profile, nor were excess histones found bound to the column post elution. To further clarify the stoichiometry, the individual dimers; xNAP₂, H2A/H2B and H3/H4 were examined by SDS PAGE with precise loadings of 60, 30 and 10 pmol. Densitometry was then carried out upon these bands to determine the staining efficiency with Coomassie Blue for each protein. The xNAP-Histone complexes were made by adding the individual components in a 1:1 stoichiometry followed by a size exclusion chromatography step to remove any excess components not involved in complexes. The fraction with the peak
absorbance (Figure 5.2, red stars) was then examined by SDS PAGE and quantified for comparison.

Figure 5.4 - Stoichiometry determination of the xNAP·Histone complexes by SDS PAGE and densitometry. 60, 30 and 10 pmol of (A) xNAP₂, H2A/H2B and 1 xNAP₂ : 1 H2A/H2B or (B) xNAP₂, H3/H4 and 1 xNAP₂ : 1 H3/H4 were examined by SDS PAGE alongside a standard protein marker (M). Densitometry of the bands corresponding to xNAP, H2A/H2B, H3 and H4 was performed, and the total integrated intensity is shown above each band in red. An example of a densitometry trace is shown in Appendix VII.

H2A and H2B have the same electrophoretic mobility therefore the band corresponding to H2A and H2B can be directly compared to the band corresponding to xNAP. When equal molar loadings are measured, the band that corresponds to xNAP (on average) has 1.14 times greater intensity than the H2A/H2B band (Figure 5.4 A). By normalising for this difference in staining efficiency, the ratio of xNAP₂:H2A/H2B was shown to be 0.95:1. This was strong evidence that the stoichiometry of binding was one xNAP dimer to one H2A/H2B heterodimer (xNAP₂·H2A/H2B).

The histones H3 and H4 have different electrophoretic mobility therefore the summation of the two bands can be compared to the band corresponding to xNAP. By comparing equal molar loadings of the xNAP₂ dimer and H3/H4 dimer, the average staining efficiencies were calculated to be 2.18:1.13:1 for xNAP:H3:H4 (Figure 5.4 B). By normalising for staining efficiency, the stoichiometry observed for xNAP:H3:H4 was 2.03:1.08:1, strongly suggesting that one xNAP dimer binds one H3/H4 dimer.

After determining this stoichiometry of binding, all of the subsequent experiments characterising the xNAP₂·H3/H4 and xNAP₂·H2A/H2B complexes were carried out by mixing the xNAP₂ with H2A/H2B or H3/H4 in a 1:1 ratio followed by dialysis into buffer containing 150 mM NaCl. Experiments in Chapter 4 showed that xNAP₂ oligomerisation has a dependence upon protein concentration (Figure 4.2). The concentration dependence of the xNAP₂·H2A/H2B and xNAP₂·H3/H4 complexes was, therefore, also examined in the range 0.5 - 2 mg/mL.
Figure 5.5 - Size exclusion chromatography of the xNAP₂·H2A/H2B and xNAP₂·H3/H4 complexes at various concentrations. Elution profiles of the (A) xNAP₂·H2A/H2B and (B) xNAP₂·H3/H4 complexes from a Superose 6 column are shown. The overlay compares the elution profiles of the complex at 0.5, 1, 1.5 and 2 mg/mL injection concentration in orange, green, purple and red respectively.

Overlaying the four elution profiles of the xNAP₂·H2A/H2B complex (Figure 5.5A) and of the xNAP₂·H3/H4 complex (Figure 5.5B) revealed no great dependence upon concentration between 1-2 mg/mL, with a predominant elution peak at 11.5 mL in each case. However, when the protein concentration was 0.5 mg/mL, the elution peak for the xNAP₂·H2A/H2B complex appeared to be split, suggesting a concentration dependent dissociation. Aggregation eluting at 7.5 mL, was observed for all eight profiles, it was approximately proportional to the protein concentration.
5.2 Light Scattering

After determining that xNAP₂ formed complexes with the H2A/H2B and H3/H4 dimers, characterisation of these complexes was initiated. The assays involving the protein concentration series, described in Figure 5.5, suggested the possibility that the predominant elution peak for the xNAP-Histone complex comprised of more than one oligomeric species. Dynamic light scattering was therefore employed to confirm if the xNAP-Histone complexes were made up of multiple species. Size exclusion chromatography with multi-angle laser light scattering (SEC MALLS) was utilised to determine the molecular weight range of species present, as they eluted from a Superose 6 column.

5.2.1 Dynamic Light Scattering

The xNAP₂·H2A/H2B and xNAP₂·H3/H4 complexes were analysed by DLS following size exclusion chromatography (section 2.9). The samples were injected onto the column at a concentration of 1 mg/mL. Peak fractions were collected and centrifuged at 16,100g for 15 min to remove any dust particulates and aggregates. The supernatant was injected into a quartz cell and the light scattering was measured for 10 min at 16°C.

The DLS of the xNAP₂·H2A/H2B and xNAP₂·H3/H4 complexes resulted in hydrodynamic radii (R_H) of 14 and 15 nm respectively (Figure 5.6). The high polydispersity of the samples, along with the broad range of R_H values obtained, suggest that the xNAP₂·H2A/H2B and xNAP₂·H3/H4 complexes form a heterogeneous mixture of oligomeric species.
5.2.2 SEC MALLS

SEC MALLS was performed on the xNAP·Histone complexes to ascertain the $M_r$ range of species present in the samples (section 2.10). The samples were prepared to a concentration of 1 mg/mL and separated using a Superose 6 column equilibrated in 150 mM NaCl buffer. Light scattering readings were taken at 11 different angles along with simultaneous refractive index (RI) measurements to follow the elution of complexes from the column.

Figure 5.7 - Size exclusion chromatography with MALLS of the xNAP-Histone complexes in 150 mM NaCl. A and B are elution profiles of xNAP₂·H2A/H2B and xNAP₂·H3/H4 respectively from a Superose 6 column, with refractive index (RI) in red and light scattering (LS) shown in blue. The $M_r$ determination curves are plotted in black and correspond to the y-axis. C and D show the region of the plot where the complexes eluted, with green arrows as a guide to determine the minima, maxima and average $M_r$ obtained.
The SEC MALLS traces for both the xNAP₂·H2A/H2B and xNAP₂·H3/H4 complexes were consistent with the DLS and AUC (section 5.3), showing that the species were not homogeneous. The profile for the xNAP₂·H2A/H2B (Figure 5.7 A+C) illustrated that the species present were in the range of 0.5 – 1 MDa. The profile for the xNAP₂·H3/H4 complex indicated a Mr range of 0.58 - 1.2 MDa (Figure 5.7 B+D). Given that the RI measures the total concentration of eluting species whereas LS is weighted for molecular mass, the misalignment of the two peaks, seen for both complexes, is another indication that the sample is polydisperse. The Mr range observed by SEC-MALLS is indicative of the formation of complex supramolecular assemblies encompassing 5-12 subunits of the stoichiometric building blocks (xNAP₂·H2A/H2B or xNAP₂·H3/H4).

5.3 Analytical Ultracentrifugation

5.3.1 Sedimentation Velocity

Sedimentation velocity experiments were performed with the xNAP·Histone complexes to determine their sedimentation coefficients and heterogeneity (section 2.11.1). Measurements were obtained for the complexes at two protein concentrations: 0.8 and 0.2 mg/mL and examined in 150 mM NaCl. The velocity scans were measured at 15,000 rpm and the best fit to the Lamm equation was calculated using a continuous c(s) distribution model in SEDFIT.
**xNAP₂·H2A/H2B**

### Figure 5.8

Sedimentation velocity of xNAP₂·H2A/H2B. The velocity scans, fit lines and residuals are shown for the sedimentation velocity scans for xNAP₂·H2A/H2B at 15,000 rpm for A) 0.8 mg/mL and B) 0.2 mg/mL. The frictional ratio and RMSD of fit are also shown. C) An overlay of the C(s) plots from the two protein concentrations measured.
xNAP$_2$·H3/H4

Figure 5.9 - Sedimentation velocity of xNAP$_2$·H3/H4. The velocity scans, fit lines and residuals are shown for the sedimentation velocity scans for xNAP$_2$·H3/H4 at 15,000 rpm for A) 0.8 mg/mL and B) 0.2 mg/mL. The frictional ratio and RMSD of fit are also detailed. C) An overlay of the C(s) plots from the two protein concentrations measured.
Table 5.1 - Sedimentation velocity of the xNAP-Histone complexes. The frictional ratio \((f/f_0)\), sedimentation coefficients \((s)\) and estimated \(M_r\) for the xNAP-Histone complexes are shown at two concentrations.

<table>
<thead>
<tr>
<th>Sample</th>
<th>(f/f_0)</th>
<th>(s)</th>
<th>(M_r) (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>xNAP₂·H2A/H2B</td>
<td>2.29</td>
<td>14.7</td>
<td>912</td>
</tr>
<tr>
<td>0.8 mg/mL</td>
<td>2.41</td>
<td>14.7</td>
<td>984</td>
</tr>
<tr>
<td>0.2 mg/mL</td>
<td>2.41</td>
<td>10.14.6</td>
<td>585,992</td>
</tr>
<tr>
<td>xNAP₂·H3/H4</td>
<td>2.41</td>
<td>9.4,14.9</td>
<td>605,1060</td>
</tr>
</tbody>
</table>

The analysis of xNAP₂·H2A/H2B at both concentrations resulted in low RMSD values of 0.005 and 0.004 indicating good fits to the experimental data (Figure 5.8 A+B). The frictional ratios obtained from these analyses were 2.29 and 2.41 for the high and low protein concentrations respectively, indicating that the species present were significantly non-spherical in shape. The \(c(s)\) plots for the two concentrations contained broad peaks suggesting that multiple species may be present (Figure 5.8 C). A very small peak was also observed at 6 S for the high concentration sample. To obtain \(M_r\) values, a continuous \(c(M)\) distribution model was used which utilises the average \(f/f_0\) to calculate \(M_r\). The estimated \(M_r\) for the NAP₂·H2A/H2B samples were 912 and 984 kDa (Table 5.1). This corresponded to between 9 and 10 of the xNAP₂·H2A/H2B stoichiometric units.

Analysis of xNAP₂·H3/H4 also resulted in low RMSD values of 0.005 and 0.004 (Figure 5.9 A+B). The frictional ratios obtained were 2.41 and 2.42 for the two concentrations measured suggesting that the xNAP₂·H3/H4 complex was also distinctly non-spherical in shape. The \(c(s)\) plots reveal two main peaks at 9.6 and 14S (Figure 5.9 C). \(M_r\) values, obtained from analysis using a continuous \(c(M)\) distribution model were 605 and 1060 kDa. These were, however, again based on a single \(f/f_0\) and were therefore only an estimate (Table 5.1).

5.3.2 Sedimentation Equilibrium

Given that size exclusion chromatography and sedimentation velocity indicated the presence of multiple species, sedimentation equilibrium studies were performed to obtain shape independent molecular weights (section 2.11.2). Data were collected on both xNAP-Histone complexes at three concentrations and three speeds so global fit analyses could be performed. The scans were analysed using ORIGIN, using a single species model, first individually and then globally to obtain the average \(M_r\) and variance for each.
5.3.2.1 xNAP₂·H2A/H2B

Figure 5.10 - Sedimentation equilibrium of xNAP₂·H2A/H2B. Analysis at three protein concentrations; 0.8, 0.4 and 0.2 mg/mL was performed, shown from left to right respectively. An overlay of final equilibrium scans are shown from rotor speeds of 3,000 (orange), 5,000 (red) and 7,000 (blue) rpm.

Figure 5.11 - Sedimentation equilibrium global analysis of xNAP₂·H2A/H2B using a single species model. The nine curves covering three speeds and three concentrations are shown, fitted to a single species model. The experimental data points of absorbance at 276 nm are shown as blue circles with the fitted lines shown in red. Above each curve is a plot of the residuals showing the deviation of the fit from the experimental data.
Table 5.2 - Sedimentation equilibrium analyses of xNAP₂-H2A/H2B using a single species model. The Mr, obtained from individual analysis of equilibrium curves are shown with kDa units (blue boxes). The variance of the fits to the experimental data is shown in brackets. Global fits of the curves at each concentration are shown in the bottom row (purple boxes). Global fits of the curves at each speed are shown in the right hand column (purple boxes). The bottom right-hand box shows the global analysis of all nine curves.

By overlaying the last scans at each speed, it was clear from the overlapping menisci that none of the sectors had leaked (Figure 5.10). Individual and global analyses of the xNAP₂-H2A/H2B equilibrium curves using a single species model, resulted in a Mr range of 921 to 1261 kDa (Table 5.2). The global analysis, including all nine curves, resulted in a Mr of 1072 kDa, with good fits to the experimental data (Figure 5.11).

