10-6-2021

A study of the mammalian high mobility group protein AT-hook 2 (HMGA2) and its interactions with DNA

Linjia Su
Florida International University, lsu008@fiu.edu

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

A STUDY OF THE MAMMALIAN HIGH MOBILITY GROUP PROTEIN AT-HOOK 2 (HMGA2) AND ITS INTERACTIONS WITH DNA

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CHEMISTRY

by

Linjia Su

2021

To: Dean Michael R. Heithaus
College of Arts, Sciences and Education

This dissertation, written by Linjia Su, and entitled A Study of the Mammalian High Mobility Group Protein AT-hook 2 (HMGA2) and Its Interactions with DNA, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_____
Yuk-Ching Tse-Dinh

_____
Jeremy Wayne Chambers

_____
Francisco Alberto Fernandez-Lima

_____
Kevin O'Shea

_____
Fenfei Leng, Major Professor

Date of Defense: October 06, 2021

The dissertation of Linjia Su is approved.

_____
Dean Michael R. Heithaus
College of Arts, Sciences and Education

_____
Andrés G. Gil
Vice President for Research and Economic Development
and Dean of the University Graduate School

Florida International University, 2021

DEDICATION

I dedicate this dissertation to my parents, grandparents, and other family members for their best loves and supports to me in my studies toward the doctoral degree in the United States.

ACKNOWLEDGMENTS

I would like to extend my sincere thanks to all the people I have met during my five years of graduate studies for their generous help and support. I am thankful to have had the chance to study and live in the U.S. over an important period of my life. It has been a wonderful journey, full of challenges, happiness, and memories.

First of all, I am extremely grateful to my advisor Dr. Fenfei Leng. My success of the five years' worth of research and the completion of my dissertation would not have been possible without the kind support and nurturing of him. Dr. Leng is an outstanding professor and mentor. Dr. Leng has provided me with encouragement, patience, and valuable advice throughout the duration of my research. With his excellent and experienced mentorship, I have gained valuable experiences in scientific research, critical thinking, writing and presentation skills.

I would also like to express my deepest appreciation to my dissertation committee, Dr. Yuk-Ching Tse-Dinh, Dr. Jeremy Wayne Chambers, Dr. Francisco Alberto Fernandez-Lima, and Dr. Kevin O'Shea for their persistent support and valuable suggestions during my graduate research in the past five years. I also wish to give special thanks to Dr. Tse-Dinh and Dr. Chambers for their very kind recommendations to apply for the Dissertation Year Fellowship. I would also like to thank Dr. Chambers for providing me the chance of learning cell experiment techniques in his lab. Many thanks to Dr. Fernandez-Lima for his very kind help regarding mass spectrometry experiments. I would also like to thank Dr. Watson Lees, former graduate program director, and Dr. Kathleen Rein, current graduate program director, for their encouragement and keeping track of my graduation progress during these years.

I am also grateful to our collaborators for their generous help with my research. I would like to thank Dr. Sabrina Battista at Istituto per l′Endocrinologia e l′Oncologia Sperimentale for her help concerning cell biology experiments. Thanks should also go to Dr. Prem P. Chapagain, Dr. Lidia Kos, and Dr. Jin He at Florida International University for their help in molecular dynamic simulation, cell experiments, and fluorescence microscopic imaging, respectively.

I would also like to acknowledge the Dissertation Year Fellowship from the University Graduate School of Florida International University for funding me in preparation of the dissertation. I would also like to acknowledge the Department of Chemistry and Biochemistry for the financial support during my Ph.D. studies. I gratefully acknowledge the funding of the dean's office of the College of Arts, Sciences and Education, and the Graduate and Professional Student Committee at Florida International University for supporting me to attend conferences. I would also like to acknowledge the 2018 Biomedical Research Initiative (BRI) Student Summer Research Awards from FIU NIGMS RISE Program for funding the initiation of my cell experimental studies.

I would like to thank all the current and former lab mates, Zifang Deng, Eddy Alfonso Maqueira, German Mejia, Daniel Alonso Boaretto, and Anthony Ramins for their kind help. I would like to recognize the help of Alejandra Velasquez and Kimberly Knox for their editing and proofreading of some chapters of my dissertation. I would also like to thank Dr. Thirunavukkarasu Annamalai and Dr. Wenjie Wang for their help on my fluorescence titration and western blotting experiments. I am grateful to some of my friends, Manqi Zhang, Wei Zhang, Chenyang Xue, etc. who I met in Miami for their help while living here.

Finally, I would like to express my deepest appreciation to my parents, grandparents, and other beloved family members for their love and support of my studies toward a doctoral degree in the United States.

ABSTRACT OF THE DISSERTATION

A STUDY OF THE MAMMALIAN HIGH MOBILITY GROUP PROTEIN AT-HOOK 2 (HMGA2) AND ITS INTERACTIONS WITH DNA

by

Linjia Su

Florida International University, 2021

Miami, Florida

Professor Fenfei Leng, Major Professor

The mammalian high-mobility-group protein AT-hook 2 (HMGA2) is a small DNA-binding protein and consists of three positively charged "AT-hooks" and a negatively charged C-terminal motif. It is a multifunctional nuclear protein linked to obesity, human height, stem cell youth, human intelligence, and tumorigenesis. Previous results showed that HMGA2 is a potential therapeutic target of anticancer and anti-obesity drugs through inhibiting its DNA-binding activities. Here a miniaturized, automated AlphaScreen ultra-high-throughput screening assay is developed to identify inhibitors targeting HMGA2-DNA interactions. After screening the LOPAC1280 library, several compounds are identified that strongly inhibit HMGA2-DNA interactions including suramin, a negatively charged antiparasitic drug. The inhibition is likely through the binding of suramin to the "AT-hooks" and therefore preventing HMGA2 from binding to the minor groove of AT-rich DNA sequences. Charge-charge interactions and hydrogen bonding between the suramin sulfonated groups and Arg/Lys residues likely play critical roles in the binding of suramin to the "AT-hooks". This study also suggests that HMGA2 may be one of suramin's cellular targets.

This dissertation also demonstrates that the negatively charged C-terminus greatly affects the DNA-binding properties of HMGA2, as the C-terminal deletion mutant HMGA2Δ95-108 binds much more tightly to the AT-rich DNA compared with the wildtype HMGA2. A synthetic peptide derived from the C-terminus of HMGA2 (CTP) strongly inhibits HMGA2 binding to AT-rich DNA through binding to the positively charged "AT-hooks", suggesting that the CTP may be used as an inhibitor to block HMGA2 binding to AT-rich DNA.

HMGA2 is also linked to human topoisomerase I and II. HMGA2 greatly reduced the chromosomal DNA damage in cancer cells caused by topoisomerase II poisons such as daunorubicin and doxorubicin. Due to the induced multidrug resistance (MDR) of cancer cells and a cumulative, irreversible cardiotoxicity, the therapeutic efficacy of them is compromised and limited. Here, four new daunorubicin and doxorubicin derivatives, daunorubicin-GTP/dGTP, and doxorubicin-GTP/dGTP conjugates, have been synthesized and characterized. These new derivatives rapidly accumulate intracellularly in human cancer cells and are cytotoxic to both doxorubicin-sensitive SKOV3 and doxorubicin-resistant NCI/ADR-RES cells. Western blotting results show that these derivatives change the expression patterns of DNA topoisomerase I and IIα.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

ABBREVIATIONS AND ACRONYMS

Arg	Arginine
ATA	Aurintricarboxylic acid.
ATP	Adenosine-5’-triphosphate
ATHP2	AT-hook peptide 2
ATHP3	AT-hook peptide 3
ATHPs	AT-hook peptides
AU	Arbitrary unit
a.u.	Arbitrary unit
BBB	Blood-brain barrier
bp	Base pairs
BRD4	Bromodomain-containing protein 4
BSA	Bovine serum albumin
BTSCs	Brain tumor stem cells
CD	Circular dichroism
cDNA	Complementary deoxyribonucleic acid
cDPCP	Cis-Diammine(pyridine)chloroplatinum (II) chloride
CEBPα	CCAAT/enhancer binding protein α
CEBPβ	CCAAT/enhancer binding protein β
ChIP	Chromatin immunoprecipitation
CK2	Casein kinase 2
CREB	cAMP response element-binding protein
CTP	Cytidine-5’-triphosphate
CTP	C-terminus peptide
Cys	Cystein
Da	Dalton
dATP	2’-deoxyadenosine 5’-triphosphate
dCTP	2’-deoxycytidine-5’-triphosphate
dGTP	2’-deoxyguanosine-5’-triphosphate
DMEM	Dulbecco's modified eagle's medium

DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNR	Daunorubicin
DNR-dGTP	Daunorubicin-dGTP conjugate
DNR/DOX-dGTPs	Daunorubicin/doxorubicin-dGTP conjugates
DNR/DOX-GTPs	Daunorubicin/doxorubicin-GTP conjugates
DNR/DOX-GTP/dGTPs	Daunorubicin/doxorubicin-GTP/dGTP conjugates
DNR-GTP	Daunorubicin-GTP conjugate
DNR-GTP/dGTPs	Daunorubicin-GTP/dGTP conjugates
dNTP	Deoxyribonucleotide triphosphate
DOX	Doxorubicin
DOX-dGTP	Doxorubicin-dGTP conjugate
DOX-GTP	Doxorubicin-GTP conjugate
DOX-GTP/dGTPs	Doxorubicin-GTP/dGTP conjugates
d.p.c.	Days post-coitum
dTTP	Thymidine-5’-triphosphate
EDTA	Ethylenediaminetetraacetic acid
EMT	Epithelial-mesenchymal transition
ESI	Electrospray ionization
FA	Fluorescence anisotropy
FBS	Fetal bovine serum
FI	Fluorescence intensity
FPLC	Fast protein liquid chromatography
Glu	Glutamic acid
GTP	Guanosine-5’-triphosphate
GWA	Genome-wide association
HMGA	Mammalian high-mobility-group protein AT-hook
HMGA1	Mammalian high-mobility-group protein AT-hook 1
HMGA2	Mammalian high-mobility-group protein AT-hook 2
HSC	Hematopoietic stem cell

IC_{50}	Half-maximal inhibitory concentration
ID2	Inhibitor of DNA binding 2
IDP	Intrinsically unstructured protein
IgG	Immunoglobulin G
ITC	Isothermal titration calorimetry
K_a	Association constant
kb	Kilo base pairs
K_d	Dissociation constant
kDNA	Kinetoplast deoxyribonucleic acid
Lys	Lysine
mAb	Monoclonal antibody
MD	Molecular dynamics
MDR	Multidrug resistance
MEFs	Mouse embryonic fibroblasts
mL	Milliliter
mM	Millimolar
MOA	Mechanism of action
mRNA	Messenger ribonucleic acid
MSC	Mesenchymal stem cell
MSI1	Musashi RNA binding protein 1
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
Mw	Molecular weight
NaCl	Sodium chloride
NCP	Nucleosome core particle
nESI	Nano electrospray ionization
nL	Nanoliter
nM	Nanomolar
NMR	Nuclear magnetic resonance
NSC	Neural stem cell
NTP	Nucleoside triphosphate

OLIG2	Oligodendrocyte transcription factor
PBS	Phosphate buffered saline
PDI-ELISA	Protein-DNA interaction enzyme-linked immunosorbent assay
PME	Particle mesh Ewald
PMSF	Phenylmethanesulfonyl fluoride
PPARγ	Peroxisome proliferator-activated receptor γ
PRI-ELISA	Protein-RNA interaction enzyme-linked immunosorbent assay
qRT-PCR	Quantitative reverse transcription polymerase chain reaction
RNA	Ribonucleic acid
Ro 90-7501	2′-(4-Aminophenyl)-[2,5′-bi-1H-benzimidazol]-5-amine
RPMI	Roswell park memorial institute
SAR	Structure activity relationship
SDS-PAGE	Sodium dodecyl sulfate -polyacrylamide gel electrophoresis
SELEX	Systematic evolution of ligands by exponential enrichment
SNP	Single nucleotide polymorphism
SSCT	Saline-sodium citrate buffer with Tween 20
TBST	Tris buffer saline with 0.1% Tween®20 detergent
TCEP	Tris-(2-carboxyethyl) phosphine hydrochloride
TLC	Thin layer chromatography
TMR	Tetramethylrhodamine-5-maleimide
TOP1	DNA topoisomerase 1
TOP2α	DNA topoisomerase 2-alpha
TR-FRET	Time-resolved fluorescence energy transfer
uHTS	Ultra-high-throughput screening assay
UTP	Uridine 5’-triphosphate
UTR	Untranslational region
UV-Vis	Ultraviolet-visible
v/v	Volume/volume
ΔG	Gibbs (free) energy change
ΔH	Enthalpy change

ΔS	Entropy change
ε	Extinction coefficient
λ	Wavelength
μg	Microgram
μL	Microliter
μM	Micromolar

CHAPTER 1: Reviews of the mammalian high mobility group protein AT-hook 2 (HMGA2): biochemical and biophysical properties, and its association with adipogenesis [1]

1.1 Abstract

The mammalian high-mobility-group protein AT-hook 2 (HMGA2) is a small DNA-binding protein and consists of three "AT-hook" DNA-binding motifs and a negatively charged C-terminal motif. It is a multifunctional nuclear protein directly linked to obesity, human height, stem cell youth, human intelligence, and tumorigenesis. Biochemical and biophysical studies showed that HMGA2 is an intrinsically disordered protein (IDP) and could form homodimers in aqueous buffer solution. The "AT-hook" DNA-binding motifs specifically bind to the minor groove of AT-rich DNA sequences and induce DNA-bending. HMGA2 plays an important role in adipogenesis most likely through stimulating the proliferative expansion of preadipocytes and also through regulating the expression of transcriptional factor Peroxisome proliferator-activated receptor γ (PPARγ) at the clonal expansion step from preadipocytes to adipocytes. Current evidence suggests that a main function of HMGA2 is to maintain stemness and renewal capacity of stem cells by which HMGA2 binds to chromosome and lock chromosome into a specific state, to allow the human embryonic stem cells to maintain their stem cell potency. Due to the importance of HMGA2 in adipogenesis and tumorigenesis, HMGA2 is considered a potential therapeutic target for anticancer and anti-obesity drugs. Efforts are taken to identify inhibitors targeting HMGA2.

1.2 Introduction

The mammalian high-mobility-group protein AT-hook 2 (HMGA2) is a non-histone

chromosome protein and belongs to the HMGA family, which includes four members: HMGA1a, 1b, 1c, and HMGA2 [2]. HMGA1a, 1b, and 1c are the different splicing products of the same gene, the HMGA1 gene [3]. HMGA2 is the product of a different gene, the HMGA2 gene [4-6]. High-mobility-group proteins were discovered, identified, and isolated by Graham H. Goodwin in E. W. Johns' lab at Chester Beatty Research Institute, UK, in the early 1970s [7, 8]. High-mobility-group (HMG) proteins simply refer to the group of fast migration, non-histone proteins in the polyacrylamide gels when calf thymus chromatin was extracted using 0.35 M NaCl and 2% trichloroacetic acid [7, 8]. Initially, only two HMG protein families, i.e., HMGB protein family (HMG-box proteins; former name HMG1/2 proteins [9]) and HMGN protein family (nucleosome binding proteins; former name HMG-14/17 proteins [9]) were identified [10]. HMGA1a/1b (former name HMG-I/Y [9]) were identified in 1983 in Hela S3 cells by Lund et al. [11]. HMGA2 (former name HMGI-C [9]) was discovered in 1985 by two different groups, Vincenzo Giancotti's group at Universita di Trieste, Italy [12], and Graham H. Goodwin' group, in the UK [13]. Interestingly, HMGA2 was only expressed in virus-transformed cells [12, 13]. The cDNA sequence and protein sequence of murine and human HMGA2 were published in 1991 [4] and 1994 [14], respectively. The protein sequences of murine and human HMGA2 are almost identical, except for five amino acid residues. None of these five amino acid residues is located in the "AT-hook" DNA-binding motifs [4, 14]. The human HMGA2 gene is located at chromosome 12, 12q14.3 [15, 16], and has five exons and four introns, occupying approximately 160 kb [17]. Intron 3 is very large ~110 kb [6] and separates the "AT-hook" DNA-binding motifs and the acidic C-terminus [17]. The 4.1 kb mRNA contains an 854 bp 5' UTR, a 330 bp coding sequence, and a 2966 bp 3' UTR [17].

The 3’ UTR carries multiple microRNA Let-7 binding sites that negatively regulate HMGA2 expression in development and tumorigenesis [18-20].

1.3 Biochemical and Biophysical Properties of HMGA2

The human HMGA2 is a small DNA-binding protein and has 109 amino acid residues (Figure 1.1). One unique feature of HMGA2 is the asymmetric charge distribution along its backbone (Figure 1.1). As a consequence, HMGA2 can form homodimers in aqueous buffer solution [21]. Early studies also showed that HMGA2 forms dimers, trimers, and tetramers, although it was attributed to the formation of a disulfide bond between the cysteine (Cys) residues of murine HMGA2 (murine HMGA2 has a Cys reside at position 41) [22]. Nevertheless, the formation of trimers and tetramers cannot be explained by the disulfide-bond formation. A different study also demonstrated that HMGA1a could interact with itself [23]. The dimerization of HMGA proteins is an unusual property because HMGA proteins, including HMGA1 and HMGA2, are intrinsically disordered/unstructured proteins (IDPs) [21]. In other words, this family of proteins does not have a secondary structure and a tertiary structure; however, it has a quaternary structure. It was initially quite a challenge to publish our results by showing that HMGA2 can form homodimers and homo-oligomers in aqueous buffer solution, although this unique feature of HMGA2 was observed in the early 2000s [24]. Nevertheless, biochemical and biophysical studies clearly demonstrated that HMGA2 can form homodimers [21]. Of course, HMGA2 is not the only IDP that can form homodimers; other IDPs can also form homodimers [25-30]. The cytoplasmic region of T-cell receptor subunit and the disordered N-terminal domain of ultraspirale from Aedes aegypti (aaUsp-NTD) can self-associate into homodimers [27]. Intriguingly, the dimerization is not accompanied by a disorder-to-order

transition [27]. Although several IDPs can self-associate into homodimers and/or homo-oligomers, two important questions are still unanswered: (1) Can one IDP interact with another IDP? The homodimerization and oligomerization of IDPs partially answered this question. (2) What forces contribute to the interaction between IDPs? As we discussed above, the dimerization of HMGA2 mainly stems from the electrostatic interactions between the positively charged "AT-hooks" and the negatively charged C-terminus, since the asymmetric charge distribution is along the HMGA2 backbone. Are hydrophobic force and hydrogen bonds also involved in the dimerization?

