Crystallographic studies of chromatin-targeting metal-based compounds and nucleosome core particles assembled with engineered cohesive end DNA fragments

Louis, DeFalco Jr.

2019

Louis, D. Jr. (2019). Crystallographic studies of chromatin-targeting metal-based compounds and nucleosome core particles assembled with engineered cohesive end DNA fragments. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/89610>

<https://doi.org/10.32657/10220/47720>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**Crystallographic Studies of Chromatin-Targeting
Metal-Based Compounds and Nucleosome Core
Particles Assembled with Engineered Cohesive End
DNA Fragments**

Louis DeFalco Jr.

School of Biological Sciences

2019

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

13-Feb-2019
.....
13-Feb-2019

.....
Louis DeFalco Jr.

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism and the research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

Feb 13, 2019
.....
13-Feb-2019

.....
Associate Professor Curtis A. Davey

**Thesis Title: Crystallographic Studies of Chromatin-Targeting Metal-Based
Compounds and Nucleosome Core Particle Assembled with Engineered Cohesive
End DNA Fragments**

I have excluded the Authorship Attribution Statement declaration because the thesis does not contain any published material.

Louis DeFalco Jr.

Associate Professor Curtis A. Davey

Date :

Date:

13-Feb-2019

Feb 13, 2019

Acknowledgements

I would like to express my deepest gratitude to Associate Professor Curtis A. Davey for giving me the opportunity to join his group and work on these projects, inspiring and encouraging curiosity and creativity throughout scientific investigation, and for being an endless source of guidance, advice, and support during my time in the lab.

I am grateful and thankful to all previous and current lab members I have worked with: Dr. Gabriela Davey, Dr. Deepti Sharma, Dr. Sivaraman Padavattan, Dr. Eugene Chua, Dr. Qiuye Bao, Dr. Zenita Adhireksan, Dr. Zhujun Ma, Mr. Wayne Shum, Dr. Sailatha Ravi, and Ms. Rachael Lee. Their insight, energy, and enthusiasm for science make the lab an outstanding working environment. I would like to extend extra thanks to Deepti, Siva, and Zenita for numerous helpful discussions and their generous advice and tutelage.

I would like to thank Professor Paul J. Dyson, Professor Ursula Roethlisberger, Dr. Lucinda K. Batchelor, and Mr. Thibaud von Erlach of École Polytechnique Fédérale de Lausanne (EPFL) and Dr. Alexey A. Nazarov of Moscow State University for providing metal-based agents and their vital feedback and involvement. I would also like to thank the Paul Scherrer Institut and the Swiss Light Source beam line scientists at PXIII for their invaluable help during data collection and the NTU Institute of Structural Biology for their X-ray facilities. Finally, I would like to thank NTU, the School of Biological Sciences, and the A*STAR SINGA scholarship for the financial support that allowed me to study in Singapore and pursue my passion for science.

I would like to specifically thank and acknowledge Professor Paul J. Dyson and Dr. Lucinda K. Batchelor of EPFL for performing, preparing, and interpreting Compound 3 cytotoxicity assay data; Dr. Deepti Sharma of Nanyang Technological University (NTU) for performing and analyzing Compound 3 tyrosine fluorescence spectroscopy experiments; Dr. Zenita Adhireksan and Dr. Sailatha Ravi of NTU for collecting Compound 1 and Compound 4 crystallographic data; and Professor Ursula Roethlisberger and Mr. Thibaud von Erlach for performing and interpreting classical molecular dynamics simulations of Compound 3–nucleosome core particle adduct structures.

Table of Contents

Statement of Originality	ii
Supervisor Declaration Statement	iii
Authorship Attribution	iv
Acknowledgements	v
Table of Contents	vi
List of Figures	vii
List of Tables	ix
Abbreviations	x
Abstract	xi
Chapter 1 – Introduction	1
1.1 Chromatin Structure and Dynamics	1
1.2 Nucleosome Core	4
1.2.1 NCP Structure	4
1.2.2 Significance of Nucleosome DNA Deformation	8
1.2.3 Octamer as peptide binding, drug target, and NCP interaction scaffold	11
1.3 Metal-based drugs in cancer therapy	14
1.3.1 Origins, chromatin targeting, and mononuclear designs	14
1.3.2 Multinuclearity, peptide-metal conjugates, and NCP site crosslink	21
1.4 Scope	23
Chapter 2 – Materials and Methods	26
2.1 DNA production and NCP construct cloning	26
2.1.1 Large scale production of DNA601L	26
2.1.2 Large scale production of DNA145	28
2.1.3 Cohesive end NCP DNA cloning	30
2.1.4 Large scale production of NCP147SK/XK/K	30
2.2 Core histone production, purification, and histone octamer refolding	30
2.2.1 Core histone and H2A.Z variant production and purification	30
2.2.2 Histone octamer refolding and purification	32
2.3 NCP reconstitution and crystallization	33
2.3.1 Human NCP145 reconstitution	33
2.3.2 Human NCP145 crystallization	34
2.3.3 Cohesive end NCP reconstitution and crystallization	34
2.4 Crystal soaking and data collection and processing	35
2.4.1 Human NCP145 crystal soaking	35
2.4.2 Data collection and processing	35
2.5 Compound III biochemical assays	36
2.5.1 Cross-linking analysis of treated nucleosome core particle	36
2.5.2 Tyrosine fluorescence spectroscopy	36

Chapter 3 – Results and Discussion	38
3.1 Cohesive end 601L cloning and non-palindromic designs	38
3.1.1 Cohesive end NCP DNA design	38
3.1.2 Palindromic cohesive end NCP DNA cloning	41
3.1.3 Non-palindromic cohesive end NCP DNA constructs	45
3.1.4 Cohesive end NCP DNA preparation	47
3.2 Preliminary models of NCP with incompatible overhand termini	50
3.2.1 Preliminary structure of human NCP with non-complimentary 3' termini	53
3.2.2 Structure of the human NCP with incompatible 5' and 3' overhangs	59
3.3 Crystal structure of human NCP with base pairing termini	63
3.3.1 Design, reconstitution, high-resolution and DNA termini structures	63
3.3.2 NCP147K DNA structure, core particle stability and oligomeric structure	68
3.4 Human NCP145 preparation and crystallization	81
3.4.1 Human NCP145 reconstitution	81
3.4.2 Human NCP145 crystallization	82
3.5 High-resolution structure of human NCP145	87
3.5.1 Comparison of human NCP145 and <i>X. laevis</i> NCP145	87
3.5.2 DNA conformational changes in human NCP145	91
3.6 Adduct structures of the human nucleosome core particle	95
3.6.1 JN128-2_KP2425 Pt(IV)–Ru(II) hetero-dinuclear NCP145 structure	97
3.6.2 EP11-019, an auranofin derivative and hetero-dinuclear precursor	102
3.6.3 A hetero-dinuclear Ru(II)–Au(I) agent targets distinct nucleosome sites	108
3.6.4 A trinuclear derivative of Compound III forms similar NCP adducts	121
3.6.5 Preliminary data of Au(I)–viral peptide conjugates bound to nucleosome	130
3.7 Concluding remarks and projected outcomes	144
Appendices	145
References	151

List of Figures

Chapter 1 – Introduction

Figure 1. Eukaryotic DNA compaction	2
Figure 2. NCP structure at 1.9 Å resolution	4
Figure 3. Schematic of the nucleosome core structure	5
Figure 4. Summary of histone octamer – NCP DNA interactions	7
Figure 5. DNA bending modes in the NCP147 structure	10
Figure 6. NCP DNA stretching	11
Figure 7. The nucleosome acidic patch	12
Figure 8. Overview of NCP crystal packing interactions	13
Figure 9. Epigenetic control of gene activity leading to tumorigenesis	14
Figure 10. Mononuclear chromatin targeting reagents	19
Figure 11. Map of metaloreagent binding sites on the nucleosome	21
Figure 12. Project overview and significance	25

Chapter 2 – Materials and Methods

Figure 13. 601L DNA production and purification	27
Figure 14. DNA145α production and purification	29
Figure 15. Human core histone production and purification	32

Figure 16. Human octamer purification	33
---	----

Chapter 3 – Results and Discussion

Figure 17. Blunt end NCP termini base stacking	41
Figure 18. Cohesive end DNA primer design	42
Figure 19. Digestion and ligation strategy for palindromic NCP DNA constructs	43
Figure 20. Cloning of cohesive end NCP DNA	45
Figure 21. Cohesive end NCP DNA design	47
Figure 22. Cohesive end NCP DNA production and purification	49
Figure 23. NCP termini base stacking and pairing	54
Figure 24. NCP147SK crystallization	55
Figure 25. NCP145SK crystallization	57
Figure 26. Preliminary termini structure of NCP147SK	58
Figure 27. NCP termini base pairing modes	59
Figure 28. NCP147XK crystallization	60
Figure 29. Preliminary termini structure of NCP147XK	62
Figure 30. NCP147K crystallization	64
Figure 31. Crystal structure of the NCP147K nucleosome	65
Figure 32. NCP147K termini electron density	66
Figure 33. NCP147K termini geometry	67
Figure 34. NCP147K DNA structure	70
Figure 35. NCP147K/NCP601 atomic displacement factors	71
Figure 36. NCP147K termini energy of stabilization	72
Figure 37. NCP147 histone octamer atomic displacement factors	73
Figure 38. Divalent cation binding at the NCP147K termini	75
Figure 39. Divalent cation coordination structure at the NCP147K terminus	76
Figure 40. NCP147K inter-particle distance	78
Figure 41. Eleven nanometer chromatin fiber	79
Figure 42. NCP147K oligomeric structure	80
Figure 43. Human NCP145 reconstitution	82
Figure 44. NCP145 crystal morphologies	83
Figure 45. NCP145 microcrystals from divalent metal screens	84
Figure 46. NCP145 microcrystals from monovalent metal screens	85
Figure 47. Free DNA affects NCP crystal morphology	87
Figure 48. High-resolution X-ray structure of human NCP145	89
Figure 49. Superposition of human NCP145 and <i>X. laevis</i> NCP145	90
Figure 50. Average RMSD between human NCP145 and <i>X. laevis</i> NCP145	91
Figure 51. Human NCP145 DNA base pair geometry	93
Figure 52. Structural determinants of human NCP145	95
Figure 53. Chemical structures of chromatin targeting reagents	97
Figure 54. Chemical structure of JN128-2_KP2425	98
Figure 55. JN128-2_KP2425 binding to the human nucleosome	101
Figure 56. JN128-2_KP2425 binding sites	102
Figure 57. EP11-019 binding to the human nucleosome	106
Figure 58. The EP11-019–RAPTA-T adducted human nucleosome structure	107
Figure 59. Bond lengths and angles of free and adducted EP11-019	108
Figure 60. Chemical structure of ‘Compound III’	109
Figure 61. Histone-histone cross-linking analysis of NCP treated with C3	112
Figure 62. Cross-linking by C3 increases nucleosome stability	113
Figure 63. ‘Compound III’ allosteric nucleosome binding sites	114
Figure 64. Evidence for ‘Compound III’ inter-particle crosslinking	115
Figure 65. ‘Compound III’ nucleosome site selectivity	117

Figure 66. ‘Compound III’ crosslinking on the occluded nucleosome face	120
Figure 67. Occluded site locations and distances for ‘Compound III’ crosslinking ..	121
Figure 68. Chemical structure of trinuclear ‘Compound III’	122
Figure 69. Trinuclear ‘Compound III’ binding and metal site locations	125
Figure 70. Extent of trinuclear ‘Compound III’–nucleosome adduct formation	126
Figure 71. Occluded binding site locations of trinuclear ‘Compound III’	128
Figure 72. Viral peptide binding to the nucleosome surface	131
Figure 73. Viral peptide–gold(I) conjugate chemical structures	132
Figure 74. Chemistry of LANA M6O	133
Figure 75. LANA peptide diffraction patterns	133
Figure 76. Inter- and intra- molecular interactions of LANA and LANA M6O	134
Figure 77. LANA M6O binding and interaction with the nucleosome	136
Figure 78. Au-LANA binding to the human nucleosome	140
Figure 79. Au-GAG binding to the human nucleosome	142
Figure 80. Gold(I)–viral peptide conjugates elicit allosteric cross-talk	143

List of Tables

Chapter 2 – Materials and Methods

Table 1. Soaking conditions for compounds tested	36
--	----

Chapter 3 – Results and Discussion

Table 2. Dinucleotide step stacking energies	46
Table 3. Data collection and refinement statistics for cohesive end NCPs	51
Table 4. Cohesive end NCP crystallographic data	51
Table 5. Crystallization conditions of nucleosome complex structures	52
Table 6. Crystallographic and geometric data for manganese cations	77
Table 7. NCP145 low divalent ion screen	84
Table 8. NCP145 low monovalent ion screen	85
Table 9. Data collection and refinement statistics for human NCP145	88
Table 10. Data collection statistics for JN128-2_KP2425–NCP145	99
Table 11. JN128-2_KP2425 metal binding sites	100
Table 12. Data collection and refinement statistics for EP11-019 treated NCP	104
Table 13. EP11-019 metal binding sites	105
Table 14. Data collection and refinement statistics for C3 treated NCP crystals	110
Table 15. C3 metal binding sites	111
Table 16. C3 metal–metal linear distances between RU1 and gold sites	116
Table 17. Difference and anomalous peaks at C3 RU1’’ and RU2’	118
Table 18. Data collection and refinement statistics for tnC3 treated NCP	123
Table 19. tnC3 metal binding sites	124
Table 20. C1 and C4 adduct sites	129
Table 21. Data collection and refinement statistics for viral peptide treated NCP	138

Abbreviations

AUF	Auranofin
CBS	Chromatin binding domain
CIP	Calf intestine alkaline phosphatase
C1	Compound I
C3	Compound III
C4	Compound IV
EDTA	Ethylenediaminetetraacetic acid
FPLC	Fast protein liquid chromatography
GAG	Group-specific antigen
HMGN	High mobility group nucleosome-binding protein
HO	Histone octamer
HSAB	Hard and soft acid – base
KSHV	Kaposi's sarcoma herpes virus
LANA	Latency-associated nuclear antigen
MPD	2-methyl-2,4-pentandiol
NBD	Nucleosome binding domain
NCP	Nucleosome core particle
Ni-NTA	Nickel-nitrilotriacetic acid
NLHA	Nucleosome-linker histone assembly
NUC	Nucleosome
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PVF	Prototype foamy virus
RMSD	Root mean square deviation
SASA	Solvent accessible surface area
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SHL	Super helix location
tnC3	Trinuclear Compound III

Abstract

In eukaryotic cells, related functions like the storage, protection, and precise regulation of genomic DNA are achieved by its compaction into protein–nucleic acid complex structures known as chromosomes. During the initial steps of compaction, naked DNA is wound around an octamer of histone proteins in a process that forms the nucleosome core. This nucleoprotein complex creates increasingly higher-order structures that, with the addition of linker histone proteins, ultimately generate densely packed chromatin. It is established that tumor cells present distinct dissimilarities in chromatin dynamics, nucleosome positioning, and epigenetic modifications like DNA methylation and post-translational modifications relative to healthy cells, and this distinction is significant in the context of genomic regulation during the eukaryotic cell cycle. With this knowledge, chromatin reactivity toward exogenous molecules is the primary principle concerning rational design of therapeutic compounds that separately, or simultaneously, target protein and DNA motifs of the nucleosome core. Beyond the nucleosome core, a comprehensive understanding of genomic structure and nucleosome–nucleosome arrangement requires the availability of minimal and reproducible systems for investigation. Understanding gained from new nucleosome structures will illuminate heretofore unknown details regarding the interplay between nucleosome dynamics and genetic regulation and their collective genome-wide impact.

Earlier studies from our collaborators have shown a synergistic effect between the antirheumatic compound auranofin (AUF), possessing a gold(I) reactive center, and the ruthenium(II)-based anticancer drug RAPTA-T. Remarkably, the synergism observed as a decrease in tumor cell viability coincides with an allosteric mechanism by which binding of RAPTA-T to its characteristic target sites on the H2A:H2B acidic patch (sites RU1 and RU2) facilitates the binding of AUF to a distinct histidine residue (H113) found on core histone H3 near the nucleosome dyad. Here in a follow-up study, Compound III (C3), a novel hetero-dinuclear compound containing distantly-linked ruthenium(II) and gold(I) metal centers, was found to form adducts at canonical nucleosome ruthenium and gold binding sites by exploiting the allosteric crosstalk between distant regions of the histone octamer surface to cross-link tetramer with dimer. Mononuclear C3 precursors, a trinuclear C3 derivative, and two viral peptide–metalloreagent conjugates possessing AUF-like gold(I) reactive centers were additionally employed in an attempt to capitalize on the dimer–tetramer allosteric

pathway towards therapeutic gain, with structural data for each presenting varying degrees of effectiveness.

Adapting established cloning techniques to produce palindromic nucleosome DNA sequences, we were able to construct nucleosomes containing linker regions that maintain the capacity to form base pairs at their termini. Controlling this process allowed for the creation of well diffracting crystals of nucleosome-linker histone assemblies. Building on this concept, several new nucleosome DNA construct designs are presented here, all of which are based on the 601 Widom high affinity sequence. This has allowed improvement of technical issues involved in nucleosome core particle (NCP) DNA production as well as the elucidation of NCP crystallization kinetic principles, which offer a degree of control over NCP crystal lattice formation. As such, this approach holds potential for facilitating acquisition of NCP structures with diverse, genetically significant DNA sequences, histone proteins and associated chromatin factors. Additionally, a high-resolution crystal structure of a cohesive-end NCP reveals a novel multimeric arrangement that may approximate specific instances of chromatin fiber packing.

Chapter 1 – Introduction

1.1 Chromatin Structure and Dynamics

Eukaryotic cells package genomic material into chromosomes. These bodies are dense, highly folded, nucleoprotein complexes that reside in the nucleus (Smith, E. and Shilatifard, A., 2010). Nearly all cellular activity is directly governed by the degree of nuclear protein access to genomic DNA. (Fierz, B., 2015; van Holde, K.E., 1989). The basic repeating unit of chromatin is the nucleosome (Fig. 1) (van Holde, K.E., 1989). In the genome, nucleosome core regions are connected to one another by a variable section of linker DNA ranging from ten to ninety base pairs in length (Cutter, A.R and Hayes, J.J., 2015). At this level of organization, the appearance of chromatin assumes a ‘beads-on-a-string’ structure with an average diameter of approximately 11 nm (Fig. 1) (Kornberg, R.D., 1974; Oudet, P. *et al.*, 1975).

Further condensation of the 11 nm structure yields the 30 nm fiber (Fig. 1), an arrangement of chromatin that is the most pervasive at physiological salt conditions *in vitro* (Woodcock, C.L. and Dimitrov, S., 2001; Dorigo, B. *et al.*, 2004; Hansen, J.C., 2002). In the fiber state the side-to-side packing of nucleosomes proceeds perpendicular to the fiber axis causing the structure to experience further compaction, transitioning through several intermediate chromosomal forms before it becomes recognizable as a mitotic chromosome (Fig. 1) (McGinty, R.K and Tan, S. 2014). These levels of structural organization and sequential condensation are abetted by various chromatin binding factors such as internally positioned linker histones (H1) (Song, F. *et al.*, 2014; Hergeth, S.P. and Schneider, R., 2015). However, the overall process remains poorly understood in terms of identified transition structures and quantified kinetic and thermodynamic parameters. There also remains a dearth of detailed knowledge regarding stable higher-order chromatin structures; a problem that persists despite high resolution NCP structural data (Fig. 2) (Davey, C.A. *et al.*, 2002; Luger, K. *et al.*, 1997a).

Figure 1. Outline of DNA's sequential folding into higher-order nucleoprotein complexes. The culmination of this process leads to a complete mitotic chromosome. Adapted from Alberts, B. *et al.*, 2002.

Upon discovery, chromatin was thought to exist only for satisfying the spatial requirements of DNA packaging; the highly convoluted structure is used to condense two meters of genetic material into a nucleus of less than 10 micrometers (Felsenfeld, G. and Groudine, M., 2003). However, further investigations revealed that chromatin is also intimately involved in gene expression, acting itself as a general transcription repressor (Fierz, B., 2015). Highly compact chromatin structures severely restrict protein factor access to DNA (Polach, K.J. and Widom, J., 1995; Li, G. and Widom, J., 2004). Regions that are condensed generally experience less active transcription due to regulatory site occlusion, while regions that are arranged more loosely are more actively transcribed (Kornberg, R.D. and Lorch, Y., 1991). This level of genomic control is a primary factor concerning cellular differentiation in multicellular eukaryotic organisms where the same genetic material is present but genes are expressed discretely (van Holde, K.E., 1989). Despite this relatively static view, chromatin is highly dynamic. Nucleosome positioning is subject to perpetual remodeling by ATP-dependent enzyme complexes, and is influenced by core histone composition, linker histone abundance, DNA methylation patterns, and other effector binding/modifications. Therefore,

chromatin exists in numerous structural states beyond merely binary ‘on’ or ‘off’ modes (Clapier, C.R. and Cairns, B.R., 2009; Fierz, B., 2015). Chromatin continuously condenses and expands causing discontinuous transcription resulting in transcriptional bursting (Clapier, C.R. and Cairns, B.R., 2009; Kaochar, S. and Tu, B.P., 2012). Chromatin dynamics also influence the regulation of other nuclear processes like DNA replication, repair and recombination (Fierz, B., 2015).

In the context of rational drug design towards anticancer therapeutics, it is critical to note that gene transcription and DNA replication, repair, and recombination are the dominant cellular processes that require nucleosome unpacking. Any disruption of the dynamic nature of chromatin will directly influence transcription and DNA replication. Increased rates of transcription and replication, and by extension cellular division, are hallmarks of oncogenic cells. Therefore, any compounds that disrupt or modulate chromatin dynamics should have a greater relative impact on tumor cells compared to normal cells. The altered gene expression profile and DNA synthesis rates of tumor cells are in part caused by characteristic nucleosome positioning and dynamic configurations. Distinct changes in core histone composition and modifications, linker histone composition and occupancy, transcription factor/activator/repressor site availability, DNA methylation and acetylation patterns, and a host of other determinants under the direct influence of nucleosomes are also not shared by normal cells (Roy, S. *et al.*, 2010; Kharchenko, P.V. *et al.*, 2010; Ernst, J. *et al.*, 2011). Therefore, the nucleosome is an ideal target for rational drug development considering the abundance of conceivable DNA heterogeneity and plethora of core histone sites.

Two key factors influence nucleosomal DNA recognition by nuclear proteins and the binding of small molecules, and may determine drug accessibility, reactivity, and efficacy; Solvent Accessible Surface Area (SASA) and conformational sampling (Pryciak, P.M. and Varmus, H.E., 1992; Makde, R.D. *et al.*, 2010; Yamada, K. *et al.*, 2011). The SASA of nucleosomal DNA is vastly reduced relative to free DNA as a result of folded chromatin’s condensed state, which causes general inhibition of effector binding (Davey, C.A. *et al.*, 2002; Fierz, B., 2015). Induced high curvature bending of the DNA double helix is initiated by superhelical wrapping around the histone octamer and forces distinct sequence and positioning dependent nucleosomal DNA conformations, which possess signature structural motifs (Richmond, T.J. and Davey, C.A., 2003; Chua, E.Y. *et al.*, 2012). Consequently, nuclear proteins and small

molecules that interact with naked DNA may not interact with the highly deformed DNA of the nucleosome core, or their binding modes may be altered. Therefore, it is more appropriate, and informative, to study these interactions in the context of the nucleosome by making use of nucleosomal DNA constructs as an investigative tool for defining molecular interactions with the eukaryotic genome.

1.2 Nucleosome Core

1.2.1 NCP Structure

In isolation, the nucleosome core particle (NCP) is the irreducible repeating unit of eukaryotic chromosomes and the basic component of nucleosomes in all forms of chromatin. The nucleosome initiates an approximately seven-fold linear compaction of genomic DNA *in vivo* through its packing, which is facilitated via nucleosome face to nucleosome face interactions (Cutter, A.R. and Hayes, J.J. 2015). The nucleosome core consists of a 145 to 147 base pair segment of DNA wound in tight association around an octamer of highly basic histone proteins in ~ 1.67 left-handed superhelical turns (McGinty, R.K. and Tan S. 2014). The octamer contains two copies of each core histone (H2A, H2B, H3, and H4) and for organization purposes, core histones are designated ‘prime’ or ‘nonprime’, to help distinguish between each copy (Fig. 2). ‘Prime’ core histones contact the ‘minus’ or negative half of the DNA, which counts down from the dyad at nucleotide position 0, and the nonprime histones contact the ‘plus’ or positive half, counting up from the dyad (Fig. 2a, red arrow).

Figure 2. Overall structure of the NCP at 1.9 Å resolution. The 147 bp strands of DNA are colored orange and cyan with core histones colored individually: H3 (blue), H4 (green), H2A (yellow), H2B (red). **(a)** View of the NCP along the DNA superhelical axis showing the pseudo two-fold axis running vertically through the central base pair demarcated by the black arrow. The red arrow indicates the nucleosome dyad at base pair 0. **(b)** and **(c)** Top and bottom perspectives are shown respectively with the pseudo-two-fold axis indicated by a black dot.

The core histone octamer and DNA form an oblate disc 5.5 nm in height and 11 nm in diameter for the average structural dimensions (McGinty, R.K. and Tan, S. 2014). This structure is nearly invariable amongst all eukaryotes (Tsunaka, Y. *et al.*, 2005). While tight association of DNA and histone octamer confers nucleosomal DNA with protection from nuclease digestion, linker DNA is rapidly digested leaving only 145 – 147 base pairs of protein bound DNA, which delineates the core particle (Fig. 2) (McGinty, R.K. and Tan, S. 2014).

To date, the highest resolution nucleosome core structure available is that of a core particle containing 147 base pairs of palindromic DNA (Fig. 2) (Davey, C.A. *et al.*, 2002). The structure was solved to 1.9 Å and reveals an extensive solvent network that helps stabilize histone–DNA interactions, and an unprecedented level of histone tail completeness (Davey, C.A. *et al.*, 2002). Structural analysis depicts a molecule with an axis of pseudo-two-fold symmetry, running through the dyad, separating the two symmetry-related NCP halves (Fig. 2). Associations between the DNA and histone octamer are semi-specific where major and minor grooves will alternate between facing inward and outward relative to the octamer surface. This groove alternating along the double helical axis allows interrogation of structural DNA motifs by specific histone

residues and results in sequence-dependent octamer positioning by the DNA. As the double helix turns around the octamer a superhelix is formed. The position of the superhelix is described as a superhelical location (SHL), a representation of the number of turns the double helix has made at any point around the octamer beginning from the central base pair designated SHL 0 (Fig. 2a, 3, & 4).

The first step of *in vitro* NCP reconstitution is the creation of histone octamer from individual core histone proteins. Pre-forming octamer is essential, as the molecule's highly basic surface is the defining factor governing overall nucleosome structure. Core histones adopt a characteristic fold termed the 'histone-fold', which is a well-conserved structural motif (Freeman, L. *et al.*, 1996). The histone-fold consists of three alpha helices ($\alpha 1$, $\alpha 2$, $\alpha 3$) connected by short loops (L1, L2) (Fig. 3). These features allow the histones to interact through an extensive protein-protein interface that guides the heterodimerization of H2A:H2B and H3:H4 (Arents, G. *et al.*, 1991; Arents, G. *et al.*, 1993). Histone dimers associate with one another through four-helix bundles at the dimer-dimer interface to produce the H3:H4 tetramer, via the H3-H3' interface, and the H2A:H2B dimers (Fig. 4). The H3:H4 tetramer and the H2A:H2B dimers interact through an additional four-helix bundle between H4 and H2B to complete the octamer thereby forming the DNA binding surface (Fig. 4). Additionally, the packing interactions of the nucleosome in chromatin are also directed in part by histone topography. Nucleosome packing into higher order chromatin states is stabilized when mediated by H4-tail-H2A:H2B dimer contact.

Figure 4. General summary of histone octamer–NCP DNA interactions based on the 1.9 Å NCP structure. Individual histones and DNA strands colored as in figure 2. Structural and chemical determinants of octamer association and positioning are represented as the sum of polar, electrostatic, and hydrophobic interactions between the octamer surface and DNA double helix.

Histone–DNA interactions occur primarily between the DNA phosphodiester backbone of the central 120 base pairs and the core histone-fold motifs (Fig. 3). The paired-loop and paired-end-of-helix motifs of the histone-fold domains contact minor groove-inward regions of the DNA backbone that face the octamer surface (Figs. 3 & 4). The side chains of twenty amino acids, primarily arginine and lysine, are inserted into all fourteen octamer facing minor grooves. Here they form polar or stacking interactions with DNA phosphodiester bonds and base edges to help stabilize DNA positioning, bending, and global conformation. DNA termini (SHLs ± 5.5 to ± 7) binding to octamer is weaker than the core (SHLs 0 to ± 5) and thus exhibit more flexibility (Battistini, F. *et al.*, 2010; Chien, F.T. and van der Heijden, T. 2014). The binding of core DNA to the octamer and its positioning around the octamer surface derives in part from interactions with H3 (α Ns and N-terminal tails), and H2B tails (not shown). These core histone domains thread through minor groove channels between neighboring DNA gyres around SHL ± 1 (H3/H3') and SHL ± 4.5 (H2B/H2B') (Fig. 4). Water molecules of the solvent

network also mediate bridging interactions between distant histone and DNA moieties (Davey, C.A. *et al.*, 2002). While many of these interactions are independent of DNA sequence, a feature that allows nucleosome formation at nearly all genomic DNA sequences, a degree of intrinsic sequence-dependency is exploited for specific functions and differential stability.

In comparison to free DNA, or the DNA of protein–nucleic acid complexes, nucleosomal DNA adopts distinct conformations with helical deformations that widely deviate from standard B-form conformation. Free B-form DNA has a persistence length of approximately 150 base pairs, whereas the 1.67 left-handed superhelical turns of nucleosomal DNA contain about 80 base pairs per turn (Richmond, T.J. and Davey, C.A. 2003; McGinty, R.K. and Tan, S. 2014). Discrepancies in persistence length represent a mechanical disadvantage when wrapping DNA around histone octamer. In order to assume the distorted superhelical configuration, nucleosomal DNA must be bent around the octamer by excessive roll into octamer facing major and minor grooves, with almost no influence from base pair tilt (Luger, K. *et al.*, 1997a; Richmond, T.J. and Davey, C.A. 2003; Tolstorukov, M.Y. *et al.*, 2007).

As the superhelix rises around the octamer, base pair slide occurs at octamer facing major and minor grooves. High roll values observed at these locations indicate points of superhelix bending. However, the distribution of curvature is unequal, which leads to particularly sharp bends at SHL ± 1.5 and SHL ± 4.5 (Luger, K. *et al.*, 1997a). Histone proteins define the overall nucleosome shape by exerting deformations across the DNA via their preference for particular elements of the nucleotide sequence (Hayes, J.J. *et al.*, 1991; Bao, Y. *et al.*, 2006). Such elements, like flexible dinucleotide steps at locations of minor groove bending, or decreases in helical twist, lead to sequence-dependent translational positioning of nucleosomes throughout the genome (Luger, K. *et al.*, 1997a; Richmond, T.J. and Davey, C.A. 2003; Hayes, J.J. *et al.*, 1991; Hayes, J.J. *et al.*, 1990).

1.2.2 Significance of Nucleosome DNA Deformation

NCP DNA is highly deformed and deviates from naked B-form DNA significantly in certain structural parameters. To achieve optimal histone octamer positioning the double helix must wrap around the octamer surface in a manner that promotes tight association. Double helical wrapping increases base pair step curvature to accommodate

the path of the superhelix and forces bending at certain superhelical locations (Richmond, T.J. and Davey, C.A. 2003). Superhelical bending creates considerable mechanical stress on base pair–base pair interactions and phosphate backbone conformation. Such stress at weakly positioning, or less stable, SHLs induces local deformations like base pair step unstacking, DNA stretching and major/minor groove kinking. The net effect of these local disruptions helps the nucleosome maintain strong contacts within the histone–DNA register (Fig. 5). Segments of NCP DNA that are bent into the minor groove experience kinking or alternate base pair shifting to alleviate mechanical stress accompanied by over-twisting (Fig. 5) (Ong, M.S. *et al.* 2007; Richmond, T.J. and Davey, C.A. 2003). Moreover, stretch deformations are more commonly observed in NCP structures assembled with shorter DNA molecules. For example, an NCP structure with 145 base pairs of DNA (NCP145) versus 147 base pairs (NCP147) displays regions of stretching associated with areas of extreme minor/major groove kinking/unstacking and DNA over-twisting (Fig. 5). These types of local structural rearrangements bear *in vivo* significance as their presence in chromatin could affect its degree of compaction or influence site recognition by DNA binding proteins (Ong, M.S. *et al.* 2007).

Insertion of arginine side chains at all minor groove-inward locations in NCP147 induces minor groove bending and causes narrowing of the groove width (Richmond, T.J. and Davey, C.A. 2003). Smooth minor groove bending across the H3:H4 tetramer occurs when blocks of four base pair steps alternate. These blocks contain at least one GC=CG, GG=CC, or AG=CT base pair step that shifts to relieve steric clash between base pair edges (Fig. 5b) (Richmond, T.J. and Davey, C.A. 2003). Kinked minor groove bending across the H2A:H2B dimer occurs at flexible CA=TG base pair steps, and accounts for the absolute curvature into the superhelix of the entire base pair step block (Fig. 5). Kinked bending into the minor groove causes a concurrent base pair step unstacking at the opposing major groove (Fig. 5c). In the cell, DNA stretching occurs at an average of one to two base pairs for every nucleosome making this a genome-wide phenomenon (Richmond, T.J. and Davey, C.A. 2003). For biological function, stretching adjusts incompatible linker lengths, eases twist angle restrictions, reduces local frame twist values, and introduces twist defects in the double helix that diffuse, in a DNA sequence dependent manner, to regions away from the nucleosome entry/exit point (Richmond, T.J. and Davey, C.A. 2003).

Figure 5. DNA bending modes in the NCP147 structure. Major and minor groove blocks are respectively colored magenta and yellow. **(a)** Smooth bending into the major groove (left) resulting in uniform curvature of major groove blocks (right). **(b)** Smooth bending into the minor groove (left) causes alternating shift in minor groove blocks (right). **(c)** Kinked bending into the minor groove localizes acute curvature at two blocks causing base edge clash and opposing major groove unstacking. Adapted from Richmond, T.J. and Davey, C.A. 2003

Interestingly, the greatest stretching observed in the NCP145 structure occurs around SHLs ± 2 , which are preceded by major groove unstacking of a CG=GC base-pair step at SHL 1.5 and minor groove unstacking of a CA=TG base-pair step at SHL -1 (Fig. 6) (Ong, M.S. *et al.* 2007). In terms of energetics, this unstacking is significant due to the high stacking energy and rigidity of the CG=GC base pair step (Fig. 6). Deformation at this SHL occurs as a result of stretching to maintain histone–DNA register (Ong, M.S. *et al.* 2007). In applied work, our lab has successfully interrogated the minor groove kink at SHL $+1.5$ of NCP145 with rationally designed platinum-based compounds decorated with intercalating functional groups (Fig. 10b) (Chua, E.Y. *et al.* 2015). The *cis*-platinum-naphthalamide (cisPtNAP) molecule was shown to be particularly effective at reducing human cancer cell viability (Chua, E.Y. *et al.* 2015). This site has now become a hot spot for our metal-based drug development efforts with ruthenium, platinum, and gold mono-, bi-, and heterodinuclear organometallics as leading candidates. Furthermore, NCP induced deformation of unique and genetically relevant DNAs could expose additional sites unique to nucleosomes that may be amenable to alternative drug treatments.

Figure 6. Stretching of shorter length NCP DNAs helps maintain precise histone–DNA register. (upper right) Schematic representation of an NCP particle. The blue circle indicates a DNA region typically associated with helical stretching. The red circle indicates the location of minor groove kinking in the NCP145 DNA structure. (lower left) Structural alignment of NCP147 DNA (green) and NCP146b DNA (blue; PDB code 3UTB) demonstrates helical stretching around SHL -5. Nucleotide register remains consistent for both DNAs until -52G of NCP146b where the register is advanced by one nucleotide, persisting until SHL -5.5 (Chua, E.Y.D. *et al.* 2012). (right) Structural alignment of NCP147 DNA (green) and NCP145 DNA (magenta; PDB code 2NZD) structures showing minor groove kinking, major groove unstacking, and backbone stretching resulting in a severe -55° bend at SHL +1.5 of the 145 base pair construct relative to the 147 base pair construct. Alignments were performed using the Needleman-Wunsch algorithm for nucleic acid matrices (Needleman, S.B. and Wunsch, C.D. 1970), part of the UCSF Chimera software package.

1.2.3 Octamer as a peptide binding, drug targeting, and NCP interaction scaffold

Aside from DNA binding, the octamer plays a critical role in chromatin condensation via a major core histone motif known as the H2A:H2B acidic patch. This patch forms a negatively charged surface that is involved in multiple facets of chromatin dynamics and protein recognition, particularly during condensation where the positive H4 N-terminal tails contacts a groove formed by the α 2-helix of H2A and the α C-helix of H2B (Fig. 7) (Kalashnikova, A.A. *et al.*, 2013). Furthermore, replacement of canonical histones with specialized variants leads to substantial chromatin structure alterations that influence numerous biological processes ranging from transcriptional regulation to genome stability (Bonisch, C. and Hake, S.B. 2012). Among core histones, H2A possess the highest sequence divergence, giving it the highest potential to regulate chromatin organization. For example, variants like H2A.Z can alter nucleosome properties such as increasing acidic patch electronegativity (Suto, R.K. *et al.*, 2000).

The acidic patch has served as the primary protein target in much of our drug development research. Of the eight histone residues comprising the acidic patch, organometallic adducts have been observed at six: E61, E64, D90, E91H2A and E102 H2B (Wu, B *et al.*, 2011; Section 3.8.3 and 3.8.4). This motif is also the target of several

viral peptides with binding motifs mimicking chromatin remodeling and architectural factors like high mobility group nucleosome-binding proteins, which contact the negatively charged groove with positively charged nucleosome-binding domains (Ueda, T. *et al.*, 2008). Among these viral peptides are the human cytomegalovirus IE1, Kaposi's sarcoma herpes virus (KSHV) latency-associated nuclear antigen (LANA), and the chromatin-binding sequence (CBS) of prototype foamy virus (PVF) structural protein GAG (Barbera, A.J. *et al.*, 2006; Fang, Q. *et al.*, 2016; Lesbats, P. *et al.*, 2017). These peptides use 'arginine anchor' motifs to recognize and bind specific acidic patch residues D90 and E92 of H2A. Both LANA and GAG are being developed by us and our collaborators as biologics capable of delivering organometallic anticancer agents to chromatin via naturally evolved, high affinity acidic patch binding.

Figure 7. The nucleosome acidic patch plays a crucial role in NCP crystallization and metalloagent targeting studies. APBS-PDB2PQR calculated electrostatic surface of the human nucleosome histone octamer (nucleic acid component in gray). The blue areas designate the positively charged octamer surface involved in DNA binding. The larger red region designates the negatively charged acidic patch. 'H2A' and 'H2B' are placed over the $\alpha 1/2$ -helix of their respective histone. The N-terminal H4-tail is outlined in green for clarity.

The acidic patch is also responsible for a major NCP crystal packing interaction that is analogous to chromatin condensation (Fig. 8a & b). During NCP crystallization this interaction helps organize other crystal contacts such interdigitating major and minor grooves at DNA–DNA interfaces, cation coordination and DNA charge neutralization, and base pair stacking at blunt end NCP DNA termini which are linker DNA entry/exit sites in chromatin (Fig. 8c & d) (Luger, K. *et al.*, 1997a; Davey, C.A. *et al.*, 2002; Richmond, T.A. and Davey, C.A. 2003; Vasudevan, D. *et al.*, 2010; Wu, B. and Davey, C.A. 2010). For crystal soaking experiments, these packing contacts dictate which

nucleosome surfaces are accessible to adducting agents as one face of the nucleosome is occluded (Fig. 8a & b). In cohesive end crystallization experiments these interactions influence crystallization time and the order of aggregation where the lattice is immutable through one crystal axis and malleable along another. Further detail regarding these crystal contacts and packing interactions will be provided in subsequent sections of the text.

Figure 8. Overview of NCP crystal packing interactions. (a) Circles identify face-to-face H4-tail-acidic patch contacts, highlighted in (b). Black arrows indicate DNA-DNA contacts where major/minor groove interdigitating and cation interactions occur, highlighted in (c). The green box outlines the NCP layer along the *a*-axis that stacks along the *c*-axis, highlighted in (d). (b) Detail of the H4-tail – acidic patch crystal contact. (c) Detail of DNA-DNA crystal contacts and cation interaction (purple sphere). (d) NCP DNA termini stacking interactions along the *c*-axis forms a layer of particles along the *a*-axis. Figure based on NCP601 crystal structure.

1.3 Metal-based drugs in cancer therapy

1.3.1 Origins, chromatin epigenetics, drug targeting, and mononuclear designs

As discussed, chromatin structure organizes the cell's genetic information into a compact state which governs gene activation and silencing. The degree of alteration to chromatin's structure by DNA methylation, covalent histone modifications, nucleosome remodeling, incorporation of histone variants, and snRNAs forms the basis of epigenetic control. This has a direct impact on DNA replication and gene transcription which are essential cellular processes that require unpacking of condensed chromatin, i.e. facultative heterochromatin, to euchromatin and other less condensed states (Collins, N. *et al.*, 2002). Locally, nucleosomes modify gene expression by obstructing or facilitating access to regulatory DNA sequences (ref). For example, a loss of nucleosomes directly upstream of transcription start sites correlates with gene activation (Yuan, G.C. *et al.*, 2005). These nucleosome free regions at the 5' and 3' ends of genes are believed to provide sites for assembly and disassembly of transcription machinery (Lin, J.C. *et al.*, 2007; Shivaswamy, S. *et al.*, 2008). As such, the structural dynamics of these processes may be sensitive to artificial interference aimed at interrupting transcription and replication machinery. In clinical contexts, carcinogenesis is typically associated with numerous genetic alterations, however, human cancer cells also harbor global epigenetic abnormalities derived from nucleosome dynamics. Moreover, tumor cells possess distinct nucleosome positioning and dynamics when compared to healthy cells due to epigenomic differences in gene expression patterns (Lin, C. *et al.*, 2007; Sharma, S. *et al.*, 2010; Gaspar-Maia, A. *et al.*, 2011) (Fig. 9).

Figure 9. Epigenetic control of gene silencing and activation leading to tumorigenesis. (top) The epigenetic status of chromatin in healthy cells. (bottom) Aberrant silencing (red arrow) of tumor suppressor gene promoters by the formation of a silent chromatin state via epigenetic factors (methylation). Aberrant activation (green arrow) of tumorigenic regions of DNA via epigenetic influence. Figure adapted from Sharma, S. *et al.*, 2010.

Therefore small molecules that can disrupt or influence nucleosome sliding, occupancy, and ejection may have an outsized effect on tumor cells relative to healthy cells. With respect to the nucleosome targeting applications, two factors influence the recognition of nucleosomal DNA and histones by nuclear proteins and small molecules. First, the condensation of DNA and histones into nucleosomes, and further compaction into heterochromatin, necessarily reduces its SASA relative to free DNA and chaperone-bound histones. Second, depending on the DNA sequence and its position around histone octamer, nucleosomal DNA assumes distinct conformations that are unattainable by free DNA. As such, ideal genomic targeting models would comprise NCP and nucleosome arrays, however most studies of DNA binding compounds have focused solely on naked DNA. Taking these mechanisms and models into consideration, and with the advantage of a broad chemical landscape that includes nucleic acid and protein moieties, the nucleosomes is a unique and viable target for anticancer drug design efforts.

Conventional pharmacological applications of transition metals encompass two broad categories; they are employed as either structural scaffolds to produce a desired bond configuration, or utilized as reactive centers in organometallic compounds. Gianferrara, T. *et al.*, 2009 propose categorization of metal compounds, specifically anticancer metal compounds, into five distinct categories based on their mode of action: (i) the metal has a functional role where it binds to a biological target, (ii) the metal has a structural role that determines the shape of the molecule and target binding is non-covalent, (iii) the metal acts as a carrier for active ligands, (iv) the metal compound is a catalyst, and (v) the metal compound is photoactive. In these roles the use of metals has significantly increased the range of biologically active therapeutic compounds available for clinical and research purposes. The following text, with respect to structural biology and rational compound design, will focus on the two broad categorizations where the metal center reacts with a biological target (the nucleosome) to form covalent bonds, and reactivities of different metal types for their target can be modulated by ligand alterations or additional metal centers.