5.3.2.1 xNAP₂-H3/H4

![Sedimentation equilibrium absorbance scans of xNAP₂-H3/H4 at three protein concentrations; 0.8, 0.4 and 0.2 mg/mL, from left to right respectively. An overlay of final equilibrium scans are shown from rotor speeds of 3,000 (orange), 5,000 (red) and 7,000 (blue) rpm.](image)
Figure 5.13 - Sedimentation equilibrium global analysis of xNAP₂-H3/H4 using a single species model. The nine curves covering three speeds and three concentrations are shown with a fit line to a single species model. The experimental data points of absorbance at 276 nm are shown as blue circles with the fitted lines shown in red. Above each curve is a plot of the residuals showing the deviation of the fit from the experimental data.

Table 5.3 - Sedimentation equilibrium analyses of xNAP₂-H3/H4 data using a single species model. The $M_r$ obtained from individual analysis of equilibrium curves are shown with kDa units (blue boxes). The variance of the fits to the experimental data is shown in brackets. Global fits of the curves at each concentration are shown in the bottom row (purple boxes). Global fits of the curves at each speed are shown in the right hand column (purple boxes). The bottom right-hand box shows the global analysis of all nine curves.

Analysis of the final scans at each speed, suggested that none of the sectors had leaked. However, at the radial position 5.95 cm, a dip in absorbance was observed at 3,000 and 5,000
rpm (Figure 5.12). This was probably due to a dust particle on the cell window. The data points within the dip, between 5.94 and 5.96 cm, were removed prior to analysis so as not to bias the results. Individual and global fits of the curves to a single species model resulted in a M, range of 939 – 1219 kDa (Table 5.3). When all nine curves were included in the analysis the average M, obtained was 1093 kDa (Figure 5.13).

5.4 Discussion

xNAP forms discrete complexes with core histones. Using size exclusion chromatography and SDS PAGE it was demonstrated that at physiological ionic strength xNAP forms large mega-Dalton sized complexes with both the H2A/H2B and H3/H4 dimers. These interactions were shown to be dependent upon ionic strength, consistent with previous data which demonstrated that histones bind to yNAP through electrostatic interactions at the concave surface of domain I (Figure 5.14) (Park and Luger, 2006a).

One xNAP dimer binds one histone fold dimer. Given the controversy surrounding the stoichiometry of NAP-Histone complexes, a systematic analysis was carried out using size exclusion chromatography and quantitative SDS PAGE analysis. Problems arose when the basic nature of the histones led to their non-specific interaction with the theoretically inert column matrix. Once this interaction had been accounted for, the stoichiometry was successfully determined. In the conditions examined, one xNAP dimer bound one histone dimer, forming complexes with the stoichiometric sub-units, xNAP₂·H2A/H2B and xNAP₂·H3/H4. These stoichiometries are consistent with the ‘prevailing opinion’ as judged by Zlatanova et al., 2007 in a recent review discussing all NAP1 homologues, including the findings of Toth et al., 2005 and McBryant et al., 2003. It is probable that the histone dimer binds to the acidic concave region of xNAP₂ (Figure 1.10) which would not disrupt the xNAP₂ oligomerisation via the β-hairpins (Figure 5.14).

The stoichiometric unit oligomerises into larger assemblies. Sedimentation equilibrium suggested that xNAP₂·H2A/H2B and xNAP₂·H3/H4 had M, of 1072 and 1090 kDa respectively, corresponding to ten stoichiometric units. Further characterisation of these complexes was accomplished by employing non-equilibrium techniques, such as sedimentation velocity and SEC-MALLS. Both of these techniques were consistent in highlighting the presence of smaller complexes than the decamer. The minimum M, obtained by SEC-MALLS for the xNAP₂·H2A/H2B complex was 500 kDa corresponding to a pentamer. These data might suggest
that a pentamer and decamer of xNAP₂-H2A/H2B are present, with the equilibrium shifted towards the decamer. A possible model for this oligomerisation could involve a pentameric ring which self-associates into a double (decameric) ring, satisfying the Mₚ range observed by SEC-MALLS (Figure 5.14).

Figure 5.14 - Hypothesis of xNAP-Histone oligomerisation. This schematic diagram illustrates the histone dimer binding to the xNAP dimer in the acidic concave region of xNAP. This stoichiometric unit can form a pentameric ring which, in turn, can self-associate into a double ringed decamer.

A similar double-ring model could be appropriate to describe the xNAP₂-H3/H4 oligomerisation. The Mₚ range determined from SEC-MALLS is slightly larger, with a Mₚ range of 0.6 – 1.2 MDa, roughly corresponding to 6 - 12 sub-units. The formation of a hexameric ring self-associating into a dodecamer is, therefore, also plausible.
6. Small Angle X-Ray and Neutron Scattering

Small angle scattering can be used to obtain low resolution structural information of molecular assemblies in solution. This is particularly useful when NMR structures cannot be obtained due to size restraints and diffracting crystals are not available. In addition, small angle X-ray scattering (SAXS) and small angle neutron scattering (SANS) are often used to complement high resolution crystallography studies. These techniques can also be used to determine oligomerisation, aggregation, and flexibility of complexes in their native solution state.

In a recent example, SAXS studies were used to characterise the histone chaperone, nucleoplasmin (Taneva et al., 2009). The technique was used to verify that the nucleoplasmin pentamer observed in the crystal lattice (PDB code 2VTX, Taneva et al., 2008) was also present in solution. Taneva et al. then used low resolution envelopes, obtained from SAXS, to create models of the Nucleoplasmin-Histone complex. This was performed using the previously published crystal structures of the individual proteins. Thus far there are no solution scattering data from any of the NAP protein family, despite their increasingly defined role as key chaperones. Building on the hydrodynamic studies in Chapter 5, solution scattering measurements were performed to verify the oligomerisation state of NAP in various ionic conditions. The xNAP-Histone complexes were also examined by SAXS and SANS to determine the size and shape of the assemblies.

6.1 Theory

Molecules in solution are present in random orientations, and therefore the diffraction data obtained represents a radial average of all the orientations. Analysing scattering at small diffraction angles yields low resolution parameters such as the radius of gyration ($R_g$) and maximum dimension ($D_{max}$), as well as molecular weight ($M_w$). To collect these data, specialist instrumentation is required. This includes highly collimated beams, long sample-detector distances (up to 20 m in SANS) and a highly sensitive detectors capable of detecting weak scattering signals. A steady beam of X-rays is required for SAXS experiments, which can be produced by a rotating anode source, or more effectively, by a synchrotron. Synchrotron facilities, such as the Diamond Light Source (Diamond) and the European Synchrotron Radiation Facility (ESRF) accelerate electrons to high energies. These electrons are constrained in a circular storage ring using magnetic fields, causing them to radiate energy including X-rays.
The X-rays are then focused into a beam which is directed at the sample (Figure 6.1A). The X-rays are subsequently scattered by the atomic electrons in the sample (Figure 6.2).

The beam of neutrons necessary for SANS experiments is produced by the fission of $^{235}$Uranium. High flux nuclear reactors, such as the one at the Institut Laue Langevin (ILL), are used to produce fast, high energy neutrons. The neutrons are guided into a beam directed at the sample (Figure 6.1B), and subsequently scattered by atomic nuclei (Figure 6.2). Neutrons are emitted in a random nuclear process; therefore, time-averaged scattering measurements are necessary to cancel out the random flow of neutrons. The contrast ($\Delta \rho$) observed in solution scattering is defined by the difference between the average scattering density of the macromolecule ($\rho_v$) and the solvent ($\rho_s$).

![Figure 6.1 - Schematic Diagram of a SAXS and a SANS beamline. A) A typical SAXS beamline is shown. The X-ray beam from the synchrotron is focused using a torroidal mirror and collimator and the wavelength is selected using a monochromator. The beam size is determined by the focus of the mirror and further defined by horizontal and vertical slits. The scattered X-rays are measured by a gas detector, for example, Rapid2 (Berry et al., 2003) which is approximately 2-3 m from the sample. B) Representation of the SANS beamline D22 at the ILL (Figure from www.ill.eu). Neutrons of a specified velocity range pass through a velocity selector into a series of guides which are used to collimate the beam. The beam dimensions are finally defined by an exchangeable diaphragm. The scattered neutrons are recorded by a 64 x 64 element area detector which can be adjusted from 2-20 m from the sample to vary the measurable q range.](image)

The use of SAXS has increased exponentially since the development of the suite of programs (ATSAS) used in data analysis, created by Prof. Dimitri Svergun and colleagues. This high
amount of interest is reflected in the conversion of MX beamlines into dedicated SAXS beamlines (e.g. ID14-3, ESRF).

Biological molecules such as DNA, lipids and proteins have significantly different neutron scattering length densities (Kneale et al., 1977). Utilising this scattering property has led to some very influential match-out experiments, such as determining the relative DNA-protein arrangement within the nucleosome (Hjelm et al., 1977). There is also a large difference in scattering length densities between hydrogen and deuterium in SANS. This difference in scattering allows contrast match experiments to be performed of protein-protein complexes if one of the subunits is selectively deuterated (described in Figure 6.3). Using a series of partially deuterated complexes another significant match-out experiment led to the first model of the 30 S ribosomal complex which included the positions of all 21 proteins (Capel et al., 1987).

![Figure 6.2 - Scattering lengths for SAXS and SANS. Relative scattering lengths are shown for hydrogen, deuterium, boron, carbon, oxygen, aluminium and silicon. Scattering lengths of X-rays are shown in red and neutron scattering in blue. The diameters of the spheres are roughly proportional to the scattering length; hydrogen has a light blue sphere to denote that the scattering length is negative.](image-url)
Figure 6.3 - Theory of a SANS contrast match experiment. A) The average scattering densities for H₂O and D₂O are shown, calculated by summation of all atomic scattering lengths in the molecule and dividing by the molecular volume (Perkins, 1988). B) Typical average scattering densities for deuterated and protonated proteins are shown. C) Schematic diagram of a partially deuterated protein complex in H₂O, 40% D₂O (60% H₂O) and 100% D₂O. Contrast can be seen for the whole complex when in H₂O; however, in 40% D₂O there is only contrast with the deuterated protein subunit and therefore the protonated protein will not contribute to the buffer subtracted scattering curve. Conversely in 100% D₂O, contrast is only seen between the solvent and the protonated subunits. By comparing the scattering curves for the partially deuterated complex in these three solvents, information about the relative subunit arrangement can be derived.

Two-dimensional solution scattering data can be radially averaged and the intensity (I) plotted as a function of the vector q (Figure 6.4A). The scattering vector q is related to the scattering angle (2θ) and is defined as \( 4\pi\sin\theta/\lambda \) (Perkins, 1988). The contribution to scattering from the solvent can be subtracted using the program PRIMUS (P.V.Konarev, 2003). At low q values, the plot of \( \ln I \) versus \( q^2 \) relates to the Guinier approximation;

\[
\ln I = \ln I_0 - RG^2q^2/3
\]

where \( R_G \) is the radius of gyration. Therefore at low q values \((qR_G \leq 1.3)\) a straight line can be plotted with a gradient directly related to the \( R_G \) (Figure 6.4B) (Guinier, 1955). The scattering intensity extrapolated to zero angle \((I_0)\), is proportional to the \( M_r \). The \( M_r \) can be derived from the ratio of \( I_0 \) to the incident beam. Alternatively the \( I_0 \) can be converted to \( M_r \) by using a well characterised protein standard such as bovine serum albumin (BSA), with the following equation:

\[
\frac{I_0(\text{BSA})}{M_r(\text{BSA}) \times \text{conc (BSA)}} = \frac{I_0(\text{Sample})}{M_r(\text{Sample}) \times \text{conc (Sample)}}
\]
Figure 6.4 - SAXS of BSA (ID14-3, ESRF). The scattering was measured to use as a standard at a concentration of 6.3 mg/mL. A) The radially averaged scattering curve following buffer subtraction. B) Guinier plot with experimental data in blue and the Guinier region defined by the red lines. The residuals in green show the deviation of the Guinier line from the experimental data. (C) The distance distribution plot p(r). (D) The back transform of the p(r) plot (red line), overlaid with the experimental data (blue circles). The corresponding $R_g$, $D_{max}$ and $I_0$ are also shown.

Fourier transform of the curve $I(q)$ results in a $p(r)$ plot, from which the particle dimensions in real space can be derived. This distance distribution function, $p(r)$, can be plotted using the program GNOM (Svergun, 1992) (Figure 6.4C). The maximum dimension ($D_{max}$) of the particle can be determined via the $p(r)$ x-axis intersection. To obtain an accurate $p(r)$ plot, a fourier transform of the data over the q range from zero to infinity is required. Cropping the curve at $q<\infty$ can therefore result in termination errors. To determine how accurate the $p(r)$ distribution is, it can be back transformed and overlaid with the experimental data to show the goodness of fit (Figure 6.4D) (Perkins, 1988).