Another unique feature of HMGA proteins is that all except HMGA1c contain three "AT-hook" DNA-binding motifs (Figure 1.1). The "AT-hook" DNA-binding motif is an 8-9 amino acid peptide that contains 5-6 positively charged amino acid residues, lysine and arginine (Figure 1.1). Specifically, this DNA-binding motif has a consensus palindromic sequence, PRGRP surrounded by one or two positively charged amino acid residues (Figure 1.1). The "AT-hook" DNA-binding motif was coined by Reeves and Nissen [31]. They demonstrated that the conformation of this consensus DNA-binding motif is similar to several typical DNA minor groove binders, such as netropsin, distamycin, and Hoeshst33258, and can preferentially bind to minor grove of AT-rich DNA sequences [31]. Indeed, nuclear magnetic resonance (NMR) and crystal structural studies showed that the "AT-hook" DNA-binding peptide specifically binds to the minor groove of AT-rich DNA [32-34] (Figure 1.2). NMR and crystal structures are quite similar with the central RGR group deeply penetrating into the minor groove of AT base pairs [33, 34]. The crystal structural study showed that the "AT-hook" also forms hydrogen bonds between the backbone NH groups of the peptide and the thymine in the minor groove [34]. It was also

discovered that the DNA is bent and the minor groove is widened [34]. The HMGA2-induced DNA bending was also observed when gel permutation assay was used [35]. The HMGA2-induced bending angle was determined to be 35 degrees, which was significantly larger than the one (24 degrees) observed in the crystal structure induced by just one "AT-hook" DNA-binding motif, suggesting that more than one "AT- hook" was involved in the DNA binding and bending [35]. Although it was suggested that "AT-hook" DNA-binding motifs adopt a defined structure upon binding to AT-rich DNA sequences, recent studies and molecular simulations do not support a disordered-to-ordered structural transition of the "AT-hook" DNA-binding motif upon DNA binding (Figure 1.2B) [36].

Figure 1.1. (A) The primary structure of the human HMGA2. The positively charged "AT-hook" DNA binding motifs and the negatively charged C-terminal motif are highlighted in

red and blue, respectively. **(B)** Sequence logo of the "AT-hook" DNA-binding motifs of HMGA1 and HMGA2. Sequence conservation, measured in bits of information, is illustrated by the height of stacking. The sequence logo was generated by WebLogo (available at https://weblogo.berkeley.edu/logo.cgi). **(C)** The C-terminal motifs of HMGA1 and HMGA2. The CK2 phosphorylation sites of the HMGA2 C-terminal motif are indicated by arrows.

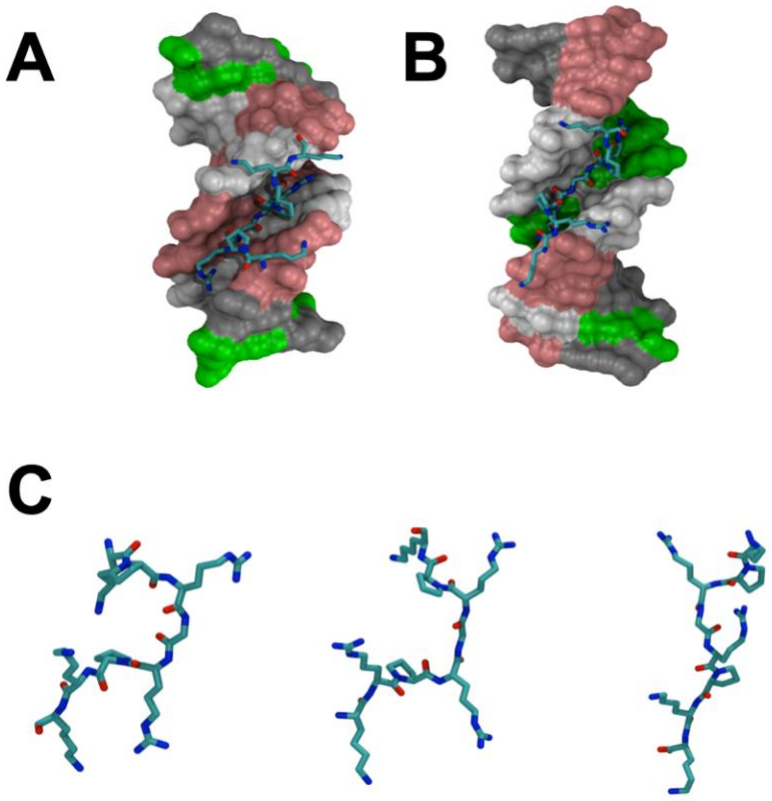


Figure 1.2. (A) The crystal structure of an "AT-hook" and DNA complex [34]. **(B)** The solution NMR structure of the complex of an "AT-hook" DNA-binding motif with DNA determined by Huth et al. [33]. **(C)** Comparison of the "AT-hook" DNA-binding motifs from the crystal structure, the NMR solution structure, and the simulation structure. The molecular dynamic simulation was performed by using NAMD with CHARMM36m force field [37-39] for 20 ns, 2 fs/time step, 310K, and 12.0 Å VDW force cutoff.

Early DNA foot-printing studies showed that HMGA proteins could bind to any stretches of 5 to 6 AT bp with similar binding affinities [40], suggesting that binding of these proteins to AT-rich DNA sequences does not have sequence specificities. However, other studies demonstrated that HMGA proteins prefer binding to two-to-three

appropriately spaced AT-rich DNA sequences with high DNA-binding affinities [41]. More importantly, HMGA proteins bind to two to three runs of AT base pairs in the promoter regions, as a transcription factor to regulate transcription in vivo [42-45]. NMR and crystal structural studies also showed that "AT-hook" DNA-binding motifs prefer certain AT DNA sequences [33, 34]. Encouraged by these results, we performed a PCR-based systematic evolution of ligands by exponential enrichment (SELEX) experiment and identified two consensus DNA-binding sequences for HMGA2, 5'ATATTCGCGAWWATT-3', and 5'-ATATTGCGCAWWATT-3', where W represents A or T [46]. This is an interesting result in that the HMGA2 preferred binding sequences contain four GC base pairs in the middle [46]. Since the minor groove of GC base pairs is crowded, it is likely that not all three "AT-hook" DNA-binding motifs bind to the DNA minor groove. Possibly one of the "AT-hook" DNA-binding motifs binds to the major groove of the middle GC-rich DNA sequence. Our recent results showed that the "AT-hook" DNA-binding motif could indeed bind to the DNA major groove (unpublished results). ChIP experiments using cancer cells overexpressing HMGA2 showed that HMGA2 prefers binding to AT-rich DNA sequences, although the center sequences are not necessarily GC-rich [47].

Another unique feature of HMGA proteins is that all contain a highly negatively charged C-terminal motif. For instance, HMGA2 has a 15 amino acid residue C-terminus, with seven glutamic acid residues and one aspartic acid residue (Figure 1.1). The C-terminus also contains three serine residues and two threonine residues that can be phosphorylated by casein kinase 2 (CK2) [48, 49]. If fully phosphorylated, the C-terminus of HMGA2 may carry up to 19 negative charges at physiological conditions (each

phosphate group introduces two negative charges). Since the electrostatic interaction is an important force for the binding of HMGA2 to AT-rich DNA [50], one possible function of the C-terminus of HMGA proteins is to regulate the DNA-binding affinity during different cellular events. Indeed, previous results showed that the negatively charged C-terminus and its phosphorylation could regulate the DNA-binding capacity of HMGA proteins [48, 49, 51-53]. The C-terminal motif of HMGA proteins may also be involved in the protein-protein interactions. The truncated HMGA2 without the C-terminal motif cannot form a homodimer [21]. The C-terminal motif may also be involved in the interaction of HMGA proteins with its protein partners [23, 54, 55]. Nevertheless, the biological functions of the C-terminal motif are still unknown, although it was implied that it might contribute to the tumorigenesis and cellular proliferation and transformation [44, 56].