Utilizing metals as scaffolding moieties offers numerous design advantages that derive from the specific properties of the chosen metal. The two major features are the number of accessible redox states a particular metal can achieve, and the variety of coordination geometries available in its complexed state (Brabec, V. and Novakova, O. 2006). These

properties allow for the synthesis of structures with unique stereochemical configurations and atypical ligand orientations that are not possible with purely organic reagents. As such, metals are frequently connected to organic ligands when functioning as structural scaffolds. In these scenarios an *in vivo* reaction between the target molecule and metal center does not occur. Instead, the metal acts to orient the compound specifically such that it is positioned for optimal target binding. This is possible because organometallic complexes can, in principle, behave like organic compounds, a characteristic that results from kinetically inert metal–carbon bonds (Meggers, E. *et al.*, 2007; Meggers, E. 2007; Bruijninx, P.C. and Sadler, P.J. 2008). An interesting example of this is the use of ruthenium complexes as inhibitors of the GSK-3 and Pim-1 protein kinases (Meggers, E. *et al.*, 2007).

Metals also confer the ability to adjust the thermodynamic and kinetic properties of ligand substitution when functioning as reactive centers. This is particularly important considering different metal–ligand complex combinations where the metal forms bonds with either organic or inorganic ligands (Gasser, G. *et al.*, 2011; Peacock, A.F. and Sadler, P.J. 2008; van Rijt, S.H. and Sadler, P.J. 2009). Therefore, designing therapeutic compounds with certain reactive metals offers the capability of tuning the reagent's chemical reactivity for different targets *in vivo*. As a reactive center, this feature is critical in directing when the metal itself will form a bond with the target molecule, act as a purely scaffolding modality, or some combination of both.

The first metal-based compound known to be effective against a specific type of cancer was cisplatin, originally discovered in the 1960s by Barnett Rosenberg and colleagues (Rosenberg, B. *et al.*, 1965). Cisplatin's clinical success led to the development of a multitude of platinum-based anticancer agents derived from its basic design; drugs like carboplatin, oxaliplatin and satraplatin (Fig. 10a) (Harrap, K.R. 1985; Kelland, L.R. *et al.*, 1993; Kidani, Y. *et al.*, 1978).

The first step of *in vivo* platinum drug–DNA binding is spontaneous exchange of the reagents's leaving group(s) with cytosolic water molecules to form an aquated and active species. The aquated species then reacts with DNA to form adducts (Davies, M.S. *et al.*, 2000; Fichtinger-Schepman, A.M. *et al.*, 1985). Specific nucleotide sequences and DNA structural features affect the site selectivity of platinum where the metal displays a preference for GpG dinucleotide steps or GpXpG motifs (Wang, D. and

Lippard, S.J. 2005). Platinum–DNA adduct formation usually takes place by exploiting existing distortions in the DNA structure, however adduct formation can also cause these distortions to occur. Adducts formed at N7 atoms of GpG or GpXpG sites cause intra- or inter-strand crosslinking of the double helix and produce a pronounced conformational change resulting in a highly kinked tertiary structure (Wang, D. and Lippard, S.J. 2005). This deformation disrupts DNA transcription and replication by sterically blocking progression of transcription and replication machinery along the active DNA.

This semi-indiscriminate binding is the primary mechanism of action for most platinum-based drugs, however, the increased activity of nuclear proteins involved in DNA repair, particularly proteins in the nucleotide excision repair pathway, can recognize structurally distorted regions of DNA and excise these lesions rendering the drug partially or completely ineffective (Wang, D. and Lippard, S.J. 2005; Spivak, G. 2015). Platinum compound binding may also be reduced through down-regulation of receptors involved in small molecule influx, or sequestration may take place through sulfur-containing proteins leading to drug efflux from the cell (Holzer, A.K. and Howell, S.B. 2006; Holzer, A.K. *et al.*, 2006; Mistry, P. *et al.*, 1991; Yang, P. *et al.*, 2006; Safaei, R. *et al.*, 2004). Increased tolerance to platinum DNA adducts through the loss of the DNA mismatch repair pathway may also arise, thereby inhibiting drug-induced apoptosis (Fink, D. *et al.*, 1996; Zdraveski, Z.Z. *et al.*, 2002). Finally, DNA polymerase may bypass adducts during replication and cells may deliberately reduce apoptosis signaling pathways (Albertella, M.R. *et al.*, 2005; Bassett, E. *et al.*, 2002; Gadducci, A. *et al.*, 2002).

Despite the clinical success of platinum-based therapeutics, issues with severe toxicity continue to be problematic. A major detracting aspect of platinum-based compounds is their nonspecific reactivity towards DNA of any origin. Therefore, they unfortunately do not have distinct activity against tumor cells. This lack of specificity means healthy cells are also compromised during treatment resulting in numerous side effects, many of which are severe e.g. nephrotoxicity, neurotoxicity, ototoxicity, myelotoxicity, and hemolytic anemia (Kelland, L. 2007; Knox, R.J. *et al.*, 1986). The toxicity issues and recurring tumor cell resistance of platinum-based agents has spurred efforts to develop novel metal-based drugs that alleviate these problems while maintaining therapeutic efficacy. For this reason, alternative transition metals are being investigated as reactive

centers in actively developed compounds with ruthenium, and more recently gold, as leading candidates (Gasser, G. *et al.*, 2011; van Rijt, S.H. and Sadler, P.J. 2009).

The clearest advantage of ruthenium-based compounds is their greater specificity for tumor cell targets over platinum-based drugs, which bestows a significant reduction in side effects. This is because blood plasma proteins like albumin and transferrin bind ruthenium compounds which may allow for their delivery to tumor cells, a theory that has gained favor in drug transport strategies (Ang, W.H. *et al.*, 2007; Aleksenko, S.S. *et al.*, 2013; Alagesan, M. *et al.*, 2014). This strategy may work because tumor cells possess increased transferrin receptor concentrations relative to normal cells due to their greater requirement for iron. Exploiting this pathway potentially yields selective delivery of ruthenium compounds to tumor cells in greater quantities. Additionally, ruthenium-based reagents are more reactive *in vivo* at lower pH values. For example, observations of the ruthenium-based RAPTA-C and KP1019 compounds show that reactivity to DNA at pH 6.0 is increased compared to pH 7.4 (Levina, A. *et al.*, 2009; Bergamo, A. and Sava, G. 2011; Bartel, C. *et al.*, 2011; Stevens, S.K. *et al.*, 2013). This feature is critical since most tumor tissues maintain hypoxic conditions while relying on glycolysis as the primary source of energy production. The accumulation of lactic acid as a result of increased glycolytic rates in cancer cells causes a reduction in cytosolic pH. The lower pH environment assists in reducing inactive Ru(III) pro-drugs to active Ru(II) metabolites. This, along with the targeted delivery described above, aids in diminishing side effects. Thus, ruthenium-based reagents are very promising for chemotherapy regimens where tumors do not respond to, or are resistant to, platinum drugs.

In addition to platinum compounds, our lab has conducted extensive structural studies on ruthenium-based organometallics, containing either a 1,3,5-triaza-7-phosphatricyclo-[3.3.1.1]decane (PTA) ligand or ethylenediamine (ED), an arene moiety, and chlorido leaving groups (Fig. 10c). These studies were carried out in the context of chromatin targeting and adduct formation with NCP through crystal soaking experiments (Wu, B. *et al.*, 2011; Adhireksan, Z. *et al.*, 2014). Evidence from RAPTA–NCP derivative studies and have shown ruthenium(II) binding at several distinct nucleosome sites (Fig. 11). These sites include core histone and DNA targets where RAPTA displays a preference for coordinating H2A E61 and E64 of the H2A:H2B acidic patch, a key binding motif for chromatin regulation (Wu, B. *et al.*, 2011; Adhireksan, Z. *et al.*, 2014).

Prototypical anti-cancer ruthenium agents, RAPTA-T and RAPTA-C are bifunctional mononuclear agents that target core histone sites at H4 E63 – K59 (RAPTA-C), and H2A E61 – E64 and H2B E102 – H106 (RAPTA-C/T) (Fig. 10c) (Wu, B. *et al.*, 2011). The monofunctional RAED-C targets both H2A E64 of the acidic patch and DNA sites at SHL ± 1.5 as well as an additional core histone site at E41/E41' of H2A/H2A' (Fig. 11) (Adhireksan, Z. *et al.*, 2014). Studies presented here will discuss dinuclear Ru(II)-based compounds, derived from the classic RAPTA structure, rationally designed to target multiple nucleosome sites simultaneously (Appendix 5). Additionally, preliminary structural data of NCP crystals derivatized with hetero-dinuclear compound, JN128-2_KP2425 possessing ruthenium(II) and platinum(IV) centers, designed to coordinate both protein and DNA sites, will also be presented (Appendix

Figure 10. Metal-based anticancer compounds. Reagents our lab has worked on in prior chromatin targeting studies **(a)** Classic cisplatin derived platinum(II) series. **(b)** Novel platinum(II)-based compounds with intercalating functional group. **(c)** Ruthenium(II)-based RAPTA and RAED series. **(d)** Gold(I)-based antirheumatic auranofin.

5), as well as hetero-bi- and trinuclear ruthenium(II)–gold(I) compounds. These reagents, and their supporting literature, will be discussed in greater depth in their respective sections (Sections 3.6.1, 3.6.3, 3.6.4).

Recent breakthroughs in organometallic compound design have focused on so called ‘noble’ metals like gold. Therapeutic compounds synthesized with gold(I) reactive centers are now gaining currency as anti-tumorigenic drugs. The first cytotoxic effects of gold(I) compounds were observed in mice with P388 leukemia (Sadler, P.J. and Sue, R.E. 1994). Since the initial cytotoxicity observations, gold compounds have been investigated as effective agents against various cancers including ovarian, breast, and lung cancer, and certain leukemias (Park, S.H. *et al.*, 2014; Varghese, E. and Busselberg, D. 2014; Fan, C. *et al.*, 2014). Unlike cisplatin and its derivatives, gold complexes primarily target sulfur and nitrogen protein substrates. With an adaptive design this mode of action could increase the specificity of these complexes, relative to cisplatin, which may reduce side effects associated with nonspecific adduct formation.

Auranofin (AUF), gold(I)-3,4,5-triacetyloxy-6-(acetyloxymethyl) oxane-2-thiolate; triethylphosphonium is a gold-based medicinal compound originally developed as an anti-rheumatic drug that is now being applied as a broad spectrum anticancer agent (Fig. 10d). As a prodrug, the gold(I) containing AUF is rapidly metabolized into a pharmacologically active species upon cellular uptake where it displays high cytotoxic efficacy (Tiekink, E.R. 2003). Elucidated antimetastatic mechanisms for gold-based drug compounds, including AUF, remain lacking. However, available research indicates that inhibition of protein targets such as thioredoxin, thioredoxin reductase, and the proteasome, are likely target pathways (Schuh, E. *et al.*, 2012; Ronconi, L. and Fregona, D. 2009). Structural and cellular localization studies performed by our lab and collaborators have identified chromatin as a novel target for auranofin when NCP crystals or A2780 cells are first treated with RAPTA-T (Adhireksan, Z. *et al.*, 2017). An extensive analysis demonstrated a synergistic effect between RAPTA-T and AUF, whereby RAPTA-T binding the H2A:H2B acidic patch causes subtle, but global, changes in nucleosome conformation that increase AUF binding at symmetry-related histidine residues on H3 and H3'. When applied in cytotoxicity assays, RAPTA-T binding induces tumor cell susceptibility to AUF reactivity.

In an attempt to enhance specificity for nucleosomes, bis-gold AUF derivatives, along with bis-ruthenium RAPTA-T derivatives, were developed in collaboration with the Dyson lab at the EPFL (Appendix 5). These dinuclear compounds contain chemically inert, variable length linkers between each metal center as their intended purpose is to crosslink two nucleosome sites. This theme is also explored with the previously mentioned hetero-dinuclear compound

JN128-2_KP2425, which possesses a platinum(IV) center in lieu of a gold(I), in addition to Compound III, a gold(I)–ruthenium(II) hetero-dinuclear agent and its trinuclear derivative, and two chromatin binding viral peptides conjugated by PEG linker to a gold(I) reactive species (Appendix 5).

Nucleosome binding Site	Metalloreagent
RU1 (H2A E61/64)	RAPTA-T/C, RAED-C, RR*, C2*, C3 (Ru)*^, tnC3 (Ru)^, C1, C4, JN128-2_KP2425
RU2 (H2B E101/H106)	RAPTA-T/C, RR*, C2*, tnC3 (Ru)^, C1, C4
AU1/1' (H3/3' H113/113')	AUF, EP11-019, C3 (Au)*, tnC3 (Au)^
H3' (M120'), H4' (M84')	cisplatin, oxaliplatin
Site 4/4' (H2A/H2A' E41/41')	RAED-C
H3/3' H39/39' H2B H46	C3 (Au)^, tnC3 (Au)^ C3 (Au)^, tnC3 (Au)^
G(-14) (SHL ±1.5)	cis/trPtNAP, cisplatin, oxaliplatin
G(-15) (SHL ±1.5)	RAED-C, RAED-THA, cisplatin, oxaliplatin

Figure 11. Map of metalloreagent binding sites on the nucleosome. Bold entries signify reagents analyzed in the present study. *Reagent forms cross-links between sites. ^Potential cross-links between sites.

1.3.2 Multinuclearity, peptide–metal conjugates, and nucleosome cross-linking

Beginning with cisplatin and cisplatin derivatives, the ability of mononuclear metalloreagents to bind nucleosomes has been invaluable as an anticancer drug development platform and a chromatin research tool (Wu, B. *et al.*, 2008; Wu, B. and Davey, C.A. 2008). The advent of the potent trinuclear platinum compound BBR3464, with IC50 values 20-fold lower than cisplatin, spurred the focus of metal-based anticancer drug design towards homo- and hetero-multinuclear reagents (Manzotti, C. *et al.*, 2000). Recently, multinuclearity as a drug design concept has grown beyond the development of homo-dimetallic complexes to explore a variety of metals types and

combinations (Batchelor, L.K. *et al.*, 2017). Synthesis of metal-based compounds continues with both rare and biologically essential metals, but the emphasis remains on platinum, ruthenium, and more recently gold as the binding mechanisms, toxicologies, and bioavailabilities of these metals are well known. In particular, auranofin is a prime candidate for drug repurposing and has undergone structural modification to achieve homo-multinuclear gold and hetero-multinuclear complexes (Roder, C. and Thomson, M.J. 2015).

Multinuclear organometallic and metalorganic anticancer compounds offer a diverse range of cytotoxicity mechanisms resulting in distinct cytotoxicity profiles in comparison to mononuclear compounds. This is due to their unique characteristics such as additional targeting modes, altered electrostatic behavior and the magnitude of charge, lipophilicity of the bridging molecule, and the ability to form specific cross-linking patterns (Pratesi, G. *et al.*, 1999; Manzotti, C. *et al.*, 2000; McGregor, T.D. *et al.*, 2002; Zhu, J.H. *et al.*, 2009). At the time of writing, researchers have synthesized thousands of organometallic and metalorganic anticancer compounds with a number of them successfully utilized in cytotoxicity assays. Some of these compounds transitioned to clinical trials where they have performed with mixed results (Raymond, E. *et al.*, 1998; Jodrell, D.I. *et al.*, 2004; Hensing, T.A. *et al.*, 2006). For example, the previously cited platinum compound BBR3464 has superseded cisplatin's potency by increasing the number of reactive metal centers per molecule but failed in clinical trials (Mazotti, C. *et al.*, 2000; Jordell, D.I. *et al.*, 2004; Hensing, T.A. *et al.*, 2006). Yet despite well-documented issues with toxicity and resistance, the efficacy of mononuclear cisplatin continues to be the reference against which metal-based anticancer candidates of any metal type are compared. Therefore a more targeted and rational approach to multinuclear design is necessary; one capable of exploiting a fundamental biological target, with distinct tumorigenic markers, and capitalize on any highly specific chemical, structural, and allosteric motifs.

Our lab's chromatin targeting work has applied multinuclear designs and initial experiments have utilized dinuclear ruthenium compounds to bind H2A and H2B histones at the acidic patch, forming intra- and inter-particle cross-links that lead to chromatin condensation (Davey, G.E. *et al.*, 2017). These observations, in conjunction with the discovery of a nucleosome allosteric cross-talk pathway, are providing inspiration for new hetero-bi- and trinuclear organometallic and metalorganic anticancer

compounds that take advantage of these discrete yet related sites (Adhireksan, Z. *et al.*, 2017). Expanding our pharmacological toolkit, the LANA and GAG viral peptides have been conjugated by PEG linkers to gold(I) containing AUF derivatives in an attempt to elicit the same conformational dynamics and synergistic response of the RAPTA-T/AUF combination (Section 3.6.5). Interestingly, this approach reverses the principle identified in category iii of Gianferrara, T. *et al.*, 2009 where the ligand, in this case a viral peptide, is employed to deliver the active organometallic compound.

1.4 Scope

The nucleosome is a thoroughly examined molecule, however, there are still gaps in knowledge concerning chromatin structure with respect to genetically relevant DNA sequences, and a similar lack of data regarding nuclear factor bound nucleosome structures. Our lab's efforts in rational molecule design have utilized the NCP crystal system as a model for chromatin targeting in anticancer drug development and the study of chromatin dynamics (Wu, B. *et al.*, 2011; Adhireksan, Z. *et al.*, 2017; Davey, G.E. *et al.*, 2017). The difficulty in obtaining chromatin factor bound structures of the NCP, and the physical crystal damage accrued during small molecule soaking studies are in part attributed to deficiencies in the NCP crystal lattice. The extent of NCP crystal damage that occurs during compound soaks can result in a loss of resolution of one angstrom or greater. This has been a stubborn technical challenge where a lack of precise information regarding NCP-adduct structures obligates use of implicit chemical knowledge instead of relying on crystallographic data. As such, the central theme throughout the thesis will be the engineering of NCP DNA sequences intended to increase crystal integrity and manipulate the crystal lattice. Cohesive end palindromic sequences have already demonstrated improved ordering of nucleosome packing and crystallization kinetics in nucleosome linker histone assemblies (Adhireksan *et al.*, under review). Furthermore, while the following projects are not directly related, they share overlap by benefiting from the development of an improved nucleosome crystal system (Fig. 12).

With the exception of diruthenium compounds, prior chromatin targeting studies performed in our lab concentrated on reagents with single metal centers. Design emphasis is now shifting to newer dimetallic compounds that have the capacity to crosslink proximal and distant areas on the nucleosome. Taking advantage of the nucleosome's broad chemical space, hetero-dinuclear organometallics are the current

focus of our chromatin targeting drug development. These reagents are designed to exploit distinct nucleosome sites with characteristic reactivity (Adhireksan, Z. *et al.*, 2017). Furthermore, efforts are underway to expand anticancer reagent development to include metal–peptide conjugates. This is a new type of approach to chromatin targeting, exploiting biologics as a drug delivery method.

The workhorse system for these studies has been the NCP145 crystal system. However, reagent soaking experiments will transition to new nucleosome crystal systems provided they perform up to expectations. Additionally, co-crystal structures of NCP in complex with factors like the architectural non-histone chromosomal protein HMGN, or its nucleosome binding domain (NBD) peptide, and the forkhead box family of transcription factors will be attempted with cohesive end DNA constructs (Fig. 12). These structures have been challenging to obtain due to their transient nucleosome binding, which may benefit from an improved crystal lattice and crystallization kinetics.

To overcome the shortcoming of blunt end NCP crystals, cohesive end sequences will be modified for NCP crystal systems to improve the quality of adduct structures and used to achieve difficult to crystallize structures of chromatin factors bound to nucleosome. With further experimentation, cohesive end DNA fragments are anticipated to offer control over NCP crystallization kinetics, which would allow crystallization of NCPs with select DNA sequences. For development purposes, initial cohesive end constructs possess a Widom 601 core sequence and either complimentary or non-complimentary cohesive ends (Fig. 12).

Figure 12. Project overview and significance. The gray box around the ‘NCP:Factor Complexes’ field indicates these are projected studies for the application of cohesive end NCP DNA fragments.

Chapter 2 – Methods and Materials

Unless otherwise stated, all chemicals were purchased from Sigma and all enzymes were purchased from New England Biolabs. P.J. Dyson provided mono, dinuclear, and hetero-dinuclear ruthenium and gold compounds and A. A. Nazarov provided JN128-2_KP2425 (Appendix 5).

2.1 DNA production and NCP construct cloning

2.1.1 Large scale production of DNA601L

Sixteen 84bp repeats of the left half of the Widom 601 sequence were cloned into pUC57 ampicillin resistant vector by GenScript (Chua, E.Y. *et al.*, 2012). This was used for production of palindromic 601L NCP DNA by a protocol adapted from Dyer, P.N. *et al.*, 2004. 601L is a strong positioning NCP DNA construct, derived from the Widom 601 sequence, created by our lab to study NCP crystal systems where nucleosome positioning is highly dependent on structural characteristics of DNA sequences (Chua, E.Y. *et al.*, 2012; Lowary, P. T. and Widom, J. 1998). Essentially, 601L is the left half of the 601 Widom sequence inverted about the dyad to generate a palindromic construct (Chua, E.Y. *et al.*, 2012). The palindromicity of 601L reveals unique and specific aspects of TA dinucleotide and G|C rich element's propensity for favorable energetic positioning at certain locations around the histone octamer (Chua, E.Y. *et al.*, 2012). Since its inception, our lab has extensively developed and modified 601L for various nucleosome and nucleosome-linker-histone-assembly applications with primary focus on structural investigation.

PAGE gels representing each step of 601L production are shown in figure 13a – f. PEG fractionation is carried out after complete EcoRV digestion of pUC57 vector containing sixteen 84 base pair half-site tandem repeats (168 base pair cloning repeats) of 601L sites A and B (Fig. 13a). A minimum of eight cloning repeats is essential to generate the necessary yield during large scale culture preparations that is required for crystallographic studies. Cut plasmid DNA is too large to enter the gel and is visualized by EtBr staining at the well (Fig. 13a). A ligation test is carried out to check for adequate 5' end dephosphorylation by calf intestinal phosphatase (CIP) (Fig. 13b). This step ensures there will be correct ligation products after removal of a nine base pair protection fragment by HinfI digestion (Fig. 13b). In each 168 base pair cloning repeat

there are two *HinfI* sites that differ by an A/T nucleotide at position three. This substitution defines the difference between the A and B fragments.

Figure 13. 601L production and processing. (a – f) Representative 10% PAGE gels of each protocol step. g & h Ion exchange chromatography to separate fragments of varying size. Blue trace = mAU UV @ 280nm; brown trace = conductivity, mS/cm; green trace = % slat concentration. (a) EcoRV digestion is followed by PEG fractionating to produce sixteen 84bp half-site fragments per plasmid. Linear plasmid DNA is too large to enter the 10% PAGE gel. (b) Ligation test demonstrating complete 5' dephosphorylation of EcoRV digested fragments will not form unwanted ligation products. (c) *HinfI* digestion of 84bp blunt end half sites removes protection sequence to produce 71+3 fragments. (d & g) Ion exchange chromatography of *HinfI* digested fragments separates 10+1 and 71+3 half sites. (e) Ligation of 71+3 sticky end fragments creating 145bp NCP DNA. Only 30 – 40% completion visible. (f & h) Separation of unligated fragments and 145bp NCP DNA yields sufficient purity of the 145bp construct.

HinfI digestion of the 84 base pair fragment produces 72 base pair fragments of A and B with three base pair 5' overhangs (71+3), and nine base pair fragments with a three base pair 3' overhang (9+3) (Fig. 13c). This mixture of fragments must be separated by ion exchange chromatography (Fig. 13d & g). Subsequent T4 ligation of the 71+3 A and B fragment *HinfI* sites yield a palindromic 145 base pair NCP DNA construct where

the nucleotide at position three of the *Hinf*I site becomes the central base pair corresponding to the NCP dyad (Fig. 13e). Additional rounds of ion exchange chromatography are necessary to separate the finished 145 base pair construct from unligated 71+3 fragments (Fig. 13f & h). Pure 145 base pair DNA is used for NCP reconstitutions.

2.1.2 Large scale production of DNA145

Human α -satellite DNA are large arrays of 170 base pair tandemly repeated non-coding palindromic sequences serving as the main component of functional centromeres and are the primary structural constituent of heterochromatin (Botchan, M. *et al.*, 1971; Gall, J.G. *et al.*, 1971; Peacock, W.J. *et al.*, 1974; Tyler-Smith, C. and Brown, W.R. 1987; Guenatri, M. *et al.*, 2004). In biological research, α -satellite DNA has been used in investigations of nucleosome and chromatin structure for decades, including the first atomic-scale resolution structure of the

NCP (Gottesfeld, J.M. and Melton, D.A. 1978; Igo-Kemenes, T. *et al.*, 1980; Luger, K. *et al.*, 1997a). In our lab, α -satellite DNA constructs are used to reconstitute nucleosome 145, 146, and 147 core particles that act as experimental controls and metal-based reagent substrates. Currently, NCP145 is the principal crystal system used for small molecule soaking experiments in our lab.

DNA145 was produced and purified using methods similar to 601L production and purification. *Eco*RV digestion, PEG fractionation, *Hinf*I digestion and CIP dephosphorylation, and Taq ligation are all identical to 601L DNA production as outlined by PAGE gels representing each step of purification (Fig. 14a - g). . Elution of 71+3 and 145 base pair fragments was carried out by ResourceQ and MonoQ anion exchange chromatography with a 1 M NaCl gradient.

Figure 14. DNA145α production and purification. (a – g) Representative 10% PAGE gels of each protocol step. For all gels, ladder in lane 1 is Promega 100bp. (a & b) EcoRV digestion (a) is followed by PEG fractionation (b) to produce sixteen 84bp half-site fragments per plasmid. Linear plasmid DNA is too large to enter the gel. (c) Hinfl digestion of 84bp blunt end half sites removes the protection sequence and yields 71+3 fragments. (d) Ion exchange chromatography of Hinfl digested half sites separates 10+1 and 71+3 fragments. Blue trace = mAU UV @ 280nm; brown trace = conductivity, mS/cm; green trace = % salt concentration. (e) 84bp (lane 2 & 3) and 71+3bp (lanes 4 & 5) controls. (f) Ligation of 71+3 Hinfl end fragments creating 145bp NCP DNA. (g) Separation of unligated 71+3 and 145 base pair fragments yields pure DNA145α (lanes 3, 4 & 5).

2.1.3 Cohesive end NCP DNA cloning

Cloning of palindromic 601L derivatives with overhang termini was attempted by following the protocol found in Dyer, P.N., *et al.*, 2004 with adaptations. The maximum tandem cloning repeat target was 8 A+B sites for a total of 16 half sites. BamHI and HindIII were chosen as cloning site enzymes with BglII as a subcloning site. PstI was chosen for NCP sticky end termini as it is half G|C with a 3' overhang. The cloning plasmid chosen was pUC19 ampicillin resistant vector. All plasmid DNA was replicated by growing DH5 α cells in liquid overnight cultures using Terrific Broth media. Plasmid isolation and insert purification were done using QIAGEN miniprep and gel extraction kits with some protocol modifications. All results were visualized on 1% agarose gels with ethidium bromide staining. Further details on the cloning process are discussed in section 3.1.

2.1.4 Large scale production of NCP147SK/XK/K

All constructs were ordered from GenScript Inc. and sequences were independently verified by 1st Base Axil Scientific Pte Ltd. Single stranded and duplex sequences are provided in Appendix 2. Further details, as well as FPLC UV traces, are given in each construct's respective section and section 3.1.4. Restriction enzyme concentrations for the release of construct from pUC19 (SacI, XhoI, KpnI) approximated 300 U $\cdot$ mg⁻¹ of plasmid DNA.

2.2 Core histone production and histone octamer refolding

2.2.1 Core histone and H2A.Z variant production and purification

Each core histone type was expressed and purified individually using the same protocol. Human core histone sequences were cloned into pET15b 6xHis-tag ampicillin resistant expression vectors at the NdeI-BamHI cloning site. Histone purification protocol was adapted from Luger, K. *et al.*, 1997a and Luger, K. *et al.*, 1999 with modifications. Clarified lysate from the insoluble fraction was solubilized in 6M guanidinium chloride and centrifuged at 20,000 g for 15 minutes to pellet debris. The supernatant contains 6xHis-tagged core histones that were initially purified by Ni-NTA affinity chromatography (IMAC[®], GE Healthcare) by an imidazole step gradient in the presence of 8 M urea. Results are visualized by SDS-PAGE. The sample is then dialyzed against Mili-Q[®] H₂O before treatment with thrombin serine protease (GE healthcare) to cleave the 6xHis-tag. Further purification to remove tightly bound contaminating DNA is done by ion exchange chromatography (ResourceS[®], GE Healthcare) with a sodium chloride

gradient. Pure samples are visualized by SDS-PAGE and their concentration is determined by UV spectroscopy, using the appropriate molecular weight and extinction coefficient, before lyophilization and long-term storage at -20°C.

Core histone genes are PCR amplified from genomic DNA and subsequently cloned into a pET15b ampicillin resistant plasmid. This vector affixes a hexahistidine tag to the N-terminal tail that allows for downstream purification by affinity chromatography. When induced to overexpress in bacterial cells, core histones aggregate as inclusion bodies that require high chaotrope concentrations during purification. For this reason, cells are pelleted and suspended in buffer containing 7 M guanidinium chloride as they undergo several rounds of lysis by moderate sonication. Each step of sonication and resuspension isolates histones located in the insoluble fraction. Clarified lysate is passed through a Ni-NTA column to trap the 6xHis-tag and histone protein is eluted by an imidazole step gradient in the presence of deionized 8 M urea where the majority of protein elutes between 50 – 250 mM imidazole (Fig. 15a). At this point during purification the positively charged core histones are highly contaminated with genomic DNA which will cause aggregation during subsequent steps if the bound DNA is too abundant. Several dialysis steps remove the urea so that the 6xHis-tag can be cleaved by thrombin digestion before the histones are finally purified and DNA is removed by ion exchange chromatography (Fig. 15c). Histone fractions are analyzed by 18% SDS-PAGE to assure purity before they can be lyophilized for long term storage or used directly for histone octamer refolding (Fig. 15b).

Figure 15. Purification of human core histones by affinity and ion exchange chromatography. (a) Elution of 6-histidine tagged H2A with imidazole step gradient by hand shows most protein eluting between 50 mM and 250 mM imidazole when visualized on a 15% denaturing gel. (b & c) Subsequent pooling, dialysis, and thrombin digestion of H2A is followed by ion exchange chromatography (NaCl gradient to 1 M) to remove contaminating DNA and proteins remaining after affinity chromatography. UV trace – blue, %B (NaCl) – green, Conductivity – brown. Pure fractions are analyzed by 15% SDS-PAGE.

2.2.2 Histone octamer refolding and purification

Octamer refolding is achieved by mixing individual lyophilized core histones in the presence of 6 M guanidinium chloride together in a H2A:H2B:H3:H4 ratio of 1:1:1.1:1.1 at a concentration of 2 mg·mL⁻¹ and dialyzing this volume against 2 M NaCl overnight. Heavy precipitate is removed by centrifugation for 10 minutes at 4°C and maximum speed. Concentration is determined by absorbance at 276 nm with UV spectroscopy using an octamer extinction coefficient of 48,900 cm⁻¹ M⁻¹ (0.45 cm² mg⁻¹) and a molecular weight of 108,000 Da. Dialyzed octamer sample is concentrated by 3K spin column concentrators at 4°C and maximum speed until the total volume is below 2 mL. Concentrated sample is injected over an S200 Superdex gel filtration column (GE Healthcare) and eluted in refolding buffer (2 M NaCl, 10 mM Tris-HCl (pH 7.5), 1 mM Na-EDTA, 10 mM βMe). Fractions under the octamer peak are isolated, pooled, and concentrated to >8 mg·mL⁻¹ before diluting the sample in half with 100% glycerol for storage at -20°C.

Human histone octamer is allowed to fold under high salt conditions which favor tight association of the individual histone proteins. Samples are dialyzed against 2 M salt to force hydrophobic interactions and maximize solvent contact with the positively charged octamer SASA. The octamer elution profile by SEC is shown in figure 16. The largest peak at approximately 70 mL represents pure fully folded octamer which is followed closely in apparent size by unincorporated H3/H4 tetramer with sufficient separation between octamer and H2A/H2B dimer (Fig. 16a). Protein aggregates and DNA contaminated histones elute around $\approx 50 - 55$ mL (Fig. 16a). Human histone octamer identity and purity is confirmed by SDS-PAGE analysis showing comigration of H3 and H2B, in our evaluations, with H2A and H4 running at lower apparent molecular weights. Faint upper bands may represent multimeric histone complexes like disulfide cross-linked H3 oligomers (Fig 16b).

Figure 16. *In vitro* octamer refolding of human core histones. (a) Elution profile of dialyzed histone octamer from size exclusion chromatography displaying different species of protein oligomers. Fractions under the octamer peak were separated and analyzed for purity by SDS-PAGE (data not shown). (b) Representative 15% SDS-PAGE gel of human histone octamer. 5 μ g loaded. H3 and H2B comigrate and appear as a single band. Lane 1 = 250kDa Precision Plus[®] marker from Bio-Rad[®]. Lanes 2 – 4 are histone octamer replicates 1 – 4.

2.3 NCP145 reconstitution and crystallization

2.3.1 Human NCP145 reconstitution

Human NCP145 was reconstituted according to adapted protocol taken from Luger, K. *et al.*, 1997a and Dyer, P.N. *et al.*, 2004 with initial dialysis screens carried out in dialysis buttons to determine the optimal DNA:HO ratio. According to results discussed below, small amounts of free DNA are needed during crystallization so the DNA:HO

ratio is usually adjusted to include slightly more DNA. Various heat shifting times and temperatures were tested until it was determined that incubation for 20 to 30 minutes at 37°C is sufficient for nearly complete centered nucleosome. Centrifugation at 4°C and maximum speed for 10 minutes helps remove insoluble precipitate from soluble NCP before concentrating. 10 kDa concentrators are preferred as they retain NCP well. NCP145 concentration was determined by measuring UV absorbance of DNA at 260 nm and multiplying by a factor of 2.216.

2.3.2 Human NCP145 crystallization

Protocol for human NCP145 crystallization followed methodology established in Ong, M.S. *et al.*, 2007. In total, more than one hundred forty conditions of varying manganese(II) chloride, potassium chloride, and K-cacodylate; pH 6.0 concentrations were tested in over two hundred thirty trials of hanging drop vapor diffusion equilibration by salting-in. The entire range of crystallization conditions screened was: 20 – 150 mM manganese(II) chloride, 10 – 100 mM potassium chloride, and 20 – 40 mM K-cacodylate; pH 6.0. Initial screens were performed with 1 μ L of NCP at 8 mg·mL⁻¹, 1 μ L of 1X buffer conditions across a wide range of concentrations, and 0.25X well conditions. For small scale screening, mixing of NCP with crystallization buffer was done on the cover slide after a final round of centrifugation for 10 minutes at 4°C and maximum speed. For large scale screens, premixing NCP (8 mg·mL⁻¹) with crystallization buffer and incubation on ice for two hours was undertaken before the final centrifugation was carried out. Preparative screens were conducted with narrower concentration ranges around 90 mM manganese(II) chloride, 80 mM potassium chloride, 40 mM (CH₃)₂AsO₂K (pH 6.0). Drop sizes in these screens totaled 8 μ L (4 μ L NCP + 4 μ L 1X buffer) to increase crystal size. Large crystals of diffraction quality size normally appeared after $\approx$ 2.5 weeks of equilibration at 18°C. Harvesting and buffer exchange in to long term storage buffer was performed in glass depression dishes by washing crystals three times with seven buffers of increasing MDP and trehalose concentrations (up to 24% v/v and 2% w/v respectively) and salt concentrations at half of the 1X concentration (45 mM manganese(II) chloride, 40 mM potassium chloride, 20 mM (CH₃)₂AsO₂K (pH 6.0).

2.3.3 Cohesive end NCP reconstitution and crystallization

Reconstitution and crystallization for cohesive end NCPs followed a protocol similar to that used for NCP145, using the same buffer conditions and NCP concentrations. A

factor of 2.2 was used to determine NCP147SK, NCP147XK, and NCP147K concentrations by UV-vis spectroscopy. Specific details on conditions screened and parameters tested for each crystal type are covered extensively in the respective section for each construct (Section 3.2.1, 3.2.2, and 3.3.1).

2.4 Crystal soaking and data collection and processing

2.4.1 Human NCP145 crystal soaking

Prior to soaking human NCP145 crystals with organometallic compounds or peptide, a final buffer exchange is necessary to replace manganese(II) with magnesium(II) (MgSO₄). For this exchange all other buffer parameters remain constant. Organometallic compounds were dissolved in magnesium buffer to the extent that they were soluble to their desired concentration (Table 1). An equivalent volume of compound is added to crystal buffer so that the final reagent concentration is 1 mM. Data collection was possible only for soluble compounds. Their analysis is discussed in section 3.6.

2.4.2 Data collection and processing

All adducted human NCP145 data was collected at beam line PXIII of the Swiss Light Source (SLS) using a PILATUS 2M-F detector. For ruthenium based agents the X-ray wavelength was tuned to 1.5 Å, 1.072 Å for Pt atoms, and 1.040 Å for gold compounds. An entire data set consists of 360 images collected at an oscillation of 0.5° with 1 second of exposure.

Data integration was performed with the iMOSFLM software package version 7.2.1 (Leslie, A.G., 2006) and data reduction was achieved with SCALA 3.3.22 of the CCP4 software suite version 7.0.053. Rigid body and restrained refinements were achieved with Refmac 5.8.0222 of the CCP4 software package using the native NCP145 model (PDB code: 2NZD). Anomalous maps were created using the Sfall and FFT modules found in CCP4. Visualization and signal analysis were done using Coot 0.8.2 (Emsley, P. and Cowtan, K. 2004) and figures were created and structural alignment and superposition, unless explicitly stated, were performed with PyMOL 2.0.6 (DeLano, W.L., 2002) and Coot.

Compound	Concentration (mM)	Duration (hrs)	Compound	Concentration (mM)	Duration (hrs)
EP11-019	1	48	Compound 1 (C1)	1	19
	1	63		1	42
	1	70		1	86
EP11-019 + RAPTA-T	1 + 0.5	20	Compound 4 (C4)	1	19
	1 + 0.5	46		1	42
	1 + 0.5	63		1	86
EP11-019 + KCl	1 + 4	89	Compound 3 (C3)	1	28
		(70 + 19)		1	54
EP11-104	1	20		1	97
EP10-069	1	26	Trinuclear Compound 3 (tnC3)	1	120
	1	63		1	28
EP10-069 + RAPTA-T	1 + 0.5	21		1	54
	1 + 0.5	63	LANA	1	96
EP10-069 + KCl	1 + 4	19		1	120
				2	15
Lucie-R33	1	22	LANA-M	2	15
	1	70	LANA M6O	2	15
Lucie-R49	1	20	Au-LANA M6O	2	29
Lucie-R50	1	20		0.5	61
	1	45		0.5	84
JN128-2_KP2425	1	20	Au-GAG CBS	2	29
	1	24		2	54
	1	47		2	83
				2	118

Table 1. Soaking conditions for all tested drug candidates, organometallic compounds, peptides, and controls.

2.5 Compound 3 biochemical assays

2.5.1 Cross-linking analysis of treated nucleosome core particle

For the nucleosome cross-linking experiments, 1 μ M NCP, in a buffer of 20 mM K-cacodylate [pH 6.0], was incubated with 2, 5, 10, 25 50 or 75 mM C3 at room temperature for 24 hours. Samples were subsequently analyzed with denaturing PAGE. 4 μ L of Laemmli 4X loading dye (Bio-Rad, Hercules, California), supplemented with the addition of 15 mM tris(2-carboxyethyl)phosphine, was added to 10 μ L of sample, followed by a 1 min incubation at 95°C. After brief centrifugation, samples were loaded onto a 15% SDS-PAGE gel, and the gel was subsequently stained with coomassie brilliant blue.

2.5.2 Tyrosine fluorescence spectroscopy

Nucleosome stability was measured by monitoring the increase in fluorescence, which results from a loss in quenching of histone tyrosine residues by proximal DNA bases as DNA-histone interactions are disrupted (Oohara, I. and Wada, A. 1987; Luger, K. *et al.*, 1997b; Chua, E.Y.D. *et al.*, 2012). For preparing cross-linked samples, 1 μ M NCP, in

a buffer of 20 mM K-cacodylate [pH 6.0], was incubated with 2 μ M or 5 μ M C3 at 21° C for $\approx$ 30 h. Fluorescence measurements were conducted with a Varian Cary Eclipse fluorescence spectrophotometer (Agilent Technologies, Santa Clara, USA) equipped with a temperature control unit. Samples of 0.5 μ M NCP were allowed to equilibrate at 21° C in a buffer of 20 mM Tris (pH 7.5) and 1 mM EDTA and NaCl ranging in concentration from 0 – 2.4 M. Fluorescence readings were taken at 21° C with samples in Teflon stopper cuvettes having 1 cm path length. Tyrosine fluorescence was measured by the emission at 305 nm through excitation at 275 nm.

Fluorescence measurements were carried out in triplicate and values for each data set were placed on a relative scale by setting the minimum and maximum values over the 0 – 2.4 M range to 0 and 1, respectively. The three normalized data sets for each construct type were then merged into one by averaging. Optimized sigmoidal profiles were fit to the merged data sets using OriginPro 8.1 (OriginLab Corp., Northampton, USA), from which 50% dissociation points with respect to NaCl concentration were derived.

Chapter 3 – Results and Discussion

3.1 Cohesive end 601L cloning and non-palindromic designs

3.1.1 Cohesive end NCP DNA design

In-lab cloning efforts to produce a palindromic NCP DNA construct with 3' overhang termini ultimately failed due to bacterial recombination and excision of multiple repeat constructs (Fig. 20a - j). The following results and discussion will present a synopsis of the strategy and rationale utilized while attempting to create palindromic overhang DNA constructs capable of producing quasi-continuous NCP crystal systems and discuss the limitations of these results. Alternative design strategies and commercial supplies were eventually successful in overcoming inherent cloning difficulties, and these designs provided sequences that granted high-resolution crystal structures. The following sections will provide analysis on NCP DNA termini conformations in crystal structures using three non-palindromic sequences, one with compatible overhangs and two with incompatible overhangs.

As of this writing, there are approximately 172 unique nucleosome structures available in the RCSB Protein data bank (<https://www.rcsb.org>). These structures include organismal homologues, small molecule adducted/derivatized NCPs, peptide/protein factor-bound nucleosomes, and octamers composed of core histone variants. Yet despite the degree of structural and core histone heterogeneity found in the database, there is a disproportionate lack of DNA sequence variety. Two distinct sequences are overwhelmingly represented; the artificially evolved Widom 601 sequence and two similar naturally occurring palindromic human α -satellite sequences (Lowary, P.T. and Widom, J. 1998; Tyler-Smith, C. and Brown, W.R. 1987). These DNA sequences also serve as progenitors for a handful of derivatives: 601L, α -147/145, 146b, and so on, each producing additional structural details though none are asymmetric (Fig. 17 table). Such a dearth of DNA sequence variety emphasizes a troubling deficit in structural information at the lowest level of chromatin hierarchy, particularly when considering absolute sequence variability across an entire genome.

The conformation of DNA bound to nucleosomes depends directly on its sequence and subtle physical differences at the structural-mechanical interface may have profound impact on molecular recognition and signal initiation (Chua, E.Y. *et al.*, 2012). A recent

crystal structure of mutant octamer in complex with a natural DNA sequence from mouse mammary tumor virus crystallized with magnesium ions reveals numerous DNA conformational differences in comparison to palindromic α -satellite DNA (Frouws, T.D. *et al.*, 2016). Encouraged by the experimental evidence of this study, and our initial results using sticky-end constructs to solve definitive structures of human nucleosome-linker-histone-assemblies, we are developing cohesive end NCP DNA constructs as tools to manipulate nucleosome crystal structures (Fig. 21). The principal goal is to create a method for employing any DNA sequence interchangeably with the eventual capability of controlling hetero-oligomeric array formation.