To determine if molecules within a sample are fully folded, a Kratky analysis can be performed by plotting $q^2 I(q)$ vs. $q$ (Figure 6.5 A) (Putnam et al., 2007). Theoretically, it is possible to obtain an absolute $M_r$ from a single SAXS scattering curve. The protein volume ($V$) can be obtained by integrating under the curve of a Kratky plot in the range $q = 0$ to $q = \infty$. Practically, scattering data is only obtained within a finite $q$ range. Therefore, the scattering curve can only be integrated between zero and $q_{max}$ (for example 0.3 Å$^{-1}$), leading to an apparent volume ($V'$). Fischer et al., calculated the discrepancy between $V$ and $V'$ for over a thousand protein
standards when different $q_{\text{max}}$ values were used. This provided a range of constants to add to the calculated $V'$ for each $q_{\text{max}}$ used, to best estimate $V$. The $M_r$ is derived by multiplying the calculated $V$ by the average protein density which is approximately 1.37 g/cm$^3$. This form of Kratky analysis can be performed using the program SANS-MoW (H. Fischer, 2010).

![Kratky plot analyses](image)

Figure 6.5 - Kratky plot analyses. Examples of an unfolded, partially folded and folded protein sample compared by Kratky analysis (Figure from Putnam et al., 2007).

Three dimensional ab initio models can be calculated from SAXS and SANS scattering curves using the program DAMMIF (Svergun, 2009). A theoretical sphere with a diameter determined from the scattering data, is filled with dummy atoms. The shape reconstruction is performed by minimising the discrepancy between the experimental scattering curve and the theoretical scattering curve of the model. Penalties are introduced to ensure the dummy atoms are connected and centralised within the volume of the model. If known, symmetry can be imposed, thus improving the data to parameter ratio and can increase the accuracy of the model. DAMAVER (Svergun, 2003) can be used to overlay multiple models in similar orientations, and to plot the volumes onto a grid with occupancy terms. An ab initio model is produced, which encompasses all of the volume from all the input models. DAMFILT is used to remove volume from the grid with low occupancy, thus removing outliers from the analysis and producing an improved model.

6.2 Small Angle Scattering of xNAP

6.2.1 SAXS

Solution scattering was carried out on xNAP alone to determine the size and shape of the oligomers formed (section 2.12). SAXS was performed in 500 and 20 mM NaCl conditions to analyse species I and IV identified in section 4.2.2. Data were also collected from BSA to provide a standard for the conversion of $I_0$ into $M_r$. Data were collected at the ESRF on their
first dedicated biological SAXS beamline, ID14-3, which at the time of collection was still in development. Samples were loaded into an in-vacuum capillary tube, manually aligned with the beam and measured. Ten measurements of one second were obtained over a q range of 0.01 to 0.48 Å⁻¹. The data were radially averaged using BsxCuBE (P Pernot1, 2010) and compared in PRIMUS to check for radiation damage. Curves with no visible damage were averaged and buffer subtracted. A q range of between 0.018 and 0.260 Å⁻¹ was used in the analysis to exclude the parasitic scattering near the beamstop at low q values and the noisy data at high q values.

The xNAP oligomers formed in 20 mM NaCl, proved too large for the experimental setup at ID14-3. The detector was at the maximum possible distance from the sample within the experimental hutch and the useful Guinier region for the sample was still obscured by the beamstop shadow.

The smaller oligomeric species of xNAP formed in high ionic conditions were, however, small enough to be measured within the geometric constraints of the beamline and useful data was

Figure 6.6 - SAXS of xNAP (ID14-3, ESRF). The scattering was measured at a concentration of 0.46 mg/mL in 500 mM NaCl, and the buffer scattering was subtracted. (A) Guinier plot, together with the estimated $R_g$; (B) the $p(r)$ plot and (C) the original scattering data compared to the back-transformed curve, showing the resultant value of $I_0$/Conc.

The smaller oligomeric species of xNAP formed in high ionic conditions were, however, small enough to be measured within the geometric constraints of the beamline and useful data was
successfully collected. No radiation damage was observed over the ten second exposures therefore all ten curves were used in analysis. The Guinier approximation of the scattering curve gave a calculated $R_G$ of $45.2 \pm 1.3 \, \text{Å}$ (Figure 6.6A). The asymmetric $p(r)$ plot with a $D_{\text{max}}$ of $155 \, \text{Å}$ (Figure 6.6B) indicated that the average xNAP conformation was elongated in shape. The back transformation of the $p(r)$ overlaid well to the experimental data, indicating that it was a good approximation for average size and shape of the molecules present (Figure 6.6C).

Due to the low signal to noise of the xNAP sample at high $q$ values, Kratky analysis of the data could not be performed with any degree of accuracy. The scattering data obtained for the BSA standard resulted in an $R_G$ of $30 \, \text{Å}$, $D_{\text{max}}$ of $90 \, \text{Å}$ and an $I_0/\text{Conc}$ of $10228$ (Figure 6.4). The results obtained correspond well to the crystal structure of the highly homologous human serum albumin (HSA) with $R_G$ of $28 \, \text{Å}$ and $D_{\text{max}}$ of $87 \, \text{Å}$ (PDB code 1E78, (Bhattacharya et al., 2000)). The $I_0/\text{Conc}$ for the xNAP sample was $13517$ and by using BSA (with a monomer $M_r$ of $66 \, \text{kDa}$) as a standard the $M_r$ of xNAP was calculated to be $87 \, \text{kDa}$.

### 6.3 Small Angle Scattering of xNAP$_2$·H2A/H2B

#### 6.3.1 SAXS

Small angle X-ray scattering was carried out to investigate the structure of xNAP in complex with H2A/H2B (section 2.12). Data was collected at the Diamond Light Source on beamline I22, covering a $q$ range of $0.005 - 0.548 \, \text{Å}^{-1}$. For improved signal to noise, 180 measurements of one second were collected for each data set. The data were radially averaged using the program DREAM (Mark Malfois, Diamond Light Source) and checked for radiation damage.

![Figure 6.7 - Comparison of SAXS curves to check for radiation-induced damage. A low angle region of the scattering curves, frames 1-10, 20 and 50 are shown overlaid for A) xNAP$_2$·H2A/H2B and B) BSA collected at ID14-3, ESRF.](image-url)

110
By overlaying the individual frames it was evident that the xNAP₂·H2A/H2B complex was resistant to radiation-induced damage and therefore all 180 curves were averaged and used for analysis (Figure 6.7 A). Conversely, BSA began to show radiation-induced damage after ten seconds of exposure, so only the first ten frames of data could be utilised (Figure 6.7 B). The data sets were averaged using EXCEL (Microsoft office), and then the buffers were subtracted from their respective samples using PRIMUS. Data in the q range of 0.01-0.26 Å⁻¹ were used for analysis.

The curved Guinier region of the scattering from the BSA sample was indicative of a sample that was partially aggregated. Small amounts of aggregation cause large effects upon the extrapolated $I_0$ as well as making the calculation of protein concentration less accurate. This BSA sample was therefore not used as a standard for converting $I_0$ to $M_r$.

![Figure 6.8 - SAXS of xNAP₂·H2A/H2B (I22, Diamond). The scattering was measured at a concentration of 1.3 mg/mL in 150 mM NaCl, and the buffer scattering was subtracted. (A) Guinier plot, together with the estimated $R_G$; (B) the $p(r)$ plot, (C) the original scattering data compared to the back-transformed curve, showing the resultant value of $I_0$/Conc and (D) the Kratky plot.](image)

Analysis of the Guinier region of the curve resulted in an $R_G$ of 80.6 ± 1.0 Å with an excellent fit. The q$R_G$ range was, however, above the typical region suggested for the Guinier approximation...
Fourier transformation of the scattering curve with an estimated $D_{\text{max}}$ of 250 Å, produced a $p(r)$ plot which, when back transformed, fitted very well with the experimental data (Figure 6.8 B+C). The $M_r$ values calculated by integrating under the Kratky plot were 1126, 1084 and 1064 kDa using $q_{\text{max}} = 0.15$, 0.20 and 0.25 Å$^{-1}$ respectively (Figure 6.8D). This suggests that ten of the stoichiometric subunits ($xNAP_2 \cdot H2A/H2B$) form a large assembly, which would have a theoretical $M_r$ of 1100 kDa.

6.3.2 SANS

Small angle neutron scattering was carried out on the $xNAP_2 \cdot H2A/H2B$ complex to obtain information regarding the relative subunit arrangement (section 2.13). The plasmid containing the $xNAP$ gene was sent to the D-lab at the ILL to be expressed by bacteria grown in deuterium based media. This expression produced 70% deuterated $xNAP$ protein ($dxNAP$), with a scattering length density close to that of a 100% D$_2$O buffer ($\sim 6.38 \times 10^{-6}$Å$^{-2}$). Using the $dxNAP$ and protonated histones, the $dxNAP_2 \cdot H2A/H2B$ complex was assembled in the usual manner (Section 2.6).

Contrast matching experiments were performed on the partially deuterated complex on the D22 beamline, at the ILL. A detector and collimation distance of 17 m was used to obtain data in the $q$ range of 0.003 – 0.04 Å$^{-1}$. The partially deuterated complex was dialysed over night into three buffers: 0, 40 or 100% D$_2$O and for each contrast the scattering was measured for 1 hr.

![Figure 6.9 - SANS of $dxNAP_2 \cdot H2A/H2B$ (D22, ILL). The scattering was measured at a concentration of 1.0 mg/mL in 150 mM NaCl, and the buffer scattering was subtracted. The Guinier plots for the partially deuterated complex in A) 0% D$_2$O and B) 40% D$_2$O are shown.](image)

The data obtained for the sample in 100% D$_2$O had extremely small signal to noise and therefore no accurate $R_G$ could be obtained. Guinier plots for the sample in 0% and 40% D$_2$O
resulted in $R_g$ values of $89.7 \pm 0.9 \text{ Å}$ and $91.1 \pm 1.9 \text{ Å}$ respectively (Figure 6.9). This indicated that there was no significant difference in $R_g$ when the histones had been matched out of the complex. The low scattering intensity in 100% D$_2$O was a result of the deuterated xNAP being matched out, leaving the relatively small histone proteins to scatter the neutrons. In the stoichiometric unit the histones make up only a quarter of the protein mass, hence the large change in scattering intensity.

### 6.3.3 Modelling

Modelling of the xNAP$_2$·H2A/H2B was performed using the SAXS data collected from I22, Diamond. Scattering data in the $q$ range of 0.01-0.26 Å$^{-1}$ which were previously used to determine a $D_{max}$ of 250 Å and an $R_g$ of 81 Å were used as the constraints for modelling. As the symmetry was unknown ten DAMMIF models were produced for each of three possible symmetries; P1, P5 or P6. The resulting models were averaged and filtered to produce one DAMFILT model for each symmetry examined.

![Figure 6.10 - Ab initio models of xNAP$_2$·H2A/H2B. Three DAMFILT models are shown with P1, P5 or P6 symmetry imposed. Two orientations are shown for each model. The average chi values for the input DAMMIN models are 1.410, 1.723 and 1.872 for P1, P5 and P6 symmetry respectively.](image)

The DAMMIN models were individually analysed to verify that the DAMMIF models produced are true representatives. The models obtained when no symmetry was imposed (P1), as expected, showed the most variation, although they were all oblate. The P5 and P6 DAMMIN
models were self-consistent with the exception of one outlier in the P6 group. The *ab initio* models produced are therefore good visual representations of the average size and shape of the molecules found in solution (Figure 6.10). The three symmetries examined were chosen to provide three possible conformations that the macromolecular assemblies could be. All three models indicate that the average shape of the xNAP₂-H2A/H2B assemblies is oblate.

### 6.4 Small Angle Scattering of xNAP₂-H3/H4

#### 6.4.1 SAXS

Scattering data were also collected at beamline I22 at the Diamond Light Source for the xNAP₂-H3/H4 complex (section 2.12). Scattering was measured for 180 frames of 1 sec each covering a q range of 0.005 - 0.548 Å⁻¹. The scattering data were radially averaged using DREAM. No radiation-induced damage was observed for the xNAP₂-H3/H4 complex so all the 180 curves were averaged in PRIMUS and cropped to a q range of 0.017-0.259 Å⁻¹.