1.4 HMGA2 in Adipogenesis

The association of HMGA2 with adipogenesis was discovered by Chada and coworkers when they studied mouse growth-hormone-independent pygmy phenotype in the early 1990s [57-59]. They showed that this mouse pygmy phenotype stems from the deletion of mouse Hmga2 gene from the chromosome, and, as a result, HMGA2 was not expressed during embryogenesis [59]. Further, they demonstrated that HMGA2 only expressed in early embryonic stage from 10.5 to 15.5 d.p.c. (days post-coitum) and did not express in mouse adult tissues [59]. They also showed that HMGA1 was predominantly expressed in 10.50-16.5 d.p.c. mouse embryos [59]. By analyzing 11.5 d.p.c. mouse embryos, they found that HMGA2 expression was observed in most tissue and organs except the brain. Only a small localized region of forebrain had HMGA2 expression [59]. The testes and adrenal gland of the mutant mice are much smaller [58, 60]. In fact, Hmga2

null mice are sterile due to the fact that germ-cell maturation was blocked in the testes [58, 60]. Nevertheless, the most noticeable phenotype of the Hmga2 deletion mice is the small size. At 10 weeks of age, the body weight of Hmga2 mutants is approximately 40% of that of the wild-type mice [58, 60]. These mutant mice have significantly reduced body fat compared to the wild-type mice [58, 60]. Additionally, the mutant mice are resistant to a high-fat diet [58, 60]. In contrast, a high-fat diet can induce the HMGA2 expression in adipose tissues and cause obesity in wild-type and leptin-deficient mice [61]. They performed an interesting experiment by using the genetic mouse model Lep^{ob}/Lep^{ob} to generate two mouse models: $Hmga2^{-/-}$ Lep^{ob}/Lep^{ob} and $Hmga2^{+/-}$ Lep^{ob}/Lep^{ob} [61]. The disruption of the Hmga2 gene caused a dramatic reduction in obesity of the leptin-deficient mice (Lep^{ob}/Lep^{ob}) in a gene-dosage-dependent manner: $Hmga2^{+/+}$ Lep^{ob}/Lep^{ob} mice weighed over three times more than $Hmga2^{-/-}$ Lep^{ob}/Lep^{ob} animals, and the weight of $Hmga2^{+/-}$ Lep^{ob}/Lep^{ob} mice was in between [61]. The adipocytes of the mutant mice are similar to those of the wild-type mice, and the expression levels and regulations of genes involved in adipogenesis are also similar [61]. The reduction of body fat is a result of a decrease of the cell numbers in the adipocyte tissues [61]. More recently, Federico et al. created Hmga1 and Hmga2 double-knockout mice that have a “superpygmy” phenotype, with 75% smaller size than that of the wild-type mice [62]. The body fat should also be greatly reduced. Several studies with transgenic mice overexpressing HMGA2 also demonstrated the association of HMGA2 with adipogenesis [56, 63, 64]. For instance, Battista et al. created a transgenic mouse model that expresses a truncated HMGA2 carrying 3 “AT-hook” DNA-binding motifs without the acidic C-terminal motif. These transgenic mice developed a giant and obese phenotype [56] with a great expansion of

adipocyte tissues. In additional to a great enhancement of abdominal fat mass, large fat pads were also associated with other organs, such as around the kidneys and at the bases of the hearts [56].

More evidence of HMGA2's association with adipogenesis comes from the studies of Lipomas, which are a type of benign tumor that is made of fat/adipocyte tissues and often found with the chromosomal arrangement at 12q14-15 [15, 16]. Early studies showed that these common mesenchymal neoplasms resulted from the expression of a chimeric protein consisting of the three "AT-hook" DNA-binding motifs fused to LIM or an acidic transactivation domain [15, 16, 65]. Further studies showed that the expression of the three "AT-hook" DNA-binding motifs alone is sufficient for the formation of lipomas [56, 64, 66]. In 2005, a case was reported for an eight-year-old boy who has a phenotype of overgrowth, advanced endochondral bone, a cerebellar tumor, and multiple lipomas [67]. Molecular analyses showed that this abnormal phenotype stems from the inversion of chromosome 12, with breakpoints at p11.22 and q14.3 that resulted in the expression of a truncated HMGA2 only with the three "AT-hook" DNA-binding motifs and lacking the negatively charged C-terminal motif [67]. This phenotype is similar to that of transgenic mice described above [56]. Genome-wide association (GWA) studies using single nucleotide polymorphism (SNP) data found that HMGA2 is associated with human height in the general population across different ethnicities or races [68-75]. Specifically, several SNPs, such as rs1042725 and rs10784502, located in the 3' UTR of HMGA2 gene, are associated with human height [68, 76]. Surprisingly, rs10784502 was also found to associate with human intracranial volumes and intelligence quotient (IQ) [77]. The association of HMGA2 with human height was further demonstrated by 12q14

microdeletion syndromes in which several genes, including HMGA2, were deleted [78, 79]. One common phenotype is the short stature and growth failure [78, 79]. For example, case #D0811079 is a boy who has a deletion that only involved HMGA2. Besides the short stature, no other anomalies were observed for this patient [79].

Adipocytes are derived from multipotent mesenchymal stem cells (MSCs) through two distinction phases: the commitment of MSCs to preadipocytes and the differentiation of preadipocytes to mature adipocytes [80-84]. The route of MSCs to preadipocytes is quite complex and can be driven by different signaling pathways [82, 83]. For the differentiation pathway of preadipocytes to mature adipocytes, two steps are involved: clonal expansion and adipocyte maturation [82, 83]. Our understanding of the molecular mechanism of preadipocytes differentiation into mature adipocytes mainly came from the studies of model preadipocyte cell lines that are committed to differentiating into adipocytes, such as 3T3-L1 and 3T3-F422A [82, 85, 86]. Several transcriptional factors, such as CREB (cAMP response element-binding) protein, CEBPβ (CCAAT/enhancer binding protein β), CEBPα, and PPARγ (peroxisome proliferator-activated receptor γ), are involved in this process [80-84]. Recent studies showed that HMGA2 is highly expressed during the exponential growth of 3T3-L1 cells [87]. Its expression is significantly reduced upon growing to confluence (the quiescent state; [87]). Interestingly, after the addition of differentiation cocktail, HMGA2's expression is induced again and reaches the highest level after two days (the mRNA reaches the highest level after a six-to-eight-hour induction [87-91]. HMGA2 graduate decreases to the basal level after the cells are differentiated into mature adipocytes [87-91]. Current evidence showed that HMGA2 functions at the clonal expansion step and regulates the expression of transcriptional factor PPARγ [91]. HMGA2 expression level is

also negatively regulated by microRNA let-7 [89], a factor that plays critical role in stem cells' self-renewal and stemness [18-20]. Some other micro RNA species, such as microRNA 33b, may also be involved in this process [92]. Although more studies are needed to determine the molecular mechanism of HMGA2 for adipogenesis, it is likely that HMGA2 affects adipogenesis through a mechanism similar to that through which it regulates other stem cells and their differentiation. Below, we briefly review HMGA2's association with cell youth and self-renewal of stem cells [93-96], one of its most intriguing functions.

As shown above, mouse HMGA2 only expressed in early embryonic stage and did not express in the adult tissues [59]. Similar to mouse HMGA2 expression pattern, HMGA2 was expressed in all human fetal tissues [97, 98]. In contrast, HMGA2 did not express in most adult tissues, except for lung and kidney [98]. These results suggest that HMGA2 is mainly expressed during embryonic and fetal development. Interestingly, HMGA2 is highly expressed in human stem cells, including human embryonic stem (hES) cells and the early differentiating embryoid bodies (EBs) [99, 100]. For instance, Nishino et al. showed that HMGA2 expression is highly expressed in neural stem cells (NSCs) and declines with age [101]. This decrease is partially caused by the increasing expression of a microRNA let-7b that targets the 3' UTR of Hmga2 mRNA [101]. They further demonstrated that HMGA2 promotes NSC self-renewal in young, but not in old, mice, most likely through a new pathway by which HMGA2 expression was inhibited by let-7b. As a result, JunB and P16Ink4a/P19Arf expression was enhanced [101]. These results are consistent with an earlier study showing that let-7 regulates self-renewal and stemness of breast cancer stem cells [19]. Since then, HMGA2 was shown to link to the stem cell youth

and self-renewal of other stem cells and progenitors [102-110]. For example, the self-renewal capacity and youth of hematopoietic stem cells (HSCs) is linked to expression of HMGA2 [104]. It was also demonstrated that the expression of HMGA2 was able to rescue the in vitro aging process of mesenchymal stem cells [111]. The self-renewal potential/capacity is determined by a unique pathway involving the RNA-binding protein Lin28, the microRNA let-7b, and HMGA2, in which Lin28 binds to let-7 pre-microRNA and inhibits the generation of let-7 [104, 110]. In 2010, Cavazzana-Calvo et al. reported a case of successful gene therapy of human β-thalassemia, a genetic disease with mutations in the β-globin gene that reduce or abolish β-globin protein production [112]. An adult patient with severe β^E/β^0-thalassaemia who was dependent on monthly transfusion became transfusion-independent after receiving the lentiviral-based gene therapy where the modified HSCs with β-globin lentiviral vector were transplanted into the patient's bone morrow [112]. Surprisingly, the therapeutic efficacy stems from the overexpression of HMGA2 in HSCs or progenitor cells to produce nucleated blood cells with overexpressed HMGA2 [112].