NCP crystals formed from cohesive end DNA constructs could serve as valuable systems for studying factor bound NCP complexes that are difficult to cocrystallize. Base pairing constructs could benefit from more controllable crystallization kinetics that are able to overcome protein factors inhibiting crystallization or molecules that rearrange the lattice or introduce defects. Furthermore, lowering free energy at crystal contact sites through better ordering could help increase lattice stability, which would be valuable during damaging metallo-reagent soaks.

The rationale for Watson-Crick pairing termini follows from observations of end-to-end nucleotide stacking interactions along the a -axis in NCP crystal lattices (Fig. 17 top). In blunt end NCP structures, base pair steps at the DNA termini participate in stacking interactions to create crystal-packing contacts that help stabilize the entire lattice (Fig. 17 top). If the stacking ends of the termini are treated as a base pair step in a continuous DNA helix, the average twist between each step from all known native NCP structures is calculated as $\approx 127^\circ$ (Fig. 17 table). Based on ideal B-form structure ($\approx 35^\circ/\text{step}$), this value necessitates 3-4 additional base pair steps to join DNA fragments from one nucleosome core to the next. However, the distance between stacked NCP termini is roughly equivalent to the average base pair step rise of an ideal helix ($\approx 3.4 \text{ \AA}$). While not mechanically coupled, rise and twist parameters were the most important determinants to consider when designing base pairing constructs.

A precise degree of twist is necessary for single stranded termini of adjacent core particles to be in phase for standard base pairing. Consequently, we placed greater emphasis on achieving agreeable twist parameters, as this has more significance over local structure than does rise. Addition of base pairs to achieve necessary twist

requirements passively increases rise values, which affect the distance between particles along the *c*-axis. Increases in rise may therefore yield greater unit cell volume leading to less dense packing and a resulting decrease in lattice stability, making the determination of optimal rise-twist interplay critical. For NCP termini, the ideal construct would require 143+4 base pairs of twist and rise (143 core base pairs and 4 nucleotide overhangs), for an effective paired length of 147 base pairs with an over-twist average of approximately 17°.

The original development goal was to create a library of base pairing NCP DNA sampling an entire turn (10 to 12 bps) to optimize the length-twist relationship and account for any variability in the crystallization process. Initial cloning attempts were conducted on a 147+4 base pair construct due to ideal calculations and for simplicity as this only required incorporation of a PstI site flanking the core sequence without base pair addition or deletion (Fig. 17). However, the failure of in-lab palindromic cloning experiments and success of the 147+4 non-palindromic construct made further screening unnecessary.

The aim of this project is two-fold. First, to design NCP DNA constructs capable of base pairing at the DNA termini crystal contacts to improve crystallization kinetics, i.e. reduce crystallization time and broaden the range of favorable conditions, increase lattice stability when perturbed by metal-based reagents, and emulate a fiber-like array throughout the crystal. Second, to use this technology as a tool for acquiring diverse nucleosome structures with a focus on heterogeneous sampling of DNA sequences, culminating in structures with direct genetic value.

bp #	Sequence	Step	Shift (Å)	Slide (Å)	Rise (Å)	Tilt (°)	Roll (°)	Twist (°)	Context
145	601 (3LZ0)	TA/TA	-0.35	0.46	3.15	1.10	-2.14	113.12	-TA-
	601L (3UT9)	AT/AT	-2.54	-0.47	3.77	9.31	2.36	106.25	-AT-
	145α (2NZD)	AT/AT	-3.68	1.10	4.14	7.80	-2.16	135.34	-AT-
146	146α (2CV5)	AT/AT	-3.53	0.66	4.12	6.23	-10.09	138.24	-AT-
	146b (3UTB)	TA/TA	-0.38	1.78	3.51	4.44	-0.80	112.30	-TA-
147	TA (3LEL)	AT/AT	-3.12	-0.56	4.63	-10.88	-22.29	136.75	-AT-
	TA2 (3UTA)	AT/AT	-3.58	0.99	4.19	8.22	-7.97	139.69	-AT-
	MMTV-A(5F99)	AT/AT	0.72	-0.83	3.26	-6.97	6.14	111.41	-AT-
	147α (1KX5)	AA/TT	-0.27	-3.39	3.80	11.88	10.85	147.06	-TT-
Avg.	--	--	-1.86	-0.03	3.84	3.46	-2.9	126.68	-

Figure 17. Rotational and translational geometry of base pair step stacking termini in blunt end NCP. Along one axis, NCPs align in a planar fashion allowing DNA termini of an adjacent particle to stack (top left). Terminal base pairs that stack, represented as base-pair step blocks colored cyan and magenta to delineate different core particles, are offset from standard helical twist by an average value of 126.68° (inset). Parameters were calculated for each terminus between two adjacent NCPs in any given native crystal structure. Base-pair step parameters were calculated for all known structures of NCP DNA using the x3dna webserver for analyses, reconstruction, and visualization of nucleic-acid structures (bottom table) (<http://w3dna.rutgers.edu>).

3.1.2 Palindromic cohesive end NCP DNA cloning

The cloning strategy employed to build base pairing NCP DNA constructs was adapted from Dyer, P.N. *et al.*, 2004, which exploits the complimentary overhangs resulting from BamHI and BglII restriction enzyme digestion (Fig. 19). For in-lab palindromic cloning trials, 71+3 base pair fragments of 601L were used as PCR templates to construct A and B half-sites with PstI restriction sites adjoining downstream EcoRV sites (Fig. 18). One forward and two reverse primers were designed that maintain the distinct HinfI sites of fragments A and B (dyad: A or T) while adding PstI sites to the termini and BamHI, BglII, and HindIII cloning sites and doubling sites (Fig. 18). Half-site A was amplified first with Forward A&B and Reverse A to preserve the GAATC HinfI site. Fragment A and pUC19 cloning vector were then digested overnight with BamHI/HindIII followed by PCR cleanup and T4 ligase treatment to produce

pUC19+A. Next, fragment B was amplified with Forward A&B and Reverse B followed by BamHI/HindIII restriction enzyme treatment and PCR cleanup (Fig. 20). pUC19+A is then sequentially digested with HindIII and BglII. BamHI digestion of fragment B produces 5'G|GATCC3' ends while BglII digestion of pUC19+A yields linear vector with 5'A|GATCT3' ends. These ends are then ligated to form 5'AGATCC3', a site that is not recognized by either enzyme in subsequent rounds of restriction digests and ligation. This step produces pUC19A+B or one tandem cloning repeat. From this point on the prior steps are repeated to double the tandem AB insert with each successive pass creating pUC19_2(A+B), then 4(A+B), and finally 8(A+B) (Fig. 20).

Figure 18. Primer design for a 147+4 base pair palindromic cohesive end NCP DNA construct based on the 601L core sequence. Construction of the final 147+4 base pair fragment follows a protocol similar to the creation of 601L. PstI sites incorporated by the forward A&B primer become NCP DNA termini with a four base pair 3' overhang while BamHI and HindIII restriction sites are used for cloning repeat doubling. The PstI site in reverse B is for downstream purification and processing only.

This strategy successfully produced a two repeat cloning insert (Fig. 20). Sufficient vector and insert purity were obtained with trace amounts of uncut super-coiled (SC), nicked conformer, and concatemerized DNA (Fig. 20b - d, & e). Ligation of pUC19_2(A+B) and the 2(A+B) insert gave four repeat clones that appeared very unstable when analyzed by gel electrophoresis (Fig. 20f; lanes 2, 3, 4, 5, 8, and 9 blue boxes). Bands appearing below SC DNA (Fig. 20f; lanes 5 & 9) are also cut when treated with either BglII (Fig. 20f; lanes 2 & 6), HindIII (Fig. 20f; lanes 3 & 7), or BamHI/HindIII (Fig. 20f; lanes 4 & 8). The presence of these bands suggests they are constructs of unknown repeat length possibly created during replication by complex nucleic acid secondary structure or recombination through the RecF pathway. After

double restriction enzyme treatment, lower bands appear below the desired insert bands. For example faint ≈ 500 and ≈ 600 bp bands below the more intense band at ≈ 800 bp in lane 4 correlate with linear constructs running at $\approx 3200 - 3300$ bp (Fig. 20f; lanes 2 and 3). These bands are additional evidence for some type of recombination/structural rearrangement event.

Figure 19. Digestion and ligation strategy for palindromic NCP DNA constructs. (a) Overview of steps involved in producing a high copy cloning vector carrying tandem non-palindromic repeats of 601L, DNA145 α , or cohesive-end NCP DNAs. (b) The principle of this strategy is based on the compatible overhangs of BamHI and BglII, which when ligated, form a site not recognized by either enzyme, making further downstream digestion-ligation steps possible.

To understand if interfering bands were a product of contamination or cell subtype background, the Dyer, P. *et al.*, 2004 cloning protocol was tested on two additional cell lines; JM109 and HB101. Superfluous bands still appeared in all cell types, further implying that contamination or cell background are not causes of extraneous band appearance. JM109 cells appear to not alter the desired construct, however insert runs at ≈ 300 bp, which is not consistent with two or four repeat lengths (Fig. 20f; lanes 10 - 13). This observation, along with linear and SC plasmid DNA running at incorrect sizes (≈ 3000 & 2700 bp respectively), suggests that the JM109 cell line may alter unwanted plasmid DNA via a different cellular pathway (Fig. 20f; lanes 10 - 13). Bands appearing above SC DNA in lanes 5, 9, and 13 are linear and nicked conformers and concatemers that disappear upon restriction enzyme treatment. Therefore, they most likely do not add to the series of lower bands found below intense vector and insert bands (Fig. 20f; lanes 2, 3, 6, 7, 10, and 11).

In an effort to completely rule out contamination and/or mitigate unwanted ligation products, gel extraction was carried out to purify four repeat insert obtained in figure

20f lanes 4 and 8, and empty pUC19 with BamHI/HindIII cohesive ends (Fig. 20g; lane 2, h lane 3). The purity of the four repeat fragment is acceptable as indicated by the absence of additional bands when visualized by ethidium bromide staining on a 1% agarose gel (Fig. 20g). BamHI/HindIII digested pUC19 purity is also satisfactory as lower bands observed in figure 20h lane 2 are removed after gel extraction, with a trace amount of the nicked DNA conformation running at a higher apparent molecular weight (Fig. 20h; lane 3).

Despite these efforts to control for unwanted contaminant introduction, lower bands still appeared after ligation. Transformation into three different cell lines, and restriction enzyme analysis, strongly indicate that constructs containing four repeats and above are subject to secondary structure rearrangements or recombination events during bacterial replication (Fig. 20i; lanes 2, 3, 4, 5, 10, and 13). Ligation of 4(A+B) and pUC19_4(A+B) was attempted with insert from figure 20g and vector from figure 20f resulting in constructs and fragments of unanticipated size (Fig. 20j). In contrast to prior ligations shown in figure 20b – i where unexpected lower bands appear, gel electrophoresis analysis of 4(A+B) ligated to pUC19_4(A+B) reveal upper bands above the most intense insert band at $\approx 750\text{bp}$. The expected insert and construct size are $\approx 1600\text{bp}$ and $\approx 4300\text{bp}$ respectively. Sequencing the “eight repeat” samples by an off-site third party (1st BASE www.base-asia.com) confirmed that incorrect product was obtained as local base pair alignments were unable to match any part of the expected sequence (data not shown).

Figure 20. Cloning of a high repeat base pairing NCP DNA construct based on the 601L core is subject to random bacterial recombination and/or secondary structure rearrangement. Representative gels of two cloning attempts. All lanes loaded with $\approx 250 - 300$ ng of DNA. First cloning trial (a – f) where green arrows represent ligation steps and unstable constructs are boxed in blue. **(a)** Successful polymerase chain reaction of 71+3 half-sites from 601L production creating A (lanes 4 & 5; 115bp) and B (lanes 6 & 7; 123bp) fragments. Lane 1 = 100bp marker, lane 2 = 71+3 control, lane 3 = 145bp control. **(b & c)** Sequential and double digests of pUC19(A+B) with HindIII (lane 2b and 3b; ≈ 2900 bp) then BglII (lane 2c and 3c; ≈ 2900 bp) and the 1 repeat fragment resulting from BamHI/HindIII double digest (lane 4b; ≈ 210 bp). Ligation of a 1 repeat fragment onto pUC19(A+B) yields pUC19_2(A+B). Lane 1b & c = 1kb marker, lane 5b = 100bp marker. **(d & e)** Sequential and double digests of pUC19_2(A+B) with HindIII (lane 2d and 3d; ≈ 3100 bp) then BglII (lane 2e and 3e; ≈ 3100 bp) and the 2 repeat fragment resulting from BamHI/HindIII double digest (lane 4b; ≈ 400 bp). Ligation of a 2 repeat fragment onto pUC19_2(A+B) yields pUC19_4(A+B). **(f)** Restriction enzyme analysis of ligation and transformation products from 2(A+B) and pUC19_2(A+B) ligation in DH5 α (lanes 2 – 5), HB101 (lanes 6 – 9), and JM109 (lanes 10 – 13) cells. Lanes 5, 9, and 13 are SC plasmid (≈ 2700 and ≈ 2000 bp respectively). BglII digestion, linear vector (lanes 2, 6, and 10; ≈ 3500 bp and 3000 bp respectively). HindIII digestion, linear vector (lanes 3, 7, and 11; ≈ 3500 bp and ≈ 3000 bp respectively). BamHI/HindIII digestion, linear plasmid and fragments of unknown repeat length (lanes 4, 8, and 12; vector at ≈ 2700 bp; fragments at ≈ 800 bp lanes 4 & 8, and ≈ 300 bp lane 12). Second cloning attempt (g – i) with identical indicators as in a – f. **(g & h)** Gel purification of 4 repeat insert (lane 2g; ≈ 800 bp) and empty pUC19 digested with BamHI/HindIII (lane 3h). Lane 1g = 100bp marker, lane 1h = 1kb marker, lane 2h = unpurified BamHI/HindIII pUC19. **(i)** Restriction enzyme analysis of ligation and transformation products from 4(A+B) and pUC19-BamHI/HindIII ligation in DH5 α (lanes 2 – 5), HB101 (lanes 6 – 9), and JM109 (lanes 10 – 13) cells. Lanes 2, 6, and 10 are SC plasmid (≈ 2500 bp; lanes 2 & 10; ≈ 4000 bp; lane 6). BglII digestion, linear vector at ≈ 3500 bp (lanes 3, 7, & 11). HindIII digestion, linear vector at ≈ 3500 bp (lanes 4, 8, & 12). BamHI/HindIII digestion, linear vector at ≈ 2700 bp and fragments ≈ 800 bp in size (5, 9, & 13). **(j)** Result from ligation of 4(A+B) (f) and pUC19_4(A+B) (f). Upper bands appear in both SC (lane 4; $\approx 6000 - 8000$ bp) and double digested samples (lanes 2 & 3; $\approx 1000 - 1500$ bp). Lanes 1 & 5, 1kb marker.

3.1.3 Non-palindromic cohesive end NCP DNA constructs

Unsuccessful in-lab cloning attempts of palindromic DNA fragments compelled *de novo* design and synthesis of an NCP DNA construct using a modified non-palindromic 601 sequence. The Widom 601 construct is an artificially evolved strong positioning NCP DNA that has been well studied and characterized (Lowary, P.T. and Widom, J. 1998; Vasudevan, D. *et al.*, 2010; Makde, R.D. *et al.*, 2010). The modified

601 sequence, denoted as 601* and outlined in figure 17, has seventeen base pair alterations in its right half and two in its left half. These alterations are mainly substitutions of more rigid CC or GG permutations ($-1.82 \text{ kcal}\cdot\text{mol}^{-1}$ stacking energy) to several more deformable TA steps ($-0.57 \text{ kcal}\cdot\text{mol}^{-1}$ stacking energy) that allow flexibility for DNA wrapping (Table 2). The reverse applies for substitutions in the terminus around SHL +6, where two TA steps are replaced with CC and GG steps, and at SHL +3 where a TT step is replaced with a GC step to foster stronger positioning (Table 2) (Krueger, A. *et al.*, 2006).

Dinucleotide step	Stacking energy $\text{kcal}\cdot\text{mol}^{-1}$
GC	-2.55
GT	-2.19
AC	-2.19
GG	-1.82
CC	-1.82
GA	-1.81
TC	-1.81
AT	-1.72
AA	-1.49
TT	-1.49
CT	-1.44
AG	-1.44
CG	-1.29
CA	-0.93
TG	-0.93
TA	-0.57

Deformability

Table 2. Dinucleotide step stacking energies. Data is taken from Krueger, A. *et al.*, 2006 and arranged in order of ascending deformability for individual base steps. The highest stacking energy GC step is the least deformable and the most stable. Likewise, TA steps are the least stable and offer the highest degree of sequence flexibility.

More than half of these substitutions are purine-to-purine, or pyrimidine-to-pyrimidine, which maintain basic chemical structure (Fig. 21a). The total effect of these changes increases DNA flexibility at key octamer contact points to aid correct positioning (Chua, E.Y, *et al.*, 2012). In fact, 601* does not form offset conformations as 601 and α -satellite sequences do and therefore does not require heat shifting (Fig. 21a) Termini truncations and additional alterations were necessary depending on the selected restriction site so that enzyme activity would not be inhibited by insufficient flanking sequence length or rigidity (Fig. 21b).

Figure 21. Design scheme for 601 derived non-palindromic base pairing NCP DNA. (a) Several base step alterations in the right half help with efficient octamer positioning of the DNA during reconstitution. No heat shifting is required as no offset conformation appears as with NCP601 or NCP145. (b) The core sequence of 601* is used for all cohesive end DNA constructs, which differ only in their termini sequences. Numbers above the

Once the desired sequence was confirmed, intervening protection fragments were added to delineate four separate repeats, features assuring high yield, and the construct was submitted to a commercial vendor (GenScript, Inc) for synthesis into a pUC19 vector. In all, four constructs were designed using the same modified 601* core that varied only in their termini orientation or sequence (Fig. 21).

3.1.4 Cohesive end NCP DNA Preparation

Preparation and purification of NCP DNA with overhang termini follows a similar albeit much simpler protocol than palindromic constructs. This is because non-palindromic constructs can be cloned as direct repeats rather than the tandem repeat half sites required of palindromic sequences. Inverted repeats are associated with genetic instability due to their propensity to form hairpin and cruciform secondary structure and are therefore unclonable (Leach, D.R. 1994; Malagon, F. and Aguilera, A. 1998). As

discussed in section 3.1.3, the core sequence of cohesive end NCP constructs (DNA147SK, XK, and K) is a modified version of the Widom 601 core. Whereas the 601L core used in palindromic cloning experiments is the left half of 601 inverted about the dyad, the 601* core differs in roughly three areas only in the right half (Fig. 21). This simple alteration decreases the number of processing and purification steps from six to two or three, which greatly reduces the amount of preparation time and allows faster screening when going to directly to experimentation.

The three constructs tested here differ only in their termini sequences, and so changes in protocol dictate the type of enzymes used during restriction digest (SacI, XhoI, KpnI) (Fig. 22a - c). As well, all three constructs are similarly designed such that their length and spacing within the plasmid result in digested fragment sizes that are consistent from one construct to the next (Fig. 22a - c). This allows for the use of a standard salt gradient, prep-to-prep discrepancy notwithstanding, during anion exchange chromatography (Fig. 22a - c). After pooling fractionated samples devoid of contamination, ethanol precipitation concentrates the desired product and the resultant purity is sufficient for crystal preparation (Fig. 22a - c). The only significant sequence specific protocol change applied to DNA147XK purification. Here, the presence of one 5' overhang required an intermediate dephosphorylation step to remove a terminal phosphate group that could interfere with end-to-end base pairing. This step was performed with CIP just before separation by anion exchange chromatography.

Figure 22. Restriction enzyme treatment and purification of cohesive end NCP DNA constructs. (a) *SacI*-*KpnI* double digest of DNA147SK yields a construct with incompatible 3' overhang termini (left). Separation of plasmid DNA and 18bp protection fragment from the 143+4 fragment by anion exchange chromatography (middle & right). (b) *XhoI*-*KpnI* double digest of DNA147XK yields a construct with one 5' and one 3' overhang (left). Separation of plasmid DNA and 18bp protection fragment from the 143+4 fragment as in a (middle & right). (c) *KpnI* digestion of DNA147K yields a construct with compatible cohesive end termini (left). Separation of plasmid DNA and 18bp protection fragment from the 143+4 fragment as in a & b (middle & right).

3.2 Preliminary models of NCP with incompatible overhang termini

Details concerning crystallization for each of the three cohesive end constructs are described in the following sections. X-ray diffraction data for each crystal type was collected at a wavelength of 1 Å using the SLS synchrotron source (NCP147K) or the NISB home source (NCP147SK and XK). In comparison to native NCP145 crystals, NCPs composed of human octamer and cohesive end DNA constructs form crystals with several distinct morphologies (Figs. 24, 25, 28, & 30). Despite these morphological discrepancies, we observed no new NCP space groups; all structures presented conform to the orthorhombic $P2_12_12_1$ space group (Table 3).

As a result, unit cell dimensions are consistent among the three cohesive end structures, displaying only small variations in cell volume ranging in difference from 0.1% to 1.4% (Table 4). Unit cell volumes of base pairing NCPs are also comparable to blunt end NCPs but diverge more in a single direction. Cell volume for sequence related NCP601 (PDB codes 3LZ0) differs by $\approx 2.6\%$, which is accounted for by a 8.07 Å increase along the *c*-axis, whereas the volume of NCP147 (PDB code 1KX5) differs by as little as 0.9% with similarity along all crystallographic axes. Helix stretching and deformation do not occur in NCP147 to the same extent as in shorter NCP DNAs, and so unit cell dimensions are not drastically different. Therefore, the slight unit cell increases of base pairing NCPs do not appreciably influence crystal-packing density relative to blunt end NCPs.

Atomic displacement parameters (B-factors), however, are starkly different between cohesive end NCP structures, particularly for the protein fraction where there is a maximum difference of 42.6% between XK and K octamers (Table 4). This is possibly a result of higher histone tail mobility and displacement in the SK and XK octamers around the DNA entry/exit sites and areas of maximum superhelix curvature. B-factors between the nucleic acid fractions of cohesive end NCPs deviate by a maximum of 34.7%, which is likely a direct result of interference caused by incompatible overhangs or partially and/or completely base paired termini from those constructs. Further structural and stability comparisons between NCP601 and NCP147K structures are presented in section 3.3.2.

Table 3. Data collection of refinement statistics for cohesive end NCPs

	NCP147SK	NCP147XK	NCP147K
Histones	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>
Data collection*			
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions			
<i>a</i> (Å)	105.17	104.92	105.31
<i>b</i> (Å)	109.74	109.78	109.66
<i>c</i> (Å)	183.79	181.8	183.82
Resolution (Å)	3.35–48.65	3.50–48.64	2.25–94.18
	(3.35–3.53)	(3.50–3.69)	(2.25–2.37)
R _{merge} (%)	15.1 (95.1)	18.0 (95.6)	9.5 (194)
R _{pim} (%)	6.9 (43.0)	8.1 (42.4)	3.0 (76.0)
<i>I</i> / <i>σI</i>	10.6 (1.8)	6.8 (1.9)	12.2 (1.0)
Completeness (%)	100 (100)	99.8 (997.)	99.7 (97.7)
CC (1/2) (%)	99.8 (80.2)	99.7 (74.8)	99.7 (61.6)
Redundancy	6.7 (6.7)	6.7 (6.8)	11.5 (7.9)
Refinement			
Resolution (Å)	3.35–48.65	3.50–48.64	2.25–94.18
No. reflections	42833	41810	107427
R _{work} / R _{free} (%)	22.6 / 28.2	22.8/31.2	24.7 / 29.7
No. atoms	12007	12007	12226
Protein	6068	6068	6127
DNA	5939	5939	6029
Solvent	--	--	53
Ions	--	--	17
B-factors (Å ²)	114.1	120.2	78.2
Protein	85.7	95.3	61.8
DNA	143.1	145.0	102.1
Solvent	--	--	50.7
Ions	--	--	98
R.m.s. deviations			
Bond lengths (Å)	0.0089	0.0090	0.0087
Bond angles (°)	1.3765	1.4279	1.4972

Single crystal data sets; values in parentheses are for the highest resolution shell.

NUC type	Unit cell volumes (Å ³)	Component	B-factors (Å ²)		
			SK	XK	K
SK	2121186	H4	83.8	93.0	59.5
XK	2093994	H3	80.3	90.6	57.4
K	2122808	H2A	78.1	95.4	63.3
601	2069315	H2B	93.0	102.4	66.9
147	2101656	DNA	143.1	145.0	102.1

Table 4. Cohesive end NCP crystallographic data.

Crystallization of all three cohesive end constructs followed traditional NCP preparation and handling protocols under monovalent, divalent, and buffer conditions first outlined by Luger, K. *et al.*, 1997a, with slight variations in concentration depending on the construct. As these salt and buffer conditions were tested first, and produced diffraction quality crystals, no further screens were attempted. However, future work involving chromatin binding factors or more complex metaloreagent designs may require broader crystal screens to account for the presence of bound protein or adduct. Therefore, future work with cohesive end NCPs will focus their robustness of the particles to form under a wide range of conditions.

It is worth noting that structures of several NCP:factor complexes have been solved under a broad range of conditions (Makde, J.R. *et al.*, 2010; Armache, K.J. *et al.*, 2011; Zhou, B.R. *et al.*, 2015) (Table 5). Procuring these structures required various molecular crowding agents and dehydration techniques to produce diffraction quality crystals yielding lower resolution structures (Table 5). Furthermore, none of the three structures presented in Table 5 possess crystal contacts at their DNA termini making it unclear how unresolved or truncated regions (e.g. linker histone tails) would bind in a more natural condensed state or interact with linker DNA. Therefore, it is anticipated that engineered cohesive end crystal contacts, in conjunction with thin slicing data

	3MVD	3TU4	4QLC
Histones	<i>X. laevis</i>	<i>X. laevis</i>	<i>D. melanogaster</i>
DNA	601 (146bp)	601 (146bp)	601 (167bp)
Space group	P2 ₁	P6 ₁	P6 ₅
Cell dimensions			
a (Å)	102.7	102.9	262.9
b (Å)	183.0	102.9	262.9
c (Å)	107.0	555.8	91.8
Resolution (Å)	2.90–50.00 (2.90–2.95)	3.00–50.00 (3.00–3.16)	3.50–50.00 (3.50–3.63)
Crystallization			
Buffer	Sodium Acetate (pH 5.5)	Hepes (pH 7.0)	Citric acid (pH 3.75)
Divalent	N/A	MgCl ₂	N/A
Monovalent	N/A	KCl	KCl
Cryo/Crowding/ Reducing/Additive/ etc.	PEG 2000-MME Sodium Citrate DTT	PEG 400	MPD
Temp (°C)	21	20	18
Conc (mg·mL ⁻¹)	≈12-18	≈8-12	≈10

Table 5. Space group, unit cell dimensions, and crystallization conditions of three prominent NCP-factor bound crystal structures (Makde, J.R. *et al.*, 2010; Armache, K.J. *et al.*, 2011; Zhou, B.R. *et al.*, 2015).

acquisition, will lend a regular and reproducible crystal lattice to nucleosome–protein complex crystallization trials.

3.2.1 Preliminary structure of human NCP with non-complimentary 3' termini

As discussed, the most dominating NCP crystal contacts are face-to-face electrostatic interactions, where basic H4-tails bind the acidic patch of adjacent particles, and planar end-to-end stacking interactions formed between DNA blunt ends of neighboring nucleosomes. Depending on the DNA sequence used, nucleosomes can form arrays with particles in mixed orientation. NCPs using the non-palindromic Widom 601 sequence are a characteristic example of mixed orientations (PDB codes 3LZ0 and 3LZ1). Like Widom 601, most DNA sequences of genetic import are non-palindromic whose NCP crystals would contain particles in one of two randomly selected orientations. However, NCPs possessing palindromes of centromeric α -satellite DNA suffer from suboptimal binding and positioning mechanics, and solution of an asymmetric NCP structure crystallized in magnesium buffer with a natural viral DNA promoter relied on an H4-tail H18R substitution to achieve asymmetry (Chua, E.Y. *et al.*, 2012; Frouws, T.D. *et al.*, 2016). These shortcomings highlight the need for a comprehensive method of attaining atomic scale structural information of DNA sequences with genetic value i.e. promoter/terminator sites, enhancer/silencer sites, etc.

Inspiration for NCP147SK design, and subsequent base pairing NCP constructs, originated from two key observations. First, early nucleosome linker histone assembly (NLHAs) structures completed by our lab showed promise for controlling crystal packing via DNA engineering. Structurally, these assemblies consist of nucleosome cores with bound human H1 histones that are paired by linker DNA. Individual NLHA components include: a novel DNA construct built from a 601L core sequence with poly A linker extensions and PstI restriction site cohesive ends (Appendix 1: CMS169s), human histone octamer, and one of eleven human H1 variants. NLHA repeating units were expected to assume an offset orientation, similar to the termini stacking interactions of NCP crystals, by Watson-Crick pairing at their cohesive end DNA termini (Fig. 23a & b). Upon crystallization these assemblies presented an unanticipated asymmetric unit arrangement where the complete structure is a crisscrossed nucleosome pair (Fig. 23c). The NLHA system's crisscrossed conformation led us to design NCP DNA constructs with non-complimentary overhangs. This design held the expectation that steric and chemical incompatibilities at the DNA termini could direct NCP

aggregation and drive particle arrangement in the asymmetric unit thereby imposing absolute particle asymmetry throughout the crystal lattice.

Figure 23. Base step stacking or Watson-Crick pairing orientations along the crystallographic *c*-axis influence nucleosome–nucleosome packing orientation along the *a*-axis. (a) Schematic cartoon of base pair step stacking interactions at blunt end NCP DNA termini that form NCP “oligomers”. This type of packing interaction is witnessed in all NCP crystals with core particle length DNAs. **(b)** Expected oligomeric nucleosome conformation based on Watson-Crick pairing at cohesive DNA termini. The constraints were determined from NCP crystal-packing interactions applied to a NLHA crystal system. **(c)** Observed DNA termini interactions as observed in a definitive NLHA structure. Here, linker DNA with cohesive end termini pair in a crisscrossed manner that closes off a duet of nucleosomes while incorporating H1 to form a complete chromatosome.

Secondly, several important nucleoprotein complexes were crystallized with the aid of base pairing DNA fragments. According to Song, T. *et al.*, 2000, three high-resolution multimeric protein/DNA structures utilized overlapping complimentary oligonucleotides to generate long, nicked DNA fragments for crystallization. These are the *E. coli* catabolite gene activator protein (CAP–DNA), the $\gamma\delta$ -resolvase–

DNA, and the integration host factor (IHF–DNA) (Schultz, S.C. *et al.*, 1990; Yang, W. and Steitz, T. 1995; Rice, P.A. *et al.*, 1996). Furthermore, a recent low-resolution structure of histone-based archaea chromatin observed *M. fervidus* histone B (HMfB) dimer subunits joined in a superhelix wrapped by 90 base pair repeats. These DNA fragments possess two base pair 5' overhangs produced by NdeI digestion (Mattioli, F. *et al.*, 2017). This is the first example of a nucleosome-like structure crystallized using cohesive end DNA fragments. Although closer inspection reveals termini base pairing in the structure is extremely troubled. Average local base step parameters for the duplex decamer TTCA|TA|TGAG with central NdeI sequence are as follows (standard deviations in parenthesis): Shift = 2.6 Å (6.4), Slide = –0.2 Å (1.5), Rise = 2.2 Å (2.1), Tilt = –10.9° (52.9), Roll = –10.6° (50.9), Twist = 39.4° (39.3).

NCP147SK is so named because it incorporates SacI and KpnI restriction sites at the 3' and 5' termini respectively of the 601* sequence (Fig. 24a). Introduction of these GC rich regions required slight modifications to the flanking sequence so that long runs of deformation resistant GC base steps would not impede enzyme activity (Fig. 24a). Initial crystallization trials were conducted using standard manganese(II) and potassium

chloride screens. Under these divalent and monovalent conditions (Luger, K. *et al.*, 1997a) NCP147SK readily formed crystals (4 – 6 days) while spheroids appeared under NLHA crystalizing conditions (Fig. 24a & b column 2).

Figure 24. Reconstitution and crystallization of NCP147SK. (a) (top) Left terminus and complete right half of NCP147SK sequence showing restriction sites in red and nucleotide substitutions in turquoise highlight. (bot left) Native PAGE gel assesses the quality of NCP147SK and human histone octamer reconstitution. There is a negligible amount of free DNA below the 200bp marker. (bot right) The range of crystal morphologies observed from the screens described in the table below. The lower left corner shows spheroids grown under NLHA conditions. The lower right panel depicts NCP-like rods while the two upper panels depict diffraction quality prisms. (b) Full range of conditions tested during NCP147SK crystallization trials. I.F. = initial final (1X).

NCP147SK was demonstrably insoluble in comparison to native NCP145 preparations (Fig. 24b). The solubility limit for NCPs is approximately 60 mM (30 mM) MnCl₂ at 1X (0.5X). In our lab, native NCP145 preparations yield maximum crystal volume around 1X (initial final, I.F.) concentrations of 90 mM MnCl₂, 80 mM KCl, and 20 mM K-cacodylate, pH 6.0 with a starting NCP concentration of 4 mg·mL⁻¹. By contrast, coarse crystallization screens of NCP147SK discovered that core particles precipitate in manganese buffer concentrations up to 75 mM at 1X (Fig. 24b). The best NCP147SK crystals were attained at I.F. conditions of 120 mM MnCl₂, 60 mM KC, and 20 mM K-

cacodylate, pH 6.0 with a starting NCP concentration of 4 mg·mL⁻¹ (Fig. 24a & b). These conditions reliably produced well diffracting crystals with prism-like morphology (Fig. 24a). Surprisingly, NCP-like rods grew in several wells but diffracted poorly (Fig. 24a). Unfortunately, subsequent fine screening was unable to maximize crystal volume to a size comparable to NCP145. The largest prism-like crystals ranged in size from approximately 400 – 500 µm in one dimension, or about half the size of standard NCP145 crystals.

Diffraction tests, data collection, data processing, molecular replacement, and refinement on the largest diameter NCP147SK crystals produced a preliminary structure good to 3.35 Å resolution (Table 3). Molecular replacement with NCP601 (orientation 1) revealed a modified spatial arrangement of stacking particles along the *c*-axis. The distance between “stacked” termini of NCP147SK increased by approximately 5.20 Å relative to the fitted model, resulting in an overall gap of 8.35 Å between base pair ends where positive electron density appears (max signal 5.44 σ) (Fig. 26b). Careful observation of the difference map suggests that the termini are not stacking but rather engaging in alternate forms of base pairing as positive density depicts a hybrid quasi-helical structure (Fig. 26a & b). From this data, it is difficult to discern clear base pairing modalities making precise nucleotide assignment impossible.

If the interacting termini were to assume ideal helical geometry, the four base pair overhangs would require ≈13.4 Å of allowance to adopt a B-form double helix, or about 5.1 Å more than available in the structure. To be certain construct length was not a mitigating factor for correct termini base pairing, a shorter construct was produced by removing two base pairs to recover an additional ≈6.7 Å of helical rise. This adjustment, however, causes under-twist by ≈51°. The NCP145SK construct was processed, purified, reconstituted, and crystallized using parameters gathered from NCP147SK preparation and crystallization. Only microcrystals and amorphous shards appeared throughout a wide divalent and monovalent screen (50 – 220 mM Mn²⁺, 20 – 100 mM K⁺, 20 mM K-cacodylate, pH 6.0), signifying the longer 147SK DNA is better suited to NCP DNA base pairing (Fig. 25). Therefore, some magnitude of over-/under-twisting inhibits base pairing and interferes with crystallization, whereas the lattice is malleable enough to absorb surfeits or deficits in length (Fig. 17). This also illustrates how critical stable NCP–NCP termini interactions are for proper crystallization.

Figure 25. A two base pair reduction in cohesive end NCP DNA length results in suboptimal crystal formation. Only microcrystals and amorphous shards not exceeding 100 μm in length formed in coarse screens of NCP145SK.

With the “stacking” distance of the SK structure acting as a physical restraint, positive electron density implies the single stranded ends create a flattened helix with A-form character (Fig. 26). The diffuse Fo-Fc map is identical yet inverted around each terminus, indicating electron density is mixed and non-canonical base pairing occurs at both termini as some percentage of particles are in the “correct” fully paired orientation and some are not. If the single stranded ends interacted as desired, the overall shape of a B-form helix, a characteristic of properly paired helices with distinct sequences, would be clearly visible in the difference and double difference maps despite low-resolution data. In the observed state, however, the low-resolution data prohibits gaining detailed information regarding specific nucleotide interactions. The pronounced flattening suggests individual bases are frequently mispairing due to a random distribution of particle orientations throughout the lattice, and that the complimentary helical structure of dual 3' overhangs are sterically insufficient to govern particle orientation during crystallization alone, regardless of sequence non-complementarity.

Furthermore, both twist and rise appear to be facilitative and/or prohibitive factors for sticky end NCP crystallization. Twist significantly influences crystal formation to a point where single stranded helices that are discordantly over-/under-twisted cannot satisfactorily base pair thereby forming lattices with high disorder (Fig. 26; Table 4). Whereas any changes in rise that maintain favorable twist relationships are more permissive and only affect unit cell volume and inter-particle distance. Moreover, manipulation of either parameter (147SK to 145SK) did not bring about the anticipated structural asymmetry. Guided by this data, we subsequently employed a different design to try to control NCP orientation within the crystal lattice.

Figure 26. Preliminary structure of NCP147SK. (a) Overview of nucleosome–nucleosome interaction showing positive density in the Fo-Fc map (red, 3.0σ) in the region between DNA fragment ends. The black arrow indicates dyad position of the main core particle. A 2Fo-Fc map is shown in grey at a contour level of 1.0σ . (b) Detailed view of nucleosome–nucleosome DNA termini interactions along the *c*-axis. Maps represented as above. Positive electron density at either terminus is an inversion of the other suggesting density is mixed resulting from two particle orientations in the lattice. Table values characterize base step parameters between the last base pair of “stacking”

3.2.2 Structure of the human NCP with incompatible 5' and 3' overhangs

The failure of the NCP147SK crystal system to achieve an asymmetric structure prompted modifications in the subsequent construct design. Although the 3' cohesive ends of 147SK DNA possess non-complimentary base pairing sequences, the single stranded ends share structural compatibility in the form of complimentary helical phase and are flexible enough to engage in non-canonical interactions. Therefore, the double 3' overhang design did not introduce sufficient constraints to force discrimination between particle orientations and control crystallization. To test this observation, we designed NCP DNA with one 5' and one 3' overhang to introduce greater steric restraints that could lead to better discrimination between particle orientations (Fig. 27). 147XK DNA replaces the *SacI* restriction site with an *XhoI* sequence while keeping *KpnI*, placing both sticky end termini on the same DNA strand.

Figure 27. Construct specific base pairing modes at overhang NCP DNA termini. Each overhang originates from a separate DNA fragment (e.g. NUC1 = yellow, NUC2 = cyan). **(a)** Watson-Crick base pairing in sticky end NCP at 3' overhangs (e.g. *SacI*-*SacI*/*KpnI*-*KpnI*). **(b)** Watson-Crick base pairing in sticky end NCP at 5' overhangs (e.g. *XhoI*-*XhoI*). **(c)** Interference of two approaching overhangs said to be “out of phase” by 180° as a result of restriction enzyme site type (e.g. *XhoI*/*KpnI*). DNA–octamer associations and the dominant face-to-face nucleosome interaction of the NCP fixes strand orientation such that it cannot rotate about the helical axis.

NCP147XK solubility was more in line with the behavior of blunt end NCP145 but demonstrated faster crystallization kinetics. An initial coarse screen led to crystallization under typical NCP buffer conditions where crystals formed after 3 to 4 days of incubation (Fig. 28c). Subsequent fine screening optimized crystal size to $\approx 350 - 400 \mu\text{m}$, slightly smaller than the largest SK crystals. XK crystals also appeared more irregular and less heterogeneous than SK crystals, forming only prism-like morphotypes with jagged edges (Fig. 28b).

Figure 28. Reconstitution and crystallization of NCP147XK. (a) Left terminus and complete right half of NCP147XK sequence showing restriction sites in red nucleotide substitutions highlighted in green. (b) (left) Native PAGE gel assesses the quality of NCP147XK and human histone octamer reconstitution. There is a negligible amount of free DNA below the 200bp marker. (right) Crystal morphologies observed from the screens described in the table below. XK crystals appear as smaller jagged versions of the prism-like morphology observed for SK crystals. (c) Full range of conditions tested during NCP147XK crystallization trials.

As with NCP147SK, diffraction tests and data collection using the largest XK crystals followed by data processing, molecular replacement, and refinement produced a preliminary structure at $3.50 \text{ \AA}$ resolution (Table 3). Nearly all measurements of the NCP147XK structure are consistent, within error, with those recorded for the SK structure. The “stacking” gap between NCP DNA termini in the XK structure is $\sim 8.7 \text{ \AA}$, an increase of only $\sim 0.35 \text{ \AA}$ over the SK structure (Fig. 29). Furthermore, base pair step parameter calculations reveal increased tilt between NCP147XK DNA ends relative

to NCP147SK (21.3° vs 12.6°) suggesting a greater occurrence of helical deformation from structurally incompatible cohesive ends (Fig. 29). Signal intensity of the positive electron density around the DNA termini is also slightly weaker for the XK structure with a maximum signal of 4.86σ . Finally, the identical inverted Fo-Fc map appearing in the gap between interacting NCP superhelices once again suggests mixed density arising from two randomly configured particle orientations within the lattice where some NCPs will pair fully, some partially, some non-canonically, and some will clash at the termini (Fig. 29).

The most noticeable difference between the two models is additional positive density located in the solvated area where three nucleosomes converge (Fig. 29). It is highly likely this density portrays displacement of an overhang at one DNA terminus, as it appears to follow the distinctive helical pattern of a single stranded phosphate backbone. If so, partial strand displacement would explain the minor differences observed between XK and SK structures; i.e. the apparent increase in inter-particle distance, exaggerated tilt at the DNA termini interface, and a less ordered and less extensive difference map. The putative displacement event may occur when the dominant and indiscriminant H4-tail–acidic patch crystal packing interaction forces two structurally incompatible single stranded helices (e.g. 5' overhang to 3' overhang, or 3' overhang to 5' overhang) to interact (Fig. 27). Weak signal implies displacement may occur intermittently during crystallization (maximum signal = 4.06σ), but frequently enough to have a subtle effect on the inter-particle structure of NCP147XK. Unfortunately, the low-resolution data, poorly ordered density and weak signal strength make explicit modeling of a single nucleotide strand unfruitful.

Figure 29. Preliminary structure of NCP147XK. (top) Overview of nucleosome–nucleosome interaction showing positive density in the Fo-Fc map (red, 3.0σ) in the region between DNA fragment ends. The black arrow indicates dyad position of the main core particle. A $2Fo-Fc$ map is shown in grey at a contour level of 1.0σ . Density is mixed as in the SK structure. (bot) Detailed view of additional electron density appearing in a large solvent channel between three adjacent core particles. Possibly a single stranded overhang shifted into the void due to steric interference from an incoming terminal strand with incompatible helical phase.

Like NCP147SK, NCP147XK failed to achieve an asymmetric crystal structure. Modified with a 3' to 5' cohesive end substitution, the XK construct was created as an

experimental method of restraining particular NCP interactions by allowing steric repulsion to guide selection of one particle orientation during nucleosome crystallization. However, even substantial interference at crucial DNA–DNA nucleosome contacts could not effectively overcome the energetic contribution of electrostatic face-to-face nucleosome interactions to crystal formation. Consequently, it may be impossible to control particle orientation within the lattice by DNA engineering alone. The following section will focus on the crystal structure of a base pairing NCP with complimentary cohesive end sequences.

3.3 Crystal structure of human NCP with base pairing termini

3.3.1 Design, reconstitution, high-resolution and DNA termini structures

Neither the 147SK nor the 147XK DNA designs could achieve fully asymmetric NCP crystal structures. However, direct experimentation with the prior two constructs established greater certainty in correct geometric values for both length and twist. As discussed, twist and rise values are the major determining factors to consider when designing base pairing NCP DNA for nucleosome crystallization. Using the same core sequence (601*) and terminal end nucleotide substitutions of 147XK DNA, the 147K NCP DNA construct incorporates two fully compatible 3' cohesive end **KpnI** sequences (Fig. 30a).