Figure 6.11 - SAXS of xNAP₂-H3/H4 (I22, Diamond). The scattering was measured at a concentration of 0.7 mg/mL in 150 mM NaCl, and the buffer scattering was subtracted. (A) Guinier plot, together with the estimated R_G; (B) the p(r) plot, (C) the original scattering data compared to the back-transformed curve, showing the resultant value of $I_0$/Conc and (D) the Kratky plot.
The xNAP₂-H3/H4 assemblies were of a similar magnitude to the xNAP₂-H2A/H2B and thus the beamline geometry was again, not ideal. A straight line was observed in the Guinier analysis which resulted in an $R_g$ of 82.1 ± 0.8 Å, but the $qR_g$ range was 1.40 – 3.14 (Figure 6.11 A). This $qR_g$ range is larger than is suggested to be valid for the Guinier approximation to hold (maximum $qR_g$< 1.3). Fourier transformation of the scattering data when $D_{max}$ was set to 265 Å led to a fairly symmetric $p(r)$ plot (Figure 6.11 B). Back transformation of the $p(r)$ plot gave an $I(q)$ curve that overlaid well with the experimental scattering curve (Figure 6.11 C). Kratky analysis of the experimental data showed that the proteins in solution were folded (Figure 6.11 D). Integration under the Kratky plot with $q_{max}$ = 0.15, 0.20, 0.25 Å⁻¹ resulted in calculated $M_r$ values of 1361, 1315 and 1290 kDa respectively. All these values corresponded well to 1320 kDa, the theoretical $M_r$ of an assembly containing 12 subunits of xNAP₂-H3/H4.

SAXS data of xNAP₂-H3/H4 were also collected at the ESRF on beamline ID2. Three, 0.1 second exposures covering a $q$ range of 0.005 – 0.332 Å⁻¹ were obtained. The data were collected and radially averaged by the beamline scientist, Dr Shirley Callow. No radiation-induced damage was seen, so the three frames were averaged in PRIMUS and the buffer was subtracted. The curve was cropped to $q$=0.008 - 0.26 Å⁻¹ to use in the analysis.
Figure 6.12 - SAXS of xNAP₂·H3/H4 (ID2, ESRF). The scattering was measured at a concentration of 1.21 mg/mL in 150 mM NaCl, and the buffer scattering was subtracted. (A) Guinier plot, together with the estimated \( R_G \); (B) the \( p(r) \) plot, (C) the original scattering data compared to the back-transformed curve, showing the resultant value of \( I_0/\text{Conc} \) and (D) the Kratky plot.

The improved beamline geometry allowed collection of a Guinier region under \( qR_G = 0.3 \) Å⁻¹ which resulted in a \( R_G \) of 85.1 ± 2.8 Å (Figure 6.12 A). The distance distribution plot with a \( D_{\text{max}} \) of 275 Å led to a good fit to the experimental data when back transformed (Figure 6.12 B+C). A BSA measurement was not obtained at the time of the experiment. However, integration under the Kratky curve was performed with a \( q_{\text{max}} \) of 0.15, 0.20, 0.25 Å⁻¹ which led to calculated \( M_r \) of 1125, 1084 and 1052 kDa respectively. All three \( M_r \) values correspond well to a 10mer of xNAP₂·H3/H4 with theoretical \( M_r \) of 1100 kDa.

### 6.4.2 SANS

Using the 70% deuterated dNAP (as described in section 6.3.2), a dxNAP₂·H3/H4 complex was prepared (section 2.13). Identical contrast matching experiments were performed to those for the dxNAP₂·H2A/H2B complex (section 6.3.2) with solvents containing 0, 40 and 100% D₂O on D22 beamline at the ILL.
Figure 6.13 - SANS of dxNAP₂·H3/H4 (D22, ILL). The scattering was measured at a concentration of 0.5 mg/mL in 150 mM NaCl, and the buffer scattering was subtracted. The Guinier plots for the partially deuterated complex in A) 0% D₂O and B) 40% D₂O are shown.

The data collection of the sample in 100% D₂O buffer had an extremely low intensity coupled with a low signal to noise, and thus no accurate $R_g$ could be determined. The $R_g$ obtained from the Guinier region of the scattering curves in 0% and 40% were 91.0 ± 1.1 and 86.8 ± 2.7 Å, respectively (Figure 6.13), indicating there was no significant difference between the two.

### 6.4.3 Modelling

Modelling of the NAP₂·H3/H4 complex was performed using the SAXS data collected on ID2 at the ESRF. Scattering data in the $q$ range of 0.008 - 0.26 Å⁻¹ which were previously used to determine a $D_{max}$ of 275 Å and an $R_g$ of 85 Å, were used as the constraints for modelling. Ten *ab initio* models were produced in DAMMIF for each symmetry tested; P1, P5 and P6. The DAMMIF models were input into DAMFILT to produce an averaged, filtered model for each symmetry imposed.
Figure 6.14 - Ab initio models of xNAP₂-H3/H4. Three DAMFILT models are shown with P1, P5 or P6 symmetry imposed. Two orientations are shown for each model. The average chi values of the input DAMMIN models are 0.294, 0.295 and 0.294 for P1, P5 and P6 symmetry respectively.

The individual DAMMIN models were analysed to verify that the DAMMIF models were a good representation of the size and shape of the molecules in solution. Where no symmetry was imposed, DAMMIN struggled to produce models with consistent shapes, although they were all oblate. More consistency was observed between the DAMMIN models with P5 and P6 symmetry where they all had the main proportion of density centralised with an overall oblate form. These DAMMIF models were therefore deemed plausible representatives of the molecules in solution (Figure 6.14). The three symmetries examined all produced oblate models including the P1 model. This suggests that the xNAP₂-H3/H4 assembly has an oblate conformation (Figure 6.14).

6.5 Discussion

xNAP₂-Histone complexes were analysed by SANS and SAXS. Utilising scattering data from ILL, Diamond and ESRF beamlines, structural parameters were successfully determined for xNAP alone and in complex with histones. Results obtained for the xNAP₂-H2A/H2B complex were highly consistent between X-ray and neutron scattering experiments, different beamlines and multiple sample preparations (Table 6.1). Although the beamline geometries of ID14-3 (ESRF) and I22 (Diamond) were not ideal for measuring such large multimeric complexes, the $R_c$ gained from Guinier analysis were the same (within error) as those obtained using data from ID2.
Table 6.1 - Summary of small angle scattering results. This table summarises the results obtained from small angle scattering experiments of xNAP and the xNAP-Histone complexes. The resultant $R_g$, $D_{max}$ and calculated $M_r$ are shown from the X-ray beamlines ID14-3, I22 and ID2 and the neutron beamline D22. The results shown for the partially deuterated dxNAP₂-Histone complexes were obtained in 100% $H_2O$ buffer. Unless otherwise stated, the samples were measured in buffers containing 150 mM NaCl.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Beamline</th>
<th>$R_g$ (Å)</th>
<th>$D_{max}$ (Å)</th>
<th>$M_r$ (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>xNAP (500 mM NaCl)</td>
<td>ID14-3, ESRF</td>
<td>45</td>
<td>150</td>
<td>87</td>
</tr>
<tr>
<td>xNAP₂-H2A/H2B</td>
<td>I22, Diamond</td>
<td>81</td>
<td>250</td>
<td>1084</td>
</tr>
<tr>
<td>dxNAP₂-H2A/H2B</td>
<td>D22, ILL</td>
<td>90</td>
<td>270</td>
<td>/</td>
</tr>
<tr>
<td>xNAP₂-H3/H4</td>
<td>I22, Diamond</td>
<td>82</td>
<td>265</td>
<td>1315</td>
</tr>
<tr>
<td>xNAP₂-H3/H4</td>
<td>ID2, ESRF</td>
<td>85</td>
<td>275</td>
<td>1084</td>
</tr>
<tr>
<td>dxNAP₂-H3/H4</td>
<td>D22, ILL</td>
<td>91</td>
<td>270</td>
<td>/</td>
</tr>
</tbody>
</table>

The $R_g$’s for the protonated complex by SAXS and the partially deuterated complex by SANS are slightly different. This must reflect the distribution of apolar amino acid side chains within the complexes. The SANS match out experiments illustrated that the histones have a relatively small contribution to scattering. Further experiments would benefit from longer exposures coupled with a smaller sample-detector distance in order to increase the signal to noise ratio when the deuterated NAP is matched out of the complex in 100% $D_2O$. These data could potentially lead to the determination of the relative sub-unit arrangement within the complexes.

**SAXS of xNAP reveals a stable dimer consistent with size-exclusion chromatography.** The structural (Figure 1.9) and sequence (Figure 1.8) homology between NAP proteins from such diverse organisms as *Saccharomyces* and *Plasmodium* strongly suggest that xNAP has a similar protein fold to yNAP1. Using PYMOL and CRYSO, the $D_{max}$ and $R_g$ of the yNAP₂ crystal structure (PDB code 2Z2R) were calculated to be 130 and 32 Å, respectively. Scattering studies of xNAP resulted in the comparatively larger $D_{max}$ of 155 Å and an $R_g$ of 45 Å. Using BSA as a standard, the $M_r$ for xNAP in solution was determined to be 87 kDa, again slightly larger than the theoretical $M_r$ for an xNAP dimer (81.4 kDa). These results are consistent with results of the size exclusion chromatography described in section 4.2.2, suggesting that in 500 mM NaCl, xNAP is predominantly a dimer with a small amount of higher oligomerisation.

**xNAP-Histone complexes form oblate assemblies with a $D_{max}$ of 250-265 Å.** The $M_r$, $D_{max}$ and $R_g$ obtained for the xNAP₂-H2A/H2B were 1084 kDa, 250 Å and 81 Å respectively. The $M_r$ gained from scattering is consistent with the sedimentation equilibrium data (section 5.3.2.), where both of the equilibrium techniques suggest the predominance of the decamer species. The *ab initio* models of the SAXS data was in agreement with sedimentation velocity data in
suggesting that the xNAP₂·H2A/H2B assemblies were non-spherical. Furthermore these 3D models are all oblate with dimensions in keeping with the double ring hypothesis proposed in chapter 5 (Figure 6.15).

Figure 6.15 - Hypothetical model of the xNAP₂-Histone complexes compared with a SAXS 3D model. A) A schematic diagram of the proposed xNAP-Histone assembly (xNAP in blue and the histone dimer in red). B) The *ab initio* DAMFILT model shown was modelled using SAXS data of xNAP₂·H2A/H2B with PS symmetry.

Scattering data from I22 at Diamond resulted in $M_r$, $D_{\text{max}}$ and $R_g$ values for the xNAP₂·H3/H4 assemblies of 1054 kDa, 275 Å and 85 Å respectively. These data were again consistent with the sedimentation equilibrium discussed in chapter 5, which suggest that the predominant species is a decamer. The oblate *ab initio* models contained dimensions which could encompass a double pentameric ring as previously hypothesised. Measurements obtained on ID2 at the ESRF for NAP₂·H3/H4 gave similar $R_g$ and $D_{\text{max}}$ values but an increased $M_r$ of 1315 kDa. Given that no significant change was seen in the $R_g$ and $D_{\text{max}}$, this is consistent with SEC-MALLS in suggesting the presence of an alternative hexamer-dodecamer double ring equilibrium.

7. Future Projects and Preliminary Results

This chapter explores some possible future projects and presents preliminary results and hypotheses. Two main topics are discussed:
1) Do natively purified histones and recombinant histones interact with xNAP in the same way?

2) Can higher resolution structures of the xNAP-Histone complexes be gained via electron microscopy and X-ray crystallography?

### 7.1 Recombinant versus Native Histones

Post translation modifications (PTM), such as acetylation and methylation, of the histone proteins are widely studied in the field of epigenetics due to their key role in regulating transcription (see section 1.1.2). Recombinantly produced proteins were used in this research due to the advantages of increased homogeneity and high production yield. However, recombinant proteins lack the PTMs that they would have in their native eukaryotic environment. The xNAP-Histone complexes containing recombinant or native histones were compared to identify any differences caused by the PTMs.

To identify if PTMs had an effect on the binding of histones to xNAP, preliminary studies with natively produced and purified *Gallus gallus* H2A/H2B (gH2A/H2B) and H3/H4 (gH3/H4) dimers were examined. The native histones were a kind gift from Dr John Baldwin, John Moores University, Liverpool, UK. The *Xenopus laevis* and *Gallus gallus* histones share 91, 90, 98 and 100% homogeneity for H2A, H2B, H3 and H4 respectively (Appendix IV). The comparison between the recombinant *Xenopus* and native *Gallus* histones cannot, therefore, be solely attributed to the PTMs. However, previous studies utilising GST pull-down assays revealed identical binding of yNAP to both *Saccharomyces* and *Xenopus* histones, suggesting that there is no observable sequence specificity (McBryant et al., 2003). The xNAP-gHistone complexes were analysed by analytical size exclusion chromatography and sedimentation equilibrium to compare them with the xNAP-xHistone complexes.

#### 7.1.1 Size Exclusion Chromatography

The xNAP-gHistone complexes containing native histones were examined by size exclusion chromatography as carried out for the fully recombinant complexes (see section 2.7 and 5.1).
Comparison of the NAP-Histone complexes containing either the native (g) or recombinant (x) histones revealed very similar elution profiles (Figure 7.1). Analysis of eluting fractions by SDS PAGE confirmed an identical stoichiometry of one xNAP dimer to one histone dimer (data not shown).

**7.1.2 Sedimentation Equilibrium**

Preliminary AUC studies were performed on the NAP-Histone complexes containing the native *Gallus* histones for comparison with the fully recombinant complexes under the same experimental conditions (see sections 2.10.2 and 5.3.2).