1.5 Conclusions and Perspectives

HMGA2 is a non-histone chromosome protein and has been linked to several phenotypic characteristics. Some of these phenotypes are reviewed here, except for its association with tumorigenesis. For the HMGA proteins' role in tumorigenesis, please refer to review articles published in the past for details [113-122]. It looks likely that the main functions of HMGA2 are promoting cell proliferation and maintaining the stemness potency of stem cells. What is still obscure is the molecular mechanism behind these phenotypes and functions. We believe that HMGA2 is an epigenetic factor that programs

or reprograms chromosomes into a “defined” state, to achieve these functions (Figure 1.3) [123, 124]. This hypothesis is in contrast with the previous belief that HMGA2 serves as a transcriptional factor or an architecture/general transcriptional factor, to promote or inhibit transcription only. Recent evidence showed that HMGA2 might also affect other cellular processes, such as DNA replication. For instance, Droge and coworkers showed that HMGA2 protects stalled DNA replication forks and prevents the forks from collapsing, to enhance stem- and cancer-cell survival when these cells are challenged with DNA-replication stress [125]. This unique property may affect the sensitivity of cancer cells to chemotherapy drugs, especially topoisomerase poisons [126, 127]. They estimated that about 10^5 to 10^6 molecules of HMGA2 exist in each human embryonic stem cell, which lead to one molecule of HMGA2 binding to 3 to 30 kb human chromosomal DNA or 10 to 100 nucleosome core particles (NCPs) on average [100]. Early studies by Goodwin and coworkers also showed that HMGA2 could compete with histone H1 for binding to nucleosomes [128]. A possible scenario is that the binding of HMGA2 to nucleosomes may “lock” chromosome into a specific state to allow the human embryonic stem cells to maintain their stemness status. Secondly, HMGA2 is regulated by microRNA let-7 [18, 20, 94]. Specifically, let-7b destabilizes HMGA2 mRNA by targeting the 3’ UTR [18, 20, 94]. As a result, HMGA2 expression is significantly reduced. Let-7 and HMGA2 play an important role in cell differentiation and should be considered as epigenetic factors.

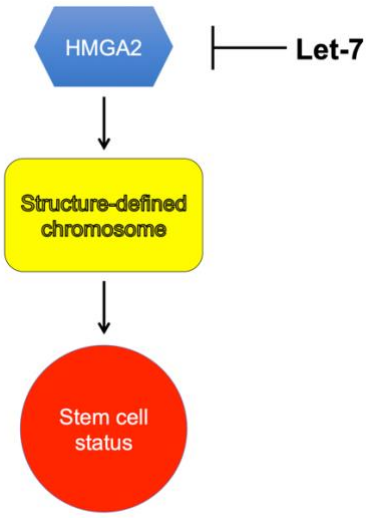


Figure 1.3. HMGA2 serves as an epigenetic or chromosome-remolding factor, to lock chromosome into a defined structure/conformation and to maintain the stem-cell status. MicroRNA let7 inhibits HMGA2 expression.

Another poorly understood area is the stability of HMGA proteins, including HMGA2, in vivo. Cao et al. showed that HMGA2 could be SUMOylated in vitro and inside cells [129]. Ubiquitin-proteasome dependent degradation may be the pathway for HMGA2's degradation [129, 130]. Apparently, more studies are needed in this field. Furthermore, the function of the negatively charged C-terminal motif of HMGA2 is still unknown. Interestingly, all HMG proteins have a highly negatively charged C-terminus [131], indicating that the negatively charged C-terminus has important functions. The C-terminal motif of HMGA2 also contains several serine and threonine residues that can be phosphorylated by CK2. One immediate consequence of the negatively charged C-terminal motif and its phosphorylation is to regulate HMGA2's binding to DNA and nucleosomes [48, 49, 51-53]. Nevertheless, more studies are needed.

As discussed above, HMGA2 plays an important role in adipogenesis and is an excellent target for the treatment of obesity [61]. Since the overexpression and/or aberrant-

expression of HMGA2 is directly linked to the formation of a variety of malignant tumors, including lung cancer [132, 133], breast cancer [134, 135], prostate cancer [136], leukemia [137], and melanoma [138-141], HMGA2 appears to be an attractive target for anticancer drugs [118, 142]. Several strategies may be used to target HMGA2 for therapeutic purposes. The first strategy is to target the AT-rich DNA-binding sequences that HMGA2 recognizes. For instance, we recently demonstrated that netropsin, a DNA minor groove binder, potently inhibits HMGA2 binding to DNA [90, 142]. Intriguingly, netropsin strongly inhibited the differentiation of mouse pre-adipocyte 3T3-L1 cells into adipocytes. It is likely that the inhibition is accomplished through the inhibition of HMGA2 binding to the target DNA sequences during differentiation [90]. Other minor groove binders can also inhibit HMGA2 binding to DNA [90]. A disadvantage of this strategy is that netropsin non-specifically binds to any five AT base pairs and displays non-specific cytotoxicity to many cell types [143-146], which prevents it from becoming an effective anticancer and anti-obesity drug. Although it is possible to design a synthetic compound that targets a specific AT sequence [147-149], the lack of knowledge about what sequences HMGA2 recognizes inside a cell makes this strategy a mission impossible. The second strategy is to identify compounds that bind to HMGA2 and prevent it from binding to AT-rich DNA sequences. This is a tough job, because DNA-binding proteins (transcriptional factors) are considered "undruggable" due to the fact that they usually do not have enzymatic activities suitable for chemical intervention [150, 151]. Additionally, HMGA2 is an IDP and lacks a deep pocket for ligand binding [21, 33]. One way to overcome these difficulties is to develop an efficient method to identify inhibitors from existing small molecule repositories. Indeed, we established a medium-throughput screening method based on the protein-DNA

interaction enzyme-linked immunosorbent assay (PDI-ELISA), to screen a small library containing 29 DNA-binding compounds, and successfully identified several small molecules that disrupt HMGA2 binding to the minor groove of AT-rich DNA sequences [90]. Recently, we developed a miniaturized automated AlphaScreen ultra-high-throughput screening (uHTS) assay to identify inhibitors targeting HMGA2-DNA interactions (unpublished results). After screening the LOPAC1280 compound library, we discovered several small molecule compounds that potently inhibit the HMGA2-DNA interaction through binding to HMGA2 (unpublished results). The third strategy is the use of the negatively charged C-terminus to inhibit HMGA2 binding to AT-rich DNA sequences. Recently, our unpublished results showed that the C-terminal motif of HMGA2 binds to the "AT-hook" DNA-binding motifs and inhibits HMGA2 binding to AT-rich DNA sequences. It is possible to synthesize the C-terminus mimics, to enhance the inhibition ability and also increase the stability in vivo [152, 153]. The fourth strategy is to target HMGA2's mRNA. Anti-sense oligomers and RNAi were used to lower the HMGA2 expression at the cellular level [19, 42, 154-157]. MicroRNA let-7b may also be used to decrease HMGA2 expression level. The delivery of these nucleic acids into cells or the target tissues may still be a challenge.

CHAPTER 2: Identification of HMGA2 inhibitors by AlphaScreen-based ultra-high-throughput screening assays [158]

2.1 Abstract

The mammalian high mobility group protein AT-hook 2 (HMGA2) is a multi-functional DNA-binding protein that plays important roles in tumorigenesis and adipogenesis. Previous results showed that HMGA2 is a potential therapeutic target of anticancer and anti-obesity drugs by inhibiting its DNA-binding activities. Here we report the development of a miniaturized, automated AlphaScreen ultra-high-throughput screening assay to identify inhibitors targeting HMGA2-DNA interactions. After screening the LOPAC1280 compound library, we identified several compounds that strongly inhibit HMGA2-DNA interactions including suramin, a century-old, negatively charged antiparasitic drug. Our results show that the inhibition is likely through suramin binding to the "AT-hook" DNA-binding motifs and therefore preventing HMGA2 from binding to the minor groove of AT-rich DNA sequences. Since HMGA1 proteins also carry multiple "AT-hook" DNA-binding motifs, suramin is expected to inhibit HMGA1-DNA interactions as well. Biochemical and biophysical studies show that charge-charge interactions and hydrogen bonding between the suramin sulfonated groups and Arg/Lys residues play critical roles in the binding of suramin to the "AT-hook" DNA-binding motifs. Furthermore, our results suggest that HMGA2 may be one of suramin's cellular targets.

2.2 Introduction

The mammalian high mobility group protein AT-hook 2 (HMGA2) is a multi-functional nuclear protein highly expressed in the early embryonic stage [59]. Early studies showed that HMGA2 is related to preadipocyte proliferation & differentiation and obesity

[61, 89, 159]. For example, Hmga2 knockout mice were severely different in fat cells and developed pygmy phenotype [59]. The disruption of Hmga2 gene dramatically reduced obesity of leptin-deficient mice (Lep^{ob}/Lep^{ob}) [61]. These results suggest that HMGA2 is a potential target for the treatment of obesity. HMGA2 is also linked to oncogenesis. The over and/or aberrant expression leads to the formation of a variety of tumors including benign tumors, such as lipomas [160] and uterine leiomyomas [161], and malignant tumors, such as lung cancer [132, 133], leukemia [137], and melanoma [138, 139]. Intriguingly, HMGA2 expression level always correlates with the degree of malignancy, metastasis, and a poor prognosis [162, 163], suggesting that this protein is also a therapeutic target of anti-cancer and anti-metastasis drugs [118, 142]. Furthermore, HMGA2 is associated with neural and hematopoietic stem cell youth [101, 104], human height [68], and human intelligence [77].