Previous trials with SK and XK constructs helped simplify preparation and setup for NCP147K crystallization. Again, an initial broad manganese screen determined solubility limits and crystallization time for the 147K nucleosome. NCP147K was noticeably more soluble than either NCP147SK or NCP147XK, forming the largest crystals of the three at approximately 600 μm under the lowest divalent concentration (Fig. 30b). Additionally, NCP147K crystal formation occurs the most rapidly out of the three constructs tested. Microcrystals appeared in the lowest divalent conditions within an hour of setup and reached terminal size within 1 – 2 days of incubation (Fig. 30b). In fact, all three cohesive end NCPs developed crystals faster than blunt end NCP145, which requires around 10 days of incubation at slightly higher divalent and monovalent concentrations. Included in the initial screen, optimal NCP145 conditions were found to yield NCP147K crystals in approximately half the incubation time (4 – 5 days), suggesting the decrease in crystallization time is exclusively due to cohesive end pairing (Fig. 30a).

Although NCP147K crystallizes under a broad range of divalent and monovalent concentrations, ideal conditions for NCP147K were determined as 60 – 80 mM manganese(II) chloride, 30 – 60 mM potassium chloride, and 20 mM K-cacodylate, pH 6.0 (Fig. 30a). Crystals grew quickly under these conditions and formed either of two well diffracting morphologies. The prism-like morphotype, observed for both SK and XK variants, appeared larger and without surface defects in 147K crystals (Fig. 30b; left). The more block-like morphology grew smaller than 147K prisms but had sharper edges (Fig. 30b; center & right). The diffraction data set used to solve the NCP147K structure to 2.25 Å was collected from a crystal with prism-like morphology.

Figure 30. Reconstitution and crystallization of NCP147K. (a) (top) Left terminus and complete right half of NCP147K sequence showing restriction sites in red with nucleotide substitutions highlighted in green. (bot left) Native PAGE gel assesses the quality of NCP147K and human histone octamer reconstitution. There is a negligible amount of free DNA below the 200bp marker. (bot right) Range of conditions tested during NCP147K crystallization trials. (b) Crystal morphologies observed from the screens described in the table above. NCP147K crystals appear as larger, more regular, and more voluminous prisms than those attained from SK and XK constructs

The same data collection, processing, and refinement workflow used for NCP147SK and XK was applied to NCP147K, assaying the largest crystals of either morphology. Atomic restrained refinement produced a final structure of high quality at 2.25 Å resolution (Table 3; Fig. 31). Very little is different about the 147K core particle structure relative to other NCP models with the noticeable exception of cohesive ends projecting from the DNA entry and exit points (Fig. 31). However, within the crystal, cohesive ends organize into B-form helices that connect one particle to the next (Fig. 31 & 32). Throughout the entire crystal, these termini interactions join all nucleosomes across a single layer into a type of oligomeric crystal structure that mimics a tightly associated version of euchromatin dubbed beads-on-a-string (Joti, Y. *et al.*, 2012; Grigoryev, S.A. 2018). Interaction between layers is also similar to states of heterochromatin with short nucleosome repeat lengths (NRLs) as in telomeres and, potentially, subtelomeric and knob regions (Ichikawa, Y. *et al.*, 2015; Matsuda, A. *et al.*, 2015). As both NCP147K cohesive ends are complimentary but the core sequence is non-palindromic, the system cannot distinguish between particle orientations. Therefore, the crystal lattice contains two orientations that differ with respect to the position of the DNA halves.

Figure 31. Refined crystal structure of NCP147K (orientation 1) to 2.25 Å. Both the histone octamer and DNA phosphate backbone are presented in standard coloring. DNA sugars and bases appear in gray. The overall structure of NCP147K is consistent with general nucleosome core structure, with the exception of two 3' overhangs protruding from the DNA entry/exit points.

Given the high-resolution and overall quality of diffraction data, both the 2Fo-Fc and Fo-Fc electron density maps are exceptionally well defined. B-form secondary structure is easily identified from the maps alone, which also allow accurate placement of individual nucleotides and their base pair complement (Fig. 32). Using compatible cohesive ends also relieved the need for utilizing more complex methods of sequence identification that are required for asymmetric structures, e.g. identification of specific metal binding sites or sequence elements.

Figure 32. NCP147K electron density maps indicate B-form helical structure at cohesive end termini. (left) The 2Fo-Fc (gray, 1.0σ) and the Fo-Fc (red, 3.0σ) maps around the NCP147K DNA entry/exit points depict a near perfect B-form DNA helix. (right) Post-refinement electron density at the termini after building additional nucleotides of the KpnI restriction site on to the NCP601 model shows a quasi-continuous double helix. NUC1

Visual evaluation of the NCP147K data readily discloses the major structural feature of the NCP147K core particles; however, precise quantitative data is always desirable for thorough analysis. B-form secondary structure and base pairing modes were confirmed by calculating base pair and base pair step parameters with the w3dna web server for analysis, reconstruction and visualization of nucleic-acid structures (<http://w3dna.rutgers.edu>) (Fig. 33). NCP147K termini generally form Watson-Crick pairs throughout the structure with the exception of a G-C pair at the start and a C-G at the end of the KpnI sequence (Fig. 33). Further analysis presented later shows these base pairs do in fact share three hydrogen bonds each, but computational errors occur

due to the w3DNA software's inability to compensate for breaks in the phosphodiester backbone.

Figure 33. Local base pair and base pair step parameters reveal near ideal B-form geometry of paired NCP147K cohesive ends. (left) 3-D structure of a ten base pair double helix formed by the cohesive ends of two unique core particles (yellow and teal). The KpnI sequence appears to the left of the helix. Arrows indicate break points in the phosphate backbone where free 5' phosphates are attached to the overhang recess. (top right) Local base pair parameters designate Watson-Crick base pairing modes for all but the two base pairs where the phosphodiester backbone is incomplete. (bot right) Local base pair step parameters designate near ideal B-form helical structure for the paired terminal end helix. Average values for each parameter are within statistical error of ideal B-form DNA. The B-form values presented are standard coordinates generated from the 'Dickerson Dodecamer' (PDB code 1BNA). <https://www.liverpool.ac.uk/~bates/B-DNAframe.html>

Average local base pair step values conclusively establish the B-form secondary structure of paired NCP termini. Rise (± 0.17 Å) and twist ($\pm 5.6^\circ$) values, the two most crucial parameters determined during cohesive end NCP development, deviate from ideal B-form structure by only 0% and 3.8% respectively (Fig. 33). The magnitude of difference across all other parameters is also equally minimal (Fig. 33). The paired termini alter the inter-particle distance, as does the partial or incomplete pairing of NCP147SK and NCP147XK's non-complimentary cohesive ends. Here remnants of the KpnI restriction site create a four-base step GTAC sequence that lends ≈ 13.4 Å of rise to the NCP147K structure. This distance is an increase of 5.1 and 5.4 Å over the "gaps" in the SK and XK structures respectively. The stable W-C pairing and base stacking of NCP147K termini likely contribute to the overall stability of the crystal lattice as evinced by markedly lower temperature factors for NCP147K's DNA component relative to NCP147SK and NCP147XK (Table 4). A detailed treatment of this topic is presented further along in the text.

The following sections will emphasize crystallographic data when comparing the DNA structures of NCP147K (orientation 1) and NCP601 (orientation 1) across geometric and stability criteria. Analysis of the key differences between each structure will be presented and discussion will focus on the nucleic acid structure of the novel 147K NCP. Thermodynamic, topological, and metal-binding features unique to the NCP147K macromolecular structure will also be discussed.

3.3.2 NCP147K DNA structure, core particle stability, and oligomeric structure

As outlined, the core 147K DNA sequence is a minimally modified version of the strong positing Widom 601 sequence with nucleotide substitutions selected to accommodate protruding overhangs (Fig. 21). The primary structural difference between each sequence is an increase in length from the 145 base pair 601 molecule to a 147 base paired length 147K molecule. This apparent addition of two base pairs could alter the superhelical path around the octamer by affecting positioning power or through stretching and over-twisting defects made necessary to maintain the fidelity of the histone-DNA register. Mechanisms for these phenomena are well established in α -satellite and 601 bearing NCPs (Ong, M.S. *et al.*, 2007; Wu, B. *et al.*, 2010; Vasudevan, D. *et al.*, 2010; Davey, G.E. *et al.*, 2010). Overall, both NCP147K and NCP601 DNA structures are highly similar, particularly throughout the left half with an RMSD of 1.003 Å vs a right half RMSD of 1.062 Å where the majority of nucleotide substitutions are located. Such similarity results in identical stretching around SHL ± 5 in both structures. For clarification, the left half of the NCP DNA construct contacts prime histones while the right half is in contact with nonprime histones.

601 nucleotide substitutions mapped on the NCP147K right half DNA structure help visualize where sequence changes might evolve structural deviations, however, structural alignment of the 601 and NCP147K DNA molecules reveals near identical agreement between NCP DNA cores (Fig. 34a). There is slight deviation near the 601 terminus from stretching at the +5 SHL with the possibility of some degree of dynamic fluctuations arising from the flexible TATAT run at SHL ± 6 (Fig. 34a).

Crystal structures of the 145 base pair NCP601 have demonstrated that stretching in both DNA halves around SHL ± 5 maintains an effective 147 base pair length and conserves positioning with respect to the histone-DNA register (Vasudevan, D. *et al.*,

2010). NCP147K DNA local base pair step parameters were calculated utilizing the w3DNA server to perform comparative analysis with the parent 601 structure. As discussed in section 3.1.3, nucleotide substitutions in the right half of NCP147K either promote greater flexibility at minor groove dinucleotide steps associated with kinking (SHL +1.5, +2.5 and +3.5), or stabilize flexible elements that disrupt nucleosome positioning (SHL +3 and 6). Three TA substitutions at SHL +1.5, +2.5, and +3.5 (CC, GG, CC in 601) decrease the energetic penalty associated with minor groove bending. This flexibility helps the superhelix assume major and minor groove facing conformations that are conducive to strong interaction with DNA binding histone motifs.

The AG=CT step at SHL ± 4.5 in NCP601 coincides with the location of extreme minor groove kinking leading to terminal stretching and is present in both NCP147K halves (Fig. 34b). Maximal negative roll values for NCP601 ($-38 \pm 8^\circ$) and NCP147K (-20.0°) occur at the dinucleotide step between position 46 and 47. Because the 147K DNA is two nucleotides longer when paired, the severe kinking observed in the 601 core particle may be modulated in the NCP147K nucleosome. This is evident from the disparate negative roll values and suggests that the forces driving the terminal base pairing interactions may influence the conformation of neighboring and even more distant DNA elements (Fig. 34b). As both constructs assume an effective 147 length, all parameters dictating nucleosomal DNA conformation are nearly identical with both DNAs positioned strongly around the octamer surface (Fig. 34b).

Overall, the two NCP DNA structures are highly similar to the point where energetic penalties or benefits incurred by nucleotide substitutions, or their associated mechanical effects, have negligible impact on global and local structural parameters. However, these substitutions, along with the individually paired termini, do influence thermodynamic characteristics of the NCP147K crystal structure.

Figure 34. Global structure and comparison of right half base pair step parameters between NCP601 orientation 1 and NCP147K DNA. (a) (top) Comparison of the right half of 601 and 147K DNA. Bolded red nucleotides are those that differ between sequences. (left) Right half SHLs where NCP601 and NCP147K sequences differ mapped in blue onto the NCP147K right half. (right) Structural alignment of NCP601 (translucent green) and NCP147K (pink) superhelices shows full agreement of superhelix register between structures despite incongruent sequence lengths. (b) Rotational (left) and translational (right) base pair step parameters for NCP147K (thick lines) and NCP601 orientation 1 (thin lines). With very few exceptions, both sequences have highly similar nucleosome

The core sequence of 147K DNA is 143 base pairs long excluding the four nucleotide overhangs. However, these extra base pairs are a key determinant of unique NCP147K thermodynamic properties. Temperature factors determined by crystallographic data suggest NCP147K nucleosomes are more stable than 601 nucleosomes in a crystalline state, leading to a better ordered, higher resolution structure. Per residue atomic displacement data indicate paired termini as a prime contribution to NCP147K's greater relative stability within the lattice. Calculated DNA temperature factors for each chain

(I and J) of the NCP601 and NCP147K crystal structures show a considerable decrease in thermal motion from blunt end to cohesive end particles ($139.0 \text{ \AA}^2$ vs $102.1 \text{ \AA}^2$). Watson-Crick base pairing of the NCP147K overhangs stabilize termini B-factor values in a tight range ($\approx 70 - 90 \text{ \AA}^2$) relative to those observed in 601 blunt end termini ($\approx 130 - 170 \text{ \AA}^2$) (Fig. 35a).

Figure 35. Atomic displacement factors indicate NCP147K is more stable than NCP601. (a) Plots of average B-factor per nucleotide residue for both DNA strands from NCP601 (black traces) and NCP147K (red traces). Paired NCP147K termini regions are remarkably ‘quieter’ than 601 termini which experience greater thermal motion. (b) Visualization of the data in (a) mapped onto NCP601 (left) and NCP147K (right) DNA structures. Warmer colors indicate greater atomic displacement.

Energetic disparities between base pairing and stacking versus base stacking alone at the terminus contribute differentially to overall NCP stability. A single NCP601 blunt end junction forms AA and TT dinucleotide steps, or one AA=TT base pair step. In a standard B-form helix these dinucleotide steps are energetically equivalent and provide $-1.49 \text{ kcal}\cdot\text{mol}^{-1}$ to duplex stability (Fig. 33). Computed from NCP601 structures, the twist value for terminal AA=TT base pair steps is 113.1° , meaning blunt ends are $\approx 77.5^\circ$ over-twisted from the ideal (35.6°). The copious over-twist leads to minimal base

stacking at the blunt end interface thereby straining pi interactions between nucleotide rings. This means the stacking energy between the blunt ends is likely even less favorable than -1.49 kcal/mol. Hydrophobic, electrostatic, and dispersive forces are also negatively affected, resulting in weakly positioned NCP DNA ends (SHLs ± 5.5 to ± 7) which in turn display high B-factors (Fig. 35b).

The B-form duplex of an NCP147K terminus creates base pairing and stacking interactions between two distinct nucleosome particles. Experimental evidence put forth by Yakovchuk, P. *et al.*, 2006 determined the extent of base stacking as the primary contribution to DNA duplex stability. Thermodynamically, the central A-T/T-A base pairs of the KpnI sequence are moderately stabilizing, while the surrounding G-C/C-G base pairs provide more pronounced stability (Yakovchuk, P. *et al.*, 2006). The uniformity of paired termini in the NCP147K structure result in closer to ideal twist values for both GG dinucleotide steps (as GG=CC and CC=GG base pair steps), thus maximizing their stacking interface (Figs. 33 & 36). The net effect of paired terminus stacking interactions create favorable and stabilizing conditions that resemble nicked duplex DNA, as opposed to the double stranded break likeness of blunt end NCPs. Therefore, the calculated values presented in figure 33 are affected by local base pair and base pair step conformations that maintain or decrease the NCP termini's contribution to overall stability.

		KpnI termini	kcal·mol ⁻¹
		GG=CC	-1.82
		GT=CA	-2.19
		TA=AT	-0.57
		AC=TG	-2.19
		CC=GG	-1.82
601 termini	kcal·mol ⁻¹	Total	-8.59
AA=TT	-1.49		
Total	-1.49		

Figure 36. Ideal inter-particle stabilization energies at blunt end (601) and cohesive end (KpnI) NCP termini. Data taken from Yakovchuk, P. *et al.*, 2006. See figure 33 for further detail.

Cohesive end pairing is the most conspicuous stabilizing feature of NCP147K; however, SHLs 0 – +3.5 also experiences a large decrease in thermal motion. Here, the L1/L2 loops of histones H3 and H4, as well as their A1 helices, contact nucleotide numbers 0 – 36 or 37 (Fig. 37a). As discussed, TA dinucleotide steps substitute CC and GG steps where these SHL minor groove blocks are located, and their flexibility relieves mechanical stress accumulated by minor groove kinking, potentially resulting in lower overall displacement factors by diminishing dynamic sampling of alternate conformational states.

Figure 37. Base pairing NCP DNA stabilizes histone octamer domains in core particle crystals. R and L designated which DNA half contacts the histone octamer. Warmer colors indicate greater atomic displacement. **(a)** The NCP601 octamer with discernable core histones labeled. The ‘mechanically challenged’ region contacting the DNA right half exhibits greater atomic displacement than the region contacting the left half or all of NCP147K’s octamer. **(b)** The NCP147K octamer displays lower overall B-factor values with the exception of the histone labeled portion of the octamer in contact with the DNA left half.

Stabilizing effects of NCP147K’s DNA are transferred to the octamer as well, particularly along binding and positioning surface areas in contact with the DNA right half (Fig. 37b). Again, these are locations of nucleotide substitution between 601 and 601*. The right half H3 N-terminal helix and tail, which bind the DNA terminus around SHLs +6 to +7, also exhibit a noticeable reduction of atomic displacement in what is likely an additive stabilizing mechanism between the DNA binding histone motif and base paired superhelix terminus. Temperature factors in the left half of the NCP147K octamer appear to increase marginally compared to NCP601’s octamer, suggesting paired cohesive end termini and right half nucleotide substitutions elicit redistribution of base level mechanical stress originating from DNA superhelix deformation during wrapping. Finally, there is greater thermal motion at the very extremities of NCP147K histone tails. The higher resolution NCP147K structure allowed for more extensive modelling tail peptides where there is greater disorder. And it is unclear to what extent specific sequence differences, almost entirely occurring in the tails, account for these increases.

Nucleosome interactions with metal cations are critical to chromatin condensation and NCP crystallization where they act as counterions to neutralize negative charge and mediate essential nucleosome–nucleosome contacts (Luger, K. *et al.*, 1997a; Hansen, J.C. 2002; Wu, B. *et al.*, 2010). More specifically, calcium and magnesium cations are responsible for any DNA charge neutralization not accounted for by octamer association (Wu, B. and Davey, C.A. 2010). Despite physiological relevance, buffer preparations for most NCP crystal applications contain manganese(II) as the prevailing divalent cation. This is because magnesium is difficult to dehydrate and binds NCPs with lower affinity due to its smaller ionic radius and higher charge density.

In the majority of NCP crystal structures, hexaaquamanganese(II) complexes preferentially recognize certain histone and minor groove structural motifs, while coordination of major groove GG and GC dinucleotides is governed by the local electrostatic environment (Wu, B. and Davey, C.A. 2010). NCP601's lone cation mediated crystal contact occurs at a histone interface where the H2B α 1 helix C-terminal carbonyl oxygen coordinates a metal cation. Acidic residues of H3'/H4' and H2A/H2B on facing nucleosomes provide a favorable electronegative environment to attract and stabilize the positive charge.

In addition to the W-C pairing of symmetry related core particles, NCP147K termini participate in metal mediated crystal contacts between offset facing nucleosomes (Fig. 38a). Electrostatics may also guide a cluster of manganese cations to a region between nucleosomes where SHL -6.5 and -7 and opposite facing SHL +2.5 and +3 interdigitate (Fig. 38a). A 5' phosphate group on the recessed side of NCP147K's cohesive end appears to coordinate a pair of manganese(II) cations that may be involved in bridging a DNA–DNA nucleosome crystal contact (Fig. 38b). Moderate to strong signal in both difference and anomalous difference maps define the cations' locations. For both manganese atoms, anomalous difference signal is slightly weaker ($Mn3 = 4.7 \sigma$, $Mn10 = 5.7 \sigma$) than signal observed in the Fo-Fc map ($Mn3 = 7.6 \sigma$, $Mn10 = 8.0 \sigma$). Therefore, metal atoms were placed at the difference peak center of mass before atomic refinement. Post refinement fractional occupancies calculated from electron density residuals for both Mn^{2+} sites are ~ 0.75 . Such high occupancy, along with intermediate atomic displacement factors, likely foster well-ordered diffraction data for implicit NCP termini modelling where manganese cation binding may also contribute to overall crystal lattice stability (Table 4).

Figure 38. A 5' phosphate at the KpnI site recess participates in ancillary cation mediated crystal contacts. The NUC1 octamer is colored teal and the NUC2 octamer is yellow. **(a)** A secondary face-to-face nucleosome crystal contact (green box) involves interdigitating DNA major/minor groove by neutralizing negative charge repulsion with positive cations. A trio of manganese atoms (purple spheres) accomplishes charge neutralization in the NCP147K structure by coordinating DNA and histone atoms. **(b)** Enlarged view of the green boxed inset shown in (a). The 5' phosphate of -71 C coordinates two manganese cations Mn3 and Mn10 that form bridging interactions between nucleosomes. Mn17 appears to coordinate D81 of the H3 L2 motif and exists only to neutralize negative phosphate charge.

The best coordination geometries for the two manganese cations were calculated with UCSF Chimera v1.10.2 (Table 6). Distances listed (Dist. Å) in table 6 are not Mn–O/N bond distances, merely distances between the metal and potential ligand atoms. Likewise, root mean square values (RMSD Å) represent the best-fit deviation between atoms in the structure and idealized atoms positions of the predicted coordination geometry (Table 6). Applying these metrics, Mn3 coordinates two phosphate oxygen atoms at positions -70A and -71C of NUC1 with near ideal tetrahedral geometry (RMSD = 0.037 Å), with square pyramidal and trigonal bipyramidal geometries as close possibilities (RMSDs = 0.090 Å) (Fig. 39b; Table 6). Consequently, a water molecule bridging Mn3 and the -70A N7 would require strict octahedral geometry (Fig. 39). However, the presence of four additional coordinating water molecules that achieve this geometry are not observed due to disordered data and insufficient electron density

Figure 39. Structural evidence for a novel cation mediated nucleosome–nucleosome crystal contact. NUC1 is colored teal and NUC2 is yellow. **(a)** Difference (red cages, 3.0 σ contour) and anomalous difference (green cages, 3.0 σ contour) maps delineate the presence of two Mn^{2+} cations interacting with the 5' cohesive end phosphate and involved in a nucleosome–nucleosome crystal contact. **(b)** Approximate coordination geometry of the cations presented in (a) as determined by UCSF Chimera. The 5' phosphate of -71 C coordinates two manganese cations with approximate trigonal prismatic, square face-bicapped (Mn3) and tetrahedral (Mn10) geometries. Direct coordinate and aqua–metal bonds between -70 A and Mn3 help position the mobile 5' phosphate of -71 C for bridging Mn10 to NUC2 27G. Potential aqua–metal bonds are colored cyan and possible coordinate bonds are shown in pink.

Moderate difference and anomalous difference signal, and high fractional occupancy, delineate Mn10's proximity to the 5' terminal phosphate of NUC1 (–71C). At this location, Mn10 binds the N7 atom of 27G (NUC2) at a distance of 2.24 Å and with approximate tetrahedral geometry (Fig. 39a; Table 6). The atomic distance calculated from the structure correlates well with experimental values for Mn–N bond distances (Avg. $\approx$ 2.24; Å Oki, A.R. *et al.*, 1997). However, the uncertain coordination geometry calculated by the UCSF Chimera software package means a bridging metal-aqua bond between Mn10 and the 26A phosphate group is possible with canonical octahedral coordination (Fig. 39b).

Strong Mn3 positioning places the phosphate oxygen atom of NUC1 –71C in line with NUC2 27G bound Mn10 where a predominantly electrostatic interaction bridges the DNA of two nucleosomes (Fig. 39b). Bond angles and a short distance between the two atoms results in high ligand set RMSD, and makes achieving a tetrahedral configuration

improbable (Table 6). Experimentally derived Mn–O bond lengths averaging ≈ 2.32 Å confirm this difficulty as well (Oki, A.R. *et al.*, 1997). Both anomalous signal and positive electron density clearly place the cation at the position observed in the structure, while the calculated ligand set RMSD value indicates an overcrowded geometry. Taken together, these parameters still place this interaction within acceptable electrostatic bond range. However, the given Mn–O/N bond lengths are taken from inorganic structures, and shorter Mn–O bonds of ≈ 1.68 Å have been observed in biological complexes like photosystem II (Gupta, R. *et al.*, 2014).

Cation	Ano Diff.	Diff.	Occu.	B-factor (Å ²)	Site	Geometry	Coord #	Dist. (Å)	RMSD (Å)
Mn3	4.7	7.6	0.75	72.9	-71C (OP2)	Tetrahedral	4	2.36	0.037
					-70A (OP2)			2.31	
					-70A (N7)	N/A	N/A	4.71	N/A
Mn10	5.7	8.0	0.75	74.5	-71C (OP1)	Tetrahedral	4	1.57	0.250
					27G (N7)			2.24	
					26A (OP2)	N/A	N/A	5.04	N/A

Table 6. Crystallographic and geometric data for manganese cations at the SHL -6.5–SHL +3 nucleosome–nucleosome interface.

The *in crystallo* distance between offset cohesive end nucleosomes is greatest in the NCP147K structure along the *c*-axis (Fig. 40). Both the SK and XK structures, and all blunt end crystals, are more compact through this dimension for reasons outlined in prior sections. The ≈ 13.4 Å of rise added by base paired termini increases the gap between particles around SHLs $\pm 0.5 - \pm 1.5$, which are near to DNA entry exit points (SHLs ± 7) (Fig. 40a). The additional helical rise converts to a maximum inter-particle distance of ≈ 12.1 Å at the nearest regions of the nucleosomes in the *c* direction without a proportional increase along the same unit cell axis (Fig. 40b; Table 3).

The dominant face-to-face crystal contacts of blunt end NCPs force DNA in close proximity where charge-charge repulsion can accrue if divalent metal occupancy is low. Such negative charge repulsion may partially account for the high atomic displacement factors observed for 601 DNA throughout SHLs $\pm 5 - \pm 7$ (Fig. 35a). A larger distance in the multimeric NCP147K structure ameliorates this situation to help yield a better ordered lattice. More importantly, the greater inter-particle distance is responsible for a defining topological feature of the NCP147K crystal structure.

Figure 40. Comparison of inter-particle distance between base paired and stacked NCPs. NCP601 DNA appears in green and NCP147K DNA is colored pink. **(a)** Linear distances between the closest inter-particle region of a nucleosome pair along the *c* unit cell axis in NCP601 and NCP147K. This inter-particle region is the major groove at SHL1/SHL -1'. The minimum distance between NCP601 particles is 4.2 Å. The minimum distance between NCP147K particles is 12.1 Å. **(b)** The gap between core particles is visibly greater in the NCP147K structure. Minimum and maximum values calculated in (a) are displayed to show scale.

NCP147K's base pairing termini ultimately create a near continuous layer of NCPs through the *a* lattice axis. NCPs organize across this layer in a manner resembling a tightly associated version of the 11 nm fiber structure, the primary level of chromatin compaction (Hansen, J.C. 2002; Sajan, S.A. and Hawkins, R.D. 2012) (Fig. 41). First observed by Oudet, P. *et al.*, 1975, and referred to as beads-on-a-string, this extended fiber structure arises when many histone octamers associate with a single DNA molecule under physiological salt conditions in the absence of histone H1 (Kornberg, R.D. 1974) (Fig. 41). The key structural difference between 11 nm fiber and the multimeric crystal structure is shorter connecting DNA and a fixed repeat length between nucleosomes (Figs. 41 & 42). In the cell, linker DNA length is variable in a range governed by nucleosome repeat length (NRL) and H1 occupancy (Luque, A. *et al.*, 2016). Conversely, the NCP147K 'linker' adds no additional DNA length beyond the NCP147 structure and exhibits NRLs at a constant interval.

Figure 41. Electron micrograph of the 11nm nucleosome fiber. Individual nucleosomes appears as bold black dots connected by thin strands of linker DNA. This structure represents the first level of chromatin organization, occurring under physiological ionic conditions and in the absence of histone H1. Figure adapted from (Palladino, M.A. and Lobo, D. 2012).

In future applications of the NCP147K crystal system, the expanded gap between particles may serve as a stable binding location for chromatin factors that have been difficult to crystallize (Fig. 42). NCP147K could also potentially serve as a model system for structural studies of DNA repair due to the double nicked DNA helix present between every core particle (Fig. 42). Finally, improved crystallization kinetics, overall lattice stability, and better ordering of flexible protein and DNA motifs could help achieve increasingly higher resolution NCP–adduct structures, where near atomic level information is invaluable for drug development.

Figure 42. The NCP147K oligomeric structure forms a near continuous layer of nucleosomes approximating a tightly packed, H1 depleted, beads-on-a-string chromatin fiber-like structure. (top) Nucleosomes shown in molecular surface representation with DNA colored dark gray and octamer in light gray. The blue circle highlights the larger inter-particle gap looking down the *b*-axis. The DNA effectively wraps around octamers as one nucleic acid molecule. (inset) Watson-Crick base pairing of NCP147K's 3' overhangs forms a double nicked B-form helix. Despite these breaks in the phosphate backbone, the gray boundary exhibits a continuous molecular surface. Bold numbers indicate average hydrogen bond length between each cohesive end base pair.

3.4 Human NCP145 preparation and crystallization

3.4.1 Human NCP145 reconstitution

Human NCP145 is reconstituted in accordance with protocol adapted from Luger, K. *et al.*, 1997a and Dyer, P.N. *et al.*, 2004. Large scale reconstitutions are first validated via small scale dialysis trials to establish DNA:HO ratios that will maximize NCP quantities with minimal free DNA contaminant. Small scale reconstitutions determined that DNA:HO ratios between 0.91 – 0.95:1 consistently provided the highest NCP yield with acceptably low amounts of free DNA (Fig. 43a). Gradually lowering potassium chloride concentrations from 2 M to 0 M during dialysis controls association of the negatively charged DNA and the positively charged octamer surface while causing little to no octamer dissociation at low salt concentrations.

Dialysis of 145 α -satellite DNA with pure human histone octamer results in two stable nucleosome conformations. The centered conformation is predominate. Here the pseudo-twofold axis runs through the central base pair and nucleosome dyad satisfying all histone–DNA interactions. The off-centered conformation occurs when interactions between a DNA terminus and H3's N-terminal tail are absent, causing the unbound terminus to behave as linker DNA (Luger, K. 2003). Heat shifting off-centered NCP facilitates repositioning of DNA around the octamer so that the thermodynamically favored centered conformation prevails (Fig. 43a). Off-centered conformations were not witnessed when reconstituting human octamer–H2A.Z and the strong positioning 601L DNA. Note that although the band in lane 4 appears to run higher, the free DNA is shifted by the same amount, indicating non-uniformities within the gel (Fig. 43b).

Figure 43. Native PAGE analyses of small- and large-scale reconstitutions of human NCPs. (a) (left) Small scale reconstitution of human NCP145 is carried out first with variable DNA:HO ratios to determine the best ratio when scaling up (middle and right). Un-shifted (US) samples containing off-centered NCP (red arrow) are treated at 37 – 50°C for 20 – 60min to obtain fully centered NCP (blue arrow) shifted (S) nonconcentrated (UC) sample. NCP samples are then concentrated (C) and analyzed for aggregate formation by PAGE (middle and right). (b) The exact same protocol is applied to NCP601L-H2A.Z except that samples do not require heat shifting.

3.4.2 Human NCP145 crystallization

Crystallization of human NCP145 by salting in progressed only after extreme difficulty obtaining the correct crystal morphology and diffraction quality size. Initial screening efforts on fully human NCP145 resulted in fast-nucleating needle-like crystals that are not suitable for crystallographic studies. Needle-like crystals appeared despite any adjustment of crystallization conditions or parameters (Fig. 44). Prior studies define optimal 0.5X condition concentrations for *X. laevis*/human α -satellite NCP146 at 35 – 37.5 mM manganese(II) chloride, 25 mM potassium chloride, , and 10 mM potassium cacodylate at NCP concentrations of 4 mg·mL⁻¹ (Luger, K. *et al.*, 1997a). These conditions, as well as broader ranges of salt concentrations, led to undesirable morphology regardless of drop size tested (Fig. 44). Needle-like crystals appeared

within 48 hours and continued to form for up to ~2.5 weeks. In certain cases, single larger crystals appeared after longer equilibration times (>3 weeks) suggesting that initial salt concentrations were too high to enable slower nucleation but initial NCP concentrations ($4 \text{ mg}\cdot\text{mL}^{-1}$) were sufficient for correct morphology. However, this observation is counter to empirical data specifying that NCP is more soluble in higher salt concentrations that cause slower nucleation where drop conditions equilibrate with well conditions giving rise to robust crystals (Luger, K. *et al.*, 1997a).

Figure 44. Initial crystal screens of NCP145 produce overwhelming quantities of needle-like morphology. Needles appeared in 100% of wells across condition ranges from 60 – 110 mM MnCl_2 , 60 – 90 mM KCl, and 20 – 40 mM $(\text{CH}_3)_2\text{AsO}_2\text{K}$ (pH 6.0), different drop sizes: 2 – 8 μL , heat shift temperatures during reconstitution: 37 – 50°C, NCP drop concentrations: 3 – 4 mg/mL, and various liquid handling parameters. Native PAGE analysis of corresponding NCP reconstitutions are displayed in figure 43.

To determine what effect divalent ion concentration has on nucleation rates, and to what degree divalent cations control crystal morphology, a larger range of salt concentrations were tested to define NCP solubility in standard crystallization buffer. NCP solubility limit was tested via a low divalent screen outlined in table 7. NCP145 was determined to be soluble in 1X manganese(II) chloride concentrations as low as 50 mM (Table 7).

		1	2	3	4	5	6
	1X Conditions	20mM MnCl ₂	30mM MnCl ₂	40mM MnCl ₂	45mM MnCl ₂	50mM MnCl ₂	55mM MnCl ₂
A	40mM KCl	Precipitate (immediate)	Precipitate (immediate)	Precipitate (immediate)	Precipitate (24hrs)	Microcrystals (72hrs)	Microcrystals (72hrs)
B	50mM KCl	Precipitate (immediate)	Precipitate (immediate)	Precipitate (immediate)	Precipitate (24hrs)	Microcrystals (72hrs)	Microcrystals (72hrs)
C	60mM KCl	Precipitate (immediate)	Precipitate (immediate)	Precipitate (immediate)	Precipitate (24hrs)	Microcrystals (72hrs)	Microcrystals (72hrs)
D	70mM KCl	Precipitate (immediate)	Precipitate (immediate)	Precipitate (immediate)	Precipitate (24hrs)	Rods (1.5wks)	Microcrystals (72hrs)

Buffer: 40mM (CH₃)₂AsO₂K; pH 6.0

Table 7. A low divalent ion (MnCl₂) screen determines the lower limit of NCP solubility. Small crystals of correct rod-shaped morphology witnessed in well D5 after ~1.5 weeks of equilibration against 25mM MnCl₂, 35mM KCl, 20mM (CH₃)₂AsO₂K; pH 6.0 at 18°C (figure 45).

Figure 45. Correct rod-like NCP morphology from the low divalent screen outlined in table 7.

The effect of monovalent ion concentration on crystal morphology, outlined in table 8, was also tested to isolate ideal parameters for correct morphology within the optimal divalent range. Here, increasing potassium chloride concentrations to certain range appeared to spur formation of proper morphology. It should be noted however, that observations were reversed in prior screens where higher potassium chloride concentrations caused thinner crystals to appear, although these screens started from a minimum of 60 mM potassium chloride.

		1	2	3	4	5	6
	1X Conditions	10mM KCl	15mM KCl	20mM KCl	25mM KCl	30mM KCl	35mM KCl
A	55mM MnCl ₂	Needles (60hrs)	Needles (60hrs)	Needles (60hrs)	Needles (60hrs)	Microcrystals (1wk)	Microcrystals (1wk)
B	65mM MnCl ₃	Needles (60hrs)	Needles (60hrs)	Needles (60hrs)	Needles (60hrs)	Rods (1.5wks)	Rods (1.5wks)
C	75mM MnCl ₄	Needles (60hrs)	Needles (60hrs)	Needles (60hrs)	Needles (60hrs)	Rods (1.5wks)	No crystals (3 weeks)
D	85mM MnCl ₅	Needles (60hrs)	Needles (60hrs)	Needles (60hrs)	Needles (60hrs)	No crystals (3 weeks)	No crystals (3 weeks)

Buffer: 40mM (CH₃)₂AsO₂K; pH 6.0

Table 8. A low monovalent ion (KCl) screen determines the limit for ideal crystal morphology. Small crystals of correct rod-shaped morphology witnessed in wells B5 & 6 and C5 after ~1.5 weeks of equilibration against 0.5X conditions (figure 46).

Figure 46. Correct rod-like NCP morphology from a low monovalent screen outlined in table 8.

Testing 1X conditions ranging from 20 – 150 mM MnCl₂, 10 – 100 mM KCl, and 20 – 40 mM (CH₃)₂AsO₂K (pH 6.0) could not conclusively identify the correct ionic strength for controlling crystal growth and optimizing morphology. Likewise, changing any other parameter during reconstitution or handling did not positively affect crystal morphology or mitigate the appearance of the fast nucleating needles. Tested crystallization parameters include: premixing NCP with crystallization buffer for equilibration at 4°C, *in situ* mixing at room temperature before sealing, the order of reagent addition (NCP to buffer vs buffer to NCP), heat shifting at 37 °C for twenty minutes vs 50 °C for one hour, different drop sizes (2 µL, 6 µL, 8 µL), and filtered vs unfiltered reagents.

Finally, magnesium(II) chloride precipitation experiment was conducted in an attempt to remove aggregates suspected of affecting NCP crystallization that appear during reconstitution. This was performed by diluting the NCP concentration from 8 mg·mL⁻¹ to 0.2 mg·mL⁻¹ in 3 mM magnesium(II) chloride and incubating the solution at room temperature before centrifugation at maximum speed. Resuspended pellet and

supernatant were analyzed by UV spectroscopy and no significant NCP was found in either fraction (data not shown). This was confirmed by PAGE analyses which found no detectable NCP (data not shown). The above criteria account for all aspects of our laboratory's established protocol for crystallizing NCP.

One noticeable discrepancy that emerged during reconstitution was the amount of free DNA found in each preparation, where it appeared that the presence of free DNA affected crystal morphology (Fig. 47a). As mentioned, reconstitutions shown in figure 43a (middle and right) gave predominantly needle-like crystals with fewer larger crystals. Direct observation by PAGE analyses show these reconstitutions contain little to no free DNA (Fig. 43a). Throughout multiple trials, reconstitutions lacking free DNA offered only needle-like crystals despite any parameter alterations discussed above (Figs. 44 & 47b). However, when reconstitutions contained visible free DNA, NCP crystals grew slower and larger. Drops from these 'contaminated' reconstitutions ultimately provided several crystals each, but required ≈ 3.5 weeks of incubation before any diffraction size crystals appeared (Fig 47a). Under these circumstances 1X salt concentrations were 90 mM manganese(II) chloride, 80 mM potassium chloride, and 40 mM K-cacodylate; pH 6.0.

When these conditions were replicated in a narrow test screen of a reconstitution without discernible free DNA, needles appeared in every well within 48 hours (Fig. 47b). Therefore, in order to isolate the effect of free DNA on crystal growth, a reconstitution lacking free DNA was split in half and pure NCP145 DNA was added to one aliquot. The amount of free DNA added to the "+Free DNA" sample was based on the NCP concentration measured as DNA, quantified by UV spectroscopy, and calculated from the difference in NCP yield and the amount of NCP145 DNA originally added to the reconstitution. This amount was divided by two as sample was split into equal volume aliquots for a final free DNA quantity of 2 μ g. Results from 10% native PAGE analyses and crystallization are depicted in figure 47c. These results suggest that the presence of free DNA in concentrated NCP samples plays some part in crystallization kinetics by controlling the rate of nucleation, which determines crystal size and morphology. If this observation is correct, the mechanism is not readily apparent, but may be due to free DNA mitigating non-specific contacts between core particles or serving as a sink for excess H2A-H2B dimer present in the reconstitutions. In NCP crystal structures, packing contacts between particles form in a manner similar to nucleosome

condensation in chromatin. One prominent interaction where the highly basic H4-tail makes an extensive electrostatic contact with the surface of H2A-H2B's acidic patch on an adjacent nucleosome is essential for stability. Any free DNA present during formation of these interactions may help shift equilibrium towards slow nucleating kinetics.

Figure 47. Free DNA during nucleosome reconstitution is necessary for diffraction quality NCP crystals. (a) Results from a narrow crystal screen (60 – 110 mM MnCl_2 , 60 – 90 mM KCl, 40 mM $(\text{CH}_3)_2\text{AsO}_2\text{K}$; pH 6.0) and 8 μL drop size setup presenting robust crystal formation. The 10% PAGE gel (top) is demarcated as in figure 43 and shows a moderate amount of free DNA (145 bp). (b) A narrow screen of an NCP reconstitution lacking free DNA around optimal salt conditions determined in (a) (60 – 110 mM MnCl_2 , 60 – 90 mM KCl, 40 mM $(\text{CH}_3)_2\text{AsO}_2\text{K}$; pH 6.0) using a 8 μL drop size produced only fast nucleating crystals with needle-like morphology. (c) Addition of $\sim 4\mu\text{g}$ of NCP145 DNA to a reconstitution lacking free DNA and set up at 8 μL drops sizes with the optimal determined salt concentrations (90 mM MnCl_2 , 80 mM KCl, 40 mM $(\text{CH}_3)_2\text{AsO}_2\text{K}$; pH 6.0) shows robust crystal

3.5 High-resolution structure of human NCP145

3.5.1 Comparison of human NCP145 and *X. laevis* NCP145

Human NCP145 (hNCP145) consists of a 145 base pair α -satellite sequence derived from human centromeres and an octamer composed of entirely human core histones. Crystallization was carried out as outlined in section 2.3 of the methods. Under optimal divalent conditions, hNCP145 crystals will grow to maximal size (≈ 1 mm) in approximately 10 to 14 days by hanging drop vapor diffusion (Fig. 47a). Crystal quality is highly reproducible with the highest resolution diffraction ranging between 2.20 – 1.99 Å using a synchrotron X-ray source. For comparison, hNCP145 crystals grown several years prior still provide high quality diffraction data (Fig. 48a; Table 9).

Table 9. Data collection and refinement statistics for human NCP145

NCP145	
Histones	<i>H. sapiens</i>
Data collection*	
Space group	P2 ₁ 2 ₁ 2 ₁
Cell dimensions	
<i>a</i> (Å)	107.56
<i>b</i> (Å)	109.71
<i>c</i> (Å)	183.48
Resolution (Å)	1.99–76.81 (1.99–2.10)
<i>R</i> _{merge} (%)	6.8 (212)
<i>R</i> _{pim} (%)	2.1 (65.2)
<i>I</i> / σ <i>I</i>	17.4 (1.3)
Completeness (%)	99.6 (97.9)
CC (1/2) (%)	99.7 (52.6)
Redundancy	12.2 (11.6)
Refinement	
Resolution (Å)	1.99–76.81
No. reflections	145312
<i>R</i> _{work} / <i>R</i> _{free} (%)	23.6 / 26.3
No. atoms	12169
Protein	6116
DNA	5939
Solvent	114
Adducts	N/A
<i>B</i> -factors (Å ²)	67.3
Protein	53.4
DNA	105.7
Solvent	42.8
Adduct	N/A
R.m.s. deviations	
Bond lengths (Å)	0.0079
Bond angles (°)	1.4744

Single crystal data set; values in parentheses are the highest resolution shell

To our knowledge, this is the first, and highest resolution, structure of an entirely human NCP with 145 base pair α -satellite DNA (Fig. 48b). As with NCP147K, the overall structure is extremely similar to other NCPs. However, detailed analysis exposes subtle changes in histone-DNA register relative to *X. laevis* NCP145 (xNCP145). In fact, despite a single base pair extension, the 145 DNA structure of hNCP145 is more similar to human NCP146 (PDB code 2CV5), suggesting minor species specific sequence differences can have a substantial impact on nucleosome DNA conformation and octamer positioning.