**7.1.3 xNAP₂·gH2A/H2B**

The equilibrium data curves were analysed with a single species model individually and globally. Due to lack of material, there was not enough xNAP₂·gH2A/H2B complex to carry out analysis at the desired concentrations, so data at 0.8, 0.2 and 0.1 mg/mL were analysed instead.
**Figure 7.2** - Sedimentation equilibrium global analysis of xNAP₂-gH2A/H2B using a single species model. The nine curves covering three speeds and three concentrations are shown with a fit line to a single species model. The experimental data points of absorbance at 276 nm are shown as blue circles with the fitted lines shown in red. Above each curve is a plot of the residuals showing the deviation of the fit from the experimental data.

**Table 7.1** - Sedimentation equilibrium analyses of xNAP₂-gH2A/H2B using a single species model. The $M_r$ obtained from individual analysis of equilibrium curves are shown with kDa units (blue boxes). The variance of the fits to the experimental data is shown in brackets. Global fits of the curves at each concentration are shown in the bottom row (purple boxes). Global fits of the curves at each speed are shown in the right hand column (purple boxes). The bottom right-hand box shows the global analysis of all nine curves.

<table>
<thead>
<tr>
<th>Speed</th>
<th>0.8 mg/mL</th>
<th>0.2 mg/mL</th>
<th>0.1 mg/mL</th>
<th>Global</th>
</tr>
</thead>
<tbody>
<tr>
<td>3,000 rpm</td>
<td>1060 (0.87)</td>
<td>1041 (0.86)</td>
<td>1052 (0.45)</td>
<td>1060 (0.73)</td>
</tr>
<tr>
<td>5,000 rpm</td>
<td>994 (2.99)</td>
<td>956 (1.19)</td>
<td>1028 (1.12)</td>
<td>992 (1.76)</td>
</tr>
<tr>
<td>7,000 rpm</td>
<td>864 (3.23)</td>
<td>951 (0.86)</td>
<td>1112 (0.91)</td>
<td>866 (1.76)</td>
</tr>
<tr>
<td>Global</td>
<td>960 (6.88)</td>
<td>978 (0.99)</td>
<td>1064 (0.84)</td>
<td>961 (2.92)</td>
</tr>
</tbody>
</table>

The range of average $M_r$ values for the native histone complex calculated from individual and global analyses was 864 – 1060 kDa (**Table 7.1**). This range was smaller than that seen for the
recombinant histone complex (921 to 1261 kDa). Analysis of all nine curves for the partially native complex resulted in an average $M_r$ of 961 kDa (Figure 7.2). This $M_r$ is again slightly smaller than the 1072 kDa calculated for NAP₂·xH2A/H2B.

7.1.4  xNAP₂·gH3/H4

An identical analysis of xNAP₂·gH3/H4 was performed including individual and global fits of the nine curves to obtain average $M_r$. xNAP₂·gH3/H4 concentrations of 0.8, 0.4 and 0.2 mg/mL were measured.

![Image](Figure 7.3 - Sedimentation equilibrium global analysis of xNAP₂·gH3/H4 using a single species model. The nine curves covering three speeds and three concentrations are shown with a fit line to a single species model. The experimental data points of absorbance at 276 nm are shown as blue circles with the fitted lines shown in red. Above each curve is a plot of the residuals showing the deviation of the fit from the experimental data.)
Table 7.2 - Sedimentation equilibrium analyses of xNAP₂·gH3/H4 using a single species model. The Mₗ obtained from individual analysis of equilibrium curves are shown with kDa units (blue boxes). The variance of the fits to the experimental data is shown in brackets. Global fits of the curves at each concentration are shown in the bottom row (purple boxes). Global fits of the curves at each speed are shown in the right hand column (purple boxes). The bottom right-hand box shows the global analysis of all nine curves.

<table>
<thead>
<tr>
<th>RPM</th>
<th>0.8 mg/mL</th>
<th>0.4 mg/mL</th>
<th>0.2 mg/mL</th>
<th>Global</th>
</tr>
</thead>
<tbody>
<tr>
<td>3,000</td>
<td>834 (0.48)</td>
<td>845 (0.69)</td>
<td>821 (2.1)</td>
<td>837 (1.1)</td>
</tr>
<tr>
<td>5,000</td>
<td>779 (3.39)</td>
<td>783 (4.35)</td>
<td>754 (4.01)</td>
<td>779 (3.93)</td>
</tr>
<tr>
<td>7,000</td>
<td>674 (2.71)</td>
<td>729 (4.11)</td>
<td>694 (3.40)</td>
<td>697 (3.96)</td>
</tr>
<tr>
<td>Global</td>
<td>750 (6.44)</td>
<td>764 (3.93)</td>
<td>740 (3.41)</td>
<td>754 (4.6)</td>
</tr>
</tbody>
</table>

The calculated average Mₗ for the partially native complex ranged from 674 to 845 kDa over the individual and global fit analyses of the xNAP₂·gH3/H4 sample (Table 7.2). This range was smaller than the range calculated for the fully recombinant complex: 939 – 1219 kDa. The global fit using all nine curves also produced a smaller average Mₗ of 754 kDa (Figure 7.3), compared to the NAP₂·xH3/H4 of 1093 kDa.

### 7.1.5 Discussion

**xNAP₂ forms large assemblies with native Gallus gallus histones.** Size exclusion chromatography resulted in comparable elution profiles for all of the NAP·Histone complexes containing either native or recombinant histones, suggesting that they form complexes with similar hydrodynamic radii. This is consistent with the findings of McBryant et al. who observed no species specificity with regards to yNAP when examining interactions with *Xenopus* and *Saccharomyces* histones (McBryant et al., 2003). Furthermore, the results indicate that the presence of PTMs on the histones do not inhibit the formation of large NAP·Histone assemblies. Evidence of the smaller NAP·Histone species, eluting at 15 mL, was also observed for the native histone complexes, as seen for the recombinant NAP₂·H2A/H2B (section 5.1.1.1).

**The equilibrium for the xNAP₂·gHistone complex is shifted towards the 500 kDa complex.** Although the xNAP₂·gHistone complexes behaved similarly to their recombinant counterparts by size exclusion chromatography, a shift in the dynamic equilibrium was observed using sedimentation equilibrium. AUC equilibrium studies for the xNAP₂·gH2A/H2B resulted in an experimentally measured Mₗ of 961 kDa, which is 10% smaller than the Mₗ found for the recombinant complex (1072 kDa). The Mₗ from sedimentation equilibrium for xNAP₂·gH3/H4 was 754 kDa, which was 31% smaller than its recombinant counterpart with an observed Mₗ of 1093 kDa. As previously mentioned in chapter 5, it is likely that the fully recombinant
complexes form a double ring structure. The data presented in this chapter suggest that this is also the case for the *Gallus* histone complexes. However, the smaller M, values obtained for the complexes containing native histones do, suggest that the equilibrium between the single ring and double ring has shifted towards the former. This shift is relatively small for the xNAP₂·gH2A/H2B complex, but for the xNAP₂·gH3/H4 a greater shift is observed.

This shift in single ring / double ring equilibrium may be explained by the PTMs of the histone tails disrupting the dimerisation of the rings, perhaps by electrostatic effects, thus making the dimerisation less favourable. It is important to clarify whether the changes in the M, measured by sedimentation equilibrium are due to the PTMs or the sequence differences between the *Xenopus* and *Gallus* histones. Further experiments, similar to the ones described in this section, are required comparing *Xenopus* histones which have been synthesised in native or recombinant systems.

### 7.2 Crystallisation

#### 7.2.1 Crystallisation trails

An aim of these studies was to produce high resolution structures of xNAP and the xNAP-Histone complexes to detail their interactions (section 2.14). Crystallisation trials were set up alone in high and low ionic strength buffers (500 and 50 mM NaCl) in an attempt to crystallise the xNAP dimer as well as the high-order oligomers. Protein concentrations of 2, 3 and 5 mg/mL were examined. Pre-made 96-condition screens were used from Molecular Dimensions including PACT premier, Structure screen I and II, Clear Strategy I and Macrosol. To date, a few crystals have formed but when tested on the in-house diffractometer it was evident that these were salt crystals due to the strong high resolution reflections and lack of lower resolution reflections. NAP homologues have been crystallised in buffers containing 0.25 M Mono-ammonium dihydrogen phosphate (γNAP) and 0.2 M di-ammonium tartrate with 20% PEG 3350 (pfNAP) precipitants (Gill et al.; Park and Luger, 2006a). Optimisation screens based on these conditions with variations in precipitant and protein concentration were also set up, but no crystals have grown thus far. Crystallisation trials were also carried out for the xNAP₂·H2A/H2B and xNAP₂·H3/H4 complexes at a concentration of 2.5 mg/mL. All of the pre-made screens used for xNAP trials were used including the additional screens: Clear strategy II, Morpheus and Proplex. To date, no protein crystals have been observed.
7.2.2  In situ Proteolysis

Within the xNAP-Histone complexes there are unstructured regions that could hinder crystallisation, such as the His-tag on the C-termini of the xNAP dimers and the unstructured N-terminal tails on the histones. Previous studies have shown in situ proteolysis (within the crystallisation droplet) to be successful in removing unstructured termini and tags, thus aiding in the formation of crystals and in improving the diffraction quality (Dong et al., 2007; Wernimont and Edwards, 2009). The xNAP-Histone complexes were therefore digested using four different proteases; α-chymotrypsin, trypsin, papain and subtilisin. A time course was carried out to determine which protease would produce the largest and most stable protein fragment (section 2.14.1). The xNAP-Histone complexes were added at 1.8 mg/mL to a 1/100 or 1/1000 dilution of the stock proteases (1 mg/mL) and incubated at 25°C. Aliquots were taken at 0, 0.5, 3 and 18 hr and put into 3 x SDS loading buffer to halt the protease activity. The resultant proteolytic fragments were examined by SDS PAGE.

7.2.2.1  α-Chymotrypsin Digestion

Digestion by α-Chymotrypsin did not have a large effect on xNAP within the xNAP₂·H2A/H2B complex. This was expected given that the xNAP being studied (xNAP1-C1) was already a product of limited proteolysis by this enzyme (see section 1.4) (Figure 7.4). Two extra bands were, however, seen in the 1/100 dilution (1c and 2c). The 1c and 2c bands were not seen in the 1/1000 dilution series suggesting these fragments were fully digested by the higher concentration of protease. Over time, the other contaminating bands in the 1/1000 dilution
were also seen to decrease in intensity, suggesting that these were being further digested over time. The predominant xNAP band, running at 55 kDa, remains throughout the proteolysis time-course. This suggests that this starting fragment (xNAP1-C1) is a relatively stable product. For xNAP₂·H2A/H2B in the 1/100 dilution series at 18 hr, the band corresponding to H2A and H2B faded and a degradation product was seen (3c). This suggested that the histone, and possibly the histone tails, were being digested (Figure 7.4A).

The digestion of xNAP within the xNAP₂·H3/H4 complex was very similar to when it was a component of the xNAP₂·H2A/H2B complex. In the xNAP₂·H3/H4 complex, the band corresponding to H3 was more readily digested than that of H4, producing a degradation product (4c) (Figure 7.4B). This suggested that H3 is more vulnerable in the complex than the H4. The stability of the xNAP band and the effects upon H3 and H2A/H2B indicated that α-Chymotrypsin protease would be a possible enzyme to use for in situ crystallisation trials.

7.2.2.2 Trypsin

Figure 7.5 - Limited proteolysis of the xNAP·Histone complexes with trypsin. A digestion series as a function of time is shown for A) xNAP₂·H2A/H2B and b) xNAP₂·H3/H4 using two dilutions of the 1 mg/mL trypsin stock solution (1/100 and 1/1000). A standard marker (M) is run alongside the complex prior to addition of the protease (pre). Aliquots are also shown of the product after 0, 0.5, 3 and 18 hr of incubation. The red arrows point to bands of interest.

Digestion of the xNAP·Histone complexes with 1/100 dilution of trypsin, led to a very fast digestion of both the xNAP and histone proteins, with virtually no intact protein remaining after 18 hr (Figure 7.5). As expected, a much slower rate of digestion of the xNAP₂·H3/H4 complex was observed with a 1/1000 dilution. No change was seen in the 1/1000 series for xNAP₂·H2A/H2B. This was probably due to a sample handling issue. Given that no stable band
was produced using trypsin, it was concluded that this would not be a good candidate to use for *in situ* crystallisation trials.