HMGA2 is a small DNA-binding protein and belongs to the HMGA family [119, 164]. This protein family has four members: HMGA1a, 1b, 1c, and HMGA2 [2]. HMGA1a, 1b, and 1c are different splicing products of the same gene, HMGA1 gene [3]. HMGA2 is the product of a separate gene, HMGA2 gene [4, 5]. All members except HMGA1c consist of three "AT-hook" DNA-binding motifs and a highly acidic C-terminal motif [46]. The "AT-hook" DNA-binding motifs contain a unique palindromic sequence, PGRGP, each side surrounded by one or two positively charged amino acids, i.e., Lysine or Arginine [165] and bind to the minor groove of AT-rich DNA sequences [31]. HMGA2 is an intrinsically disordered protein (IDP) [21]. When it binds to AT-rich DNA sequences, the "AT-hook" DNA binding motifs adopt defined structures [33]. This disordered-to-ordered conformational transition allows HMGA2 to adapt to different AT-rich DNA sequences

and to participate in different nuclear activities [119, 164]. These results suggest that HMGA2-DNA interactions could be chemically intervened for therapeutic purposes [115]. Utilizing a systematic evolution of ligands by exponential enrichment (SELEX) method, we previously identified two consensus DNA sequences for HMGA2 binding: 5'-ATATTCGCGAWWATT-3' and 5'-ATATTGCGCAWWATT-3', where W is A or T [46]. This result and a following study [35] suggests that HMGA2 binds and bends specific DNA sequences.

With the identification of HMGA2 as a potential target for the treatment of obesity and cancer, the next step is to search for chemical compounds that prevent HMGA2 binding to its target DNA sequences. For instance, utilizing protein-DNA interaction enzyme-linked immunosorbent assays (PDI-ELISA), we recently found several DNA-binding inhibitors including netropsin that disrupt HMGA2-DNA interactions [90]. Intriguingly, our results showed that netropsin strongly inhibited the differentiation of the mouse pre-adipocyte 3T3-L1 cells into adipocytes most likely through a mechanism by which netropsin inhibits HMGA2-DNA interactions [90]. However, DNA-binding compounds are usually too toxic to be used as anti-obesity and anticancer drugs. Novel inhibitors of HMGA2-DNA interactions that are not cytotoxic and do not directly bind DNA are required before such approaches can be considered as therapeutic applications. A high-throughput screening (HTS) strategy is needed to identify novel compounds that disrupt HMGA2-DNA binding. To achieve this, we developed an AlphaScreen-based assay of HMGA2 binding to DNA that is amenable to automated ultra HTS (uHTS) in a high-density plate format. Here we report the establishment of the assay, and the identification of several HMGA2 inhibitors including suramin, a century-old antiparasitic drug.

2.3 Materials and Methods

2.3.1 Materials

The His-tagged mammalian HMGA2 was purified using a Ni-NTA agarose column followed by an SP Sepharose fast flow column as described previously [46]. An extinction coefficient of 5810 M^{-1} cm^{-1} at 280 nm was used to determine its concentration [46]. Biotin-labeled DNA oligomer FL814 carrying a specific SELEX binding site of HMGA2 [46, 90] was purchased from Eurofins MWG Operon, Inc. AlphaScreen histidine (nickel chelate) detection kits containing nickel chelate acceptor beads and streptavidin donor beads (#6760619M), LANCE Ultra ULight-anti-6xHIS (#TRF0105) and LANCE Eu-W1024 Streptavidin (#AD0062), as well as 1536 well Optiplates plates (#6004290) were purchased from Perkin Elmer. Netropsin was purchased from Sigma and used without further purification. Suramin and NF023 were purchased from MilliporeSigma, Inc. Sodium 1-Naph-thalenesulfonate was obtained from TCI America, Inc. NF110, NF340, NF449, and NF546 were purchased from Tocris Biosciences. His-tagged BRD4-BD1 was purchased from BPS Bioscience (#31042). Pre-acetylated Biotin-Histone 4 Peptide was purchased from AnaSpec (#64989-025). Bovine serum albumin fraction V (#A7888), CHAPS (#C3023-25G), and LOPAC1280 library (#LO1280) were purchased from Sigma Aldrich. Human epidermoid carcinoma A-431 cell line was purchased from the ATCC (ATCC-CRL-1555). DMEM media was purchased from Thermo Fisher (Gibco #11995), 100 × Penicillin/Streptomycin solution (#30-002-CI), DPBS (#21- 031-CV), 0.25% Trypsin in HBSS (#25-050-CI), 200 mM L-glutamine (#25-005-CI), white high base 1536 well plates (#4570), and sterile white high base 1536 well plates (#4571) were purchased

from Corning. The control compound MG-132 was purchased from Promega (#G932B). ATP-Lite 1-step was purchased from Perkin Elmer (#6016739).

2.3.2 HMGA2 AlphaScreen ultra HTS assay.

Using a Labcyte Echo 555 acoustic dispenser 5 nL of DMSO were added to columns 1-4 and 45-48 of a white 1536 well plate, while 5 nL of 2 mM compounds in DMSO were added to columns 5-44. Using a Beckman BioRAPTR FRD bulk reagent dispenser 1 μL of assay buffer (30 mM Citrate, 300 mM NaCl, and 0.005% Tween 20) was added to columns 1 and 2. Next 1 μL of assay buffer containing 125 nM HMGA2 was added to columns 3-48 with a Beckman BioRAPTR FRD bulk reagent dispenser. Finally, 1 μL of assay buffer containing 25 nM FL814 was added to every well of the plate using a Beckman BioRAPTR FRD bulk reagent dispenser. The plate was then centrifuged at 200 × *g* for 1 min. After 30 min at room temperature, 2 μL of bead buffer (10 mM HEPES, 300 mM NaCl and 0.005% Tween 20) containing 20 μg/mL anti-6xHIS acceptor beads and 20 μg/mL streptavidin donor beads were dispensed into every well using a Beckman BioRAPTR FRD bulk reagent dispenser. The plates were then centrifuged at 200 × *g* for 1 min. After 1 h at room temperature the plates were read on a Perkin Elmer Envision multimode plate reader in AlphaScreen mode.

2.3.3 HMGA2 Lance assay.

Using a Labcyte Echo 555 acoustic dispenser 5 nL of DMSO were added to columns 1-4 and 45-48 of a white 1536 well plate, while 5 nL of 2 mM compounds in DMSO were added to columns 5-44. Using a Beckman BioRAPTR FRD bulk reagent dispenser 1 μL of assay buffer (10 mM Tris, 300 mM NaCl and 0.005% Tween 20) was added to columns 1 and 2. Next, 1 μL of assay buffer containing 125 nM HMGA2 was added to columns 3-48

with a Beckman BioRAPTR FRD bulk reagent dispenser. Finally, 1 μL of assay buffer containing 25 nM FL814 was added to every well of the plate using a Beckman BioRAPTR FRD bulk reagent dispenser. The plate was then centrifuged at 200 × *g* for 1 min. After 30 min at room temperature 1 μL of assay buffer containing 250 nM LANCE Ultra ULight-anti-6xHIS and 1 μL of assay buffer containing 12 nM LANCE Eu-W1024 Streptavidin was dispensed into every well using a Beckman BioRAPTR FRD bulk reagent dispenser. The plates were then centrifuged at 200 × *g* for 1 min. After 1 h at room temperature the plates were read on a Perkin Elmer Envision multimode plate reader in TR-FRET mode (excitation @340 nm, first emission at @665 nm, second emission at 615 nm).

2.3.4 BRD4 AlphaScreen assay.

Using a Labcyte Echo 555 acoustic dispenser 5 nL of DMSO were added to columns 1-4 and 45-48 of a white 1536 well plate, while 5 nL of 2 mM compounds in DMSO were added to columns 5-44. Using a Beckman BioRAPTR FRD bulk reagent dispenser 1 μL of assay buffer (50 mM HEPES, 100 mM NaCl, 0.1% BSA and 0.0005% CHAPS) was added to columns 1 and 2. Next, 1 μL of assay buffer containing 50 nM BRD4 was added to columns 3-48 with a Beckman BioRAPTR FRD bulk reagent dispenser. Finally, 1 μL of assay buffer containing 50 nM peptide was added to every well of the plate using a Beckman Bio- RAPTR FRD bulk reagent dispenser. The plate was then centrifuged at 200 × *g* for 1 min. After 60 min at room temperature 2 μL of assay buffer containing 20 μg/mL anti-6xHis acceptor beads and 20 μg/mL streptavidin donor beads was dispensed into every well using a Beckman BioRAPTR FRD bulk reagent dispenser. The plates were then centrifuged at 200 × *g* for 1 min. After an overnight incubation at room temperature the plates were read on a Perkin Elmer Envision multimode plate reader in AlphaScreen mode.

2.3.5 Cell viability assay.

A-431 cells were grown in media (DMEM + 10% FBS + 1×Penicillin/Streptomycin + 2 mM L-glutamine) at 37 °C in a humidified incubator with 5% CO_2 to 70% confluency. They were then washed with DPBS and trypsinized. For the assay the cells were resuspended in media (DMEM + 10% FBS + 1×Penicillin/Streptomycin + 2 mM L-glutamine) at 125,000 cells/mL. Using a Thermo multidrop combi bulk reagent dispenser 4 μL of cell suspension (500 cells) were dispensed into every well of a 1536 well plate. The plates were centrifuged at 200 × g for 1 min, relidded and incubated overnight at 37 °C with 5% CO_2. The next day using a Labcyte Echo 555 acoustic dispenser 5 nL of MG-132, final in well concentration of 25 μM, were added to columns 1-2, 5 nL of DMSO were added to columns 3-4 and 45-48 and 5 nL of 2 mM compounds in DMSO were added to columns 5-44. The plates were then centrifuged at 200 × g for 1 min, relidded and returned to the incubator. After 48 h 4 μL of room temperature ATP-Lite were added with a Beckman BioRAPTR FRD bulk reagent dispenser. The plates were centrifuged at 200 × g for 1 min and then incubated at room temperature. After 10 min the plates were read using a Perkin Elmer Viewlux multimode plate reader in luminescence mode.