Figure 48. X-Ray crystal structure of hNCP145 at 1.99 Å resolution. (a) Representative crystals of hNCP145 at 1X – 2X magnification. Using the ocular scale at this power crystal size is measured to the nearest μm giving a long axis of approximately 1mm for the displayed crystals. The large size mitigates proliferation where all available NCP material in the drop is used in a slow single nucleation event controlled by the divalent concentration. Crystals shown in (a) were grown approximately two years apart (top: 4/11/17; bottom: 8/8/15) demonstrating high reproducibility for achieving the best diffracting size crystals. (b) Cartoon schematic of the hNCP145 crystal structure at 1.99 Å resolution. Core histones are shown in standard coloring: blue, green, yellow, and red for H3, H4, H2A, and H2B respectively. DNA phosphate backbones are colored orange and cyan to differentiate each strand with ribose sugars and nitrogenous bases depicted as gray rings.

Despite identical DNA length and sequence, structural alignment and superpositioning of hNCP145 and xNCP145 reveals a shift in DNA register in the particles' positive halves (the DNA half which contacts nonprime histones). An RMSD of 0.253 Å after octamer backbone alignment is negligible (Fig. 49a). For DNA sugar-phosphate backbone and nucleotide bases however, average RMSD is ≈ 0.99 Å with a 0.814 Å and 1.182 Å variation between negative halves (the DNA half contacting prime histones), both backbone and base pair step positions are essentially identical. In the positive half,

backbone deviation is significantly more noticeable between SHLs +3.5 and +5 with sizeable jumps in base pair step agreement (Fig. 49b).

Figure 49. Structural superposition of hNCP145 and xNCP145. (a) Structural alignment of hNCP145 (blue) and xNCP145 (magenta) showing the (left to right) face, top, and bottom views of the nucleosome particles. The CCP4 Superpose task was used for fitting in which core histone tails and both DNA strands were excluded. This treatment results in an RMSD of 0.253 Å. (b) Schematic representation of the structural superposition in (a) of hNCP145 (cyan) and xNCP145 (magenta) of the positive half DNA structures. Bold numbers represent SHLs that differ significantly between each molecule. Phosphate backbones are represented as ribbons and their positions vary between hNCP145 and xNCP145 from SHL +2 to SHL +5. Minor groove kinking in xNCP145 occurs at SHL +1.5 and causes stretching through SHL +5. (c) Schematic representation of the structural superposition in (a) of the hNCP145 (blue) and xNCP145 (light pink) of the negative half DNA structures. The phosphate backbones and base pair steps are closely aligned indicating no difference in DNA register for the negative half.

Maximum DNA RMSD between *X. laevis* and human core particles reaches ≈ 3.66 Å and occurs around SHL +1.5 – +2, signifying the beginning of xNCP145 stretching. Deviation values return closer to average between SHLs +3 and +4 before increasing again to ≈ 3.55 Å around SHL +4.5 – +5, signifying a return to agreement in DNA register between hNCP145 and xNCP145 (Fig. 50).

Figure 50. The average RMSD between hNCP145 and xNCP145 DNA. Quantification of the visual representations in figure 49 shows a sharp increase in average deviation starting at SHL +1.5. Average RMSD for the entire DNA molecule is ≈ 0.99 Å.

3.5.2 DNA conformational changes in hNCP145 lead to change in register.

Human NCP145 and *X. laevis* NCP145 sequences are approximately 97% similar possessing identical palindromic DNA fragments while differing only in the core histone sequences, where most amino acid substitutions occur in tail regions. The large shift in DNA register witnessed in the positive hNCP half persists from SHL +2 to SHL +5.

Full analysis of hNCP145 and xNCP145 rotational and translational base pair step parameters show the structures are highly similar with two noticeable discrepancies (Fig. 51). *X. laevis* NCP145 maintains proper histone-DNA register through super helix deformation. A minor groove kink at SHL +1.5 of xNCP145 leads to stretching around SHL +2 and gives the NCP an effective 147 base pair length. Structurally, the kink manifests as a large negative roll value between the fourteenth and fifteenth base pairs (GG=CC) (Fig. 51 top). This does not occur in the hNCP145 structure. As a result of xNCP145's kink at SHL +1.5, rise values between the two DNA structures become staggered starting around SHL +2 and lasting until SHL +5 (Fig. 51 bottom).

Roll and slide parameters between each structure are fairly consistent with the exception of moderate negative roll values found at three base pair steps, one in the hNCP145 negative half and two in the positive half (Fig. 51 top). At hNCP145 SHL -1.5 nucleotide -14G experiences approximately -27.6° degrees of roll whereas xNCP145 experiences only -1.3° . Similarly, 47G and 57G of hNCP145 experience -22.9° and $-$

23.4° of negative roll respectively. In comparison, xNCP145 experiences only -6.9° and -13.7° of negative roll at the equivalent locations. These two locations coincide with 'pressure points'; minor groove-inward regions with respect to the octamer surface where there is maximal mechanical strain and deformation (Chua, E.Y. *et al.*, 2012). Smaller negative roll values, associated with helical bending around the octamer, are distributed throughout hNCP145 DNA, particularly in regions of xNCP145 stretching, which may help alleviate mechanical strain without the energetic penalty of severe minor groove kinking.

The hNCP145 GG=CC base pair step analogous to the xNCP145 step that undergoes minor groove kinking has a negative roll value of -14.6°, while the xNCP145 step is deformed by approximately -53.9°. It is likely that some mechanical strain is transferred across these two base pair steps since they exist at a 'hinge' where bound single stranded DNA regions allow conformational variation between nucleic acid sequences (Chua, E.Y. *et al.*, 2012). The much smaller negative roll at this hinge is indicative of the conformational freedom DNA possesses in a nucleosome bound context. However, what is surprising in this study is the fact that both hNCP145 and xNCP145 include the exact same DNA molecule. This could be an indication of the energetic degeneracy within a set of alternate conformations and registers accessible to the nucleosomal double helix. Presumably this attribute has in any case evolved into the system to foster repositioning and dynamics characteristics necessary for genomic function.

Figure 51. Rotational and translational base pair step parameters for hNCP145 and xNCP DNA. (upper) Rotational base pair step values plotted after the first A-T base pair. The xNCP145 plots are displayed as thin lines and hNCP145 are thick. Bold numbers above the 145 α -satellite sequence designate super helix location. Minor groove blocks are shaded in gray. Large negative roll indicates minor groove kinking at SHL +1.5 in xNCP occurring at a CG=GC step. (lower) Translational base pair step plots were prepared exactly as above. Large positive rise at SHL +1.5 in xNCP indicates minor groove kinking. Such a magnitude in rise does not occur at this location in hNCP DNA. After this point, rise values between hNCP and xNCP are staggered from SHL +2 to +5.

Aside from structural variations that regulate disagreement between hNCP145 and xNCP145 DNA conformation, the lone chemical differences between *X. laevis* and human NCP145 exist in the histone octamer. Several sequence changes between amino

acids in core histone-tail domains may influence DNA conformation and positioning in ways that make hNCP145 unique (Appendix 4). In a human nucleosome structure comprising human core histones and 146 base pair α -satellite DNA, bulky amino acid substitutions in the H2B N-terminal tail clash with the DNA backbone at SHL -2, -5, and +5 to alter the nucleosomal DNA path such that stretching is observed at SHLs ± 5 (Tsunaka, Y. *et al.*, 2005). Human NCP146 is asymmetric with respect to its DNA halves. Excluding the dyad, the negative half includes 72 base pairs while the positive half retains 73. The extra nucleotide shifts the sequence ahead one base pair relative to hNCP145, but the numbering remains constant such that both particles possess the same histone-DNA register in the positive half (Fig. 52a). This suggests that human H2B-tail substitutions similarly effect the positive DNA half of both human core particles.

The relevant H2B-tail changes from *X. laevis* to human include triplet residues R27, K28, T29 to K27, R28, S29. These substitutions swap steric bulk at position 28 (Lys to Arg) and 29 (Thr to Ser). In human NCP146, an R28 omega nitrogen forms a hydrogen bond with the S29 gamma oxygen locking the arginine's aliphatic side chain in a position that clashes with the DNA phosphate backbone resulting in its shift towards the octamer surface (Fig. 52b). In the hNCP145 structure, this clash occurs only at SHL +5 to shift the hNCP145 DNA backbone, relative to xNCP145, by the same mechanism observed in human NCP146. However, the bulky R28 of hNCP145 does not interact with the smaller S29, therefore the side chain clash with DNA backbone is less prevalent (Fig. 52c). Stretching due to energetic degeneracy in hNCP145 and xNCP145 around SHL -2 moves the phosphate backbone away from H2B' R28 and alleviates potential steric clash resulting in matching negative-half registers.

Figure 52. The structural determinants of hNCP145 DNA conformation are similar to human NCP146. (a) Superposition of the positive half of xNCP145 (magenta), hNCP145 (cyan), and human NCP146 (translucent yellow; PDB code 2CV5) in block representation shows SHL +2 through SHL +5 (black numbers). Enumerated nucleotide pairs indicate where base pair steps agree (doublet), where they begin to disagree (singlet), and where they return to agreement. Text color matches block color. The helical paths are traced as thick lines and correspond with block color. (b) Structural alignment and superpositioning of hNCP145, xNCP145, and human NCP146 octamer cores. Amino acid changes between *X. laevis* and human histone sequences in the H2B-tail introduce a bulky arginine (R28) that clashes with the DNA backbone around SHL +5 in both hNCP145 and human NCP146. The DNA backbone shifts as a result where hNCP145 backbone is intermediate between the extent of human NCP146 and no shift in xNCP145 (black arrow). (c) R28 and S29 interactions with the DNA phosphate backbone in hNCP145. Potential hydrogen bonds and electrostatic interactions appear as blue dashes

3.6 Adduct structures of the human nucleosome core particle

After rectifying crystallization issues, NCP crystals were prepared for data collection by buffer exchange to substitute MnCl_2 with MgSO_4 and introduce MDP as a cryoprotectant. Crystals were soaked with the specified concentrations of relevant reagents for the indicated lengths of time, or until diffraction quality deteriorated. Most of the dinuclear compounds presented in figure 53 and Appendix 5 proved insoluble in

24% MDP stabilization buffer, except for the compounds discussed in the following sections. Compounds soluble under standard buffer treatment, or compounds solubilized with adjusted crystal treatment, bound the NCP sufficiently for structure solution and analysis. In cases where a small molecule adduct structure could not be built, or the interpretation of electron density was problematic, difference and anomalous difference maps were created and adequate signal intensity ($\geq 3.5 \sigma$) was used to deduce adduct binding modes and locations. Several insoluble compounds (Lucie-R33/R49/R50, EP10-069, and EP11-104) showed no appreciable signal in the difference or anomalous difference maps even after lengthy soaks, increased soaking buffer MPD concentrations, diffraction data collection, rigid body refinement, and preliminary map generation (Appendix 5). We recorded these negative observations as no adduct binding.

In metal-based drug development research, both rare (Ir, Os, Rh) and biologically significant (Cu, Fe) metals have gained favor as therapeutic possibilities. Despite these considerations, the primary metal candidates used as reactive groups remain platinum, as derivatives of cisplatin, ruthenium, particularly the RAPTA series, and more recently gold with interest in repurposing the anti-rheumatic drug auranofin. As successful mononuclear organometallic drugs, these compounds are excellent precursors for multinuclear derivatization. Creating homo- and hetero-multinuclear derivatives of well-established mononuclear drugs is an effective way to redesign their structure for extensible function. With additional metal centers, new binding modes become available through various types of cross-linking, increased polar and electrostatic interactions, intercalation, topological recognition, altered ligand exchange kinetics, and so on. The following sections will discuss these concepts in the framework of nucleosome targeting, emphasizing hetero-dinuclear designs and protein-protein cross-links on the octamer surface.

Figure 53. Chemical structures of multinuclear reagents and precursors used in chromatin targeting studies. (a) JN128-2_KP2425 contains a Pt(IV) center with a bidentate diamine ligand and two Cl⁻ leaving groups and an Ru(II) center in complex with two imidazole groups, a cymene ligand, and a single Cl⁻ leaving group. The two metal centers are bridged by two alkyl linkers culminating in an amide-ethyl-carboxyl chain connected to the platinum center for a distance of ~25 Å based on C-C bond lengths (~154pm). (b) EP11-019 is an AUF derivative possessing an Au(I) reactive center with the same sugar-thiol leaving group. Two phenyl groups and a benzoic acid replace the phosphine ligand of AUF. (c) Chemical structure designs for “Compound III” and Tri-nuclear “Compound III” capitalize on AUF+RAPTA-T synergy by joining RAPTA and AUF/EP11-019 derivatives with a PEG linker. (d) PEG linkers also connect acidic patch targeting viral peptide-metalloreagent conjugates GAG (top) and LANA M60 (bot) to AUF/EP11-019 derivatives. A methionine to ornithine mutation at position six increases LANA’s positive

3.6.1 JN128-2_KP2425 Pt(IV)–Ru(II) hetero-dinuclear NCP145 structure

Our first foray into hetero-dinuclear chromatin targeting reagents began with crystal soaking tests of platinum–ruthenium reagents (JN series; Appendix 5). These compounds were designed to build on the success of previous studies involving mononuclear cisplatin and the RAPTA family of anticancer drugs by introducing a DNA–protein cross-linking mode of action to chromatin targets (Rosenberg, B. *et al.*,

1969; Allardyce, C.S. *et al.*, 2001; Houten, B.V. *et al.*, 1993). JN128-2_KP2425 is a novel hetero-dinuclear compound, designed by our collaborator Dr. Alexey Nazarov, possessing ruthenium(II) and platinum(IV) reactive groups (Fig. 54) connected by alkyl linkers that adopt the imidazole carrier ligand features of established NAMI-A and KP1019 Ru(III) anti-cancer drugs (Hartinger, C.G *et al.*, 2008; Sava, G *et al.*, 2002).

Figure 54. The chemical structure of JN128-2_KP2425 is shown with approximate linear linker length represented in angstroms and based on experimental values for C-C (154pm), C-N (147pm), C=N (135pm), C-O (143pm), Ru-N (210pm) (Wishart, J.F. *et al.*, 1986) bond lengths. Chloride leaving groups are highlighted in red.

Previous efforts with hetero-dinuclear platinum–ruthenium compounds focused on a ruthenium(III)–platinum(II) complex (Herman A. *et al.*, 2008), a complex with reversible or irreversible oxidation states ($\text{Ru}^{\text{II}} \rightleftharpoons \text{Ru}^{\text{III}}$, $\text{Pt}^{\text{II}} \rightarrow \text{Pt}^{\text{IV}}$; Higgins S.L. *et al.*, 2011), or a ruthenium(II)–platinum(II) compound (Houten, B.V. *et al.*, 1993; Prussin, A.J. *et al.*, 2009). However, the structures of these molecules all include aromatic or polyaromatic bridging ligands. JN128-2_KP2425’s monofunctional ruthenium(II), with adorning cymene, targets core histone residues on the nucleosome surface at the acidic patch, and the bifunctional Pt(IV), with a bidentate ethylene diamine ligand, is predicted to coordinate the DNA component of NCPs (Wu, B. *et al.*, 2011; Pinato, O. *et al.*, 2013). The alkyl linkers afford the molecule a degree flexibility between reactive centers, a design that capitalizes on the two modes of targeting to enhance site specificity for DNA and influence the thermodynamics of adduct formation.

Table 10. Data collection statistics for JN128-2_KP2425 treated human NCP145 crystals.

JN128-2_KP2425	1.5 Å	1.072 Å	1.5 Å	1.072 Å
[Agent] (mM)	1	1	1	1
Duration (hrs)	24	24	47	47
Histones	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>
Data collection				
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions				
<i>a</i> (Å)	107.49	107.31	107.24	107.49
<i>b</i> (Å)	109.69	109.72	109.71	109.78
<i>c</i> (Å)	183.29	182.83	182.85	183.25
Resolution (Å)	2.47–48.85	2.45–48.85	2.51–48.84	2.30–48.89
	(2.47–2.60)	(2.45–2.59)	(2.51–2.64)	(2.30–2.43)
<i>R</i> _{merge} (%)	4.3 (68.6)	4.1 (74.0)	3.9 (72.2)	4.7 (66.3)
<i>R</i> _{pim} (%)	2.1 (30.4)	1.9 (32.6)	2.5 (57.8)	2.4 (42.0)
<i>I</i> / <i>σI</i>	23.5 (2.7)	26.7 (2.8)	24.6 (1.7)	18.2 (1.7)
Completeness (%)	98.4 (95.2)	99.6 (97.96)	99.1 (95.0)	97.6 (86.4)
CC (1/2) (%)	100 (81.6)	100 (82.6)	100 (61.7)	99.9 (77.8)
Redundancy	6.5 (6.3)	6.6 (6.5)	6.1 (3.8)	5.8 (3.5)

Column headings indicate data collection X-ray wavelength.

Single crystal data sets; values in parentheses are for highest resolution shell.

Crystals soaked with 1 mM JN128-2_KP2425 exhibited zero (24hr) or low (47hr) occupancy at the tested soaking times as determined by zero or low anomalous signal for both platinum(IV) and ruthenium(II) atoms at 1.5 Å and 1.072 Å (Pt edge) wavelengths. Thus, difference and anomalous signals appearing at the longer time point required us to use the 47hr data sets for analysis (Table 10). Anomalous signal for the ruthenium atom is not detectable above background in either the 1.5 Å or 1.072 Å data sets and so is recorded as zero. Therefore, the moderate positive electron density (max signal 5.51 σ at 1.5 Å) located around E61 and E64 of H2A, likely corresponding to the cymene and imidazole ligands, is taken as evidence of ruthenium coordination at one of the glutamine residues of the canonical ruthenium(II) binding site RU1 (Table 10; Fig 54). JN128-2_KP2425's ruthenium(II) center possesses a cymene ligand that imparts upon it a hydrophobic surface similar to RAPTA-C's from which this part of the reagent is derived. Presumably, as with RAPTA-C the cymene group aids in molecular recognition of the RU1 binding site by facilitating hydrophobic contacts with non-ruthenium(II) coordinating histone protein side chains. Wu, B. *et al.*, 2011 were the first to observe this mode of core histone targeting.

Table 11. Ruthenium(II) and platinum(IV) binding sites in human NCP145 treated with JN128-2_KP2425.

Data Set	Difference* or Anomalous Difference Peak height (σ)		
	RU1*	PT1	PT2
1.5 Å	5.5	3.9	4.1
1.072 Å	5.2	4.1	4.5

Sites correspond to coordination of the ruthenium carrier group (RU1) or platinum atom locations (PT1/2). Values correspond to difference ($F_o - F_c$; RU1 only) or anomalous difference electron density map peak magnitudes.

The identification of platinum atom locations was more difficult for several reasons discussed below. Whereas the reactive platinum center of cisplatin is in a Pt^{2+} oxidation state with square planar coordination geometry, JN128-2_KP2425's platinum center is in the more chemically inert Pt^{4+} state with accompanying octahedral geometry (Luzyanin, K.V. *et al.*, 2005). These chemical and structural features, combined with a linker of inadequate length for intra-particle DNA binding, likely alter the platinum group's preferred reactivity towards GG or GXG motifs and makes placement of the Pt(IV) atoms at empirical sites ambiguous. Nonetheless, a weak anomalous signal, although sufficiently above background, in the platinum edge data set appears near the phosphate backbones of adjacent core particles (Fig. 55; Table 11). These sites are designated PT1 (a CT=AG step) and PT2 (a GG=CC step) with PT2 exhibiting slightly stronger signal over PT1 (Fig. 55; Table 11).

The platinum(IV) center's distance from the *N*-heterocyclic bases of DNA along with its proximity to the negative phosphate groups imply the metal – DNA interaction is characteristically electrostatic where the Pt(IV) center is not biologically reduced to a platinum(II) prodrug (Shi, Y. *et al.*, 2012). Such binding modes are also observed in a trinuclear copper(II) complex that cleaves DNA upon electrostatic interaction, and noncovalently interacting trinuclear platinum compounds where cationic Pt(II) centers can interact with backbone phosphates to produce cytotoxic effects (Harris, L.A. *et al.*, 2005; Komeda, S. *et al.*, 2006; Chen, J.W. *et al.*, 2007).

Finally, the additional elongated density emerging from the difference peak's center of mass around RU1 and extending towards sites PT1 and PT2 designates the presence of the aliphatic linker (Fig. 55). Density extending towards PT2, appearing as two tendrils, may be indicative of the two linkers connecting each metal center while the density

extending towards PT1 is less distinguishable as linker and may instead be parts of the cymene group (Fig. 55). The positive electron density extending towards PT1 seems to breach a narrow pocket of aliphatic carbons formed by histone side chains to create potential hydrophobic contacts in the manner of RAPTA-C's cymene.

Figure 55. Preliminary anomalous and Fo – Fc maps of the JN128-2_KP2425 adducted human NCP145. The anomalous signal shown in green, contoured at 3.0 σ , specifies the position of two Pt(IV) atoms. Platinum atoms appear to be interacting with the DNA phosphate backbone of neighboring NCP molecules (NUC1 – PT1: C -20 & T -21 ; NUC2 – PT2: G -63 & G -64). The Fo – Fc map, shown in red and contoured at 3.0 σ , displays moderate positive electron density around E61 and E64 of core histone H2A, signifying the presence of the Ru(II) atom, cymene, and imidazole moieties, with evidence of the linkers extending towards sites PT1 and PT2. H2A E61 and E64 are acidic patch residues and a known Ru(II) binding site (Wu, B. *et al.*, 2011; Adhireksan, Z. *et al.*, 2014).

Metal atoms placed at the centers of the strongest difference or anomalous peaks determine the approximate distances between reactive groups (Fig. 56a & b). The distances measured between heavy atoms coincide approximately with the length of JN128-2_KP2425's linker as determined by its chemical structure; further supporting that, despite weak signal due to low occupancy, the observed peaks are in fact platinum and ruthenium (Figs. 54 & 56b).

Figure 56. Binding site and locations of JN128-2_KP2425 heavy atoms in the adducted human NCP145. This structure is based on X-ray data collected at 1.072 Å (Pt L-III edge). **(a)** NCP is colored as in figure 2 with Pt(IV) atoms appearing as gray spheres ($r_w = 1.75$ Å) and the blue-green Ru(II) center ($r_w = 2.05$ Å) of JN128-2_KP2425 displayed near H2A E64 of the acidic patch. The position of the strongest anomalous signals and positive electron density are taken as the position of the corresponding metal atoms. **(b)** Detailed structure of Pt(IV) and Ru(II) atoms with measured distances between metal centers that correlate well with JN128-2_KP2425's approximate linker length. The representation here is rotated 180° about the vertical axis from the representation in (a) such that PT1 and PT2 appear in the opposite orientation.

JN128-2_KP2425 is a promising addition to the JN series of platinum – ruthenium hetero-dinuclear chromatin targeting compounds (Appendix 5). However, subsequent optimization is needed to adjust linker length and platinum center reactivity and coordination geometry so that intra- and inter-particle NCP DNA binding is possible. Once achieved, JN128-2_KP2425 may be a viable drug candidate for *in vivo* assays and efficacy experiments against cancer cells.

3.6.2 EP11-019, an auranofin derivative and hetero-dinuclear precursor.

EP11-019 is a metalorganic compound, synthesized in the lab of Prof. P.J. Dyson by Ms. L. Batchelor, whose design is derived, both structurally and chemically, from the disease-modifying antirheumatic drug auranofin (AUF) (Batchelor, L.K. *et al.*, 2017). Like AUF, EP11-019 bears a gold(I) reactive center but departs from its parent structure by replacing the triethylphosphine group with a 4-(diphenylphosphanyl)benzoic acid ligand. The leaving group chemistry between both compounds is the same where each prodrug possesses a β -D-thioglucosetetraacetate ligand (Figs. 10 & 53). Based on earlier studies performed in our lab, AUF reactivity towards a pair of H3/H3' histidine residues near the nucleosome dyad was recognized as a function of allosteric synergism with preceding RAPTA-T binding at RU1 and RU2 of the acidic patch (Adhireksan, Z. *et al.*, 2017). These findings advance prior research establishing gold–organometallic and metalorganic reactivity towards imidazole nitrogen atoms, whether through covalent bonding to imidazole ligands (Bovio, B., *et al.*, 1984; Berners-Price, S.J. *et al.*, 1985) or by targeting histidine residues when forming protein adducts (Zou, J. *et al.*, 2000; Messori, L. *et al.*, 2013).

EP11-019 is a mononuclear precursor to a series of homobimetallic gold compounds, where monofunctional gold(I) moieties are connected by variable length PEG spacers (Appendix 5; EP10-069) (Batchelor, L.K. *et al.*, 2017). The dinuclear gold series was created with the intent to improve the potency and efficacy of gold(I) metalorganic drugs by building on the multinuclear concept of introducing new modes of action to surmount drug resistance in chemoresistant cancers (Hartinger, C.G. *et al.*, 2011). In a chromatin targeting context, we are interested in dinuclear gold designs aimed at cross-linking H113 of H3 and H3'. This targeting mode could reduce the required dose while achieving the same therapeutic response. However, structures of homo-dinuclear gold–NCP adducts were unobtainable due to pervasive insolubility regardless of PEG linker length.

Exploiting structural changes arising from the RAPTA-T–AUF allosteric mechanism also aids EP11-019 nucleosome targeting. RAPTA-T adducts at RU1 and RU2 cause systematic changes in nucleosome octamer conformation leading to greater hydrophobic contact potential around H3–H113 (Adhireksan, Z. *et al.*, 2017). The new phenyl–benzoic acid ligand design tests if site occupancy improves when further increasing hydrophobic contacts in the AU1/AU1' binding channels and adjusting this design to be more lipophilic could help improve current cytotoxic limitations (Mendoza–Ferri, M.G. *et al.*, 2008; Barreiro, E. *et al.*, 2014). Cytotoxicity assays applying a series of novel gold(I) reagents against an ovarian carcinoma cell line, cisplatin resistant cells, and nontumoral HEK-293 cells found that dinuclear versions of EP11-019 are significantly more cytotoxic than AUF by approximately one order of magnitude. Mononuclear EP11-019, however, is nearly as effective against A2780 cells as AUF and about 6 to 8 fold less effective in HEK-293 and cisplatin resistant cells respectively (Batchelor, L.K. *et al.*, 2017).

Initial diffraction tests of human NCP145 crystals soaked with 1 mM EP11-019 for ~24 hours in the absence of RAPTA-T exhibited no detectable adduct formation as per lack of visible impact on the crystals, consistent diffraction patterns during the course of treatment, and zero signal at H3/H3' H113/H113' in the anomalous map. Indeed, crystal diffraction quality as a function of resolution remains nearly constant even up to 70 hours of continued reagent exposure (Table 11). As stated, diffraction quality and the overall resolution limit are indicators of nucleosome adduct formation where

reagent solubility in soaking buffer generally dictates the extent of uptake by NCP crystals. Diffraction data sets were collected at three time points at $\lambda = 1.0402 \text{ \AA}$ after the introduction of 0.5 mM RAPTA-T (Table 12). Crystals soaked with the EP11-019 – RAPTA-T combination experienced physical degradation with a concomitant loss in resolution, largely from sudden RAPTA-T uptake where occupancy at RU1 approaches one (Fig. 57b; Table 13). EP11-019 adducts began forming at the anticipated histidine residues as determined by moderate peaks in the anomalous map, and RAPTA-T binding is determined by electron density corresponding to PTA and arene ligands present in the Fo–Fc map at RU1 but not RU2 (Fig. 57b; Table 13). The best combination of resolution and signal was present in the 46-hour data set including RAPTA-T, which was used for structure refinement and solution.

Table 12. Data collection and refinement statistics for EP11-019, EP11-019 + KCl*, and EP11-019 + RAPTA-T treated human NCP145 crystals.

	EP11-019 + 0.5 mM RAPTA-T			EP11-019			
[Agent] (mM)	1	1	1	1	1	1	1
Duration (hrs)	20	46	63	48	63	70	89*
Histones	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>
Data collection							
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions							
<i>a</i> (Å)	107.33	106.43	107.28	107.32	107.23	107.27	107.29
<i>b</i> (Å)	109.75	110.04	109.89	109.74	109.69	109.81	109.72
<i>c</i> (Å)	182.87	181.82	179.91	182.99	182.80	183.06	182.96
Resolution (Å)	2.50–48.86 (2.50–2.64)	2.75–48.88 (2.75–2.90)	3.45–48.90 (3.45–3.64)	2.40–48.86 (2.40–2.53)	2.40–48.83 (2.40–2.53)	2.50–48.88 (2.50–2.64)	2.25–48.85 (2.25–2.37)
<i>R</i> _{merge} (%)	4.3 (76.9)	4.4 (79.2)	4.4 (60.7)	4.6 (67.3)	4.2 (65.8)	5.1 (77.8)	4.2 (68.8)
<i>R</i> _{pim} (%)	2.0 (35.4)	2.0 (34.8)	2.1 (27.5)	2.1 (30.9)	1.9 (31.8)	2.4 (36.3)	2.0 (42.8)
<i>I</i> / σ <i>I</i>	24.0 (2.4)	25.8 (2.5)	21.2 (3.5)	22.3 (2.8)	26.1 (2.6)	23.9 (2.3)	22.1 (1.8)
Completeness (%)	99.6 (97.9)	99.6 (97.9)	99.6 (97.7)	99.8 (98.5)	99.8 (99.0)	99.3 (96.4)	99.8 (98.9)
CC (1/2) (%)	100 (81.6)	100 (79.2)	100 (92.3)	99.9 (98.5)	100 (85.2)	100 (68.6)	100 (98.9)
Redundancy	6.5 (6.5)	6.5 (6.3)	6.4 (6.3)	6.6 (6.1)	6.6 (6.1)	6.4 (5.9)	6.2 (4.3)
Refinement							
Resolution (Å)	2.75–48.88						
No. reflections	55820						
<i>R</i> _{work} / <i>R</i> _{free} (%)	21.4 / 26.5						
No. atoms	12128						
Protein	6116						
DNA	5939						
Solvent	9						
Adducts	64						
<i>B</i> -factors (Å ²)	116.4						
Protein	74						
DNA	135.6						
Solvent	57.8						
Adduct	198.2						
R.m.s. deviations							
Bond lengths (Å)	0.0092						
Bond angles (°)	1.4892						

Single crystal data sets; values in parentheses are for highest resolution shell. Inset: comparison of auranofin (top) and EP11-019 (bottom) chemical structures.

EP11-019–NCP structural data reveal that the allosteric mechanism by which AUF coordinates buried histidine residues is identical for EP11-019. However, with EP11-019 the minimal conditions for nucleosome binding suggest that preceding RAPTA-T adduct formation at RU1 alone is sufficient for gold coordination at AU1/AU1' (Fig. 57a & b), as was predicted from the AUF-RAPTA-T study (Adhireksan et al., 2017). Unfortunately, as with JN128-2_KP2425, occupancy remained low owing to persistent insolubility and peak anomalous signals do not approach those of AUF (Adhireksan, Z. *et al.*, 2017).

Table 13. Ruthenium(II) and gold(I) binding sites in human NCP145 treated with EP11-019 + RAPTA-T

Data Set	Difference* or Anomalous Difference Peak height (σ)			
	RU1*	RU2*	AU1	AU1'
20 hr	7.03	--	5.20	4.90
46 hr	4.39	--	7.50	5.80
63 hr	8.57	--	6.50	5.10

Sites correspond to coordination of RAPTA-T (RU1/2) or EP11-019 (AU1/1'). Values correspond to difference ($F_o - F_c$; RU1/2) or anomalous difference electron density map peak magnitudes.

Based on HSAB theory, leaving group chemistries for RAPTA-T and EP11-019 are decidedly different. Gold–thiol bonds require greater aquation time whereas gold – chloride bonds are more labile and will aquate swiftly (Pearson, R.G. 1963; Ayers, P.W. *et al.*, 2006). Due to this principle, RAPTA-T chloride ligands released into soaking buffer may exchange with EP11-019's thioglucosetetraacetate ligand subsequently creating the reactive species. If so, this would potentially lead to gold(I) activation and histone binding without allosteric conformational changes induced by RAPTA-T binding at RU1. To determine whether rapidly liberated RAPTA-T chlorides could cleave EP11-019 thioglucosetetraacetate groups, a control using crystal samples soaked only with EP11-019 for up to 70 hours were supplemented with 4 mM potassium chloride and incubated for an additional 19 hours. These results determined that even in excess quantities, with soaking times of up to 19 hours beyond maximum reagent exposure, free chloride anions alone do not promote EP11-019 binding to the nucleosome (Table 12).

Figure 57. Fo – Fc omit and anomalous maps of EP11-019 and RAPTA-T adducted human NCP145. Maps are colored as in figure 26. **(a)** Moderate anomalous signal, contoured to 3.0σ , observed around a pair of symmetry related histidine residues on core histones H3 and H3' indicate the presence of Au(I) heavy atoms for two EP11-019 adducts. The Fo-Fc omit map, contoured at 3.0σ , also supports the location of the Au(I) centers and indicates some additional density for EP11-019's phenyl or benzoate ligands. **(b)** The Fo – Fc omit map, contoured at 3.0σ , shows moderate positive electron density around E61 and E64 of H2A which correlates with RAPTA-T binding at the RU1 site.

The site specificity of EP11-019 for H113 is essentially indistinguishable, and possibly enhanced, relative to AUF. This occurs despite a significant change to its chemical structure as EP11-019 and AUF binding locations on the nucleosome surface are identical (Fig. 58a). The principal difference between AUF and EP11-019 is the replacement of ethyl groups with bulkier phenyl and benzoate ligands. The increase in size prompted concerns over EP11-019's ability to coordinate at the AU1/AU1' binding channel due to anticipated steric hindrance between the phenyl–benzoate ligand and histone residues around H113 (Figs. 58b & 53b). However, this was not an obstacle as evidenced by the clear presence of heavy atoms at AU1/AU1' in the 1.040Å data set anomalous map (Fig. 57a).

In fact, EP11-019 appears to benefit from the 4-(diphenylphosphanyl)benzoic acid's increased hydrophobic surface area over AUF's triethylphosphine as the bulkier ligand makes more extensive hydrophobic contacts with aliphatic side chain carbons in the AU1/AU1' binding channel (Fig. 58c). Furthermore, RU1, RU2, AU1, and AU1' binding site conformations are conserved, with only slight deviations present in core histone backbones likely attributed to *H. sapiens* – *X. laevis* sequence differences and RAPTA-T binding only at RU1 (Fig. 58d). Indeed, hydrogen bonding between H113 residues and carboxylate groups at AU1/AU1' in the EP11-019–RAPTA-T is no different when compared to the AUF–RAPTA-T structure (Fig. 58b). This confirms low occupancy is primarily due to solubility issues and not steric clash between the

phenyl-benzoic acid ligand and histone residues or a drastically altered octamer conformation.

Figure 58. X-ray crystal structure of EP11-019 and RAPTA-T adducted NCP at 2.75 Å resolution. (a-d) Structures of EP11-019 and RAPTA-T adduct sites. H2A, H2B, H3, and H4 histone protein backbones are colored yellow, red, blue, and green respectively. (a) Overview of EP11-019 and RAPTA-T binding to NCP, depicting the >30 Å distance between each adduct site. (b) Structure of histone adducts formed between H3/H3' H113 and EP11-019. Dashes represent bifurcated hydrogen bonding between imidazole epsilon nitrogen atoms, carboxylate delta oxygen groups, and main chain cysteine oxygen atoms. (c) The van der Waals environment of an EP11-019 adduct shown in space-filling representation. The primary hydrophobic interactions exist between EP11-019 phenyl ligands and aliphatic side chain carbons of the binding channel. (d) Structural alignment of the EP11-019-RAPTA-T adducted nucleosome and the AUF-RAPTA-T adducted nucleosome (PDB code 5DNN) with and RMSD of 0.244 Å. Subtle structural changes arising from slight rearrangements in core histone backbone conformation due to species related sequence differences and exclusive RU1 occupancy do not alter the conformation of the RU1 site or the AU1/AU1' binding channel.

Consistent with prior studies of AUF–protein adducts, EP11-019's gold(I) atom covalently coordinates imidazole nitrogen atoms with approximate linear geometry (Figs. 57a, 58b, & 59; top) (Zou, J. *et al.*, 2000; Messori, L. *et al.*, 2013; Adhireksan Z. *et al.*, 2017). The bond lengths and angles of adducted EP11-019 are also congruent with un-adducted AUF and compound 9c, where EP11-019 and compound 9c share the 4-(diphenylphosphanyl)benzoic acid ligand but possess different leaving groups (Batchelor, L.K. *et al.*, 2017). Moreover, a sample of existing structural parameters for gold(I)–protein adducts shows there is good agreement between experimentally derived bond geometries and those of EP11-019–NCP (Fig. 59; bot).

Bond (Å, °)	AUF	Bond (Å, °)	9c
Au–P	2.259	Au–P	2.233
Au–S	2.293	Au–Cl	2.296
P–Au–S	173.60	P–Au–Cl	178.07

Bond (Å, °)	AUF–cyclophilin	Cl–Au(I)–lysozyme*	AUF–NCP**	EP11-019–NCP**
Au–P	2.270	N/A	1.890	2.285
Au–N	2.060	2.200	2.120	2.075
P–Au–N*	167.70	174.00	176.10	174.56

Figure 59. Distances (Å) and angles (°) around gold(I) centers for AUF (Hill, D. T. and Sutton, B. M. 1980) and 9c (Batchelor, L.K. *et al.*, 2017) (top) and metallodrug–protein adducts: AUF–cyclophilin (Zou, J. *et al.*, 2000), Cl–Au(I)–lysozyme (Messori, L. *et al.*, 2013), AUF–NCP (Adhireksan, Z. *et al.*, 2017), and EP11-019–NCP (bot). *Angle stated is for Cl–Au–N. Average values of three Au(I)–His15 lysozyme adducts. Two gold(III) prodrugs, Auoxo6 and Au₂phen, and gold(I) prodrug AuSac₂. **Average values for adducts formed at histidine residues of AU1 and AU1’.

On its own EP11-019 is not effective against ovarian cancer cells in cytotoxicity assays due to insufficient lipophilicity (Batchelor, L.K. *et al.*, 2017), and difficulties with solubility make high resolution NCP adduct structures difficult to attain. Nevertheless, with data presented here the reagent offers a useful base structure for a more sophisticated compound design discussed in the next section.

3.6.3 A hetero-dinuclear Ru(II)–Au(I) agent targets distinct nucleosome sites

Building on concepts introduced in the prior two sections: multi-site targeting, allosteric cross-talk, drug-drug synergy, hetero/multinuclearity (Manzotti, C. *et al.*, 2000; Mendoza–Ferri, M.G. *et al.*, 2009; Hartinger, C.G. *et al.*, 2011), HSAB theory, and druggable site recognition, we tested a hetero-bimetallic compound, synthesized in the lab of Prof. P.J. Dyson by Ms. L. Batchelor, designed to cross-link nucleosome tetramer with dimer by targeting and coordinating discrete histone sites.

Human NCP145 crystals were treated with a novel gold(I)–ruthenium(II) hetero-dinuclear agent to determine if the reactive centers would simultaneously coordinate distinct histone sites. Adduct formation was characterized by X-ray crystallography in which data collected at the gold edge (L-III, 11.9187 keV, 1.0402 Å) was used for structure solution and analysis. “Compound III” (C3) consists of a RAPTA-like moiety with a ruthenium(II) reactive center joined by a ~38 Å PEG linker to a gold(I) containing AUF/EP11-019 derivative (Fig. 60). The reactive center of the AUF derivative is more similar to EP11-019 with a 4-(diphenylphosphanyl)benzoic acid

ligand where the thioglucosetetraacetate leaving group is replaced with a chloride (Figs. 53b & 60).

Based on preceding data we anticipated the ruthenium(II) center of C3's RAPTA-like moiety would bifunctionally coordinate to the nucleosome acidic patch (Wu, B. *et al.*, 2011; Meier, S.M. *et al.*, 2013; Adhireksan, Z. *et al.*, 2014; Davey, C.A., 2015; Palermo, G. *et al.*, 2016). As discussed, RAPTA compounds are known to form adducts at the carboxylate groups of E61 and E64 of H2A (RU1) and the imidazole and carboxylate groups of H106 and E106 of H2B (RU2) through substitution of its chloride leaving groups (Wu, B. *et al.*, 2011; Adhireksan, Z. *et al.*, 2014; Adhireksan, Z. *et al.*, 2017). In the present study the RAPTA-like component of C3 binds at RU1 as indicated by strong signal in the $F_o - F_c$ difference map around H2A E61 and E64 (Fig. 63a; Table 15). RU2 however remains empty (no signal) at all tested soaking durations, where it lacks any globular density characteristic of cymene or PTA ligands (Fig. 63a; Table 15). This is likely due to steric interference from the flexible PEG linker at RU1 blocking access to RU2 and entropic effects governing binding order.

The prior study cited in section 3.6.2 discovered an allosteric cross-talk mechanism in which RAPTA-T binding to RU1 and RU2 causes systematic changes in the inter-helical orientations of the H2A and H3/H3' α -helices. Repositioning of these helices generates subtle rearrangements in the H3–H4 tetramer near the nucleosome dyad. These structural changes result in AU1/AU1' binding channel compaction which in turn promotes adduct formation at a pair of symmetry related histidine residues on H3/H3'. Here the gold(I) center of antirheumatic agent AUF monofunctionally coordinates the H113/H113' imidazole delta nitrogen atoms of the tetramer (Adhireksan, Z. *et al.*, 2017). Functionally, this allosteric cross-talk increases RAPTA-T and AUF efficacy and leads to cytotoxic drug-drug synergism whereby viability of an ovarian cancer cell line significantly decreases when cells are concurrently treated with both drugs.

Figure 60. The ideal chemical structure of C3 includes a PEG linker approximately 38.4 Å in length. Distance measurement performed in PyMOL (DeLano, W.L., 2002).

Table 14. Data collection and refinement statistics for Compound III treated human NCP145 crystals.

	C3			
[Agent] (mM)	1	1	1	1
Duration (hrs)	28	54	97	120
Histones	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>
Data collection				
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions				
<i>a</i> (Å)	107.46	107.21	107.17	107.31
<i>b</i> (Å)	109.80	109.66	109.67	109.74
<i>c</i> (Å)	183.12	182.88	183.11	183.25
Resolution (Å)	2.15–48.89	2.14–48.82	2.35–48.82	2.25–48.85
	(2.15–2.27)	(2.14–2.26)	(2.35–2.48)	(2.25–2.37)
<i>R</i> _{merge} (%)	3.7 (140)	5.9 (207)	5.3 (204)	5.1 (234)
<i>R</i> _{pim} (%)	1.1 (59.8)	1.8 (69.5)	1.7 (73.5)	1.6 (76.3)
<i>I</i> / <i>σI</i>	33.3 (1.3)	20.4 (1.1)	27.1 (1.1)	25.8 (1.1)
Completeness (%)	94.4 (68.6)	97.7 (85.7)	99.4 (96.1)	99.7 (97.9)
CC (1/2) (%)	100 (56.4)	99.9 (51.8)	100 (56.5)	100 (52.6)
Redundancy	11.9 (6.9)	12.9 (10.1)	12.9 (11.2)	13.0 (11.0)
Refinement				
Resolution (Å)				2.25–48.85
No. reflections				100758
<i>R</i> _{work} / <i>R</i> _{free} (%)				23.2 / 25.7
No. atoms				12152
Protein				6098
DNA				5939
Solvent				68
Adducts				43
<i>B</i> -factors (Å ²)				110.1
Protein				70.1
DNA				130.3
Solvent				56.5
Adduct				183.6
R.m.s. deviations				
Bond lengths (Å)				0.0086
Bond angles (°)				1.392

Single crystal data sets; values in parentheses are for highest resolution shell.

The preceding section establishes EP11-019–NCP adduct formation as nearly identical to NCP adducts formed by its parent structure AUF. This is based on difference and anomalous difference peaks observed at AU1/AU1' indicating site recognition and reactivity are conserved despite structural changes (Fig. 63; Table 14). However, since the metal centers of C3 are tethered as a single molecule and because RAPTA-T adducts must form prior to AUF and EP11-019 binding, we were uncertain if gold accumulation would occur at AU1/AU1' during C3 treatment.

Evidence for both RAPTA and gold group binding to RU1 and AU1/AU1' appears at the earliest time point (Fig. 63a; Table 14). Metal occupancies at these sites increase

as reagent exposure continued, culminating in a 2.25 Å data set at 120 hours of total treatment. At this time point strong difference and anomalous difference peaks permitted modeling and refinement of C3's AUF/EP11-019-like and RAPTA moieties (Fig. 63; Table 15). As with EP11-019, the coordination of C3's gold atom to the imidazole delta nitrogen atom of H3 H113 is similar to AUF, and consistent with other gold–protein adducts with a coordination number of 2 and nearly linear geometry (Au–P = 2.26 Å; Au–N = 2.11 Å; P–Au–N = 171.7°; Fig. 59).