7.2.2.3 Papain

Digestion of the xNAP-Histone complexes using papain occurred very quickly, even in the apparent time-zero aliquots (Figure 7.6). This suggested proteolytic activity occurred more quickly than the SDS could bind to the protease and quench activity. The 1/100 dilutions degraded the xNAP and the histone proteins into multiple unstable fragments. However, in the 1/1000 dilution a stable band appeared for the xNAP protein (1p), and the contaminating bands faded. After 18 hr the band corresponding to H2A and H2B within the xNAP₂·H2A/H2B complex faded, whilst another band corresponding to a smaller fragment appeared (2p). This new band may correspond to the histone tails being cleaved (Figure 7.6A). For the xNAP₂·H3/H4 complex, the band of H3 faded over the time course, once again suggesting that H3 is in a more vulnerable position within the complex (Figure 7.6B). The production of a stable xNAP band made Papain a suitable enzyme to use in crystallisation trails but given its potency, a higher dilution factor of 1/10,000 would be adequate.

Figure 7.6 - Limited proteolysis of the xNAP-Histone complexes with papain. A digestion series as a function of time is shown for A) xNAP₂·H2A/H2B and b) xNAP₂·H3/H4 using two dilutions of the 1 mg/mL papain stock solution (1/100 and 1/1000). A standard marker (M) is run alongside the complex prior to addition of the protease (pre). Aliquots are also shown of the product after 0, 0.5, 3 and 18 hr of incubation. The red arrows point to bands of interest.
Subtilisin at a 1/100 dilution also showed a high potency, almost completely degrading the xNAP and histones. H4 however, was relatively stable with the main band remaining even after 18 hr (Figure 7.7). With the 1/1000 dilution, the predominant xNAP band running to 55 kDa appeared to be stable, and the bands corresponding to the contaminating protein faded. Extra bands appeared (1s) beneath the H2A and H2B band, whilst the main band faded, suggesting some digestion of the histones (Figure 7.7A). In the xNAP₂-H3/H4 complex the band that corresponded to H3 was, once again, the most digested (Figure 7.7B). Given that stable bands were observed for xNAP, H2A/H2B and H4, Subtilisin would also be an interesting protease to use for in situ crystallisation. However, a higher dilution factor such as 1/10,000 would be sufficient.

Using the results from the limited proteolysis time course experiments as a guide, crystallisation trials were performed using in situ proteolysis with α-chymotrypsin and papain. The two xNAP-Histone complexes were examined at two protein concentrations 1.8 and 0.45 mg/mL with the PACT screen. Immediately before loading the protein into the crystallisation drops, the α-chymotrypsin or papain were added at 1/1000 and 1/10,000 dilutions to the protein solution respectively. The trays were sealed and stored at 4°C and checked regularly. To date no protein crystals have been seen in the drops.
7.2.3 Discussion

Stable proteolytic domains of the xNAP-Histone complexes were identified. Extensive crystallisation trials did not result in any protein crystals of xNAP or the xNAP-Histone complexes. Limited proteolysis trials indicated that smaller stable domains of xNAP and histones could be produced. Easily accessible cleavage sites include the flexible histone tails and the His-tag of xNAP. Further digestion could, however, lead to cleavage in other regions within the proteins, such as flexible loop regions. Although in situ proteolysis during crystallisation has been shown to be effective for certain proteins (Wernimont and Edwards, 2009), the crystallisation solution is rendered heterogeneous by the very nature of the experiment, which may hamper crystallisation.

Proposed optimisation of crystal trials. The dynamic equilibrium in which xNAP and the xNAP-Histone complex exists may be hindering crystallisation due to the heterogeneity of the solutions. Mutations to the β-hairpins within γNAP successfully prevented the oligomerisation of γNAP₂ (Park et al., 2008). Given the sequence homology of γNAP and xNAP, the same mutations in the β-hairpins may lead to a homogeneous solution of the xNAP dimer, allowing for crystallisation. The smallest oligomer observed for the xNAP-Histone complex eluted at 15 mL from a superose 6 column (Figure 7.1 A), and is most likely an xNAP dimer bound to a histone dimer. The structure of this stoichiometric sub-unit would provide valuable information about the xNAP-Histone interactions. The same β-hairpin mutations would presumably inhibit the xNAP-Histone sub-unit from oligomerising, again reducing heterogeneity within the sample and aiding crystallisation. Further attempts to increase the purity of xNAP would also be beneficial, as the degradation bands present in the sample prior to crystallisation may also be inhibiting the crystallisation process. Furthermore, experiments using limited proteolysis could be performed followed by a subsequent purification step to remove the enzymes and cleaved regions from the solution, again increasing homogeneity of the sample prior to crystallisation trials.

7.3 Electron Microscopy

Preliminary EM studies of the xNAP₂·H3/H4 complex were performed in collaboration with Prof. Carlo Pertosa and Dr. Guy Schoen at the Institut Biologie Structurale (IBS). Samples were purified and diluted for electron microscopy using uranyl acetate as a negative stain.
Figure 7.8 shows a representative selection of particles from the raw images obtained of xNAP₂·H3/H4. The assemblies observed have a maximum diameter of approximately 220 Å and appear vaguely hexagonal. For a more rigorous analysis of the xNAP-Histone complexes, samples were sent to Prof Raimond Ravelli and colleagues in Leiden University Medical Centre, The Netherlands, where negative stain with uranyl acetate and 2D cryo EM were performed. Data analysis including class sums and averaging of the xNAP₂-H2A/H2B complex is currently in progress in the Ravelli laboratory. So far, they have identified multiple species in the samples and are collecting further images for averaging. In this preliminary analysis, the 1400 particles imaged were averaged using 5, 10 and 20 classes to allow for various levels of heterogeneity within the sample.

These preliminary class averages for xNAP₂-H2A/H2B (Figure 7.9) reveal particles with an average $D_{max}$ of approximately 250 Å. Some images, e.g. 3, 8, 12, 18 and 25 reveal a common
feature of a central region of high density surrounded by 4-9 other areas of high density in a ring-like formation.

7.3.1 Discussion

Electron microscopy images have been obtained for the xNAP-Histone complexes. The negative stain and 2D cryo-EM images collected of the xNAP-Histone complexes have a D\textsubscript{max} of approximately 250 Å. This is consistent with the D\textsubscript{max} obtained by small angle scattering, described in chapter 6. The negative stain images of xNAP-H3/H4 obtained at the IBS, although somewhat indistinct, could be compatible with six-fold symmetry, which would be in keeping with the double hexameric ring hypothesis based upon SEC-MALLS and SAXS data. The 2D cryo-EM images collected in Leiden show multiple units of distinct densities within the particles, which are the correct order of magnitude to correspond to the xNAP\textsubscript{2}-H2A/H2B stoichiometric unit. Clear heterogeneity exists between the images, showing particles with 5-10 of these individual units. Given the heterogeneity observed, more cryo-tomograms are currently being collected to provide a 3D reconstruction of a single particle.
8. Final Discussion

The overall aim of this research was to provide a biophysical and structural insight into the interactions between the highly conserved histone chaperone NAP and its biological relevant binding partners, the histones. A complex and dynamic equilibrium has been observed and successfully characterised using a broad range of complementary techniques (Summarised in Table 8.1). Sedimentation equilibrium, and small angle scattering were employed to characterise the complexes at equilibrium whilst sedimentation velocity and SEC-MALLS were used to identify the distinct species within this multi component system. Preliminary electron microscopy of the xNAP-Histone complexes was performed in order to visualise individual particles.

<table>
<thead>
<tr>
<th></th>
<th>xNAP (41 kDa)</th>
<th>xNAP₂-H2A/H2B (110 kDa)</th>
<th>xNAP₂-H3/H4 (109 kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEC-MALLS</td>
<td>Mₐ, range</td>
<td>96, 165, 400-720 kDa</td>
<td>500 – 1000 kDa</td>
</tr>
<tr>
<td>Sedimentation</td>
<td>NaCl</td>
<td>20 mM</td>
<td>1 M</td>
</tr>
<tr>
<td>Velocity</td>
<td>Mₛ</td>
<td>8 – 18 S</td>
<td>4.4 S</td>
</tr>
<tr>
<td></td>
<td>f/f₀</td>
<td>/</td>
<td>1.65</td>
</tr>
<tr>
<td>Sedimentation</td>
<td>NaCl</td>
<td>20 mM</td>
<td>1 M</td>
</tr>
<tr>
<td>Equilibrium</td>
<td>Mₛ</td>
<td>83 kDa</td>
<td>1072 kDa</td>
</tr>
<tr>
<td>Small Angle X</td>
<td>NaCl</td>
<td>500 mM</td>
<td>I22</td>
</tr>
<tr>
<td>Ray Scattering</td>
<td>Beamline</td>
<td>ID14-3</td>
<td>1084 kDa</td>
</tr>
<tr>
<td></td>
<td>Mₛ</td>
<td>87 kDa</td>
<td>81 Å</td>
</tr>
<tr>
<td></td>
<td>Dₘₐx</td>
<td>150 Å</td>
<td>250 Å</td>
</tr>
<tr>
<td></td>
<td>Rₛ</td>
<td>45 Å</td>
<td>82 Å</td>
</tr>
</tbody>
</table>

Table 8.1 - Summary of Results. Experimental results are shown for xNAP, xNAP₂-H2A/H2B and xNAP₂-H3/H4. The theoretical Mₛ value of the subunit is in brackets in the top row. The buffer conditions were 3 mM MgCl₂ and 20 mM Tris HCl, pH 7.5 and (unless otherwise stated) 150 mM NaCl. The Rₛ shown were calculated from the Guinier plot.

8.1 xNAP

Analytical ultracentrifugation of xNAP revealed that in high ionic strength, a stable dimer was formed, consistent with the published data describing γNAP1 (Toth et al., 2005). In more physiological ionic conditions, SEC-MALLS demonstrated that xNAP₂ forms multiple large assemblies. This oligomerisation of xNAP₂ was similar to that observed for γNAP₂ (Toth et al., 2005) and, given their sequence similarity, it is likely that it also occurs via the formation of β-sheets between β-hairpins within adjacent xNAP molecules.

Sedimentation equilibrium of xNAP in low ionic strength was consistent with the findings of other researchers (Friedeberg et al., 2006) (Toth et al., 2005), suggesting that the sample was
heterogeneous. An in-depth systematic analysis of the AUC data revealed that this delicate equilibrium was being disrupted by sedimentation which was selectively removing the larger species from solution at increasing speeds. Analysis with complementary SEC-MALLS was thus used to identify the species within this heterogeneous sample.

### 8.2 xNAP·Histone Complexes

Characterisation of the xNAP·Histone complexes was also performed to identify their binding stoichiometries and oligomeric status. Using analytical size exclusion chromatography and SDS PAGE, the controversial binding stoichiometry was clarified to be one xNAP dimer to one H2A/H2B or H3/H4 dimer. In physiological conditions these stoichiometric subunits oligomerised into large assemblies. Using SEC-MALLS these assemblies were shown to contain between 5-10 and 6-12 subunits for xNAP₂·H2A/H2B and xNAP₂·H3/H4 respectively. Equilibrium techniques, such as sedimentation equilibrium and small angle scattering, suggested that the predominant species at physiological ionic strength was a decamer. Small angle scattering revealed that the complexes were oblate in shape with a maximum dimension of approximately 250 Å. This was further supported by electron microscopy experiments which imaged individual particles of similar maximum dimensions.

The largest species observed from any of the techniques was 1200 kDa, suggesting the presence of a defined oligomerisation rather than non-specific aggregation. To accommodate the results from all of the techniques employed, a model has been formulated in which 5 or 6 stoichiometric subunits oligomerise into a primary complex, which in turn self-associates into a decamer or dodecamer respectively. With this number of subunits, it is likely that a ring-like complex is present which can then self-associate into a double ring, such as that observed for another histone chaperone, nucleoplasmin (Dutta et al., 2001). A double-ring conformation would be consistent with the small angle scattering data satisfying a $D_{\text{max}}$ of 250 Å whilst accounting for the $M_r$. Table 8.1 illustrates that all of the techniques employed consistently described the xNAP₂·H2A/H2B complex as being predominantly a decamer. Results for the xNAP₂·H3/H4 complex were less decisive, but provided evidence that the predominant species in solution was either a decamer or a dodecamer.
8.3 Hypothetical Model
Interestingly, a pentamer-decamer equilibrium is observed for nucleoplasmin, and these high-order structures are thought to facilitate its main function as a histone storage molecule (Dutta et al., 2001). Given that the proposed model of a double ring assembly for xNAP resembles nucleoplasmin oligomerisation, histone storage could also be a possible function of the high-order xNAP assemblies. Many NAP binding partners have been identified, suggesting that NAP is highly controlled and influenced by the cellular environment. The high abundancy of xNAP within the cell and the multitude of signalling factors, gives rise to its ability to be involved in multiple cellular activities. Further experiments are required to identify affects of phosphorylation of NAP, PTMs of histones and the binding of non-histone proteins. Further use of the complementary techniques described in this thesis would be ideal for such experiments. It would also be interesting to determine if NAP interacts with the whole histone octamer and if it can bind a mixture of H2A/H2B and H3/H4 dimers.