2.3.6 Protein-DNA interaction ELISA (PDI-ELISA) assay.

The PDI-ELSA assays were described previously [90] and used to determine the inhibition IC_{50} against HMGA2-DNA interactions. The apparent inhibitory IC_{50} values were obtained using the following equation: $y = min + (\frac{(max-min)}{1+10^{((x-logIC_{50})\times Hill\ slope}})$, where x, y, max, and min represent the inhibitor's concentration, the inhibition level, the maximum inhibition value, and the minimum inhibition value, respectively.

2.3.7 Isothermal titration calorimetry (ITC).

ITC experiments were conducted using a VP-ITC titration calorimeter (MicroCal, Inc., Northampton, MA) interfaced to a PC. Origin 7.0, supplied by the manufacturer was used for data acquisition. For a typical ITC experiment, the titration was set up so that 10 µL of 0.2 mM suramin or analogs was injected every 120 seconds, up to a total of 29 injections, into an HMGA2 sample (1.44 mL of 5 µM) or ATHP3 sample (1.44 mL of 15 µM) in the sample cell in 10 mM Tris-HCl pH 8.0 and 1 mM EDTA. The heat liberated or absorbed is observed as a peak corresponding to the power required keeping the sample and reference cells at identical temperatures. The peak produced during the injection is converted to heat output by integration and corrected for cell volume and sample concentration. Control experiments were also carried out to determine the contribution of the heats of dilution arising from (1) suramin or analogs into buffer and (2) buffer into HMGA2 or ATHP3. The net enthalpy for the titration reaction was determined by subtraction of the component heats of dilution.

2.3.8 Native mass spectrometry.

A custom-built nano electrospray unit (nESI) was coupled to a Maxis Impact HD Q-TOF mass spectrometer (Bruker, Billerica, MA) for all the native mass spectrometry analysis. Quartz glass capillaries (O.D.: 1.0 mm and I.D.: 0.70 mm) were pulled utilizing a P-2000 micropipette laser puller (Sutter Instruments, Novato, CA) and loaded with 10 µL aliquot of the sample solution. Sample solutions consisted of 1-10 µM HMGA2 in 10 mM ammonium acetate solution at physiological pH (pH = 6.7). For the observation of the HMGA2-Ligand complexes, a 1:1, 1:3 and 1:10 ratio of 5 µM concentration of the HMGA2 and Ligand (suramin) were prepared in 10mM ammonium acetate and let it rest

for 10 minutes prior infusion. A typical nESI source voltage of +/- 600-1200 V was applied between the pulled capillary tips and the MS instrument inlet. Ions were introduced via a stainless-steel tube (1/16 x 0.020'', IDEX Health Science, Oak Harbor, WA) held at room temperature into the TIMS cell. Solvents, methanol, and ammonium acetate salts utilized in this study were analytical grade or better and purchased from Fisher Scientific (Pittsburgh, PA). A Tuning Mix calibration standard (G24221A) was obtained from Agilent Technologies (Santa Clara, CA) and used as received. Mass spectra were processed using Bruker Compass Data Analysis version 5.1 (Bruker Daltonik GmbH).

2.3.9 Cell cultures and growth curves.

BTSC#83 and BTSC#30p and their culture conditions have been described [166]. For growth curves, suspension cultures were mechanically disaggregated and filtered through 37 μm filters (StemCell Technologies, Vancouver, Canada) in order to obtain *bona fide* single-cell suspension. 1×10^3 cells were plated in U-bottom 96-well plates and treated or not with 100, 200 and 400 μM Suramin. Cell growth was assayed at the indicated time using the CellTiter Assay System (Promega). The growth conditions of 8505c anaplastic thyroid carcinoma cells and mouse Embryonic Fibroblasts (MEFs) knock-out for HMGA1, HMGA2, and HMGA1/A2 double mutant have been described previously [62, 167]. P values were calculated by using two-way ANOVA followed by Sidak's multiple comparison test.

2.3.10 Protein extraction and Western blot.

The single-cell suspension was obtained by mechanical disaggregation of the spheres and total proteins were extracted 48 h later, as previously described [166]. After separation by SDS-polyacrylamide gel electrophoresis, proteins were blotted on nitrocellulose

membranes (GE Healthcare Europe Gmb) and hybridized with the following antibodies: anti-HMGA1, anti-HMGA2 (Genetex) or affinity-purified anti-HMGA2, anti-α-actin (SantaCruz Biotechnologies).

2.3.11 RNA extraction and qRT-PCR analyses.

The single-cell suspension was obtained by mechanical disaggregation of the spheres and RNA was extracted 48 h later, by using Direct-zol RNA MiniPrep (Zymo Research). One μg of RNA was retrotranscribed, by using QuantiTect® Reverse Transcription Kit (Qiagen) and qRT-PCR was performed as described [166]. Primers used in the qRT-PCR experiments are listed in Table S2.1. The $2^{-\Delta\Delta Ct}$ formula was used to calculate the differential gene expression.

2.3.12 Chromatin immunoprecipitation (ChIP) assay.

Chromatin immunoprecipitation was performed as described [168]. Briefly, for each sample, 5×10^6 cells were collected by trypsinization and cross-linked with 1% formaldehyde; after sonication, the samples were immunoprecipitated with 10 μL anti-HMGA2 antibody (Genetex) or with control normal rabbit IgG and the chromatin was extracted. qPCR amplification was performed on 6 μL of immunoprecipitated DNA, using primers amplifying regions A and B of the ID2 promoter (Table 2.1). C_t values of samples precipitated with anti-HMGA2 and IgG antibodies were normalized with C_t values of input DNA. Fold change enrichment was calculated in comparison to normalized IgG, using the $2^{-\Delta\Delta Ct}$ formula. P values were calculated by using one-way ANOVA followed by Sidak's multiple comparison test.

2.3.13 Molecular dynamics simulation.

Since HMGA2 is shown to be an intrinsically disordered protein, we first performed molecular dynamics (MD) simulations to generate multiple disordered conformations of the suramin-interacting segment of HMGA2 chain (-GEKRPRGRPRKW-). Using the Charmm-Gui web interface [169], the peptide was solvated in a cubic water box with TIP3 water and the system was neutralized by adding five Cl^- ions. The final system contained ~16,000 atoms. NAMD 2.12 [39] was used to perform all-atom molecular dynamics with CHARMM36 force field [170]. The particle mesh Ewald (PME) method [171] was used for calculating the long-range ionic interactions. The system was minimized for 10,000 steps, followed by a 100 ps equilibration at 300 K with 1 fs time step. A 100-ns production simulation with 2-fs time step was performed at a constant pressure of 1 atm. and T = 300 K. The Nose-Hoover Langevin-piston method [172] was used for pressure coupling, with a piston period of 50 fs and a decay of 25 fs, and the Langevin temperature coupling with a friction coefficient of 1 ps^{-1} was used for maintaining the temperature. From the 100-ns simulation trajectory, 1000 protein pdb frames were extracted using Visual Molecular Dynamics (VMD) [173].

2.3.14 In silico docking studies.

Suramin was docked to 1000 MD-generated conformations using AutoDock Vina 1.1.2 [174]. The protein pdb files and the suramin compound structure were first converted to pdbqt format for docking. Using custom scripts, Suramin was screened against the protein conformations and the resulting scores of the complexes were sorted and ranked according to their binding affinities.

2.4 Results

2.4.1 An automated uHTS assay to identify HMGA2 inhibitors.

We previously used a PDI-ELISA assay to screen a small library containing 29 DNA-binding compounds and successfully identified several small molecules that disrupt HMGA2 binding to the minor groove of AT-rich DNA sequences [90]. Although this assay performs well in a 96 well plate format, it is not suitable for automated ultrahigh throughput screening of 100,000s of compounds due to the need for streptavidin-coated assay plates and multiple wash steps. To address these limitations, we established a new assay using AlphaScreen technology. The assay entails binding a biotin-labeled DNA oligomer FL814 and His-tagged HMGA2 to streptavidin-coated donor beads and nickel chelate (Ni-NTA) acceptor beads, respectively (Figure 2.1). A series of DNA binding studies were performed to determine the optimal conditions for the AlphaScreen Primary Assay (Figure S2.1 A-D). After these experiments, 12.5 nM of FL814 and 62.5 nM of HMGA2 were chosen for the assay. The assay tolerated up to 1% DMSO without any significant change in signal. We have previously identified two commercially available compounds netropsin and WP631 that strongly inhibit HMGA2 binding to FL814 [90]. These two inhibitors are readily available for purchase and served as positive controls for HMGA2-DNA interaction inhibition in the assay. Results in Figure 2.2 E demonstrate that netropsin and WP631 potently inhibit HMGA2-DNA interactions with an IC_{50} of 22 and 48 nM, respectively.