Anomalous difference peak height is approximately 60% higher at AU1 relative to AU1' at the longest time point, where AU1 is on the same face of the octamer as RU1/RU2 (Fig. 63b). Also, AU1' shows a decrease in anomalous peak height at the longest soak time while AU1 peak height steadily increases throughout treatment (Table 15). Symmetry related ruthenium(II) site RU1' is obscured by NCP crystal packing contacts between acidic patches and H4-tails of adjacent core particles. This NCP–NCP packing interaction restricts appreciable ruthenium(II) coordination and in doing so likely limits gold(I) binding to AU1'. However, in the EP11-019–NCP structure AU1/AU1' difference and anomalous difference peak heights are roughly equivalent suggesting most adducts form when a single C3 molecule coordinates the RU1 and AU1 sites (Fig. 63; Table 15).

Table 15. Ruthenium(II) and gold(I) binding sites in human NCP145 treated with C3

Data Set	Difference or Anomalous Difference Peak height (σ)									
	RU1*	AU1	AU1'	H3 H39	H3' H39	H2B' H46	-2G(J)	-2G(I)	29G(I)	51G(I)
28 hr	5.1	7.3	6.5	--	6.4	--	--	--	--	--
54 hr	6.1	14.8	10.9	5.8	10.4	5.6	--	--	--	--
97 hr	7.1	22.7	16.7	17.2	25.5	8.6	4.4	4.2	--	--
120 hr	7.7	23.9	14.7	24.0	37.9	11.0	5.7	6.6	5.6	6.6

Sites correspond to coordination of RAPTA (RU1) or the gold group (all other locations). Values correspond to difference ($F_o - F_c$; RU1/2) or anomalous difference electron density map peak magnitudes. "G" sites refer to guanine nucleotide sites near the nucleosome center (-21/J) and 29 or 51 base pairs distant (29I/51I).

As noted, RAPTA-T binding at RU1 induces global nucleosome conformational changes that augment hydrophobic surface area around the AU1/AU1' binding channel. Extensive hydrophobic interactions between AUF's triethylphosphine ligand, or EP11-019's diphenylphosphanyl-benzoic acid ligand, and residues surrounding H3 H113/H113' provide ligand recognition that allows metal cation access to the partially buried histidine ring. By this mechanism, AUF and EP11-019 are highly specific for their target site (Fig. 63c; Adhireksan, Z. *et al.*, 2017). Likewise, the aromatic rings of C3's AUF/EP11-019-like moiety increase the group's hydrophobic surface area to help

maximize van der Waals contacts between the reagent and aliphatic side chain carbons in the AU1/AU1' binding channel. This is precisely the targeting and interaction witnessed in the EP11-019–NCP structure (Fig. 63c).

Finally, residual electron density corresponding to PEG linker extends from the RU1 adduct site towards AU1, further suggesting that a single C3 molecule spans the octamer surface to crosslink dimer with tetramer (Fig. 64). Unfortunately, linker flexibility hinders detection of continuous and well-defined electron density, which rules out clear positions for linker atoms and prohibits explicit modelling. Our collaborators, Prof. U. Röthlisberger and Mr. T. von Erlach, conducted molecular dynamics simulations on the system, which showed that the PEG linker is completely disordered and tends to even avoid interacting with the octamer surface. This is unsurprising considering the large intramolecular distance the linker must traverse in order to cross-link sites (~ 36.7 Å; H3 H113 ND1 – H2A E64 OE2). Even in structures of nucleosomes adducted with short linker bimetallic ruthenium compounds, where metal adducts at RU1 and RU2 cross-link dimeric H2A and H2B, long flexible linkers are not visible (Davey, G.E. *et al.*, 2017).

Figure 61. Histone-histone cross-linking analysis of NCP treated with C3. SDS-PAGE analysis shows cross-linked histone protein species arising from compound treatment. Samples include molecular weight protein marker (m), untreated control (NCP) and NCP treated with six different concentrations of C3. Note that the H3 and H2B

Biochemical data asserts the existence of H3-H2A crosslinks, but does not distinguish which H3 site is cross-linked with H2A. However, this is sufficient evidence that the linker does not breakdown upon adduct formation. Incubating native NCP145 with a range of C3 concentrations shows the disappearance of H2A and H3 bands and the

presence of dimeric and multimeric species upon SDS-PAGE analysis (Fig. 61). Bands demarcated as dimeric migrate at H3 (15.3 kDa) + H2A's (14.0 kDa) approximate molecular weight (≈ 29.3 kDa). Human H3 and H2B comigrate on a 15% acrylamide gel, however, H3 exists at the upper portion of the merged band and its intensity is reduced with increasing treatment of C3 (Fig. 61).

Histone octamer dissociates from DNA by initial and successive loss of the H2A-H2B dimers, ultimately followed by the release of the H3-H4 tetramer (Burton, D.R. *et al.*, 1978). This biochemical property of NCP allowed us to test the impact of C3 on nucleosome stability (Fig. 62). Tyrosine fluorescence spectroscopic analysis shows that NCP treated with C3 displays significantly higher salt stability relative to native NCP (Fig. 62). The salt-induced dissociation midpoint for native NCP is 0.95 M NaCl, versus 1.05 M and 1.06 M for NCP treated with 2:1 and 5:1 molar stoichiometry of C3, respectively. Thus tethering the H2A-H2B dimers to the H3-H4 tetramer, C3 cross-linking of the nucleosome inhibits dissociation of the histone octamer from the DNA.

Figure 62. Crosslinking by C3 increases nucleosome stability. Salt-induced DNA-histone dissociation profiles were measured by tyrosine fluorescence spectroscopy. Values given in molarity correspond to dissociation midpoints. Samples include untreated NCP (native) and NCP treated with 2:1 and 5:1 molar stoichiometry of C3. R-squared parameters correspond to adjusted values.

Despite lack of clear structural data, studies involving homo-dinuclear ruthenium and gold agents show that compounds with flexible bridging linkers possess greater cytotoxicity than homo-bunuclear ruthenium and gold compounds with rigid linkers (Murray, B.S. *et al.*, 2014; Pettinari, R. *et al.*, 2015; Batchelor, L.K. *et al.*, 2017). Furthermore, lipophilicity related to linker length is also a cytotoxicity factor (Mulyana, Y. *et al.*, 2011). Alkyl-linked dinuclear gold compounds displayed higher cytotoxicity with increasing lipophilicity in B16 melanoma cells, and alkyl-linked dinuclear polypyridyl ruthenium complexes and PEG-linked dinuclear ruthenium RAPTA derivatives showed cytotoxicity increases are correlated to longer linker length (Mirabelli, C.K. *et al.*, 1987; Pisani, M.J. *et al.*, 2010; Batchelor, L.K. *et al.*, 2017).

These conditions are favorable in a therapeutic context where C3's long flexible lipophilic linker could be an asset in cytotoxicity assays. C3 cellular cytotoxicity assays were carried out in the lab of our collaborator, P.J. Dyson. C3 cytotoxicity was assessed against cisplatin sensitive and cisplatin resistant human ovarian carcinoma (A2780 and A2780cisR) and non-tumoral human embryonic kidney (HEK-293) cell lines using the 3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) assay. The cytotoxicity of C3 is in the low micromolar range with an IC_{50} of $6.9 \pm 1.3 \mu M$ recorded against the A2780 cells. A two-fold selectivity was observed towards the A2780 cells with an IC_{50} of $13.6 \pm 1.5 \mu M$ observed against non-tumoral HEK-293 cells.

Figure 63. C3 reactive group binding locations on the nucleosome core. (a) Detail of adduct coordination for each reactive group. (Left) Gold(I) attacks H113 of AU1. The green cage represents signal in the anomalous map contoured at 3.0σ . (Right) Positive electron density (red cages) contoured at 3.0σ in the $F_o - F_c$ map designates ruthenium(II) group coordination at E61 and E64 of RU1. Density distal to the ruthenium(II) metal, and trailing off towards AU1, suggests the presence of PEG linker joining both reactive moieties. Black chevrons indicate the atoms from which the PEG linker emerges. (b) Relative locations of each metal center on the nucleosome. The distance between the gold atom at H113 and the ruthenium atom at E61–E64 is $\approx 35 \text{ \AA}$, approximately the length of the PEG linker based on C3's chemical structure.

Figure 64. Residual electron density as evidence for cross-linking at C3 metal adducts sites. Residual electron density around AU1 and RU1 binding sites is depicted as cyan cages contoured at 3.0σ . Black arrows indicate the direction of the dyad and chevrons illustrate PEG linker locations. **(a)** Post refinement areas of positive electron density at the AU1 binding channel suggest the presence of ordered PEG linker sections that are constrained by binding at AU1 or extending beyond H113 to contact guanine nucleotides near the dyad. **(b)** PEG linker originating at RU1 travels in multiple directions towards any one of the gold binding locations designated in Table 14, most noticeably H46 of H2B. Density around the RAPTA carbon, which is the connection point for linker, is positioned towards AU1 and sterically blocks RU2. There is no evidence of ruthenium coordination at RU2 as no difference peak develops after refinement once again asserting that C3 adducts are formed by a single molecule.

In addition to AU1/AU1' sites, gold adducts formed at several previously unwitnessed histidine residues and guanine nucleotides. Gold cations have a predilection for solvent exposed imidazole nitrogen atoms, such as those found in core histone tails. (Bovio, B., *et al.*, 1984; Berners-Price, S.J. *et al.*, 1985; Zou, J. *et al.*, 2000; Messori, L. *et al.*, 2013) (Table 15). Replacing AUF/EP11-019's thiolate leaving group with a chloride allows C3 to form gold adducts more readily. This is possible since bonds between hard chloride anions and soft gold cations are more ionic and labile in nature than soft-soft (Au–S) acids and bases which have a propensity for covalent bonding and therefore greater stability. Such a chemical change allows for faster metal aquation and subsequently faster coordination to the intermediate imidazole nitrogen (Pearson, R.G. 1963; Ayers, P.W. *et al.*, 2006).

X-ray data sets of C3 adducted NCP, collected at $1.042 \text{ \AA}$, show anomalous difference peaks at several previously unobserved histidine residues (Fig. 65a). The most notable of which are the exposed H39 residues of H3/H3' where a maximum peak height at H39 H3' surpasses peak heights at C3–AU1 and AUF–AU1' (Table 15; Adhireksan, Z. *et al.*, 2017). H39 is located in the N-terminal H3 tail that is involved in DNA binding. In a core particle context, these locations are at superhelix termini where NCP DNA experiences its greatest thermal motion (Blossey, R. and Schiessel, H. 2011). Such

action could further expose the histidine residue leading to greater metal occupancy over time. Another solvent exposed histidine site adjacent to RU1 also shows evidence of gold binding. After initial treatment H46 of H2B' develops adducts to a moderate degree, likely the result of its proximity to ruthenium binding sites where the nearest gold cation will be from a C3 molecule bound at RU1 (Fig. 65a; Table 15). By comparison, X-ray data sets of EP11-019 adducted NCP exhibit gold anomalous peaks only at AU1/AU1' that appear only in the presence of RAPTA-T (Table 15). This observation is consistent with NCP adducts formed by AUF, albeit with weaker signal. Moreover, at the longest soaking time gold–DNA adducts begin to appear at several guanine sites (Fig. 65a; Table 15). Two adducts at guanine nucleotides near the nucleosome dyad suggest DNA-octamer crosslinks may also form, although to a far lesser degree than histone–histone adducts (Fig. 65a). Again, gold cations have the propensity to coordinate imidazole nitrogen atoms but here adducts form at N7 of guanine nucleotide rings (Table 15).

New adduct sites, both protein and DNA, may be the result of better targeting from ligand structural changes in conjunction with more labile leaving group chemistry, or entropic advantages or penalties arising from PEG linker flexibility and order of adduct formation (Chung, S. *et al.*, 2009). Where the RAPTA group coordinates histone first, the degrees of freedom the gold(I) group can sample will be restricted due to rotational and translational entropy losses (Chung, S. *et al.*, 2009). Gold sites H46 H2B, 29G(I), and 29G(I) may be the result of this phenomenon (Fig. 65a; Table 15). If occupancy at RU1 approaches 1 and RAPTA adducts do not form at other nucleosome sites, free C3 with aquated reactive gold cations may bind more promiscuously to NCP. Such capacity is apparent at longer soak times when new gold binding sites H39 H3, H39 H3', -2G(I), and -2G(J) appear and anomalous peaks increase (Table 15). Although, with the exception of H3 H39', these sites may also be cross-linked to RU1 as the approximate linear distance between C3 metal centers is 53.1 Å (Table 16). This may explain why the anomalous difference peak at maximum soaking time is 58% stronger at H39' than at H39 as doubly adducted C3 molecules may form less readily (Table 15).

Histone - a.a.	Dist (Å)	DNA - n.t.	Dist (Å)
H3 H39	47.1	51G(I)	32.1
H3 H39'	65.5	29G(I)	41.3
H2B H46	12.9	-2G(I)	47.7
H3 H113	34.9	-2G(J)	49.7

Table 16. Linear distances between the ruthenium metal center at RU1 and gold atoms at the indicated site.

Figure 65. Site selectivity and adduct structure of C3–NCP at 2.25 Å. (a) X-ray structure of C3–NCP with the helical axis perpendicular and the pseudo-twofold axis parallel (arrow) to the viewing plane. Histones and DNA are in standard coloring. Sticks represent the C3 adduct that links sites RU1 and AU1. Red cages denote the anomalous difference electron density map contoured at 3.0 σ and specify gold(I) binding locations. (b) The AU1 binding channel, shown in space filling representation, forms a shallow hydrophobic cavity that makes extensive contacts with C3's diphenylphosphanyl-benzoic acid gold carrying group (sticks). Histone side chains contributing to the hydrophobic environment are labeled. Gold(I) ion $r_w = 1.66$ Å.

However, because the distance between C3's metal centers exceeds the distance necessary to cross-link residues at RU1 and AU1, early cross-linking between RU1 and H3 H39 is probable, with RU1–DNA cross-links at –2 positions forming later. The same conformational changes that give rise to allosteric cross-talk between H2A and H3 may also destabilize the nucleosome sufficiently to induce high mobility in DNA termini allowing greater access to the H3 N-terminal tail. Therefore the large anomalous peaks at H3/H3' H39 may result from a combination of the two binding modes previously described, where some C3 molecules are RU1–H39 cross-linked adducts and some where the activated gold group of C3 coordinates H39 alone.

The excessive distance between RU1 and H39 H3' makes C3 cross-linking problematic and the nearly 40 σ anomalous peak cannot be accounted for by non-cross-linked gold adducts alone; the later point being particularly significant when considering dense packing in NCP crystals block access to the H3' bound DNA terminus by three adjacent particles. Face-to-face NCP crystal packing involves electrostatic interaction between the acidic patch of one nucleosome and positively charged H4-tail of an adjacent particle. This strong interaction normally disallows adduct accumulation at the symmetry related RU1' and RU2' sites. Yet in the C3–NCP adduct structure, evidence exists for RAPTA group binding on the opposite face of the nucleosome in a way that is modified by interference from the H4 N-terminal tail of an interacting core particle (Fig. 66; Table 17).

Site	Diff Peak hgt (σ)	Ano Peak hgt (σ)	Dist to AU1' ($\text{\AA}$)	Dist to H3' H39 ($\text{\AA}$)
RU1' '	5.0	4.5	30.6	45.5
RU2'	5.8	4.3	34.8	53.0

Table 17. Difference and anomalous difference peak heights at RU1'' and RU2' and linear distances between ruthenium atoms at these sites and gold atoms at AU1' and H3' H39.

Positive electron density that is nearly equivalent to difference peak heights at RU1 and anomalous signal, which RU1 lacks, are present at RU2' and in the vicinity of RU1' (Fig. 66a & b; Table 17). N-terminal residues of the adjacent nucleosome's H4-tail obstruct RU1' and prevent ruthenium binding at typical H2A residues E61 and E64 (Fig. 66c). This NCP crystal packing interaction introduces a histidine ring between RU1' residues H2A E61 and E64 in an area where there is both extra electron density and anomalous signal (Fig. 66a, b, & c). Negatively charged residues E61, D90, E91,

and E92 encircle this area, along with H18 H4' of the adjacent NCP, offering a potential coordination sphere for a RAPTA ruthenium atom. The H4-tail may direct adduct formation to these new residues, which include half of the established RU1 site (E61), and may promote adduct formation by presenting a new ligand not available at symmetry related RU2. This new reputed site is therefore referred to as RU1''.

Ruthenium atoms are placed at the difference peak center of mass (RU1'') or maximum anomalous peak height at a distance corresponding to empirical metal-coordinating atom bond lengths (RU2'). Gold atoms placed at maximum anomalous peak height, at a distance of 2.05 Å from the H39' and H113' delta nitrogen atoms, serve as reference points for distance measurements (Figs. 66d & 59). RU1'' and RU2' are roughly the same distances to AU1' and H39' as symmetry related RU1 is to AU1 and H39 (Fig. 67; Table 17). A strong difference peak of 22.7σ at H39' shows electron density pointing through the DNA strands towards a clear path to RU1''/RU2' (Fig. 66d). Therefore, with the exception of weak gold-DNA adducts at the dyad, all adducts forming on the opposite face are likely RU1''/RU2'-H39' or RU1''/RU2'-AU1'. AU1' occupancy begins to decrease at the latest time point suggesting more gold adducts accumulate at H39', where increases in absolute peak height are the largest for any peripheral gold site, culminating in the highest signal observed at these sites.

Figure 66. Evidence for cross-linking between H39 H3' and RU1''/RU2'. Histones and DNA are shown in standard coloring with adjacent nucleosome in dark red and ruthenium and gold atoms in blue-green and yellow respectively. **(a).** Putative binding location for C3 RAPTA groups on the NCP face occluded by nucleosome–nucleosome packing interactions and opposite the symmetry related adduct sites. Red cages denote the Fo-Fc difference map contoured at 3.0σ where the center of mass specifies ruthenium(II) binding locations. **(b)** Anomalous difference peaks, represented as green cages contoured at 3.0σ , help position C3 ruthenium(II) metal centers at or near putative coordinating residues. **(c)** RU1', RU1'', and RU2' in detail showing the neighboring nucleosome's H4-tail obfuscating canonical ruthenium binding site RU1' (E61 & E64). H18 H4' may also coordinate a C3 heavy atom, along with D90 and E92. RU2' is typical of RAPTA–NCP adducts with E102 and H106 H2B participating in ruthenium coordination. **(d)** H39 H3' gold binding site in detail. A gold atom was placed at maximum anomalous peak height $2.05 \text{ \AA}$ away from the H39 delta nitrogen atom. Red cages contoured at 3.0σ denote positive electron density around the coordination site. The orientation of the image shows ample clearance for the gold moiety between DNA strands and depicts a clear path for the PEG linker to sites RU1'' and RU2'.

Figure 67. Intra-particle distances between RU1''/RU2' and H39' and allosterically related AU1'. Histones and DNA are shown in standard coloring and ruthenium and gold atoms in blue-green and yellow respectively. The putative binding location for a C3 RAPTA group on the NCP face occluded by nucleosome–nucleosome packing interactions is denoted RU1''. Canonical symmetry related ruthenium site RU2' coordinates a metal atom at E102 and H106 of histone H2B. Gold atoms are shown at H3' H39 and AU1' with Au–N bond lengths of 2.05 Å. RU1''–H39' = 45.5 Å; RU2'–H39' = 53.0 Å; RU1''–AU1' = 30.6 Å; RU2'–AU1' = 34.8 Å.

C3 simultaneously coordinates distinct histones sites on nucleosomes by exploiting an allosteric pathway relating histones H2A–H2B of the octamer dimer and H3/H3' of the tetramer. To our knowledge, this is the first example of the nucleosome tetramer cross-linked to its dimer. The following section presents preliminary data on a trinuclear derivative of C3 and offers supporting evidence for the newly identified nucleosome gold binding sites and modes of adduct cross-linking.

3.6.4 A trinuclear derivative of Compound III forms similar NCP adducts

A trinuclear version of C3 possessing two bifunctional ruthenium(II) RAPTA groups opposite a monofunctional gold(I) AUF/EP11-019 group was also tested on human NCP145 crystals to observe reactivity towards chromatin substrate (Fig. 68; Table 18). The design of trinuclear C3 (tnC3) is intended to advance the dinuclear structure of C3 so that it can form simultaneous adducts at all previously established nucleosome histone sites (RU1, RU2, AU1, AU1'), and build upon nucleosome cross-linking methods our lab has used to cross-link octamer dimer (Davey, G.E. *et al.*, 2017). The

previous section presented structural data on a gold–ruthenium dinuclear compound that cross-links a heterogeneity of nucleosome sites. The following section will offer more support for C3 binding modes by exploring the reactivity of the tnC3 analogue and attempt to strictly rationalize the gold(I)–H3' H39 interaction.

Figure 68. Measured linear distances of tnC3's PEG linkers based on an ideal structure and quantified by PyMOL (Delano, W.L. 2002).

Aside from adding a RAPTA group, tnC3 trades C3's $\sim 38 \text{ \AA}$ base PEG linker for a branched one that extends an extra $\sim 28 \text{ \AA}$ (Fig. 68). All targeting modes and leaving group chemistries present in C3 are retained by tnC3, and adduct formation on the nucleosome surface is generally similar (Figs. 68; 62). As with C3, the best data set, collected at $1.0402 \text{ \AA}$, combining high resolution with strong difference and anomalous difference peak heights was available after 120 hours of compound treatment and used to generate $2F_o - F_c$, $F_o - F_c$, and anomalous difference maps for analysis (Tables 18 & 19). Possibly due to the additional arene ring and extended PEG chain, tnC3 was markedly less soluble than C3 in standard stabilization buffer. Dissolution of tnC3 required resuspension in stabilization buffer at a final MPD concentration of 47% w/v and prolonged vortexing. NCP crystals tolerated the higher MPD concentration well with no losses in diffraction quality or resolution limit even after 5 days of continuous soaking (Table 18).

Table 18. Data collection and refinement* statistics for trinuclear compound III treated human NCP145

	Tri-nuclear CIII				120-h		
[Agent] (mM)	1	1	1	1	Refinement		
Duration (hrs)	28	54	96	120*			
Histones	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	Resolution (Å)	2.19–48.57	
Data collection					No. reflections	103482	
					R_{work} / R_{free} (%)	23.2/27.2	
	Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	No. atoms	12152
	Cell dimensions					Protein	6098
	<i>a</i> (Å)	106.57	106.41	106.25	106.34	DNA	5939
	<i>b</i> (Å)	109.18	109.18	109.16	109.19	Solvent	--
	<i>c</i> (Å)	174.90	174.82	174.92	175.07	Adducts	5
	Resolution (Å)	2.25–48.59	2.40–48.57	2.20–48.55	2.19–48.57	<i>B</i> -factors (Å ²)	100.8
		(2.25–2.37)	(2.40–2.53)	(2.20–2.32)	(2.19–2.31)	Protein	73.9
	R_{merge} (%)	4.7 (178)	4.8 (182)	4.3 (186)	4.7 (180)	DNA	133.1
R_{pim} (%)	1.4 (79.3)	1.5 (75.8)	1.3 (67.0)	1.4 (66.1)	Solvent	--	
$I/\sigma I$	26.1 (1.0)	28.9 (1.0)	27.1 (1.2)	24.8 (1.1)	Adduct	95.5	
Completeness (%)	98.1 (87.6)	96.6 (80.4)	99.5 (96.7)	98.7 (91.3)	R.m.s. deviations		
CC (1/2) (%)	99.9 (58.9)	100 (60.2)	99.9 (70.7)	99.9 (65.7)	Bond lengths (Å)	0.0095	
Redundancy	11.9 (6.6)	12.1 (7.4)	12.7 (9.2)	12.6 (8.9)	Bond angles (°)	1.4405	

Single crystal data sets; values in parentheses are for highest resolution shell.

In our studies the H2A–H2B acidic patch is the most consistently active location for RAPTA analogue binding. Trinuclear C3's RAPTA groups predictably coordinate known glutamate and histidine residues as demonstrated by a large area of positive electron density around RU1 and RU2 (Fig. 69a). Here difference peak heights remain constant throughout the treatment suggesting site occupancy quickly and stably nears 100% (Table 19). X-ray diffraction data collected at 1.0402 Å does not reliably expose anomalous scattering from ruthenium atoms, yet faint peaks in the anomalous difference map between RU1 H2A–E61/E64 and RU2 H2B–E102/H106 residues allowed the placement of heavy atoms for further analysis (RU1 = 4.1 σ and RU2 = 3.9 σ at 120 hours) (Fig. 69b). Anomalous signals at AU1/AU1' did not exceed background levels making tnC3 gold activity difficult to evaluate (Table 19). Even with identical ligands and leaving groups, gold binding at AU1/AU1' is remarkably low relative to C3 data (~15% vs ~60%). Poor gold coordination at this site is likely a consequence of the overall insolubility of the compound.

Table 19. Ruthenium(II) and gold(I) binding sites in human NCP145 treated with trinuclear compound

Data Set	Difference* or Anomalous Difference Peak height (σ)								
	RU1*	RU2*	AU1*	AU1'*	H3 H39	H3' H39	H2B' H46	-2G(I)	-2G(J)
28 hr	6.06	4.66	5.66	4.46	--	--	4.30	--	--
54 hr	5.76	4.16	5.36	--	5.10	8.10	6.90	4.10	--
96 hr	5.56	4.16	4.96	4.16	5.70	9.30	5.30	4.50	--
120 hr	5.82	4.62	5.42	4.02	7.30	11.90	6.70	--	4.62

Sites correspond to coordination of RAPTA (RU1/2) or the gold group (all other locations). Values correspond to difference ($F_o - F_c$; RU1/2 and AU1/1') or anomalous difference electron density map peak magnitudes. "G" sites refer to guanine nucleotide sites near the nucleosome center.

In a recent study of PEG linked dinuclear ruthenium compounds, the authors noticed increasing linker length and corresponding hydrophobicity is associated with greater antiproliferative activity towards human ovarian cancer and cisplatin resistant cells lines (Batchelor, L.K. *et al.*, 2017). For tnC3, the increase in hydrophobicity imparted by its branched PEG linker may lead to better performance in cytotoxicity assays and thereby raise its potential for therapeutic application. However, because the linker is excessively long and contains an additional arene group, solubilization in partial organic solvent is necessary, and this possibly affects the phenyl-benzoic acid ligand's capacity to target buried histidine residues at AU1/AU1' by diminishing the hydrophobic effect. Still, bulk solvent dehydration does not explain the gold group's ability to target solvent exposed histidine residues or histidine and guanine residues in high charge density environments. At these types of residues, dehydration from the organic solvent would decrease the dielectric constant, which would enhance electrostatic interactions and mitigate hydrophobic interactions.

Similar to C3's mode of action, tnC3 gold attack at H2B H46 is likely the result of the side chain's proximity to RU1/RU2. But unlike C3, the excessive linker length causes unstable binding as demonstrated by fluctuating peak signal over the complete compound soaking time course (Table 19). For peripheral gold sites H3/3' H39 and -2G(I/J), gold adduct formation may rely on the targeting specificity of C3's phenyl-benzoate ligand. Although bulky and hydrophobic, its ringed structure possesses multiple delocalized pi systems that could theoretically participate in DNA base stacking interactions that help stabilize adduct formation. Therefore, in order to model structural results, positive electron density with approximate shape and geometry of phenyl rings and Au-P-Phe bond angles is used to identify the presence of tnC3's AUF-like moiety around AU1/AU1' (Table 19; Fig. 69c). At peripheral gold sites, anomalous peak height alone was sufficient for further analysis (Table 19).

Figure 69. Overview of tnC3 interaction with human NCP145 at observed C3 sites. (a) A large swath of positive electron density, contoured at 3.0σ , signifies the presence of the two RAPTA-C analogues and bifurcated PEG linker around RU1 and RU2 coordinating residues. (b) Anomalous signal, contoured at 3.0σ , shows Ru(II) heavy atoms centers at canonical NCP ruthenium binding sites RU1 and RU2 with appropriate coordinating residues (RU1 = 4.10σ ; RU2 = 3.90σ). (c) Minimal positive electron density with approximate diphenyl ligand shape and weak anomalous signal around gold(I) coordinating sites of the AU1 and AU1' binding channels indicates the presence of Trinuc C3's AUF/EP11-019 group. (d) Relative locations on the NCP surface of tnC3's three heavy atoms placed at corresponding difference and anomalous difference peak heights illustrate the nucleosome intraparticle distances between AU1 and RU1 ($35.0 \text{ \AA}$) and AU1 and RU2 ($34.4 \text{ \AA}$) ($\text{Au } r_w = 1.66 \text{ \AA}$) ($\text{Ru } r_w = 2.05 \text{ \AA}$). The black arrow points to the central base pair with core histones in standard coloring.

Not surprisingly, several of the stronger gold adducts first identified in C3 data also appear in tnC3 anomalous maps, albeit with lower anomalous signal (Fig. 70; Table 19). Average relative change in anomalous peak height between peripheral gold-histidine sites in trC3-NCP and C3-NCP is -0.7 , and approximately -0.3 at gold-guanine sites (Fig. 70). Relative change is calculated as the actual change in anomalous signal between data sets (tnC3 – C3) divided by a reference anomalous signal, here chosen to be C3 anomalous signal at the corresponding time point. The difference in peak height magnitude at gold peripheral sites and reduced binding at AU1/1' may be due to some negative effect from the reagent's structure or linker hydrophobicity since the chemistry of tnC3/C3 reactive groups is the same. It is more likely that tnC3 cross-linking adducts do not readily form because of the greater degree of dehydration from increased MDP in the soaking buffer that is necessary to solubilize the reagent. This

may affect targeting to AU1/AU1' where the C3/tnC3 gold moiety relies on hydrophobic interactions with binding channel residues. But this is not a factor for adducts at H3/H3' H39, H2B H46, and -2G(I/J), where imidazole rings are solvent exposed or surrounded by a charged environment, making compound insolubility the prime factor for reduced binding at these sites. For the RU1/RU2 sites, solubility may also be an issue in adduct formation, however, double the number of ruthenium atoms per reagent molecule may lead an increase in probability of metal binding at these sites that gives the appearance of greater coordination relative to the gold moiety.

Figure 70. Trinuclear C3 gold(I) adduct locations on NCP145 and their comparative peak magnitude to C3. (left) Several gold(I) adducts identified in the C3–NCP adduct structure are also found in tnC3–NCP data. Anomalous difference peaks at each gold binding site is shown as red cages contoured at 3.0 σ with the exception of AU1/AU1' where the Fo–Fc map is shown as red cages of equal contour level. (right) All tnC3 gold adduct sites are lower occupancy, expressed as significantly weaker peak heights, compared to C3 data. Relative change (R.C.) for late appearing guanine adducts is less significant since they have had insufficient time to form to the same extent as histone adducts. Adducts at sites 29G(I) and 51G(I) do not form even after 120 hours of continuous

Surprisingly, tnC3 RAPTA adducts appear to form on the positive face of the nucleosome where the strongest electron density and anomalous signals are found at RU1'' and RU2' (Fig. 71a & b). In fact, based on absolute peak height, these sites seem to develop adducts to a greater extent than unobstructed symmetry related nucleosome sites RU1/RU2. Ruthenium atoms were placed in the tnC3 structure at maximum anomalous peak height between RU1'' and RU2' coordinating residues and link records were created with experimentally derived bond lengths for Ru(II)–O, Ru(II)–N, and Au(I)–N before atomic refinement and visual bond representation (Fig. 71: Table 19). The putative H4 H18 interaction deduced in the C3–NCP adduct structure is more clearly defined in the tnC3–NCP structure where the histidine side chain moves into range of metal atom and coordinating residues (Fig. 71c). Here the imidazole delta

nitrogen of H4' H18 is within 3.0 Å of the RU1'' ruthenium atom bound to the delta oxygen of H2A D90, and its epsilon nitrogen is 2.6 Å from the epsilon oxygen of H2A E61 (Fig. 71c). The refined double difference map around the H4-tail of an adjacent nucleosome places H18's side chain in a favorable position to interact with the E61 bound ruthenium atom via the imidazole delta nitrogen (Fig. 71c). The 3.0 Å distance is too long to represent a dative covalent bond, but is within both electrostatic and dipole range. Were the metal center to lose one of its ligands, these additional modes of interaction could become available.

In the native 1.99 Å resolution human NCP145 structure, H4 H18 is part of the H4-tail–H2A-H2B acidic patch crystal packing interaction and is in proximity to H2A E61 where, depending on the ionization state of the imidazole, salt bridge formation with the glutamate side chain is feasible. Structural alignment of three crystal structures, C3–NCP, tnC3–NCP, and native NCP145, shows H4 H18 moving away from both RU1' and RU1''. This side chain repositioning is clear evidence for C3's RAPTA group coordination that leads to RU1''–H39' cross-linking. From the six molecule superposition H18 exclusion from RU1'/RU1'' can easily be observed across structures and is likely occurring due to steric influence from PEG linker bulk (Fig. 71d).

For the intermediate case of tnC3–NCP, anomalous scattering may be the result of aqua–ruthenium complexes where one or more carrier ligands have been cleaved (the arene ligands are likely first), allowing them to interact more readily with the H18 side chain. This would in part explain why the extensive area of electron density witnessed at RU1/RU2 is not present at RU1'/RU1''/RU2' where positive density is localized only to coordinating residues. The loss of arene and PEG chain would substantially reduce the degree of X-ray scattering at these sites, but possibly also reduce thermal motion enough to allow strongly positioned ruthenium atoms to contribute greater anomalous scattering at a 1.0402 Å wavelength. The exact opposite could be true for C3, where the presence of the highly disordered arene and PEG linker attenuate X-ray scattering resulting in weaker peak heights in both difference and anomalous difference data (Fig. 66). However, if the bridging linker is retained, the apparently obstructed site would become partially unobstructed via H4' H18 clash with a much longer branched PEG molecule. Additionally, better difference and anomalous signal may be attributed to less disorder in the region, and not necessarily greater metal binding.

Trinuclear C3 is a novel organometallic reagent that binds to chromatin substrate. Its inception builds upon C3's novel hetero-multinuclear design but requires further development concerning linker length and arrangement of metal groups. It is made clear from C3/tnC3–NCP adduct structures that both the chemistry and size of the bridging linker exert critical influence on the kinetic, stoichiometric, and thermodynamic interplay of adduct formation in chromatin. Therefore, another effective design could be to utilize a shorter unbranched PEG linker to join the AUF/EP11-019 derivative to dinuclear ruthenium compounds that are proven to cross-link H2A and H2B (Davey, G.E. *et al.*, 2017).

Figure 71. Symmetry related ruthenium sites RU1'' and RU2' show increased RAPTA coordination relative to C3 RU1'' and RU2' and tnC3 RU1 and RU2. (a) The Fo-Fc omit map, as red cages contoured to 3.0 σ , shows a focused area of positive electron density around RU1'' and RU2'. (b) Anomalous peaks (green cages at 3.0 σ contour) are also strongest at RU1'' and RU2'. (c) A refined double difference map around the H4-tail of an adjacent nucleosome (blue cages contoured at 1.0 σ) shows H18 interacting with a ruthenium atom coordinated by D90 of H2A at a distance of 3.0 Å and E61 of H2A at a distance of 2.6 Å (black dashes). (d) Structural superposition of the coordinating environment around RU1', RU1'', and RU2' of C3 (cyan), tnC3 (magenta), and native NCP145 (yellow); C α RMSD = 0.144 Å.

Moreover, RU1/2–AU1/1' allosterism may largely depend on the presence of a complete RAPTA group at ruthenium binding sites, while RAPTA analogues are insufficient. Compounds I (C1) and IV (C4) substitute the PTA ligand with a diphenylphosphanyl benzoic acid group that connects directly to the PEG linker (Appendix 5). Crystallographic data from small molecule soaking experiments of C1 and C4 on human NCP145 show strong binding of RU1/RU2 and RU2' sites with no gold binding site activation (Table 20). Furthermore, RAED-C–AUF crystal soaking experiments elicit only a low level of allosteric impact at AU1 after 121 hours of treatment with difference and anomalous difference peaks of 5.39 σ and 6.20 σ respectively. This is in contrast to AUF and C3 AU1 binding where anomalous signal is approximately 5-fold greater.

Table 20. Ruthenium(II) and gold(I) binding sites in human NCP145 treated with compound I and compound IV

	Duration	Res. (Å)	Difference and (Anomalous) Signal				
			RU1	RU2	RU1'	RU2'	AU1
C1	20-h	2.20	0.0	0.0	0.0	0.0	0.0
	43-h	2.48	4.4	4.8	4.6	4.4	0.0
	63-h	3.34	8.3	8.3	0.0	6.1	0.0
	86-h	3.10	7.0	13.0(4.8)	0.0	7.4	0.0
C4	20-h	2.70	5.9	4.9	0.0	0.0	0.0
	43-h	3.00	9.0(4.7)	10.5(4.8)	0.0	5.5	0.0
	63-h	3.91	5.5	4.4	0.0	4.6	0.0
	86-h	3.55	7.6	15.4(4.7)	0.0	10.8	0.0

Sites correspond to coordination of ruthenium(II) carrier group (RU1/2/1'/2') or the gold group (AU1/1'). Values correspond to difference ($F_o - F_c$) or anomalous difference (in parentheses) electron density map peak magnitudes.

Although linker length for C1 is too short to reach AU1/1' (≈ 20.6 Å), C4 linker length is sufficient at approximately 31.4 Å indicating the potential for RU1/2–AU1 cross-linking. These observations suggests two outcomes. First, when ruthenium(II) coordinates RU1/RU2, a PTA ligand is required at these sites to evoke the RU1/2–AU1/1' allosteric pathway as C1 and C4 form no AU1/1' gold adducts and no C4 cross-linking is witnessed. Second, the fact that no peripheral gold adducts are present, even at the longest soaking duration, suggests no extraneous C1 or C4 molecules bind the nucleosome. This supports single molecule cross-link formation in the C3 and tnC3 data sets.

3.6.5 Preliminary data of Au(I)–viral peptide conjugates in complex with NCP

In our studies of organometallic nucleosome adducts, the H2A–H2B acidic patch is the primary target for ruthenium-based RAPTA adduct formation. Here negatively charged acidic residue side chains act as hard bases and offer lone pair electrons to Ru^{2+} atoms thus forming dative covalent bonds. By HSAB theory and experimental determination, ruthenium(II) acts as a hard acid in this context (Hoggard, P.E. and Porter, G.B. 1981).

The acidic patch is a unique motif that plays a major role in the structural and biological regulation of chromatin and its association with H4-tails of neighboring nucleosomes plays a crucial role in condensation (Chodaparambil, J.V. *et al.*, 2007). Furthermore, with its abundance of negatively charged residues, the acidic patch can act as a scaffold for numerous structurally disparate architectural proteins and remodeling factors that are unrelated in function but retain conserved nucleosome binding motifs (Kalashnikova, A.A. *et al.*, 2013). In addition to chromatin regulating proteins, several viral peptides compete with H4-tail binding to access the acidic patch, a fundamental mechanism for their survival and replication. The following study will focus on the Kaposi's sarcoma herpes virus (KSHV) latency-associated nuclear antigen (LANA) and the chromatin-binding sequence (CBS) of a prototype foamy virus (PFV) structural protein termed group-specific antigen (GAG).

LANA is an N-terminal peptide fragment of the KSHV that tethers the viral episome to mitotic chromosomes (Ballestas, M.E. *et al.*, 1999). The first twenty two residues of LANA comprise the chromosome association domain, a highly charged peptide essential for viral persistence (Barbera, A.J. *et al.*, 2006a). Residues 5 – 16 are required for peptide–nucleosome interaction and contain R9, an anchor motif common among proteins that vie for exclusive binding to the acidic patch (Piolot, T. *et al.*, 2001; Barber, A. *et al.*, 2004; Lesbats, P. *et al.*, 2017). The PFV GAG CBS is a necessary component of spumavirus retroviral genome integration into host chromosomes (Lesbats, P. *et al.*, 2016). GAG CBS is located in a glycine/arginine rich box of the viral protein's C-terminal region. Found within this region, R540 acts as the conserved anchor motif to tether GAG protein to host chromatin in a gene-independent manner (Lesbats, P. *et al.*, 2017). Although both peptides utilize the same motif to recognize the acidic patch, they originate from vastly different viral protein structures and attendant mechanisms.

Whereas LANA adopts a hairpin conformation when associated with nucleosomes, GAG CBS extends across the entire octamer surface (Fig. 72).

Figure 72. Overview of LANA and GAG CBS peptide binding to the nucleosome acidic patch. (a) Close up detail of LANA (magenta) and GAG CBS (cyan) arginine anchors interacting with the same negatively charged residues of the acidic patch (red). (b) The nucleosome acidic patch acts as a binding scaffold for an assortment of chromatin factors while providing access to a large surface area capable of accommodating various structures. LANA (magenta) folds into a tight hairpin while GAG CBS (cyan) runs along the entirety of the octamer face.

To exploit the affinity of LANA and GAG CBS for the acidic patch, our collaborators (Prof. P.J. Dyson and Ms. L. Batchelor) fashioned viral peptide–metalloreagent conjugates by attaching AUF derivative reactive groups to the peptide’s N-terminus via an unbranched PEG linker (Fig. 73). These peptide–metalloreagent conjugates are referred to as Au-LANA and Au-GAG. With this quasi-biopharmaceutical design we will attempt to capitalize on a naturally evolved pathogenic mechanism by using it as a delivery method for organometallic therapeutics. If successful, these designs could potentially bridge the gap between traditional biologic and small molecule organometallic anticancer therapeutics.

Depicted in figure 73 are the full chemical structures for Au-LANA and Au-GAG. The reactive gold(I) groups for both peptide–metalloreagent conjugates are identical to the one found in C3 and tnC3. Joining the metal center to the peptide’s N-terminus is a PEG linker of approximately 70 Å in linear length.

Figure 73. Full chemical structures with key features of gold(I)–viral peptide conjugates. As a result of the purification process, the C-terminus of both peptides have been modified with an amine group. Green boxes indicate the sequence position of the arginine anchor motif. The N-terminus of each peptide is conjugated to a ~70 Å PEG linker ending with a reactive AUF/EP11-019 derivative. **(top)** Residues 2 – 22 of the LANA peptide with a Met to Ornithine substitution at position six (red box). **(bottom)** CBS residues 535 – 551 of the GAG peptide conjugated

M6 and L8 of native LANA are buried in a hydrophobic region of the acidic patch (⁴⁹VYLAADVLEYL⁵⁸) where they form hydrophobic contacts with V54, Y50, and Y57 of H2A (Barbera, A.J. *et al.*, 2006a). Specifically, the M6 terminal methyl is situated in a narrow groove between H2A L23, A53, and Y57. Here the methionine rotamer observed in the LANA–NCP crystal structure places the epsilon carbon approximately 2.9 Å away from E110 of H2B. This is an ideal distance for salt bridge formation between the negatively charged carboxylate acceptor group of E110 and a positively charged donor group in place of methionine’s terminal methyl. LANA M6O substitutes an ornithine for methionine at position six in a truncated version of LANA. This mutation, introduced to eliminate nucleophilic sulfur groups from the system, maintains the peptide’s relative net charge at pH 7.0 while conserving the side chain’s general structure (Fig. 74).

Figure 74. An ornithine to methionine substitution at position six of LANA alters peptide chemistry. (left) Comparative chemical structures of methionine (top) and ornithine (bot) showing approximate side chain length and the transition from a sulfide to a primary amine at the terminal atom position. (right) Biochemical and biophysical parameters of acidic patch binding peptides (Chodaparambil, J.V. *et al.* 2007). Despite the loss of 9 amino acids, that include a proline and an arginine, LANA M6O maintains the same pI and net charge as full length LANA.