8.4 Biological Relevance
Although these large assemblies have been characterised in vitro, it is still unclear if they are present in vivo and what function they may have within the cell. The NAP1 β-hairpins are responsible for oligomerisation and also contain a putative NLS region (Park and Luger, 2006a; Park et al., 2008). It is therefore likely that histone binding and nuclear import are performed by the NAP dimer. However, the nuclear pore complex (NPC), through which the NAP-Histone complex must travel to enter and exit the nucleus, has a central pore of 50 nm (Frenkiel-Krispin et al., 2010) making it theoretically large enough to transport the high-order NAP-Histone oligomers. As the complex enters the nucleus, the local protein concentration increases by a factor of four (Toth et al., 2005), perhaps favouring oligomerisation into the high-order species. These large NAP-Histone complexes may have a role in histone storage, as previously mentioned, and/or could act as a scaffold for nucleosome assembly and disassembly. The NES region within the dimerisation helix of NAP has been shown to be involved in NAP nuclear export (Miyaji-Yamaguchi et al., 2003) and is thought to be modulated by the accessory domain (Park and Luger, 2006a). The oligomerisation of the NAP-Histone subunit may also lead to a conformational change within NAP, altering the proximity of the accessory domain to the NES, and thus effecting the sub-cellular localisation.
8.5 Future Work

As discussed in chapter 7, further structural studies using X-ray crystallography and electron microscopy will be important for high resolution characterisation of the NAP-Histone complex \textit{in vitro}. Electron microscopy could also be used to determine the relative subunit arrangement using antibodies raised against specific domains, as successfully used by (Wilkens et al., 2000). To investigate the biological function of these large multi-subunit complexes, future experiments to determine their sub-cellular localisation are required. Recent advances of \textit{in vivo} super resolution microscopy have included techniques such as interferometric photoactivated localisation microscopy (iPALM) (Shtengel et al., 2009), which have been used to provide highly detailed images of fluorescence-tagged molecules with 3D nanoscale resolution (Kanchanawong et al., 2010). Using iPALM, it may now be possible to identify the oligomerisation state of the NAP-Histone complexes in combination with determining their sub-cellular localisation. Where and when NAP interacts with various non-histone binding proteins could also be established by this super resolution microscopy technique. Such insight into the biological function of NAP would be invaluable and would give insight into the diverse mechanisms of NAP function and greatly improve our knowledge of how chromatin structure is regulated.
9. References


McRorie D. K., V.P.J. Self Associating Systems in the Analytical Ultracentrifuge (Research and Applications Department, Spinco Business Unit, Beckman Instruments, Inc.Palo Alto, California).


P Pernot1, P.T., T Giraud1, R Nogueira Fernandes1, D Nurizzo1, D Spruce1, J Surr1, S McSweeney1, A Round2, F Felisa2, L Foedinger2, A Gobbo2, J Huet2, C Villard2 and F Cipriani2 (2010). New beamline dedicated to solution scattering from biological macromolecules at the ESRF *Journal of Physics: Conference Series* **247**.


Ralston, G. Introduction to Analytical Ultracentrifugation (Department of Biochemistry, The University of Sydney, Sydney, Australia).


Appendix I - BenchMark™ Protein Ladder.

The pre-mixed solution is used as a standard for SDS PAGE containing proteins ranging from 10 – 220 kDa. The 20 and 50 kDa bands have a stronger intensity than the others for ease of identification.
Appendix II – xNAP1-C1

i) pET28 plasmid map

<table>
<thead>
<tr>
<th>Sequence/Annotation</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>T7 promoter</td>
<td>270-306</td>
</tr>
<tr>
<td>T7 transcription start</td>
<td>360</td>
</tr>
<tr>
<td>His+Tag coding sequence</td>
<td>270-287</td>
</tr>
<tr>
<td>T7 Tag coding sequence</td>
<td>207-236</td>
</tr>
</tbody>
</table>

Multiple cloning sites:
- BamHI I - XhoI: 158-203
- His+Tag coding sequence: 140-157
- T7 terminator: 26-72
- JspI coding sequence: 773-1852
- pBluescript II origin: 2286
- Kan coding sequence: 1995-4807
- T1 origin: 4903-5358

The maps for pET-28b(+) and pET-28c(+) are the same as pET-28a(+) (shown) with the following exceptions: pET-28b(+) is a 5368bp plasmid: subtract 2bp from each site beyond BamHI at 108. pET-28c(+) is a 5367bp plasmid: subtract 2bp from each site beyond BamHI at 108.

ii) pET28 plasmid sequence and restriction sites

Using the highlighted Ncol and Xhol restriction sites, the xNAP1-C1 gene shown below was ligated into the pET-28b plasmid. The plasmid was sent for sequencing using the T7 promoter and terminator primers. The resulting sequence was translated into the protein sequence below. The restriction sites and the subsequently incorporated His-tag are also highlighted within the sequence below.
i) xNAP1-C1 DNA and protein sequence

```
ACCATGCTGTCTGGCAAGGAAAGGATGACCTAGTGCGGACAATTCAAGGATACATT
MAALQLRLDDDLVGTPTGYI
GAAGGCTTACCCCAAGTAGAAAGGAGGATGACCTTAAATATCTTACAGTGAA
ESLPDRVVKRVRNVNLKQLNVK
TGGTACAGTAAAGCAATATCTTCAAGAATAGGAGGATATGCGTACAGTGCA
CAQUIAEHKFYEEVEHELEYK
GGCCCTCTATCAACCCCTTTTGGGCAAGGAAGCAGATAATATAATATCTTACATAGAACT
ALYQPLFDKRSIDINATYEPT
ACAGAAGGAAATGTGGAAGTAGAAGGAGGATATTTCTTGGGGAACCTGAGGAA
TEEZCEWERVEEEDISGDLDKE
AAAAGCCCAACTGCGAGGGAAGGAAAAATGGAAGGAGGAGGATCCTAAAGGAGGAT
EFLWLTVPKNVDDLSSDMLQEH
AGTGAACCAATTAAAAATCTTGGAAGACATAAAAGTGAATATTTCTAGATGTGGCGAG
DEIFIKLHKLKDIKVKFSDAGQQ
CCTATGAGTTTCTACCTAGAATTCTTTCTTGAACCAATAGAATCTTTCAAAAT
PNSFTLEFYESPEFNRBFTEV
TTGCACAAGACATACAAAATAGGTCAAGCCGTAGTCAAGACACCTCTCTCTTCGAT
LTHTYKMRSEPDFEDSFSD
GGACCTGAATAATTGAGTCACTGGGCTTGTATGCTGGACTGGAAGAGAAAGAAGATGTG
GPEIMGCTGGCLIDWKUNKV
ACTTTAAAAACCTAAAGAAGGAAAGCATAAAGGTCTCTGGAACCTGTAAGACTTTT
TLKTKEKHKHRHKRGSTVYRTV
ACAATAAATGTTCCTAAGAGGACTCTTCTTCTACCTACCTCAACCTCTCTACAGTACCTGA
TKTVPNDSFPNFPTFEPVE
AATCTGTGAGCTGGTAGTACATGTGCGAAGGACCTCTCTACTTCGAGTTCAATTGACACAC
NGELDDDAEAAILTADFEIGH
TCCTGAGTCTAGCATATTCTAAGACAGCTGATGTGATTTCTCTGGTGAAAOGAATTGAA
FLRERIIPRSLIFZIGEAITE
GATGACAGATGATTATGAGAAAGAGGAGGAGGATGAGGGAAGGAGGAGGAGGAGGAGGAGG
DDDDDDYDEGGEAADDEEGEK
GAAAGCTGATGAAGACATAATGAGCTGACAAAGAACTGTGGCAAGCAAGAGGATCCTAAACCCACGCGAG
EADENDNPDYEPKKGQNPME
TGCAAATGCAACATCTGAGTTACCCCAACCAACCCACAC
CRRQQQLHHHHHHHH
```
Appendix III - Histones

i) pET3 plasmid map

Histones H2A and H2B were cloned into pET3a using restriction sites NdeI and BamHI. The NdeI restriction sites can be seen at the start of the protein sequence whereas the protein stop codons are much before the BamHI site shown below. Histone H3 was also cloned into pET3a but with restriction sites Ncol and BgIII. These restriction sites can be identified flanking the protein residues. The histone H4 was cloned into pET3d using the restriction sites NdeI and BgIII. The plasmids were sent for sequencing, the resulting DNA sequences was translated into protein sequences and are shown below, the natural stop codons are highlighted by grey boxes and the restriction sites are also highlighted.
iii) Histone H2A DNA and protein sequence

```
catatgcacagggaaaacaggctgttaaccaccggct gaagacagctctgca
```

```
K T E S S K S A K S K
```

iv) Histone H2B DNA and protein sequence

```
catatgcacagggggcccggcgtaaccaccggctgaagacagctctgca
```

```
M A K S A P A P K G S K K A V T K T
```

v) Histone H3 DNA and protein sequence

```
accatgcaccccttaccaacagaccccggctgaacctccgggctgaagacagctctgca
```

```
M A R T K Q T A R K T S T C G K A P R K C
```

vi) Histone H4 DNA and protein sequence

```
catatgcctcgtgtgtgtaaaaggtgtgaaactgctacaaccggctgaagacagctctgca
```

```
K V F L E N V I R D A V T Y H A K R
```

```
G F G G
```

## Appendix IV - Histone Alignment

Alignments were performed using Vector NTI (Invitrogen).

### i) Histone H2A

<table>
<thead>
<tr>
<th></th>
<th>Gallus gallus H2A</th>
<th>Xenopus laevis H2A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MSGRGRKGGKVRRAKAKSRSSRAGLQFPGRVRHRLRLKGNYAERVGAG</td>
<td>MSGRGRKGGKVRRAKAKSRSSRAGLQFPGRVRHRLRLKGNYAERVGAG</td>
</tr>
</tbody>
</table>

### ii) Histone H2B

<table>
<thead>
<tr>
<th></th>
<th>Gallus gallus H2B</th>
<th>Xenopus laevis H2B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FPEPSAFAFKRGSKTAKTKQRKKRKRKRESVSYTVYKLKQTHF</td>
<td>FPEPSAFAFKRGSKTAKTKQRKKRKRKRESVSYTVYKLKQTHF</td>
</tr>
</tbody>
</table>

### iii) Histone H3

<table>
<thead>
<tr>
<th></th>
<th>Gallus gallus H3</th>
<th>Xenopus laevis H3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HATYQZAKSTZGGAPKQLATKAAKSSAPACGGVKKHFRYPRGTVLRE</td>
<td>HATYQZAKSTZGGAPKQLATKAAKSSAPACGGVKKHFRYPRGTVLRE</td>
</tr>
</tbody>
</table>

### iv) Histone H4

<table>
<thead>
<tr>
<th></th>
<th>Gallus gallus H4</th>
<th>Xenopus laevis H4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MSGRGRKGGKVRRAKAKSRSSRAGLQFPGRVRHRLRLKGNYAERVGAG</td>
<td>MSGRGRKGGKVRRAKAKSRSSRAGLQFPGRVRHRLRLKGNYAERVGAG</td>
</tr>
</tbody>
</table>

Key: ![Identical]
Appendix V– Beginners Guide to SAXS Analysis

Firstly sort out all the raw data files into one folder and have to hand the sample numbers and concentrations in relation to each number e.g. :

<table>
<thead>
<tr>
<th>Run Number</th>
<th>Sample</th>
<th>conc- OD 280</th>
<th>mg/mL</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>BSA</td>
<td>4.13</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>BSA buffer</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Download Primus from the ERSF website [http://www.embl-hamburg.de/ExternallInfo/Research/Sax/manual_primus.html](http://www.embl-hamburg.de/ExternallInfo/Research/Sax/manual_primus.html) (it’s in the ATSAS package). MAKE SURE YOU SAVE IT TO THE C DRIVE otherwise half the applications won’t work. Primus is a program where you can subtract the buffer scattering from the sample scattering. You can obtain radii of gyration, maximum dimensions and molecular weights from the scattering data.

- Start up Primus and click on the tools tab, a data processing (dp) window will appear, this is where you control what is shown in the graphics window (gw).
- Click on range in the dp window and the range parameter window will open.
- Click select, you probably have multiple files for each sample and buffer, select the first file for your sample, this will load all of the files for that sample.
- Click apply in the range parameter window.

All your files should now be shown in the dp window, by clicking plot in the dp window the files will be plotted in the gw. Samples after radiation can become damaged so to remove the data collected after damage has occurred you scale the samples.

- Click scale on the dp window

The multiplier column on the dp window will indicate how divergent the curves are from the file in position #1. A general rule is to remove any over 5% different (> 1.05 and < 0.95). To remove the files that have this difference, simply uncheck the box on the left of the row, only the checked the files will be used in the gw.

To get the best data, you can use all of the undamaged sample curves and merge them into one.
A new merged curve will be plotted over the top of the other ten.