Derivatization of human NCP145 with viral peptides and viral peptide–metalloreagent conjugates was carried out according to soaking conditions stipulated in Barbera, A.J. *et al.*, 2006a and Lesbats, P. *et al.*, 2017. The seemingly minimal length, 14 amino acid, peptide (LANA-M), based on visualized (ordered) amino acids in the LANA-NCP structure, and this ‘minimal’ peptide with a M6O substitution (LANA M6O) served as controls for the LANA–metalloreagent conjugates, which are all composed of a near full-length (21 amino acid) LANA peptide with the M6O substitution. Upon soaking, detrimental physical impact to NCP crystals occurred in both the LANA (full-length peptide) and the LANA-M data sets. Here, peptide binding rapidly compromised diffraction quality (Fig. 75). In the LANA M6O data set, a discernible physical impact occurred but diffraction quality remained high (Fig. 75; Table 21).

Figure 75. Representative diffraction patterns for full length LANA peptide and engineered controls. Introduction of peptide to soaking buffer damaged crystals and affected diffraction quality in both LANA – full length and LANA-M data sets. The diffraction patterns for these two data sets show poor resolution limits, anisotropy, and high mosaic spread. Conversely, the LANA M6O data set demonstrates good resolution with a

The LANA M6O structure was solved to 2.50 Å resolution and used as a model to analyze Au-LANA–NCP X-ray data. At the earliest time point (15 hrs), both full length

and LANA-M badly damaged NCP crystals suggesting that the peptide displaces H4-tail-acidic patch crystal contacts just as it outcompetes H4-tail governed face-to-face nucleosome packing in chromatin during a critical step in KSHV's chromosomal integration mechanism (Barbera, A.J. *et al.*, 2006b; Chodaparambil, J.V. *et al.*, 2007). This damaging effect of LANA M6O on human NCP145 crystals is not surprising considering LANA M6O interactions with human NCP145 particles is highly similar to full length LANA interactions with *X. laevis* NCP146 (Figs. 76 & 77b).

Intermolecular				Intramolecular			
Histone	Residue	LANA	LANA M6O	LANA		LANA M6O	
H2A	E92	R9	R9	M6(mc) - T14(mc)	R7	-	G11(mc)
	D90	R9	R9	M6(mc) - T14	L8(mc)	-	G11(mc)
	E61	R9	R9	L8(mc) - G11(mc)	L8(mc)	-	R12(mc)
	E61	R9(mc)	R9(mc)	L8(mc) - R12(mc)	S10	-	G11(mc)
	E64	S10	S10	S10 - R12(mc)	S13	-	G15(mc)
H2B	E110	R7(mc)	O6	S13	-	G15(mc)	
	V45(mc)	R12	R12				
	K113	G5(mc)	P4(mc)				
	Q44	G15(mc)	--				

Figure 76. Summary of inter- and intramolecular interactions for full length LANA and LANA M6O peptides. (left) Peptide side chain and mainchain (mc) interactions between *X. laevis* NCP146 histone sides chains and full length LANA peptide, and human NCP145 histone side chains and engineered LANA M6O. **(right)** Intramolecular interactions between peptide residue side chain and main chain atoms for full length LANA and LANA M6O that help fold the peptide into the correct hairpin conformation.

LANA M6O binds to the nucleosome surface in the same location as native viral peptide as designated by positive electron density in the Fo-Fc difference map around acidic patch residues (Fig. 77a). Specific interactions and the overall hairpin fold of the M6O peptide are similar to the native peptide despite being seven amino acids shorter with a drastic change in side chain chemistry at position six (Fig. 77b). This change in side chain chemistry allows for the formation of a salt bridge between LANA O6 and H2B E110 (Fig. 77c). In terms of free energy contribution to binding stabilization, the addition of a salt bridge could compensate for the loss of hydrophobic and van der Waals interactions between LANA M6's terminal methyl and H2A residues L23, A53, and Y57. However, quantifying the stability of an individual intermolecular salt bridge is difficult, and more complex in this case, as the side chains involved are not adequately solvated due to the local hydrophobic environment. The energetic penalty accrued for burying a charge (O6) where relative permittivity is high (L23, A53, Y57) may offset any stabilizing free energy contribution from salt bridge formation, and this effect could be significant on a small peptide. Nevertheless, the 2Fo-Fc map

confirms the correct orientation of H2B E110 and LANA O6 side chains for interaction where the distance between E110 carboxylate oxygen atoms and ornithine's epsilon nitrogen is 2.6 Å. (Fig. 77c).

It is worth noting that the LANA O6–E110 H2B interaction replaces an existing salt bridge between the main chain amide of LANA R7 and H2B E110. In all, LANA M6O loses one intermolecular interaction with H2B, Q44–G15, and shifts a salt bridge between the K113 of H2B and the main chain carboxyl oxygen of G5 one residue over to the main chain carboxyl oxygen of P4 (Figs. 76 & 77b). In this region the engineered peptide is more compact due to a general rearrangement of intramolecular interactions and the exclusion of a positive charge by hydrophobic contacts as a result of the M6O substitution.

Intramolecular interactions involved in folding LANA M6O into its hairpin conformation are also altered relative to the native peptide. The main chain amine of G11 now forms a bifurcated h-bond with the S10 side chain and L8 main chain oxygen, the G11 main chain oxygen forms a salt bridge with the R7 side chain, and both interactions between M6 and T14 of native LANA are lost in addition to an S10 to R12 H-bond (Figs. 76 & 77b). How these changes effect LANA M6O binding affinity to the nucleosome acidic patch require further investigation and quantitative biochemical and biophysical characterization.

Figure 77. Overview of LANA M6O binding and interaction with human NCP145. Core histones appear in standard coloring. **(a)** (left) The binding location of LANA M6O on the nucleosome surface is identical to full length LANA in the *X. laevis* NCP146 structure (PDB code 1ZLA). (right) The Fo-Fc omit map shows positive electron density as red cages contoured at 3.0 σ with LANA M6O represented as black sticks and the strongly positioned R9 arginine anchor fitting in well-defined density and interacting with acidic patch residues. **(b)** Black dashes represent polar and electrostatic interactions between LANA M6O and residues of the H2A–H2B acidic patch while red dashes represent the peptide's intramolecular interactions. **(c)** (left) Detail of ornithine's interaction with E110 of H2B in the LANA M6O structure and its comparison to the analogous location in 1ZLA by structural superposition. LANA M6O is shown as black sticks with the corresponding E110 H2B residue in red. The distance between ornithine's epsilon nitrogen atom and E110's carboxylate oxygen atom is approximately 2.6 Å. LANA from 1ZLA is shown as translucent blue sticks with the corresponding E110 H2B residue in gray. (right) The 2Fo-Fc difference map contoured at 1.0 σ shows X-ray data supports the O6–E110 interaction.

NCP145 crystals survived Au-LANA treatment for 3.5 days and maintained good diffraction with high-resolution limits after a reduction in peptide-metallureagent conjugate concentration from 2 mM to 0.5 mM (Table 21). Crystals appeared to melt in the presence of reagent even after the 75% reduction in reagent concentration. This effect was not observed with unconjugated peptide, mononuclear EP11-019, or any PEG- or alkyl-linked multinuclear compounds. Therefore, crystal melting is likely attributed to the concurrence of specific nucleosome site targeting by the peptide and the duration of reagent soaking where the longer exposure results in higher peptide occupancy at the acidic patch to increasingly interfere with nucleosome–nucleosome interactions thus damaging crystals.

Table 21. Data collection for LANA M6O, Au-LANA M6O, and Au-GAG and refinement statistics for LANA M6O treated human NCP145 crystals

	LANA-O		Au-LANA-O		Au-GAG			
[Agent] (mM)	2	2	0.5	0.5	2	2	2	2
Duration (hrs)	15	29	61	84*	29*	54	83	118
Histones	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>
Data collection*								
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions								
<i>a</i> (Å)	109.71	107.43	107.21	107.29	107.17	108.11	107.95	107.71
<i>b</i> (Å)	109.88	109.69	109.64	109.47	109.64	110.42	110.35	107.87
<i>c</i> (Å)	184.88	183.28	183.16	183.48	183.56	178.77	178.93	179.44
Resolution (Å)	2.50–49.13	2.19–48.85	2.11–48.81	2.10–48.85	2.17–48.80	3.00–93.94	3.10–93.92	3.20–92.45
	(2.50–2.64)	(2.19–2.31)	(2.11–2.23)	(2.10–2.21)	(2.17–2.29)	(3.00–3.16)	(3.10–3.27)	(3.20–3.37)
<i>R</i> _{merge} (%)	5.8 (283)	5.1 (106)	6.2 (186)	5.3 (177)	6.4 (186)	15.1 (174)	14.4 (161)	12.5 (152)
<i>R</i> _{pim} (%)	1.7 (84.4)	1.6 (34.3)	1.9 (67.8)	1.6 (57.6)	2.0 (64.2)	5.0 (57.6)	4.6 (53.9)	4.0 (45.6)
<i>I</i> / σ <i>I</i>	23.0 (1.1)	25.3 (2.3)	20.6 (1.1)	23.6 (1.3)	20.3 (1.2)	7.7 (1.1)	7.4 (1.0)	10.1 (1.4)
Completeness (%)	99.8 (98.9)	99.7 (98.1)	97.9 (85.5)	99.7 (98.0)	98.2 (88.1)	99.4 (99.0)	99.4 (98.1)	99.9 (99.4)
CC (1/2) (%)	100 (66.7)	99.9 (83.4)	99.9 (51.8)	100 (58.6)	100 (57.3)	99.8 (61.2)	99.8 (57.9)	99.9 (62.0)
Redundancy	13.3 (12.7)	13.0 (11.1)	12.5 (8.9)	13.1 (11.2)	12.2 (9.7)	10.7 (10.6)	10.8 (10.0)	11.1 (10.1)
Refinement								
Resolution (Å)	2.50–49.13							
No. reflections	73641							
<i>R</i> _{work} / <i>R</i> _{free} (%)	21.2 / 26.4							
No. atoms	12154							
Protein	6116							
DNA	5939							
Solvent	--							
Peptide	99							
<i>B</i> -factors (Å ²)	121.9							
Protein	88.0							
DNA	145.8							
Solvent	--							
Adduct	131.8							
R.m.s. deviations								
Bond lengths (Å)	0.0089							
Bond angles (°)	1.4759							

Single crystal data sets; values in parentheses are for the highest resolution shell. “*” indicates data set used for map generation and analysis. The highlighted fields in the Au-GAG data sets indicate a doubled *b*-axis length.

Despite this troubling observation, Au-LANA treated crystals consistently diffracted to high resolution throughout the treatment time course. At the gold edge, the 84-hour time point produced a 2.10 Å diffraction data set that was used to generate a primitive structure and difference and anomalous difference maps (Fig. 78a; Table 21). Although resolution remained high while NCP crystals sustained deleterious physical damage, Au-LANA did not seem to bind nucleosome as well as the M6O control. The difference peak height at R9 for LANA M6O is 12.2 σ in the refined omit map versus 6.9 σ for Au-LANA suggesting lower occupancy of the peptide–metalloreagent conjugate.

A Fo-Fc map, displayed in figure 78a contoured at 3.0 σ , on the nucleosome surface depicts the peptide binding location of Au-LANA. Positive electron density also

indicates that the conjugated peptide assumes the native/engineered peptide conformation and maintains most of the former's interaction profile. Superposition of the LANA M6O peptide structure onto the Au-LANA nucleosome structure with Fo-Fc map places the R9 anchor motif in clearly defined density surrounded by H2A E61, D90, and E92 (Fig. 78b).

Evidence for PEG linker emerging from the location indicated by black arrows (P4) extends towards the AU1 gold site (Fig. 78a & b). However, there is no anomalous peak associated with gold(I) adduct formation at H113 of H3. In fact, the entire data set does not register anomalous signal above the 3.5σ background cutoff. The Au-LANA peptide adopts a native fold which naturally orients PEG linker, attached to the N-terminal APP sequence, away from AU1 and the majority of peripheral gold sites. This fold also places LANA P4 and H3 H133 about 50 Å apart. PEG linker is sufficiently long and flexible enough to accommodate this site at this distance. Therefore, lack of reactivity is more likely the cause of the peptide's inability to elicit the RAPTA–AUF allosteric effect. Furthermore, the PEG linker may also be influenced by attractive/repulsive interactions between peptide and octamer surface.

Figure 78. Au-LANA binds to the nucleosome surface. (a) Au-LANA binding location on the nucleosome surface as depicted by positive electron density in the Fo-Fc map contoured at 3.0 σ , with evidence for linker extending towards AU1 (black arrow). Core histones in standard coloring. (b) Detail of the Au-LANA peptide's putative interaction with H2A–H2B residues based on the superimposed LANA M6O structure. P4 and P17 are terminal LANA M6O residues. In Au-LANA, the PEG–AUF/EP11-019 group is attached to an N-terminal alanine-proline sequence that precedes P4. Beyond P17, the Au-LANA C-terminal sequence concludes with LTRGS-NH₃⁺.

Au-GAG data was processed exactly as Au-LANA data in order to generate maps and preliminary structures. With Au-GAG however, longer incubation times badly damaged NCP crystals leading to a precipitous drop in resolution between the 29- and 54-hour data sets. In fact, the *b*-axis appeared to double in the 54-, 83-, and 118-hour data sets when processed with go.com. Therefore, an alternative method of data processing was employed using iMOSFLM and the CCP4 scala routine, which accounts for higher upper resolution limits observed in the later time point data sets (Table 121). Difference maps were generated for all time points, but only the 29-hour data set presented peptide binding as extensive electron density extending across the octamer surface (Fig. 79). Consequently, the 29-hour data set, where difference peak height is $\approx 9.2 \sigma$ around R540, was used to create a preliminary structure at 2.17 Å resolution and Fo-Fc maps for initial analysis.

Linker electron density is not detectable in Au-GAG difference maps as density related to the peptide stretches across the entire face of the nucleosome (Fig. 79a). A superposition of the PVF GAG CBS structure (PDB code 5MLU) with the Au-GAG nucleosome structure and difference map shows excellent agreement between peptide–metalloreagent conjugate X-ray data and published viral peptide coordinates (Fig. 79b). As with Au-LANA data, there is no indication of gold adduct formation at AU1. It appears that GAG CBS residue Y549 in any case blocks access to H3 H113 (Fig. 79b). And once again, there is no anomalous signal above the 3.5 σ cutoff in the Au-GAG data set, serving as further evidence negating the notion of free gold atoms forming adducts at AU1 or peripheral gold sites.

Figure 79. Au-GAG binds to the nucleosome surface. (a) Au-GAG binding location on the nucleosome surface as depicted by positive electron density in the Fo-Fc map contoured at 3.0 σ . Core histones in standard coloring. The black arrow indicates the point of PEG linker attachment (b) Detail of the Au-GAG peptide's putative interaction with H2A, H2B, and H3 residues based on the superimposed spumavirus PFV GAG CBS structure (PDB code 5MLU). The PEG-AUF/EP11-019 group is attached to the N-terminal G535 residue of CBS GAG.

Moreover, it is unclear if viral peptide binding to the nucleosome acidic patch elicits the allosteric cross-talk pathway observed between RAPTA-T and AUF. Peptide backbone alignments of LANA-NCP (1ZLA) and GAG-NCP (5MLU) with AUF-NCP (5DNN) and LANA-NCP and GAG-NCP with *X. laevis* NCP145 (2NZD) reveal similar subtle main chain conformational changes between organometallic adducted and peptide bound nucleosomes (Fig. 80a). Conformational variation occurs to a slightly higher degree in backbone alignments between native and peptide bound NCP (Fig. 80b). This is surprising considering the arginine anchors of LANA and GAG CBS recognize and bind noncovalently to different acidic patch residues than those coordinated by ruthenium(II) atoms of RAPTA family compounds. The peptides also contact larger surface areas on the nucleosome than small molecule reagents, which introduce a number of steric effects and changes in chemistry that may not elicit RU1/2–AU1/1' allosterism.

Figure 80. Allosteric cross-talk between dimer and tetramer persists in peptide-metalloreagent conjugate treated NCP crystals. (a) Superposition of core histone octamers (tails excluded) between AUF-NCP (yellow) and LANA-NCP (magenta), or GAG-NCP (cyan) show slight rearrangements in backbone conformation with more pronounced changes in side chain positions. RMSD values are 0.262 and 0.340 Å respectively. (b) Peptide bound histone octamer conformations deviate slightly more relative to native (empty) NCP. RMSD values are 0.280 Å for LANA to xNCP145 and 0.346 Å for GAG to xNCP145.

Viral peptide-metalloreagent conjugates are a novel entry to the array of organometallic compounds our lab has studied. As focus shifts from mononuclear reagents to drug-drug combinations and multinuclear designs, any feature that can leverage experimental knowledge and biomimicry towards better therapeutic outcomes is a feasible candidate for research inquiry. The substituted LANA M6O is interesting because it presents a way of tuning chromatin targeting and association for drug delivery purposes. The GAG CBS peptide is also promising as it binds rapidly and tightly to nucleosome based on well-defined electron density and its accelerated impact on crystal stability.

However, both peptide–metalloreagent conjugates need further development concerning metal center reactivity as it relates to optimal linker design.

3.7 Concluding remarks and projected outcomes

Preliminary and final results presented in the above text, from crystallization experiments and structural adduct studies, exemplify promising opportunities in the development of novel NCP crystal systems and their application. These systems can be applied to chromatin-targeting metal-based anticancer drug discovery and development, NCP-factor bound structure crystallization and analysis, and elucidation of higher-order chromatin structures.

Recent studies have demonstrated, for the first time, crystallization of a nucleosome with a natural DNA sequence (Frouws, T.D. *et al.*, 2016). In the context of chromatin, this finding carries significance as it more closely represents a native *in vivo* nucleosome structure. Our intention moving forward is to expand on this result by continuing with the development of cohesive end NCP crystals as a blunt end replacement. By engineering DNA fragments designed to increase crystal integrity while more closely mimicking endogenous human sequences, we hope to crystallize increasingly heterogeneous nucleosome arrangements that will lead to greater understanding of the structural requirements for nucleosome positioning and expand our knowledge of overall chromatin structure, particularly as it pertains to *in vivo* states.

NCP147K stabilizes crystal contact points which improves lattice order. Moving forward, this could yield more detailed structural information at higher resolution which could help elucidate the biomolecular interactions of factor bound NCP complexes. Finally, by adding DNA sequence heterogeneity to the NCP147K crystal system we will be able to explore a larger pharmacological space, in terms of DNA targeting, with greater emphasis on the *in vivo* relevance of chromatin-targeting drug discovery. Intramolecular cross-linking of nucleosome dimer with tetramer by C3 increases the particle's stability in terms of DNA binding. This raises the potential to use small molecule organometallics as a tool for revealing further biologically active allosteric signaling mechanisms within chromatin and develop molecular tools for chromatin research.

Appendix 1

	NCP145	1.5 Å	1.072 Å	EP11-019	C3				Tri-nuclear CIII				LANA M60	Au-LANA-O				Au-GAG			
[Agent] (mM)	N/A	1	1	1	1	1	1	1	1	1	1	1	2	2	0.5	0.5		2	2	2	2
Duration (hrs)	N/A	47	47	46*	63	28	54	97	120*	28	54	96	120*	15	29	61	84	29	54	83	118
Histones	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>	<i>H. sapiens</i>
Data collection																					
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions																					
<i>a</i> (Å)	107.56	107.24	107.49	106.43	107.28	107.46	107.21	107.17	107.31	106.57	106.41	106.25	106.34	109.71	107.43	107.21	107.29	107.17	107.82	107.86	107.90
<i>b</i> (Å)	109.71	109.71	109.78	110.04	109.89	109.80	109.66	109.67	109.74	109.18	109.18	109.16	109.19	109.88	109.69	109.64	109.47	109.64	220.60	220.50	220.81
<i>c</i> (Å)	183.48	182.85	183.25	181.82	179.91	183.12	182.88	183.11	183.25	174.90	174.82	174.92	175.07	184.88	183.28	183.16	183.48	183.56	178.85	178.84	179.16
Resolution (Å)	1.99–76.81 (1.99–2.10)	2.51–48.84 (2.51–2.64)	2.30–48.89 (2.30–2.43)	2.75–48.88 (2.75–2.90)	3.45–48.90 (3.45–3.64)	2.15–48.89 (2.15–2.27)	2.14–48.82 (2.14–2.26)	2.35–48.82 (2.35–2.48)	2.25–48.85 (2.25–2.37)	2.25–48.59 (2.25–2.37)	2.40–48.57 (2.40–2.53)	2.20–48.55 (2.20–2.32)	2.19–48.57 (2.19–2.31)	2.50–49.13 (2.50–2.64)	2.19–48.85 (2.19–2.31)	2.11–48.81 (2.11–2.23)	2.10–48.85 (2.10–2.21)	2.17–48.80 (2.17–2.29)	2.95–49.10 (2.95–3.11)	3.28–49.09 (3.28–3.45)	3.22–48.47 (3.22–3.40)
<i>R</i> _{merge} (%)	6.8 (212)	3.9 (72.2)	4.7 (66.3)	4.4 (79.2)	4.4 (60.7)	3.7 (140)	5.9 (207)	5.3 (204)	5.1 (234)	4.7 (178)	4.8 (182)	4.3 (186)	4.7 (180)	5.8 (283)	5.1 (106)	6.2 (186)	5.3 (177)	6.4 (186)	9.0 (199)	8.3 (72.4)	11.1 (115)
<i>R</i> _{pim} (%)	2.1(65.2)	2.5 (57.8)	2.4 (42.0)	2.0 (34.8)	2.1 (27.5)	1.1 (59.8)	1.8 (69.5)	1.7 (73.5)	1.6 (76.3)	1.4 (79.3)	1.5 (75.8)	1.3 (67.0)	1.4 (66.1)	1.7 (84.4)	1.6 (34.3)	1.9 (67.8)	1.6 (57.6)	2.0 (64.2)	2.7 (58.4)	2.5 (21.3)	3.3 (33.6)
<i>I</i> / <i>σ</i>	17.4 (1.3)	24.6 (1.7)	18.2 (1.7)	25.8 (2.5)	21.2 (3.5)	33.3 (1.3)	20.4 (1.1)	27.1 (1.1)	25.8 (1.1)	26.1 (1.0)	28.9 (1.0)	27.1 (1.2)	24.8 (1.1)	23.0 (1.1)	25.3 (2.3)	20.6 (1.1)	23.6 (1.3)	20.3 (1.2)	16.8 (1.5)	19.2 (3.8)	13.3 (2.2)
Completeness (%)	99.6 (97.9)	99.1 (95.0)	97.6 (86.4)	99.6 (97.9)	99.6 (97.7)	94.4 (68.6)	97.7 (85.7)	99.4 (96.1)	99.7 (97.9)	98.1 (87.6)	96.6 (80.4)	99.5 (96.7)	98.7 (91.3)	99.8 (98.9)	99.7 (98.1)	97.9 (85.5)	99.7 (98.0)	98.2 (88.1)	99.6 (97.6)	99.2 (94.5)	99.6 (97.2)
CC (1/2) (%)	99.7 (52.6)	100 (61.7)	99.9 (77.8)	100 (79.2)	100 (92.3)	100 (56.4)	99.9 (51.8)	100 (56.5)	100 (52.6)	99.9 (58.9)	100 (60.2)	99.9 (70.7)	99.9 (65.7)	100 (66.7)	99.9 (83.4)	99.9 (51.8)	100 (58.6)	100 (57.3)	99.8 (53.3)	99.9 (82.6)	99.9 (68.3)
Redundancy	12.2 (11.6)	6.1 (3.8)	5.8 (3.5)	6.5 (6.3)	6.4 (6.3)	11.9 (6.9)	12.9 (10.1)	12.9 (11.2)	13.0 (11.0)	11.9 (6.6)	12.1 (7.4)	12.7 (9.2)	12.6 (8.9)	13.3 (12.7)	13.0 (11.1)	12.5 (8.9)	13.1 (11.2)	12.2 (9.7)	13.3 (13.2)	13.2 (13.1)	13.1 (13.1)
Refinement																					
Resolution (Å)	1.99–76.81			2.75–48.88					2.25–48.85				2.19–48.57	2.50–49.13							
No. reflections	145312			55820					100758				103482	73641							
<i>R</i> _{work} / <i>R</i> _{free} (%)	23.6 / 26.3			21.4 / 26.5					23.2 / 25.7				23.2 / 27.2	21.2 / 26.4							
No. atoms	12169			12128					12152				12152	12154							
Protein	6116			6116					6098				6098	6116							
DNA	5939			5939					5939				5939	5939							
Solvent	114			9					68				--	--							
Adducts	N/A			64					43				5	99							
<i>B</i> -factors (Å ²)	67.3			116.4					110.1				100.8	121.9							
Protein	53.4			74					70.1				73.9	88.0							
DNA	105.7			135.6					130.3				133.1	145.8							
Solvent	42.8			57.8					56.5				--	--							
Adduct	N/A			198.2					183.6				95.5	131.8							
R.m.s. deviations																					
Bond lengths (Å)	0.0079			0.0092					0.0086				0.0095	0.0089							
Bond angles (°)	1.4744			1.4892					1.392				1.4405	1.4759							

*Dataset used for refinement

Appendix 2

NCP145 α -satellite

-70 -60 -50 -40 -30 -20 -10 0 10 20 30 40 50 60 70
 ATCAATATCCACCTGCAGATACTACCAAAAGTGTATTTGGAAACTGCTCCATCAAAGGCATGTTCACTGAATCAGCTGAACATGCCTTTTGATGGAGCAGTTTCCAAATACACTTTTGGTAGTATCTGCAGGTGATATTGAT

601

-70 -60 -50 -40 -30 -20 -10 0 10 20 30 40 50 60 70
 ATCAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCCTTAAACGCACGTACGCGCTGTCCCCCGCGTTTAAACCGCCAAGGGGATTACTCCCTAGTCTCCAGGCACGTGTGAGATATATACATCGAT

601L

-70 -60 -50 -40 -30 -20 -10 0 10 20 30 40 50 60 70
 ATCACAAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCCTTAAACGCACGTACGCGAATCCTGACGTGCGTTTAAACCGCGTCTAGAGCTGTCTACGACCAATTGAGCGGCCCTCGGCACCGGGATTGTGAT

NCP147K

-70 -60 -50 -40 -30 -20 -10 0 10 20 30 40 50 60 70
 CATATATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCCTTAAACGCACGTACGCGCTGTCTACCGCGTTTAAACCGCCACTAGAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCATATATGGTAC

NCP147SK

-70 -60 -50 -40 -30 -20 -10 0 10 20 30 40 50 60 70
 CCAGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCCTTAAACGCACGTACGCGCTGTCTACCGCGTTTAAACCGCCACTAGAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATATCAGGTAC

NCP145SK

-70 -60 -50 -40 -30 -20 -10 0 10 20 30 40 50 60 70
 CATATATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCCTTAAACGCACGTACGCGCTGTCTACCGCGTTTAAACCGCCACTAGAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCATATAGGTAC

NCP147XK

-70 -60 -50 -40 -30 -20 -10 0 10 20 30 40 50 60 70
 TCGAGATATATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCCTTAAACGCACGTACGCGCTGTCTACCGCGTTTAAACCGCCACTAGAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCATATATGGTAC

Appendix 3

NCP147K duplex

```
1      10      20      30      40      50      60      70      80      90      100     110     120     130     140     147
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CATATATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCTACCGCGTTTTAACCGCCACTAGAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCATATATGGTAC
CATGGTATATAGGGCCACGGCTCCGGCGAGTTAACCAGCATCTGTCTGAGATCGTGGCGAATTTGCGTGCATGCGCGACAGATGGCGCAAAATTGGCGGTGATCTTCGCGAATGATCAGAGGTCCGTGCACACTCTGGCCGTATATAC
```

NCP147SK duplex

```
1      10      20      30      40      50      60      70      80      90      100     110     120     130     140     147
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CCAGCATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCTACCGCGTTTTAACCGCCACTAGAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCACATCAGGTAC
TCGAGGTCGTAGGGCCACGGCTCCGGCGAGTTAACCAGCATCTGTCTGAGATCGTGGCGAATTTGCGTGCATGCGCGACAGATGGCGCAAAATTGGCGGTGATCTTCGCGAATGATCAGAGGTCCGTGCACACTCTGGCCGTGTAGTC
```

NCP145SK duplex

```
1      10      20      30      40      50      60      70      80      90      100     110     120     130     140     145
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
CATATACCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCTACCGCGTTTTAACCGCCACTAGAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCATATAGGTAC
TCGAGTATATGGGCCACGGCTCCGGCGAGTTAACCAGCATCTGTCTGAGATCGTGGCGAATTTGCGTGCATGCGCGACAGATGGCGCAAAATTGGCGGTGATCTTCGCGAATGATCAGAGGTCCGTGCACACTCTGGCCGTATATAC
```

NCP147XK duplex

```
1      10      20      30      40      50      60      70      80      90      100     110     120     130     140     151
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
TCGAGATATATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCTCTAGCACCGCTTAAACGCACGTACGCGCTGTCTACCGCGTTTTAACCGCCACTAGAAGCGCTTACTAGTCTCCAGGCACGTGTGAGACCGGCATATATGGTAC
CTATATAGGGCCACGGCTCCGGCGAGTTAACCAGCATCTGTCTGAGATCGTGGCGAATTTGCGTGCATGCGCGACAGATGGCGCAAAATTGGCGGTGATCTTCGCGAATGATCAGAGGTCCGTGCACACTCTGGCCGTATATAC
1      10      20      30      40      50      60      70      80      90      100     110     120     130     140     143
```

Appendix 4

Core Histone Sequences

xH2A	SGRGKQGGKTRAKAKTRSSRAGLQFPVGRVHRLLRKGNYAERVGAGAPVYLAHVLEYLTAEILELAGNAARDNKKTRIIPRHLQLAVRNDEELNKLGRVTIAQGGVLPNIQSVLLPKKTESSKSAKSK
hH2A	SGRGKQGGKARAKAKTRSSRAGLQFPVGRVHRLLRKGNYSERVGAGAPVYLAHVLEYLTAEILELAGNAARDNKKTRIIPRHLQLAIRNDEELNKLGRVTIAQGGVLPNIQAVLLPKKTESHKAKGK
xH2B	PEPAKSAPAPKKGSKKAVTKTQKKDGKKRRKTRKESYAIYVYKVLKQVHPDTGISSKAMSIMNSFVNDVFERIAGEASRLAHYNKRSTITSREIQTAVRLLLPGELAKHAVSEGKAVTKYTSK
hH2B	PEPAKSAPAPKKGSKKAVTKAQKKDGKKRRSRKESYSVYVYKVLKQVHPDTGISSKAMGIMNSFVNDIFERIAGEASRLAHYNKRSTITSREIQTAVRLLLPGELAKHAVSEGKAVTKYTSK
xH3	ARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKSTELLIRKLPFQRLVREIAQDFKTDLRFSQSAVMALQEASEAYLVLFEDTNLCAIHAKRVTIMPKDIQLARRIRGERA
hH3	ARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKSTELLIRKLPFQRLVREIAQDFKTDLRFSQSAVMALQEACEAYLVGLFEDTNLCAIHAKRVTIMPKDIQLARRIRGERA
xH4	SGRGKGGKGLGKGGAHRKRVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGLVKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRPLYGFGG
hH4	SGRGKGGKGLGKGGAHRKRVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGLVKVFLENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRPLYGFGG

Linker Histone Sequences

hH1.2	MSETAPAAPAAAPAEKAPVKKKAACKAGGTTPRKASGPPVSELITKAVAASKERSGVSLAALKKALAAAGYDVEKNNSRIKLGKSLVSKGTLVQTKGTGASGSFKLNKKAASGEAKP KVKKAGGTGPKKPVGAACKPKKAAGGATPKKSACKTPKKAKKPAATVTKKVAKSPKKAKVAKPKKAASAAKAVKPKAAKPKVVKPKKAAPKKK
hH1.4	MSETAPAAPAPAEKTPVKKKARKSAGAAKRKASGPPVSELITKAVAASKERSGVSLAALKKALAAAGYDVEKNNSRIKLGKSLVSKGTLVQTKGTGASGSFKLNKKAASGEAKP KAKKAGAAKAKKPAAGAAKPKKATGAATPKKSACKTPKKAKKPAAGAAKAKSPKKAKAAKPKKAPKSPAKAKAVKPKAAKPKTAKPKAAKPKKAAAKK
hH1t	MSETVPAASASAGVAAMEKLPTKKRGRKPAGLISASRKVPNLSVSKLITEALSVSQERVGMSLVALKKALAAAGYDVEKNNSRIKLSLKS LVNKGILVQTRGTGASGSFKLSKKVIPKST RSKAKKSVSATKKLVLSDSKSPKTAKTNKRAKKPRATTPKTVRSGRKAGKAGKQKQKSPVKARASKSLTQHHEVNVRKATSKK
hH1T2	MEQALTGEAQSRRPWRGGSGAMAEAPGPSGESRGHSATQLPAEKTGGPSRGCSVSLRVSQLVLAISTHKGTLAALKKELRNAGYEVRRKSGRHEAPRGQAKATLLRVSGSDAAG YFRVWKVPKPRRKPGRARQEEGTRAPWRTPAAPRSSRRRRQLRKAARKAREVWRRNARAKAKANARARRRRARPRAKEPPCARAKEEAGATAADEGRGQAVKEDTTPRSGDKRRS SKPREEKQEPKKPAQRITQ
hH1oo	MAPGSVTSDISPSTSTAGSSRSPSEKPGPSHGVPVGGPSHSSLPVGRRHPPVLRMVLEALQAGEQRRGTSVAAIKLYILHKYPTVDVLRFKYLLKQALATGMRRGLLARPLNS KARGATGSFKLVPHKKKKIQPRKMAPATAPRRAGEAKGKPKKPS EAKEDPPNVGKVKKAAKRPKVQKPPPKPGAATEKARKQGGAAKDTRAQSGEARKVPPKPKDKAMRAPSSAGGLS RKAKAKGSRSSQGDAAEYRKTKAESKSSKPTASKVKNGAASPTKKKVVAKAKAPKAGQGPNTKAAAPAKGSGSKVVPVPHLSRKTEAPKGPRAKGLPIKASSSKVSSQRAEA
hH1s1	MLHASTIWHLRSTPPRRKQWGHCDPHRILVASEVTTEITSPTAPRAQVCGGQPVWTVLDPLSGHTGREAEERHFATVSIASVELKYCHGWRPAGQRVPSKTATGQRTCAKPCQKPST SKVILRAVADKGTCKYVSLATLKKAVSTTGDMARNAYHFKRVLKGLVDKGSAGSFTLGKKQASKSKLVKVRQRQRWRSGQRPFQQRSLLGSKQGHKRLIKGVRRVAKCHCN

Appendix 5

Appendix 6

Viral Peptides

LANA ¹MAPPGMRLRSGRSTGAPLTRGSC²³
LANA-M ⁴PGMRLRSGRSTGAP¹⁷
LANA-O ⁴PGORLRSGRSTGAP¹⁷
GAG ⁵³⁵GGYNLRPTYQPQRYGG⁵⁵¹

HMGN Peptides

hHMGN1 ¹³KEEPKRRSARLSAKP-PAKVEAKPKKAAAK⁴²
hHMGN2 ¹⁷KDEPQRRSARLSAKPAPPKPEPKPKKAPAK⁴⁷

HMGN Full Length

hHMGN1 MPKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAKGKQAEVANQETKEDLPAENGETKTEESPASDEAGEKEAKSD
hHMGN2 MPKRKAEGDAKGDKAKVKDEPQRRSARLSAKPAPPKPEPKPKKAPAKKGEKVPKGKKGKADAGKEGNNPAENGDAKTDQAQKAEGAGDAK

References

- Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B, Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., and Zwart., P.H. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst.* **D66**, 213-221.
- Adhireksan, Z., Palermo, G., Riedel, T., Ma1, Z., Muhammad, R., Rothlisberger, U., Dyson, P. J., and Davey, C. A. 2017. Allosteric cross-talk in chromatin can mediate drug-drug synergy. *Nat. Commun.* **8**:14860.
- Adhireksan, Z., Davey, G.E., Campomanes, P., Groessl, M., Clavel, C.M., Yu, H., Nazarov, A.A., Yeo, C.H., Ang, W.H., Dröge, P., Rothlisberger, U., Dyson, P.J., and Davey, C.A. 2014. Ligand substitutions between ruthenium-cymene compounds can control protein versus DNA targeting and anticancer activity. *Nat. Commun.* **5**:3462 - 3475.
- Alagesan, M., Bhuvanesh, N.S., and Dharmaraj, N. 2014. An investigation on new ruthenium(II) hydrazone complexes as anticancer agents and their interaction with biomolecules. *Dalton Trans.* **43(16)**:6087 - 6099.
- Albertella, M. R., Green, C. M., Lehmann, A. R., and O'Connor, M. J. 2005. A role for polymerase eta in the cellular tolerance to cisplatin-induced damage. *Cancer Res.* **65**:9799 - 9806.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. 2002. *Molecular Biology of the Cell*. New York, NY: Garland Science.
- Aleksenko, S.S., Matczuk, M., Lu, X., Foteeva, L.S., Pawlak, K., Timerbaev, A.R., and Jarosz, M. 2013. Metallomics for drug development: an integrated CE-ICP-MS and ICP-MS approach reveals the speciation changes for an investigational ruthenium(III) drug bound to holo-transferrin in simulated cancer cytosol. *Metallomics.* **5(8)**:955 - 963.
- Allan, J., Hartman, P.G., Crane-Robinson, C., and Aviles, F.X. 1980. The structure of histone H1 and its location in chromatin. *Nature.* **288(5792)**:675 – 679.
- Allardyce, C. S., Dyson, P. J., Ellis, D. J., and Heath, S. L. 2001. [Ru(eta(6)-p-cymene) Cl-2(pta)] (pta $\frac{1}{4}$ 1,3,5-triaza-7-phosphatricyclo[3.3.1.1]decane): a water soluble compound that exhibits pH dependent DNA binding providing selectivity for diseased cells. *Chem. Commun. (Camb).* 1396–1397.
- Ang, W.H., Daldini, E., Juillerat-Jeanneret, L., and Dyson, P.J. 2007. Strategy to tether organometallic ruthenium-arene anticancer compounds to recombinant human serum albumin. *Inorg. Chem.* **46(22)**:9048 - 9050.
- Arents, G. Burlingame, R.W., Wang, B.C., Love, W.E. and Moudrianakis, E.N. 1991. The nucleosome core histone octamer at 3.1Å resolution: a tripartite protein assembly and left-handed superhelix. *Proc. Natl. Acad. Sci. U.S.A.* **88**:10148 – 10152.

- Arents G., and Moudrianakis, E.N. 1993. Topography and the histone octamer surface: repeating structural motifs utilized in the docking of nucleosomal DNA. *Proc. Natl. Acad. Sci. U.S.A.* **90**:10489 – 10493.
- Armache, K.J., Garlick, J.D., Narlikar, G.J., and Kingston, R.E. 2011. Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. *Science*. **334(6058)**:977 – 982.
- Ayers, P.W., Parr, R.G., and Pearson, R.G. 2006. Elucidating the hard/soft acid/base principle: a perspective based on half-reactions. *J. Chem. Phys.* **124(19)**:194107.
- Ballestas, M.E., Chatis, P.A., and Kaye, K.M. 1999. Efficient persistence of extrachromosomal KSHV DNA mediated by latency-associated nuclear antigen. *Science*. **284(5414)**:641 – 644.
- Bao, Y., White, C.L., and Luger, K. 2006. Nucleosome core particles containing a poly(dA.dT) sequence element exhibit a locally distorted DNA structure. *J. Mol. Biol.* **361**:617 – 624.
- Barbera, A.J., Ballestas, M.E., and Kaye, K.M. 2004. The Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen 1 N terminus is essential for chromosome association, DNA replication, and episome persistence. *J. Virol.* **78(1)**:294 – 301.
- Barbera, A.J., Chodaparambil, J.V., Kelley-Clarke, B., Joukov, V., Walter, J.C., Luger, K., and Kaye, K.M. 2006a. The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science*. **311(5762)**:856 – 861.
- Barbera, A.J., Chodaparambil, J.V., Kelley-Clarke, B., Luger, K., and Kaye, K.M. 2006b. Kaposi's sarcoma-associated herpesvirus LANA hitches a ride on the chromosome. *Cell Cycle*. **5(10)**:1048 – 1052.
- Barreiro, E., Casas, J.S., Couce, M.D., Sánchez, A., Sánchez-Gonzalez, A., Sordo, J., and Vázquez-López, E.M. 2014. Mono and dinuclear phosphinegold(I) sulfanylcarboxylates: Influence of nuclearity and substitution of PPh₃ for PET₃ on cytotoxicity. *J. Inorg. Biochem.* **138**:89 – 98.
- Bartel, C., Egger, A.E., Jakupec, M.A., Heffeter, P., Galanski, M., Berger, W., and Keppler, B.K. 2011. Influence of ascorbic acid on the activity of the investigational anticancer drug KP1019. *J. Biol. Inorg. Chem.* **16(8)**:1205 – 1215.
- Bassett, E., Vaisman, A., Tropea, K. A., McCall, C. M., Masutani, C., Hanaoka, F., and Chaney, S. G. 2002. Frameshifts and deletions during in vitro translesion synthesis past Pt-DNA adducts by DNA polymerases beta and eta. *DNA Repair (Amst)*. **1**:1003 - 1016.
- Batchelor, L.K., Păunescu, E., Soudani, M., Scopelliti, R., and Dyson, P.J. 2017. Influence of the linker length on the cytotoxicity of homobinuclear ruthenium(II) and gold (I) complexes. *Inorg Chem.* **21;56(16)**:9617-9633.

- Battistini, F., Hunter, C.A., Gardiner, E.J., and Packer, M.J. 2010. Structural mechanics of DNA wrapping in the nucleosome. *J. Mol. Biol.* **396(2)**:264 – 279.
- Becker, P.B. 1995. *Drosophila* chromatin and transcription. *Semin. Cell Biol.* **6(4)**:185 - 190.
- Biondi, O., Motta, S., and Mosesso, P. 2002. Low molecular weight polyethylene glycol induces chromosome aberrations in Chinese hamster cells cultured in vitro. *Mutagenesis*. **17(3)**:261 – 264.
- Bergamo, A. and Sava, G. 2011. Ruthenium anticancer compounds: myths and realities of the emerging metal-based drugs. *Dalton Trans.* **40(31)**:7817 – 7823.
- Berners-Price, S.J., DiMartino, M.J., Hill, D.T., Kuroda, R., Mazid, M.A., and Sadler, P.J. 1985. Tertiary phosphine complexes of gold(I) and gold(III) with imido ligands: ¹H, ³¹P, and ¹⁵N NMR spectroscopy, antiinflammatory activity, and x-ray crystal structure of (phthalimido)(triethylphosphine)gold(I). *Inorg. Chem.* **24(21)**:3425 – 3434.
- Blank, T.A. and Becker, P.B. 1995. Electrostatic mechanism of nucleosome spacing. *J. Mol. Biol.* **252**:305 – 313.
- Blossey, R. and Schiessel, H. 2011. The dynamics of the nucleosome: thermal effects, external forces, and ATP. *FEBS J.* **278(19)**:3619 – 3632.
- Bonisch, C. and Hake, S.B. 2012. Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Res.* **40(21)**:10719 – 10741.
- Botchan, M., Kram, R., Schmid, C.W., and Hearst, J.E. 1971. Isolation and chromosomal localization of highly repeated DNA sequences in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* **68(6)**:1125 – 1129.
- Bovio, B., Bonati, F., Burini, A., and Pietroni, B.R. 1984. Linear gold(I) derivatives with bulky phosphine and imidazolato ligands. X-ray crystal structure of 1-(C₆H₁₁)₃Pau-2-i-Pr-Imidazole · C₆H₆. *Zeitschrift für Naturforschung B.* **39(12)**:1747 – 1754.
- Brabec, V. and Novakova, O. 2006. DNA binding mode of ruthenium complexes and relationship to tumor cell toxicity. *Drug Resist. Updat.* **9(3)**:111 – 122.
- Brujininx, P.C. and Sadler, P.J. 2008. New trends for metal complexes with anticancer activity. *Curr. Opin. Chem. Biol.* **12(2)**:197 - 206.
- Burton, D.R., Butler, M.J., Hyde, J.E., Phillips, D., Skidmore, C.J., and Walker, I.O. 1978. The interaction of core histones with DNA: equilibrium binding studies. *Nucleic Acids Res.* **5**:3643 – 3663.
- Chen, J.W., Wang, X.Y., Shao, Y., Zhu, J.H., Zhu, Y.G., Li, Y.Z., Xu, Q., and Guo, Z.J. 2007. A trinuclear copper(II) complex of 2,4,6-tris(di-2-pyridylamine)-1,3,5 triazine shows prominent DNA cleavage activity. *Inorg. Chem.* **46(8)**:3306 – 3312.