- Click **Clear** on the dp window, this will clear everything from the dp window.
- Now do the same for the scattering data for the appropriate buffer. Click **range** on dp, **select** on range parameter window then select first file for buffer and **apply**.
- Scale and Merge the buffer files by again clicking **scale**, unchecking damaged samples, then typing in an appropriate name for it to be saved as and click **merge** on dp as before.

You now have two merged curves saved, one for your sample and the other for the buffer. Clear the dp window before going on to the next stage – subtraction!

To subtract the Buffer from the sample you first have to load the sample and the buffer into the dp window, this is done slightly differently to loading ranges.
- **Click Select** on the dp window in row #1

- Locate and select your merged sample curve by double clicking.

- Click **Select** on the dp window in row #2 and double click on your merged buffer curve.

- **Plot**

These two curves also need to be scaled however you only want to scale the last half of the curves which are protein independent, not the guinier region. To do this you select only data points 500 onwards. In the dp window, two of the columns are headed nBeg and nEnd, changing these values for the files before a command means that only the data points set will be changed. So....

- In the dp window change the nBeg values in row #1 AND in #2 to 500 and click **Scale**.

You will again be able to see how far off each other the curves are by the multiplier value in row #2. It is less important to be under 5% similar.

Now the buffer curve can be subtracted away from the protein curve to remove any buffer scattering from the analysis. The precise concentration (in mg/mL) of the sample needs to be taken into consideration at this point.

- Change the nBeg for the two rows back to 1.

- Click **Plot**, this will show the whole scaled curve so you can check it looks correct.
- Enter in the concentration of your sample into the Conc column, in the sample row (#1)
- Type in what you want you want the subtracted curve to be saved as in the text box by the out box again and click Subtract.

A subtracted curve will overlay the previous two curves. Clear the dp window and load the subtracted curve before performing the guinier analysis.
- Click Clear in the dp window.
- Click the Select in the dp window for the first row (#1) and load the subtracted curve by double clicking it.
- Click autoRg on the gw tool bar.

The program will produce another window with the guinier analysis in (note- if you have not saved the ATSAS in the C drive the auto Rg function is unlikely to work). The data points and sRg limits it used for the analysis are shown as well as the Rg. This function is a very useful start to guinier analysis but it is important to do it manually as well. This is achieved by closing the autoRg windows so you have the gw and dp window back in action. You can then change the data region to find the guinier region of the curve.
- Close auto Rg windows.
- Click Guinier in the dp window.

A guinier line (red) will appear in the gw along with a residual line (green). The guinier region is generally between sRg values of 0.9-1.3 however can be longer. The guinier region is linear so to aid in finding the extent of the region the green residuals must be evenly spaced either
side of the pink line, you will see if you go outside of the guinier region the green residuals will curve meaning you’ve gone too far. To change the data points under Guinier analysis use the arrows to the right of the nBeg and nEnd. First reduce the nEnd number until about 200, then keep reducing until the guinier line covers a straight part of the data. The beam stop covers some of the scattering so some of the beginning points must also be removed; the number of points is dependant on the beam stop that was used in the experiment. You can see where the beam stop is by eye.

This Rg should be noted, how similar is it to the autoRg?? The data points used should be noted as the beginning point will be needed in the GNOM analysis later. In the data collection, a standard should also be measured, in this case it is the standard BSA, it is known to provide good scattering data as well as the Mw being known. The Io from the standard can be used to calibrate the beamline so a Mw weight for the unknown sample can be determined:

\[
\frac{\text{Standard Mw}}{\text{Standard Io}} = \text{Conversion factor}
\]

\[\text{Conversion factor} \times \text{sample Io} = \text{sample Mw}\]
GNOM

Reset primus to the subtracted, scaled curve with all the points.

- Change the nBeg to 1 and the nEnd to 9999 to include all data points again.
- Click Plot in the dp window.
- Click Gnom on the gw tool bar.

This will open a GNOM parameter window. The input is done automatically by primus and should be your subtracted curve. You can change the name that the output file will be saved by typing in the Output box. The other parameters which you set are much to do with trial and error. The first step is to uncheck the P(Rmax)=0 tick box, also enter the first data point for the guinier analysis found earlier in the nBeg box.

To begin with chose an Rmax which will be far too large, this is to give you an estimate of the true Rmax value. The P(R) curve will cross the x-axis around the true Rmax value (the Rmax you insert has to be in the same ball park for this to happen, with 10 fold differences the curve will look bazaar and may not cross the x-axis). Clicking Run on this GNOM parameter window will run the program GNOM with the set parameters.

A curve with a fit line will appear alongside the GNOM text window. The closer you are to having the correct parameters the closer the fit line will be to the curve. The trail and error part therefore is changing the parameters until you have the best fit. By pressing return on your keyboard the P(R) plot with open. This also indicates how close your parameters are to being optimal. If the curve seems to continue upwards at high x values, try a higher Rmax. If it...
goes into negative x-axis then try a lower Rmax. Ideally you want it coming down and just crossing the x-axis at the highest value of x, this is done via trial and error.

By decreasing the nEnd you can remove the noisy data from the analysis, which will help fit the line to the curve, again trial and error here too.

Reciprocal and Real space
Rg and Io given

If you press return again on your keyboard it will return to the GNOM parameter box allowing you to alter them and rerun. If you leave the output name the same the file will be overwritten.
Appendix VI– Beginners Guide to SANS Analysis

Data was collected from the ID22 at the ILL, Grenoble with beamline scientist Dr Phil Callow. This help is based on version Super GRASP beta V5.09.

Upon loading the program opens two windows, the main super grasp window and the text based box. The text based box logs everything you do within the user interface window, this is a useful to confirm when an action has been performed.

Setting up GRASP

1- Go to the instrument tab and make sure it is set to the correct beamline, in my case D22 (not D22raw). A tick will appear beside the set beamline. Click on your beamline if the correct one is not checked.

2- Under the File tab set the file and data directory, you should make a folder in which ALL of your raw data is, go to this and double click on just one of the files. Grasp will then take the numbers from this directory.

3- In Data – Data normalisation, you must select how you want your data to be normalised, usually the default monitor would suffice, however at the time of my analysis the monitor was not working so the data was normalised to time instead. Check with your beamline scientist who can advise you which is best at the time.

Loading Files

You need to load in your files whilst telling the program what the files are. You do this by

1. Choose the correct name for the file in the drop-down box in the Foreground.

2. Type in the file number into the Numbr box.

3. Press Get it!
First load the 17 meter (or 10 meter) data. The drop-down menu is written for physicists so just not directly translate for biological samples. Here is a list of what should be entered into the boxes for biological samples:

<table>
<thead>
<tr>
<th>GRASP Menu</th>
<th>Biological meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
<td>Sample scattering</td>
</tr>
<tr>
<td>Empty Cell</td>
<td>Buffer scattering</td>
</tr>
<tr>
<td>Cadmium</td>
<td>cadmium / B$_4$C scattering</td>
</tr>
<tr>
<td>Trans sample</td>
<td>Sample transmission</td>
</tr>
<tr>
<td>Trans empty cell</td>
<td>Buffer transmission</td>
</tr>
<tr>
<td>Trans empty Beam</td>
<td>Empty beam transmission</td>
</tr>
</tbody>
</table>

These are the only measurements you need to load. Which ever is selected in the foreground box, will be displayed above pictorially. After loading these data you must let the program know where the beam centre is.

1. Select Trans empty beam in the **Foreground** box.
2. Zoom in around the beam centre.
The \( c_x \) and \( c_y \) shown are the coordinates where the program thinks the centre of the beam is, check by eye that this is correct. You can zoom out by clicking on zoom out button and clicking in the viewing window.

**Transmissions**

By monitoring the transmissions you can obtain approximate D\(_2\)O concentration of your buffer and check your sample has been properly dialysed into the correct buffer.

1. Load the trans empty cell into the **Foreground**.
2. For the **Background**, there is another drop down menu, use the default of empty beam.
3. Press **Calc Trans**, this will divide the number of counts from the foreground sample by the number of counts in the background sample, in this case:
4. \[ T_e = \frac{\text{No. of buffer counts}}{\text{No. of counts in empty beam}} \]
5. Check the lock box next to \( T_e \) to keep this transmission setting throughout this sample.
If the $T_e$ is around 0.5 your buffer contains no $D_2O$ and is likely to be a $H_2O$ based buffer. As the $D_2O$ concentration increases in your buffer, the scattering decreases ($H_2O$ scatters more than $D_2O$) so the $T_e$ approaches 1. In my (limited) experience, 100% $D_2O$ buffer has a $T_e$ of 0.85.

To look at the sample vs. buffer transmission it is similar to before:

1. Change the **Foreground** to Trans sample.
2. The **Background** automatically defaults to Trans empty cell.
3. Clicking **Calc Trans** this will then work out the $T_s$.
4. Check the lock $T_s$ box.

$$T_s = \frac{\text{No. of sample counts}}{\text{No. of buffer counts}}$$

The protein being a relatively low concentration will not contribute significantly to the transmission of the sample, therefore the sample and buffer counts should be approximately the same, resulting of a $T_s$ of $\sim 1$. If this is not the case your dialysis may not have completed or the $D_2O$ may have exchanged with the air over time.
Masking the Beam Stop

Due to the collimation and detector moving such great distances from 2 to 17 meters, over the course of the experiment the beam position can alter very slightly. This means that sometimes it has moved between your sample and buffer readings causing inaccuracies close to the beam after subtraction. To overcome this you can mask the beam stop and some of its surrounding area to remove faulty subtractions. Load sample into the foreground and zoom in around the beam like before.

1. Under Analysis tab, click Mask Editor
2. This window will appear. Click sketch
3. With your mouse outline a box ~2 squares around the apparent beam stop.
4. Under Lines, type x1,x2,x126,x127,y1,y2,y126,y127 and click the + box next to it.

This will block out the area around the beam stop and take it out of analysis. The detector is made up of tubes, which when in close contact with other tubes optimises their working capacity. The tubes therefore on the outside of the detector are less efficient than the others and are also removed from analysis. To remove these lines you type in x1,x2,x126,x127,y1,y2,y126,y127 in the lines box, and click +.

To check the masking you have done has been effective in blocking out the unwanted data, you can plot the data intensity against Q.
1. Check the boxes next to **Background** and **Cadmium**; this subtracts the background (buffer) and the cadmium scattering from the **Foreground** (sample).

2. **Analysis – Averaging: Radial and Azimuthal**

3. In new Averaging window click **I vs. (q)**

A new window will appear displaying the graph, an example of an unmasked (A) and masked (B) and curve are shown below. The low readings at low Q around the beam stop have been removed in B.

![Graph A](image1.png)

![Graph B](image2.png)

You have finished loading in your 17m data, now for the 2m.

1. In the **Foreground** change the number to worksheet 2 and uncheck the **Background** and **Cadmium** subtractions (uncheck the boxes).

2. As you did for the 17m data you need to load in the files. You don’t need to look at the transmissions again though, so load the 2m data for:

   Sample
   
   Empty Cell (buffer)
   
   Cadmium
   
   Empty Beam

This is performed in the same way as before, load the **Foreground**, type in the file number and **Get It!** The Beam Centre needs to be found again, so as before go to empty beam, zoom in
and centre calc. Check by eye that also looks correct. You need to mask this sample as well in the same way as for the 17m data (see above).

**Calibrate.**

The last thing to do is calibrate the 17m and 2m data. Check the **Calibrate box** in the main window, **Calibration Options** window will appear, choose **direct beam** (this may change depending on how you normalise your data, check with your beam line scientist).

Now with the **Background** and **Cadmium** subtracted you can plot an **I vs. (q)** plot for your 2m data, click **hold** in the graph window. Go back to your 17m data (back to worksheet 1) and click **I vs.(q)** again and you should now see both of your graphs in the same window, they should meet each other.

**Export the data**

1. **File – Export Data - Export Data Options** – uncheck **Auto filenames**

2. **File – Export Data - Export Data**

3. **Name curves.**

**Analysing Curves**

Perfect data can be loaded straight into Primus and curves can be merged. However, I have found that if your buffer subtraction is not perfect you get some negative points at high Q values which primus doesn’t like. The merging can also go wrong due to a bug in primus. The best thing to do is to load both the curves into Excel (delimited). Find your lowest negative number and add this constant to the entire x axis (2 and 17m), this will normally be approx 0.005. You can now load the curves into Primus.

Work out how many points you need to remove from the end of the 17 metre data so the curves don’t overlap. Re load in Excel, delete appropriate points and add on the 2m data to the end of the 17m, save as one curve. This can now be used in primus to obtain guinier regions and P(r) plots.
Appendix VII – Densitometry Trace

Below is an example of the densitometry traces used for determination of stoichiometries of the xNAP-Histone complexes. The corresponding lane within the SDS PAGE of NAP₂·H3/H4 is shown on the left with the densitometry of the lane on the right. The xNAP, H3 and H4 are highlighted by red arrows.

[Image of densitometry trace with red arrows indicating xNAP, H3, and H4]
Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.