- Chien, F.T. and van der Heijden, T. 2014. Characterization of nucleosome unwrapping within chromatin fibers using magnetic tweezers. *Biophys. J.* **107**(2):373 – 383.
- Chua, E.Y, Vasudevan, D., Davey, G.E., Wu, B., and Davey, C.A. 2012. The mechanics behind DNA sequence-dependent properties of the nucleosome. *Nucleic Acids Res.* **40**(13):6338 - 6352.
- Chua, E.Y., Davey, G.E., Chin, C.F., Dröge, P., Ang, W.H., and Davey, C.A. 2015. Stereochemical control of nucleosome targeting by platinum-intercalator antitumor agents. *Nucleic Acids Res.* **43**(11):5284 - 5296.
- Chung, S., Parker, J.B., Bianchet, M., Amzel, L.M., and Stivers, J.T. 2009. Impact of linker strain and flexibility in the design of a fragment-based inhibitor. *Nat. Chem. Biol.* **5**(6):407 – 413.
- Collins, N., Poot, R.A., Kukimoto, I., Garcia-Jimenez, C., Dellaire, G., and Varga-Weisz, P.D. 2002. *Nat. Genet.* **32**(4):627 – 632.
- Clapier, C. R. and Cairns, B. R. 2009. The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* **78**:273 - 304.
- Cutter, A.R. and Hayes, J.J. 2015. A brief review of nucleosome structure. *FEBS Lett.* **589**(20 Pt A):2914 – 2922.
- Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W., and Richmond, T. J. 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. *J. Mol. Biol.* **319**:1097 - 1113.
- Davey, C. A. 2015. Exposure to metals can be therapeutic. *Chimia.* **69**:125–130.
- Davey, G.E., Wu, B., Dong, Y., Surana, U., and Davey, C.A. 2010. DNA stretching in the nucleosome facilitates alkylation by an intercalating antitumour agent. *Nucleic Acids Res.* **38**(6):2081 – 2088.
- Davey, G.E., Adhireksan, Z., Ma, Z., Riedel, T., Padavattan, S., Rhodes, D., Ludwig, A., Sandin, S., Murray, B.S., Dyson, P.J., and Davey, C.A. 2017. Nucleosome acidic patch-targeting binuclear ruthenium compounds induce aberrant chromatin condensation. *Nat. Commun.* **8**(1):1575.
- Davies, M. S., Berners-Price, S. J., and Hambley, T. W. 2000. Rates of platination of -AG- and -GA- containing double-stranded oligonucleotides: effect of chloride concentration. *J. Inorg. Biochem.* **79**:167 - 172.
- DeLano, W. L. 2002. The PyMOL Molecular Graphics System.
- Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R. R., and Richmond, T. J. 2004. Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science.* **306**:1571 - 1573.

- Dyer, P.N., Edayathumangalam, R.S., White, C.L., Bao, Y., Chakravarthy, S., Muthurajan, U.M., Luger, K. 2004. Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* **375**:23 – 44.
- Emsley, P. and Cowtan. K. 2004. Coot: model-building tools for molecular graphics. *Acta. Crystallogr. D. Biol. Crystallogr.* **60(Pt 12 Pt 1)**:2126 - 2132.
- Ernst, J., Kheradpour, P., Mikkelsen, T.S., Shores, N., Ward, L.D., Epstein, C.B., Zhang, X., Wang, L., Issner, R., Coyne, M., Ku, M., Durham, T., Kellis, M., and Bernstein, B.E. 2011. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature.* **473(7345)**:43 - 49.
- Fan, C., Zheng, W., Fu, X., Li, X., Wong, Y.S., and Chen, T. 2014. Enhancement of auranofin-induced lung cancer cell apoptosis by selenocystine, a natural inhibitor of TrxR1 in vitro and in vivo. *Cell Death Dis.* **5**:e1191.
- Fang, Q., Chen, P., Wang, M., Fang, J., Yang, N., Li, G., and Xu, R.M. 2016. Human cytomegalovirus IE1 protein alters the higher-order chromatin structure by targeting the acidic patch of the nucleosome. *Elife.* doi:10.7554/eLife.11911.
- Felsenfeld, G. and Groudine, M. 2003. Controlling the double helix. *Nature.* **421**:448-53.
- Fierz, B. 2015. Dynamic Chromatin Regulation from a Single Molecule Perspective. *ACS Chem. Biol.* doi: 10.1021/acscchembio.5b00832
- Fichtinger-Schepman, A. M., van der Veer, J. L., den Hartog, J. H., Lohman, P. H., and Reedijk, J. 1985. Adducts of the antitumor drug cis-diamminedichloroplatinum(II) with DNA: formation, identification, and quantitation. *Biochemistry.* **24**:707 - 713.
- Fink, D., Nebel, S., Aebi, S., Zheng, H., Cenni, B., Nehme, A., Christen, R. D., and Howell, S. B. 1996. The role of DNA mismatch repair in platinum drug resistance. *Cancer Res.* **56**:4881 - 4886.
- Freeman, L., Kurumizaka, H., and Wolffe, A.P. 1996. Functional domains for assembly of histones H3 and H4 into the chromatin of *Xenopus* embryos. *Proc. Natl. Acad. Sci. U.S.A.* **93(23)**:12780 – 12785.
- Frouws, T.D., Duda, S.C., and Richmond, T.J. 2016. X-ray structure of the MMTV-A nucleosome core. *Proc. Natl. Acad. Sci. U.S.A.* **113(5)**:1214 - 1219.
- Gadducci, A., Cosio, S., Muraca, S., and Genazzani, A. R. 2002. Molecular mechanisms of apoptosis and chemosensitivity to platinum and paclitaxel in ovarian cancer: biological data and clinical implications. *Eur. J. Gynaecol. Oncol.* **23**:390 - 396.
- Gall, J.G., Cohen, E.H., and Polan, M.L. 1971. Repetitive DNA sequences in drosophila. *Chromosoma.* **33(3)**:319 – 344.
- Galloway, S.M., Deasy, D.A., Bean, C.L., Kravayak, A.R., Armstrong, M.J., and Bradley, M.O. 1987. Effects of high osmotic strength on chromosome

- aberrations, sister-chromatid exchanges and DNA strand breaks, and the relation to toxicity. *Mutat. Res.* **189(1)**:15 – 25.
- Gaspar-Maia, A., Alajem, A., Meshorer, E., and Ramalho-Santos, M. 2011. Open chromatin in pluripotency and reprogramming. *Nat. Rev. Mol. Cell Biol.* **12(1)**:36 – 47.
- Gasser, G., Ott, I., and Metzler-Nolte, N. 2011. Organometallic anticancer compounds. *J. Med. Chem.* **54(1)**:3 - 25.
- Gerbi, S.A. 2018. Chromocenters: Bundling up DNA. *Elife.* 7. pii: e37234. doi: 10.7554/eLife.37234.
- Gianferrara, T., Bratsos, I., and Alessio, E. 2009. A categorization of metal anticancer compounds based on their mode of action. *Dalton Trans.* **37**:7588 – 7598.
- Gottesfeld, J.M. and Melton, D.A. 1978. The length of nucleosome-associated DNA is the same in both transcribed and nontranscribed regions of chromatin. *Nature.* **273(566)**:317 – 319.
- Grigoryev, S.A. 2018. Chromatin higher-order folding: A perspective with linker DNA angles. *Biophys. J.* **114(10)**:2290 – 2297.
- Guenatri, M., Bailly, D., Maison, C., and Almouzni, G. 2004. Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J. Cell Biol.* **166(4)**:493 – 505.
- Gupta, R., Taguchi, T., Lassalle-Kaiser, B., Bominaar, E.L., Yano, J., Hendrich, M.P., and Borovik, A.S. 2014. High-spin Mn-oxo complexes and their relevance to the oxygen-evolving complex within photosystem II. *Proc. Natl. Acad. Sci. U.S.A.* **112(17)**:5319 – 5324.
- Hansen, J.C. 2002. Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. *Annu. Rev. Biophys Biomol. Struct.* **31**:361 - 392.
- Harrap, K. R. 1985. Preclinical studies identifying carboplatin as a viable cisplatin alternative. *Cancer Treat. Rev.* **12 Suppl A**:21 - 33.
- Harris, A.L., Yang, X., Hegmans, A., Povrik, L., Ryan, J.J., Kelland, L., and Farrell, N.P. 2005. Synthesis, characterization, and cytotoxicity of a novel highly charged trinuclear platinum compound. Enhancement of cellular uptake with charge. *Inorg. Chem.* **44(26)**:9598 – 9600.
- Hartinger, C.G., Jakupec, M.A., Zorbas-Seifried, S., Groessi, M., Egger, A., Berger, W., Zorbas, H., Dyson, P.J., and Keppler, B.K. 2008. KP1019, a new redox-active anticancer agent--preclinical development and results of a clinical phase I study in tumor patients. *Chem Biodivers.* **5(10)**:2140-2155.
- Hartinger, C.G., Philips, A.D., and Nazarov, A.A. 2011. Polynuclear ruthenium, osmium, and gold complexes. The quest for innovative anticancer chemotherapeutics. *Curr. Top. Med. Chem.* **11(21)**:2688-2702.

- Hayes, J.J., Tullius, T.D. and Wolffe, A.P. 1990. The structure of DNA in a nucleosome. *Proc. Natl. Acad. Sci. U.S.A.* **87**:7405 – 7409.
- Hayes, J.J., Clark, D.J., and Wolffe, A.P. 1991. Histone contributions to the structure of DNA in the nucleosome. *Proc. Natl. Acad. Sci. U.S.A.* **88**:6829 – 6833.
- Hayes, J.J., Bashkin, J., Tullius, T.D., and Wolffe, A.P. 1991. The histone core exerts a dominant constraint on the structure of DNA in the nucleosome. *Biochemistry.* **30**:8434 – 8440.
- Hensing, T.A., Hanna, N.H., Gillenwater, H.H., Camboni, M.G., Allievi, C., and Sociniski, M.A. 2006. Phase II study of BBR 3464 as treatment in patients with sensitive or refractory small cell lung cancer. *Anticancer Drugs.* **17(6)**:697 – 704.
- Hergeth, S.P. and Schneider, R. 2015. The H1 linker histones: multifunctional proteins beyond the nucleosome core particle. *EMBO Rep.* **16(11)**:1439 – 1453.
- Herman A., Tanski, J.M., Tibbetts, M.F., and Anderson, C.M. 2008. Synthesis, characterization, and in vitro evaluation of a potentially selective anticancer, mixed-metal [ruthenium(III)-platinum(II)] trinuclear complex. *Inorg. Chem.* **47(1)**:247 – 280.
- Higgins S.L., White, T.A., Winkel, B.S., and Brewer, K.J. 2011. Redox, spectroscopic, and photophysical properties of RU-Pt mixed-metal complexes incorporating 4,7-diphenyl-1,10-phenanthroline as efficient DNA binding and photocleaving agents. *Inorg. Chem.* **50(2)**:463 – 470.
- Hill, D. T. and Sutton, B. M. 1980. (2, 3, 4, 6-TETRA-O-ACETYL-1-THIO-BETA-D-GLUCOPYRANOSATO-S)-(TRIETHYLPHOSPHINE) GOLD, C₂₀H₃₄AUO₉PS. *Cryst. Struct. Commun.* **9**:679.
- Hoggard, P.E. and Porter, G.B. 1981. Is ruthenium(II) hard or soft? *J. Inorg. Nucl. Chem.* **43(1)**:185 – 186.
- Holzer, A. K. and Howell, S. B. 2006. The internalization and degradation of human copper transporter 1 following cisplatin exposure. *Cancer Res.* **66**:10944 - 10952.
- Holzer, A. K., Manorek, G. H., and Howell, S. B. 2006. Contribution of the major copper influx transporter CTR1 to the cellular accumulation of cisplatin, carboplatin, and oxaliplatin. *Mol. Pharmacol.* **70**:1390 - 1394.
- Houten, B.V., Illenye, S., Qu, Y., and Farrell, N. 1993. Homodinuclear (Pt,Pt) and heterodinuclear (Ru,Pt) metal compounds as DNA–protein cross-linking agents: potential suicide DNA lesions. *Biochemistry.* **32(44)**:11794 – 11801.
- Ichikawa, Y., Nishimura, Y., Kurumizaka, H., and Shimizu, M. 2015. Nucleosome organization and chromatin dynamics in telomeres. *Biomol. Concepts.* **6(1)**:67 – 75.

- Igo-Kemenes, T., Omori, A., and Zachau, H.G. 1980. Non-random arrangement of nucleosomes in satellite I containing chromatin of rat liver. *Nucleic Acid Res.* **8(22)**:5377 – 5390.
- Israelachvili, J. 1997. The different faces of poly(ethylene glycol). *Proc. Natl. Acad. Sci. U.S.A.* **94(16)**:8378 – 8379.
- Jodrell, D.I., Evan, T.R., Steward, W., Cameron, D., Prendiville, J., Aschele, C., Noberasco, C., Lind, M., Carmichael, J., Dobbs, N., Camboni, G., Gatti, B., and DeBraud, F. 2004. Phase II studies of BBR3464, a novel tri-nuclear platinum complex, in patients with gastric or gastro-oesophageal adenocarcinoma. *Eur. J. Cancer.* **40(12)**:1872 – 1877.
- Joti, Y., Hikima, T., Nishino, Y., Kamada, F., Hihara, S., Takata, H., Ishikawa, T., and Maeshima, K. 2012. Chromosomes without a 30-nm chromatin fiber. *Nucleus.* **3(5)**:404 – 410.
- Kalashnikova, A.A., Porter-Goff, M.E., Muthurajan, U.M., Luger, K., and Hansen, J.C. 2013. The role of the nucleosome acidic patch in modulating higher order chromatin structure. *J. R. Soc. Interface.* **10(82)**.
- Kaochar, S. and Tu, B.P. 2012. Gatekeepers of chromatin: Small metabolites elicit big changes in gene expression. *Trends Biochem Sci.* **37(11)**:477 - 483.
- Kelland, L. R., Abel, G., McKeage, M. J., Jones, M., Goddard, P. M., Valenti, M., Murrer, B. A., and Harrap, K. R. 1993. Preclinical antitumor evaluation of bis-acetato-ammine-dichloro-cyclohexylamine platinum(IV): an orally active platinum drug. *Cancer Res.* **53**:2581 - 2586.
- Kelland, L. 2007. The resurgence of platinum-based cancer chemotherapy. *Nat. Rev. Cancer.* **7**:573 - 584.
- Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., Linder-Basso, D., Plachetka, A., Shanower, G., Tolstorukov, M.Y., Luquette, L.J., Xi, R., Jung, Y.L., Park, R.W., Bishop, E.P., Canfield, T.K., Sandstrom, R., Thurman, R.E., MacAlpine, D.M., Stamatoyannopoulos, J.A., Kellis, M., Elgin, S.C., Kuroda, M.I., Pirrotta, V., Karpen, G.H., and Park, P.J. 2011. Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature.* **471(7339)**:480 - 485.
- Kidani, Y., Inagaki, K., Iigo, M., Hoshi, A., and Kuretani, K. 1978. Antitumor activity of 1,2-diaminocyclohexane--platinum complexes against sarcoma-180 ascites form. *J. Med. Chem.* **21**:1315 - 1318.
- Knox, R. J., Friedlos, F., Lydall, D. A., and Roberts, J. J. 1986. Mechanism of cytotoxicity of anticancer platinum drugs: evidence that cis-diamminedichloroplatinum(II) and cis-diammine-(1,1-cyclobutanedicarboxylato)platinum(II) differ only in the kinetics of their interaction with DNA. *Cancer Res.* **46**:1972 - 1979.

- Komeda, S., Moulaei, T., Woods, K.K., Chikuma, M., Farrell, N.P., and Williams, L.D. 2006. A third mode of DNA binding: Phosphate clamps by a polynuclear platinum complex. *J. Am. Chem. Soc.* **128(5)**:16092 – 16103.
- Kornberg, R. D., 1974. Chromatin structure: a repeating unit of histones and DNA. *Science*. **184**:868-71.
- Kornberg, R.D. and Lorch, Y. 1991. Irresistible force meets immovable object: transcription and the nucleosome. *Cell*. **67(5)**:833 - 836.
- Kostova, I. 2006. Ruthenium complexes as anticancer agents. *Curr. Med. Chem.* **13(9)**:1085-1107.
- Krueger, A., Protozonva, E., and Frank-Kamenetskii, M.D. 2006. Sequence-dependent base pair opening in DNA double helix. *Biophys. J.* **90(9)**:3091 – 3099.
- Leach, D.R. 1994. Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. *Bioessays*. **16(12)**:893 – 900.
- Lesbats, P., Engelman, A.N., and Cherepanov, P. 2016. Retroviral DNA integration. *Chem. Rev.* **116(20)**:12730 – 12757.
- Lesbats, P., Serrao, E., Maskell, D.P., Pye, V.E., O'Reilly, N., Lindemann, D., Engelman, A.N. and Cherepanov, P. 2017. Structural basis for spumavirus GAG tethering to chromatin. *Proc. Natl. Acad. Sci. U.S.A.* **114(21)**:5509 – 5514.
- Leslie, A.G. 2006. The integration of macromolecular diffraction data. *Acta Crystallogr. D. Biol. Crystallogr.* **62(Pt 1)**:48 - 57.
- Levina, A., Mitra, A., and Lay P.A. 2009. Recent developments in ruthenium anticancer drugs. *Metallomics*. **1(6)**:458 – 470.
- Li, G. and Widom, J. 2004. Nucleosomes facilitate their own invasion. *Nat. Struct. Mol. Biol.* **11(8)**:763 - 769.
- Li, T., Stark, M.R., Johnson, A.D., and Wolberger, C. 1995. Crystal structure of the MATa1/MAT alpha 2 homeodomain heterodimer bound to DNA. *Science*. **270(5234)**:262 – 269.
- Lin, J.C., Jeong, S., Liang, G., Takai, D., Fatemi, M., Tsai, Y.C., Egger, G., Gal-Yam, E.N., and Jones, P.A. 2007. Role of nucleosomal occupancy in the epigenetic silencing of the MLH1 CpG island. *Cancer Cell*. **12(5)**:432 – 444.
- Lowary, P.T. and Widom, J. 1998. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276(1)**:19 - 42.
- Luger, K., Mader, A.W., Richmond, R.K. Sargent, D.F., and Richmond T.J. 1997a. Crystal structure of the nucleosome core particle at 2.8Å resolution. *Nature*. **389**:251 – 260.

- Luger, K., Rechsteiner, T.J., Flaus, A.J., Wayne, M.M., and Richmond, T.J. 1997b. Characterization of nucleosome core particles containing histone proteins made in bacteria. *J. Mol. Biol.* **272**:301 – 311.
- Luque, A., Ozer, G., and Schlick, T. 2016. Correlation among DNA linker length, linker histone concentration, and histone tails in chromatin. *Biophys. J.* **110**(11):2309 – 2319.
- Luzyanin, K.V., Kukushkin, V.Y., Ryabov, A.D., Haukka, M., and Pombeiro, A.J. 2005. Kinetic and mechanistic study of the Pt(II) versus Pt(IV) effect in the platinum-mediated nitrile-hydroxylamine coupling. *Inorg. Chem.* **44**(8):2944-2953.
- Makde, R.D., England, J.R., Yennawar, H.P., and Tan, S. 2010. Structure of RCC1 chromatin factor bound to the nucleosome core particle. *Nature.* **467**(7315):562-6.
- Malagon, F. and Aguilera, A. 1988. Genetic stability and DNA rearrangements associated with a 2 x 1.1-Kb perfect palindrome in *Escherichia coli*. *Mol. Gen. Genet.* **259**(6):639 – 644.
- Manzotti, C., Pratesi, G., Menta, E., Di Domenico, R., Cavalletti, E., Fiebig, H.H., Kelland, L.R., Farrell, N., Polizzi, D., Supino, R., Pezzoni, G., and Zunino, F. 2000. BBR 3464: a novel triplatinum complex, exhibiting a preclinical profile of antitumor efficacy different from cisplatin. *Clin. Cancer Res.* **6**(7):2626 – 2634.
- Markowska, A., Kasprzak, B., Jaszczynska-Nowinka, K., Lubin, J., and Marjowska, J. 2015. Noble metals in oncology. *Contemporary Oncology* **19**(4):271 – 275.
- Matsuda, A., Chikashige, Y., Ding, D.Q., Ohtsuki, C., Mori, C., Asakawa, H., Kimura, H., Haraguchi, T., and Hiraoka, Y. 2015. *Nat. Commun.* **6**:7753.
- Mattioli, F., Bhattacharyya, S., Dyer, P.N., White, A.E., Sandman, K., Burkhart, B.W., Byrne, K.R., Lee, T., Ahn, N.G., Santangelo, T.J., Reeve, J.N., and Luger, K. 2017. Structure of histone-based chromatin in Archaea. *Science.* **357**(6351):609 – 612.
- Maze, I., Noh, K.M., Soshnev, A.A., and Allis, C.D. 2014. Every amino acid matters: essential contributions of histone variants to mammalian development and disease. *Nat. Rev. Genet.* **15**:259 – 271.
- McGinty, R.K. and Tan, S. 2014. Nucleosome structure and function. *Chem. Rev.* **115**:2255 – 2273.
- McGregor, T.D., Hegmans, A., Kaspárková, J., Nepelchová, K., Nováková, O., Penázová, H., Vrána, O., Brabec, V., and Farrell, N. 2002. A comparison of DNA binding profiles of dinuclear platinum compounds with polyamine linkers and the trinuclear platinum phase II clinical agent BBR3464. *J. Biol. Inorg. Chem.* **7**(4 – 5):397 – 404.

- Meggers, E., Atilla-Gokcumen, G.E., Bregman, H., Maksimoska, J., Mulcahy, S.P., Pagano, N., and Williams, D.S. 2007. Exploring chemical space with organometallics: ruthenium complexes as protein kinase inhibitors. *Synlett*. **8**:1177 – 1189.
- Meggers, E. 2007. Exploring biologically relevant chemical space with metal complexes. *Curr. Opin. Chem. Biol.* **11**(3):287 – 292.
- Meier, S.M., Hanif, M., Adhireksan, Z., Pichler, V., Novak, M., Jirkovsky, E., Jakupec, M. A., Arion, V. B., Davey, C. A., Keppler, B. K., and Hartinger, C. G. 2013. Novel metal(II) arene 2-pyridinecarbothioamides: a rationale to orally active organometallic anticancer agents. *Chem. Sci.* **4**:1837–1846.
- Mendoza-Ferri, M.G., Hartinger, C.G., Eichinger, R.E., Stolyarova, N., Severin, K., Jakupec, M.A., Nazarov, A.A., and Keppler, B.K. 2008. Influence of the spacer length on the in vitro anticancer activity of dinuclear ruthenium–arene compounds. *Organometallics*. **27**(11):2405 – 2407.
- Mendoza-Ferri, M.G., Hartinger, C.G., Mendoza, M.A., Groessl, M., Egger, A.E., Eichinger, R.E., Mangrum, J.B., Farrell, N.P., Maruszak, M., Bednarski, P.J., Klein, F., Jakupec, M.A., Nazarov, A.A., Severin, K., and Keppler, B.K. 2009. Transferring the concept of multinuclearity to ruthenium complexes for improvement of anticancer activity. *J. Med. Chem.* **52**(4):916 – 925.
- Mendoza-Ferri, M.G., Hartinger, C.G., Nazarov, A.A., Eichinger, R.E., Jakupec, M.A., Jakupec, M.A., and Keppler, B.K. 2009. Influence of the arene ligand, the number and type of metal centers and the in vitro leaving group on the anticancer activity of polynuclear organometallic compounds. *Organometallics*. **28**(21):6260 – 6265.
- Messori, L., Scaletti, F., Massai, L., Cinellu, M.A., Gabbiani, C., Vergara, A., and Merlino, A. 2013. The mode of action of anticancer gold-based drugs: a structural perspective. *Chem. Commun. (Camb)*. **49**(86):10100 – 10102.
- Milacic, V., Fregona, D., and Dou, Q.P. 2008. Gold complexes as prospective metal-based anticancer drugs. *Histol Histopathol*. **23**:101 – 108.
- Mirabelli, C.K., Hill, D.T., Faucette, L.F., McCabe, F.L., Girard, G.R., Bryan, D.B., Sutton, B.M., Barus, J.O.L., Crooke, S.T., and Johnson, R. K. 1987. Antitumor activity of bis (diphenylphosphino)alkanes, their gold (I) coordination complexes, and related compounds. *J. Med. Chem.* **30**(12):2181 – 2190.
- Mistry, P., Kelland, L. R., Abel, G., Sidhar, S., and Harrap, K. R. 1991. The relationships between glutathione, glutathione-S-transferase and cytotoxicity of platinum drugs and melphalan in eight human ovarian carcinoma cell lines. *Br. J. Cancer*. **64**:215 – 220.
- Mulyana, Y., Webber, D.K., Buck, D.P., Motti, C.A., Collins, J.G., and Keene, F.R. 2011. Oligonuclear polypyridylruthenium(II) complexes incorporating flexible polar and non-polar bridges: synthesis, DNA-binding, and cytotoxicity. *Dalton Trans.* **40**(7):1510 – 1523.

- Murray, B. S., Menin, L., Scopelliti, R., and Dyson, P. J. 2014. Conformational control of anticancer activity: the application of arene-linked dinuclear ruthenium(II) organometallics. *Chem. Sci.* **5**(6):2536 – 2545.
- Needleman, S.B. and Wunsch, C.D. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**(3):443 – 453.
- Oki, A.R., Gogineni, P., Yurchenko, M., and Young Jr., V.G. 1997. Manganese(II) complex of the ‘tripod’ ligand tris(2-benzimidazolylmethyl)amine. Seven-coordinate Mn(II) in the crystal structure. *Inorganica Chim. Acta.* **257**(2):279 – 283.
- Ong, M.S., Richmond, T.J., and Davey, C.A. 2007. DNA stretching and extreme kinking in the nucleosome core. *J. Mol. Biol.* **368**:1067 – 1074.
- Oohara, I. and Wada, A. 1987. Spectroscopic studies on histone-DNA interactions. I. The interaction of histone (H2A, H2B) dimer with DNA: DNA sequence dependence. *J. Mol. Biol.* **196**:389 – 397.
- Oudet, P., Gross-Bellard, M., Chambon, P., 1975. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell.* **4**:281 – 300.
- Palermo, G., Magistrato, A., Riedel, T., von Erlach, T., Davey, C. A., Dyson, P. J., and Rothlisberger, U. 2016. Fighting cancer with transition metal compounds: from naked DNA to protein and chromatin targeting strategies. *ChemMedChem* **11**:1199 – 1210.
- Palladino, M.A. and Lobo, D. 2012. *Biology of Cancer, 2nd edition*. San Francisco, CA: Pearson Education.
- Park, S.H., Lee, J.H., Berek, J.S., and Hu, M.C. 2014. Auranofin displays anticancer activity against ovarian cancer cells through FOXO3 activation independent of p53. *Int. J. Oncol.* **45**(4):1691 – 1698.
- Peacock, W.J. Brutlag, D., Goldring, E., Appels, R., Hinton, C.W., and Lindsley, D.L. 1974. The organization of highly repeated DNA sequences in *Drosophila melanogaster* chromosomes. *Cold Spring Harb. Symp. Quant. Biol.* **38**:405 – 416.
- Peacock, A.F., Melchart, M., Deeth, R.J., Habtemariam, A., Parsons, S., and Sadler P.J. 2007. Osmium(II) and ruthenium(II) arene maltolato complexes: rapid hydrolysis and nucleobase binding. *Chem. J. Eur.* **13**(9):2601 – 2613.
- Peacock, A.F. and Sadler, P.J. 2008. Medicinal organometallic chemistry: designing metal arene complexes as anticancer agents. *Chem. Asian J.* **3**(11):1890 – 1899.
- Pearson, R.G. 1963. Hard and soft acids and bases. *J. Am. Chem. Soc.* **85**(22): 3533 – 3539.

- Pettinari, R., Marchetti, F., Pettinari, C., Petrini, A., Skelton, B.W., White, A.H., Bonfili, L., Cuccioloni, M., and Eleuteri, A.M. 2015. Dinuclear (η^6 -arene) ruthenium(II) acylpyrazolone complexes: Synthesis, characterization and cytotoxicity. *J. Organomet. Chem.* **791**:1 – 5.
- Pinato, O., Musetti, C., Farrell, N.P., and Sissi, C. 2013. Platinum-based drugs and proteins: reactivity and relevance to DNA adduct formation. *J. Inorg. Biochem.* **122**:27 – 37.
- Piolot, T., Tramier, M., Coppey, M., Nicolas, J.C., and Marechai, V. 2001. Close but distinct regions of human herpesvirus 8 latency-associated nuclear antigen 1 are responsible for nuclear targeting and binding to human mitotic chromosomes. *J. Virol.* **75**(8):3948 – 3959.
- Pisani, M.J., Weber, D.K., Heimann, K., Collins, J.G., and Keene, F.R. 2010. Selective mitochondrial accumulation of cytotoxic dinuclear polypyridy ruthenium(II) complexes. *Metallomics.* **2**(6):393 – 396.
- Polach, K.J. and Widom, J. 1995. Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. *J. Mol. Biol.* **254**(2):130 – 149.
- Pratesi, G., Perego, P., Polizzi, D., Righetti, S.C., Supino, R., Caserini, C., Manzotti, C., Guilianì, F.C., Pezzoni, G., Tognella, S., Spinelli, S., Farrell, N., and Zunino, F. 1999. *Br. J. Cancer.* **80**(12):1912 – 1919.
- Prussin, A.J., Zhao, S., Jain, A., Winkel, B.S., and Brewer, K.J. 2009. DNA interaction studies of tridentate bridged Ru(II)-Pt(II) mixed-metal supramolecules. *J. Inorg. Biochem.* **103**(3):427 – 43.
- Pryciak, P.M. and Varmus, H.E. 1992. Nucleosomes, DNA-binding proteins, and DNA sequence modulate retroviral integration target site selection. *Cell.* **69**(5):769 – 780.
- Raymond, E., Chaney, S.G., Taamma, A., and Cvitkovic, E. 1998. Oaliplatin: a review of preclinical and clinical studies. *Ann. Oncol.* **9**(10):1053 – 1071.
- Rice, P.A., Yang, S., Mizuuchi, K., and Nash, H.A. 1996. Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell.* **87**(7):1295 – 1306.
- Richmond, T.A. and Davey, C.A. 2003. The structure of DNA in the nucleosome core. *Nature.* **423**:145 – 150.
- Roder, C. and Thomson, M.J. 2015. Auranofin: repurposing an old drug for a new golden age. *Drugs R. D.* **15**(1):13–20.
- Ronconi, L., and Fregona, D. 2009. The midas touch in cancer chemotherapy: from platinum- to gold-dithiocarbamate. *Dalton Trans.* **28**:10670 – 10680.
- Rosenberg, B., VanCamp, L., and Krigas, T. 1965. Inhibition of Cell Division in Escherichia Coli by Electrolysis Products from a Platinum Electrode. *Nature.* **205**:698 - 699.

- Rosenberg, B., VanCamp, L., Trosko, J.E., and Mansour, V.H. 1969. Platinum compounds: a new class of potent antitumour agents. *Nature*. **222(5191)**:385 – 386.
- Roy, S., Ernst, J., Kharchenko, P.V., Kheradpour, P., Negre, N., Eaton, M.L., Landolin, J.M., Bristow, C.A., Ma, L., Lin, M.F., Washietl, S., Arshinoff, B.I., Ay, F., Meyer, P.E., Robine, N., Washington, N.L., Di Stefano, L., Berezikov, E., Brown, C.D., Candeias, R., Carlson, J.W., Carr, A., Jungreis, I., Marbach, D., Sealfon, R., Tolstorukov, M.Y., Will, S., Alekseyenko, A.A., Artieri, C., Booth, B.W., Brooks, A.N., Dai, Q., Davis, C.A., Duff, M.O., Feng, X., Gorchakov, A.A., Gu, T., Henikoff, J.G., Kapranov, P., Li, R., MacAlpine, H.K., Malone, J., Minoda, A., Nordman, J., Okamura, K., Perry, M., Powell, S.K., Riddle, N.C., Sakai, A., Samsonova, A., Sandler, J.E., Schwartz, Y.B., Sher, N., Spokony, R., Sturgill, D., van Baren, M., Wan, K.H., Yang, L., Yu, C., Feingold, E., Good, P., Guyer, M., Lowdon, R., Ahmad, K., Andrews, J., Berger, B., Brenner, S.E., Brent, M.R., Cherbas, L., Elgin, S.C., Gingeras, T.R., Grossman, R., Hoskins, R.A., Kaufman, T.C., Kent, W., Kuroda, M.I., Orr-Weaver, T., Perrimon, N., Pirrotta, V., Posakony, J.W., Ren, B., Russell, S., Cherbas, P., Graveley, B.R., Lewis, S., Micklem, G., Oliver, B., Park, P.J., Celniker, S.E., Henikoff, S., Karpen, G.H., Lai, E.C., MacAlpine, D.M., Stein, L.D., White, K.P., and Kellis, M. 2010. Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science*. **330(6012)**:1787-97.
- Sadler, P.J. and Sue, R.E. 1994. The chemistry of gold drugs. *Met Based Drugs*. **1(2-3)**:107 - 144.
- Safaei, R., Holzer, A. K., Katano, K., Samimi, G., and Howell, S. B. 2004. The role of copper transporters in the development of resistance to Pt drugs. *J. Inorg. Biochem.* **98**:1607 - 1613.
- Sajan, S.A. and Hawkins, R.D. 2012. Methods for indentifying higher-order chromatin structure. *Annu. Rev. Genomics Hum. Genet.* **13**:59 – 82.
- Sava, G., Bergamo, A., Zorzet, S., Gava, B., Casarsa, C., Cocchietto, M., Furlani, A., Scarcia, V., Serli, B., Iengo, E., Alessio, E., and Mestroni, G. 2002. Influence of chemical stability on the activity of the antimetastasis ruthenium compound NAMI-A. *Eur. J. Cancer*. **38(3)**:427-435.
- Schuh, E., Pfluger, C., Citta, A. Folda, A., Rigobello, M.P. Bindoli, A., Casini, A., and Mohr, F. 2012. Gold(I) carbene complexes causing thioredoxin 1 and thioredoxin 2 oxidation as potential anticancer agents. *J. Med. Chem.* **55**:5518 – 5528.
- Schultz, S.C., Shields, G.C., and Steitz, T.A. 1990. Crystallization of *Escherichia coli* catabolite gene activator protein with its DA binding site. The use of modular DNA. *J. Mol. Biol.* **213(1)**:159 – 166.
- Sharma, S., Kelly, T.K., and Jones, P.A. 2010. Epigenetics in cancer. *Carcinogenesis*. **31(1)**:27 – 36.

- Sheth, S.R. and Leckband, D. 1997. Measurements of attractive forces between proteins and end-grafted poly(ethylene glycol) chains. *Proc. Natl. Acad. Sci. U.S.A.* **94(16)**:8399 – 8404.
- Shi, Y., Lui, S.A., Kerwood, D.J., Goodisman, J., and Dabrowiak, J.C. 2012. Pt(IV) complexes as prodrugs for cisplatin. *J. Inorg. Biochem.* **107(1)**:6 – 14.
- Shivaswamy, S., Bhinge, A., Zhao, Y., Jones, S., Hirst, M., and Iyer, V.R. 2008. Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biol.* **6(3)**:e65
- Smith, E. and Shilatifard, A. 2010. The chromatin signaling pathway: diverse mechanisms of recruitment of histone-modifying enzymes and varied biological outcomes. *Mol Cell.* **40(5)**:689 – 701.
- Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.M., Zhu, P., and Li, G. 2014. Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science.* **344(6182)**:376 – 380.
- Spivak, G. 2015. Nucleotide excision repair in humans. *DNA Repair (Amst).* **36**:13 – 18.
- Stevens, S.K., Sterhle, A.P., Miller, R.L., Gammons, S.H., Hoffman, K.J., McCarty, J.T., Miller, M.E., Stultz, L.K., and Hanson, P.K. 2013. The anticancer ruthenium complex KP1019 induces DNA damage, leading to cell cycle delay and cell death in *Saccharomyces cerevisiae*. **83(1)**:225 – 234.
- Suto, R.K., Clarkson, M.J., Tremethick, D.J., and Luger, K. 2000. Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat. Struct. Biol.* **7(12)**:1121 – 1124.
- Tan, S. and Davey, C.A. 2011. Nucleosome structural studies. *Curr. Opin. Struct. Biol.* **21(1)**:128 – 136.
- Tan, S., Hunziker, Y., Pellegrini, L., and Richmond, T.J. 2000. Crystallization of the yeast MATalpha2/MCM1/DNA ternary complex: general methods and principles for protein/DNA cocrystallization. *J. Mol. Biol.* **297(4)**:947 – 959.
- Tan, S. and Richmond, T.J. 1998. Crystal structure of the yeast MATalpha2/MCM1/DNA ternary complex. *Nature.* **391(6668)**:660 – 666.
- Tiekink, E.R. 2003. Phosphinegold(I) thiolates--pharmacological use and potential. *Bioinorg. Chem. Appl.* **2003**:53 – 67.
- Tolstorukov, M.Y., Colasanti, A.V., McCandlish, D.M., Olson, W.K., and Zhurkin, V.B. 2007. A novel roll-and-slide mechanism of DNA folding in chromatin: implications for nucleosome positioning. *J. Mol. Biol.* **371**:725 – 738.
- Tsunaka, Y., Naoko, K., Tate, S., and Morikawa, K. 2005. Alteration of the nucleosomal DNA path in the crystal structure of a human nucleosome core particle. *Nucleic Acids Res.* **33(10)**:3424 – 3434.

- Tyler-Smith, C. and Brown, W.R. 1987. Structure of the major block of alphoid satellite DNA on the human Y chromosome. *J. Mol. Biol.* **195(3)**:457 – 470.
- Ueda, T., Catez, F., Gerlitz, G., and Bustin, M. 2008. Delineation of the protein module that anchors HMGN proteins to nucleosomes in the chromatin of living cells. *Mol. Cell Biol.* **28(9)**:2872 – 2883.
- Ussery, D.W. 2002. DNA structure: A-, B- and Z-DNA helix families. In: eLS. John Wiley & Sons Ltd, Chichester. <http://www.els.net>.
- van Holde, K.E. 1989. Chromatin, Springer Verlag, New York.
- van Rijt, S.H. and Sadler, P.J. 2009. Current applications and potential for bioinorganic chemistry in the development of anticancer drugs. *Drug Discov. Today*. **14(23-24)**:1089-97.
- Varghese, E. and Büsselberg, D. 2014. Auranofin, an anti-rheumatic gold compound, modulates apoptosis by elevating the intracellular calcium concentration ([Ca²⁺]_i) in mcf-7 breast cancer cells. *Cancers (Basel)*. **6(4)**:2243 - 2258.
- Vasudevan D., Chua, E.Y., and Davey, C.A. 2010. Crystal structure of nucleosome core particles containing the '601' strong positioning sequence. *J. Mol. Biol.* **403(1)**:1 – 10.
- Vock, C. A., Renfrew, A. K., Scopelliti, R., Juillerat-Jeanneret, L. and Dyson, P. J. 2008. Influence of the diketonato ligand on the cytotoxicities of [Ru(Z6-p-cymene)-(R2acac)(PTA)] complexes (PTA = 1,3,5-triaza-7-phosphaadamantane). *Eur. J. Inorg. Chem.* 1661–1671.
- Wang, D. and Lippard, S.J. 2005. Cellular processing of platinum anticancer drugs. *Nat. Rev. Drug Discov.* **4(4)**:307 – 302.
- Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G. J. 2009 Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. **25(9)**:1189 – 1191.
- Wishart, J.F., Bino, A., and Taube, H. 1986. A very short ruthenium(II)-nitrogen heterocycle bond: crystal structures of pentaammine(*N*-methylpyrazinium)ruthenium(II) Iodide and pentaammine(*N*-methylpyrazinium)ruthenium(III) *p* -toluenesulfonate pentahydrate. *Inorg. Chem.* **25**:3318-3321.
- Wolberger, C., Vershon, A.K., Lui, B., Johnson, A.D., and Pabo, C.O. 1991. Crystal structure of a MAT alpha 2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell*. **67(3)**:517 – 528.
- Woodcock, C. L., Dimitrov, S. 2001. Higher-order structure of chromatin and chromosomes. *Curr. Opin. Genet. Dev.* **11**:130 - 135.
- Wu, B., Droge, P., and Davey, C.A. 2008. Site selectivity of platinum anticancer therapeutics. *Nat. Chem. Biol.* **4(2)**:110 – 112.

- Wu, B. and Davey, C.A. 2008. Platinum drug adduct formation in the nucleosome core alters nucleosome mobility but not positioning. *Chem. Biol.* **15(10)**:1023 – 1028.
- Wu, B., Mohideen, K., Vasudevan, D., and Davey, C.A. 2010. Structural insight into the sequence dependence of nucleosome positioning. *Structure.* **18(4)**:528 – 536.
- Wu, B. and Davey, C.A. 2010. Using soft X-rays for a detailed picture of divalent metal binding in the nucleosome. *J. Mol. Biol.* **398(5)**:633 – 640.
- Wu, B., Ong, M.S., Groessl, M., Adhireksan, Z., Hartinger, C.G., Dyson, P.J., and Davey, C.A. 2011. A ruthenium antimetastasis agent forms specific histone protein adducts in the nucleosome core. *Chemistry.* **17(13)**:3562 - 3566.
- Wu, B., Davey, G.E., Nazarov, A.A., Dyson, P.J., and Davey, C.A. 2011. Specific DNA structural attributes modulate platinum anticancer drug site selection and cross-link generation. *Nucleic Acids Res.* **39(18)**:8200 - 8212.
- Yakovchuk, P., Protozanova, E., and Frank-Kamenetskii, M.D. 2006. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res.* **34(2)**:564 – 574.
- Yamada, K., Frouws, T. D., Angst, B., Fitzgerald, D. J., DeLuca, C., Schimmele, K., Sargent, D. F., and Richmond, T. J. 2011. Structure and mechanism of the chromatin remodelling factor ISW1a. *Nature.* **472**:448-53.
- Yang, P., Ebbert, J. O., Sun, Z., and Weinshilboum, R. M. 2006. Role of the glutathione metabolic pathway in lung cancer treatment and prognosis: a review. *J. Clin. Oncol.* **24**:1761 - 1769.
- Yang, W. and Steitz, T.A. 1995. Crystal structure of the site-specific recombinase gamma delta resolvase complexed with a 34 bp cleavage site. *Cell.* **82(2)**:193 – 207.
- Yuan, G.C., Liu, Y.J, Dion, M.F., Slack, M.D., Wu, L.F., Altschuler, S.J., and Rando, O.J. 2005. Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science.* **309(5734)**:626 – 630.
- Zdraveski, Z. Z., Mello, J. A., Farinelli, C. K., Essigmann, J. M., and Marinus, M. G. 2002. MutS preferentially recognizes cisplatin- over oxaliplatin-modified DNA. *J. Biol. Chem.* **277**:1255 - 1260.
- Zhou, B.R., Jiang, J., Feng, H., Ghirlando, R., Xiao, T.S., and Bai, Y. 2015. Structural mechanisms of nucleosome recognition by linker histones. *Mol. Cell.* **59(4)**:628 – 638.
- Zou, J., Taylor, P., Dornan, J., Robinson, S.P., Walkinshaw, M.D., and Sadler, P.J. 2000. First crystal structure of a medically relevant gold protein complex: unexpected binding of [Au(PEt₃)]⁺ to histidine. *Angew. Chem. Int. Ed. Engl.* **39(16)**:2931 – 2934.

Zhu, J.H., Zhao, Y.M., Zhu, Y.Y, Wu, Z.Y., Lin, M.X., He, W.J., Wang, Y., Chen, G.J., Dong, L., Zhang, J.F., Lu, Y., and Guo, Z.J. 2009. DNA cross-linking patterns induced by an antitumor active trinuclear platinum complex and comparison with its dinuclear analog. *Chem. Eur. J.* **15(21)**:5245 – 5253.