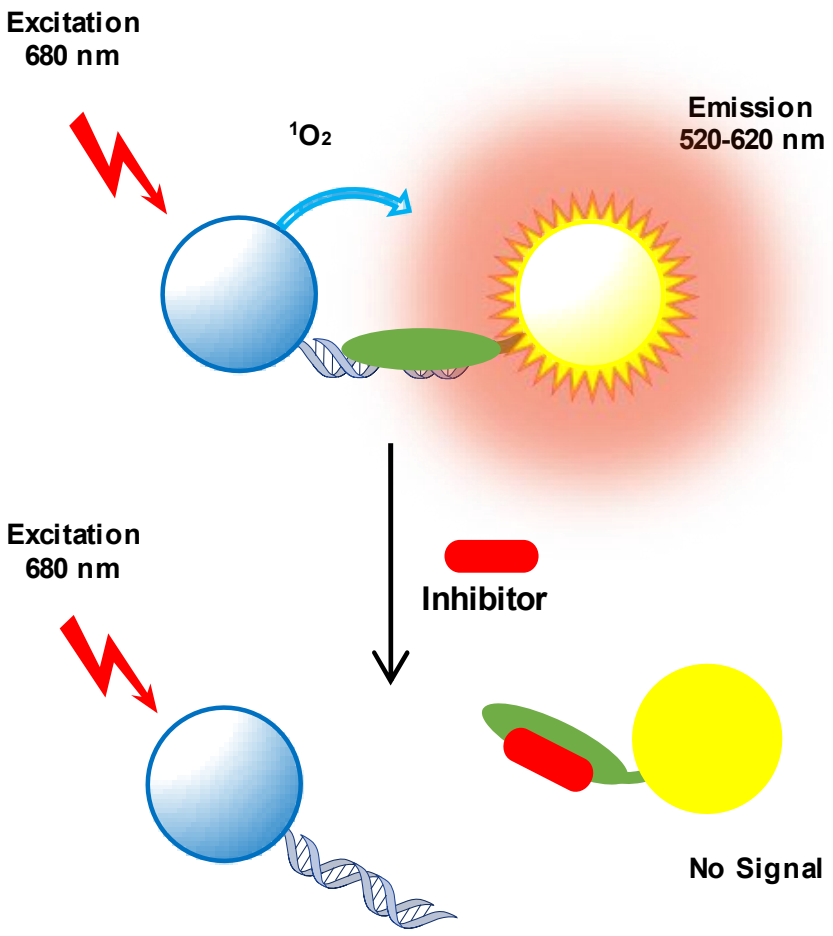


Figure 2.1. The AlphaScreen primary assay for HMGA2-DNA interaction. The biotin-labeled FL814 (double-stranded DNA) and the His-tagged HMGA2 (green oval) were immobilized to streptavidin-coated donor beads and nickel chelate (Ni-NTA) acceptor beads, respectively.

In this study, we also developed a LANCE time-resolved fluorescence energy transfer (TR-FRET) assay as a secondary assay for our screening. The His-tagged HMGA2 and biotinylated oligomer FL814 were linked to Europium-labeled anti-6 × His antibody and APC-labeled streptavidin, respectively. A series of DNA binding studies were performed to determine the optimal conditions for the secondary Assay (Figure S2.2). After these experiments, 12.5 nM of FL814 and 62.5 nM of HMGA2 were chosen for the assay, which

is the same for the primary assay. The assay tolerated up to 0.25% DMSO without any significant change in signal (Figure S2.2 C). For a counterscreen assay to exclude those compounds that nonspecifically bind to the protein surface, an AlphaScreen assay of H4 peptide binding to bromodomain-containing protein 4 (BRD4) was adopted (Figure S2.2 E). Since we are looking for compounds with anti-metastasis or anti-obesity activities, compounds with high levels of cytotoxicity are undesirable. A cytotoxicity assay using the ATPlite Luminescence Assay System and the human epidermoid carcinoma cell line, A431, was developed to eliminate compounds that exhibit cytotoxicity. Compounds that potently and selectively inhibit HMGA2 binding to FL814, and display an $IC_{50} > 50$ μM are will be prioritized for further characterization. Those that do not will be excluded. Furthermore, potential DNA-binding compounds, i.e., DNA intercalators, minor groove binders, and DNA alkylating agents, can be identified by analyzing their chemical structures and will be excluded as well.

2.4.2 Screen the LOPAC1280 compound library.

With the establishment of the miniaturized, automated uHTS assays, we screened the Sigma LOPAC1280 collection of pharmacologically active chemical compounds at a final concentration of 5 μM (Figure S2.3 A). Figure 2.2, Figure S2.3 B, and Table S2.2 show our results and parameters of the AlphaScreen primary uHTS assay. The following 16 compounds showed ≥ 50% inhibition of HMGA2-FL814 interactions: cisplatin, cDPCP, carboplatin, mitoxantrone, Ro 90-7501, aurintricarboxylic acid, GW5047, indirubin-3'-monoxime, 6-hydroxy-DL-DOPA, methyl-3,4-dephostatin, tyrphostin 51, (2'Z,3'E)-6-bromoindirubin-3'-oxime, reactive blue-2, JFD00244, steviol, and suramin. These 16 compounds were cherry picked and then subjected to testing in the LANCE TR-FRET

secondary assay. The following seven compounds demonstrated ≥ 50% inhibition of HMGA2-DNA interactions in both uHTS assays: cisplatin, cDPCP, carboplatin, mitoxantrone, Ro 90-7501, aurintricarboxylic acid, and suramin. Dry powders of these seven compounds were purchased for additional testing in these assays. The identity of the compounds was confirmed by LC/ MS, and fresh stock solutions (10 mM) were prepared in 100% DMSO. We next performed a series of titration experiments and determined the potencies (IC_{50}) values of these seven hits in the primary AlphaScreen assay, and the secondary LANCE TR-FRET assay, as well as the counter screen assay, and cytotoxicity assay. Table S2.3 summarizes our results. Cisplatin, cDPCP, carboplatin, and mitoxantrone are known DNA-binding agents and likely represent compounds that inhibit HMGA2-DNA interactions by binding to the AT-rich DNA sequence of FL814. Additionally, these DNA-binding compounds may also inhibit other essential cellular functions, which prevent them from further investigation. Intriguingly, although Ro 90-7501, aurintricarboxylic acid, and suramin do not bind to DNA due to their chemical properties, these three compounds strongly inhibit HMGA2-DNA interactions. Of particular interest is suramin, a highly negatively charged antiparasitic drug [175] (Figure S2.4) that potently inhibits HMGA2-FL814 interactions with an IC_{50} of 2.58 μM (Figure 2.3.). Additionally, suramin did not inhibit H4 peptide binding to BRD4 in the counter screen assay (Figure 2.3. C) and is not cytotoxic to human A-431 cells (Figure 2.3. D, Table S2.3). As a final validation of suramin as a "hit" from the HTS, we confirmed its inhibitory effect on HMGA2 interactions with DNA using the PDI-ELISA assay (Table 2.1). Thus, suramin meets all criteria that we set for the identification of novel HMGA2-DNA inhibitors.

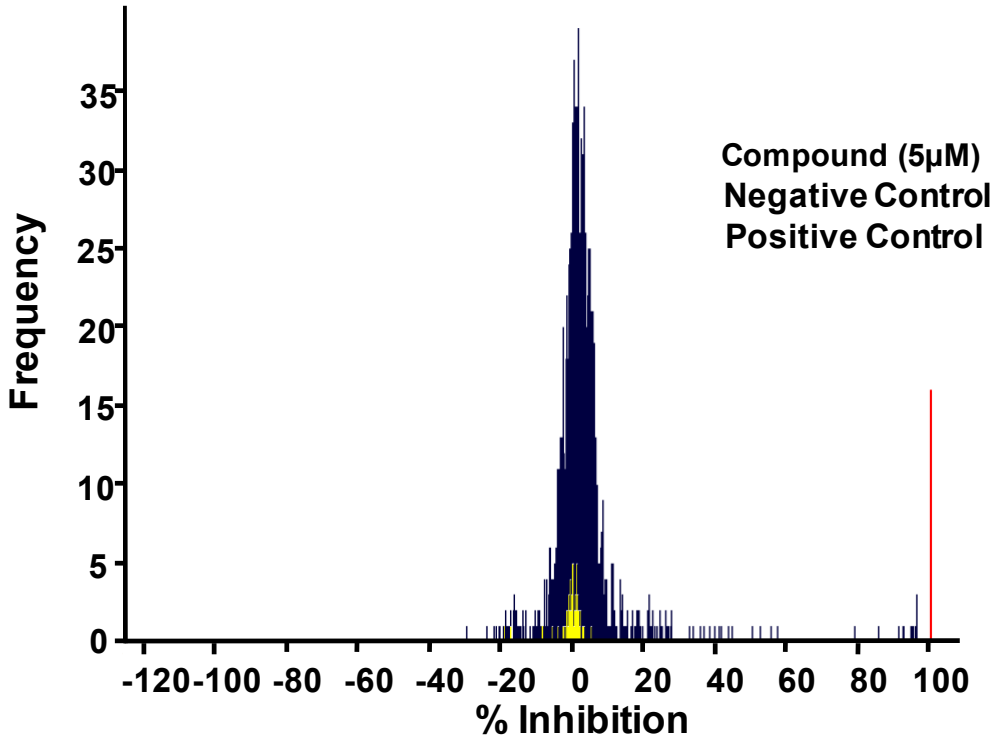


Figure 2.2. Results of HMGA2-DNA pilot screens using the Sigma LOPAC1280 compound library. Netropsin was used as positive controls.

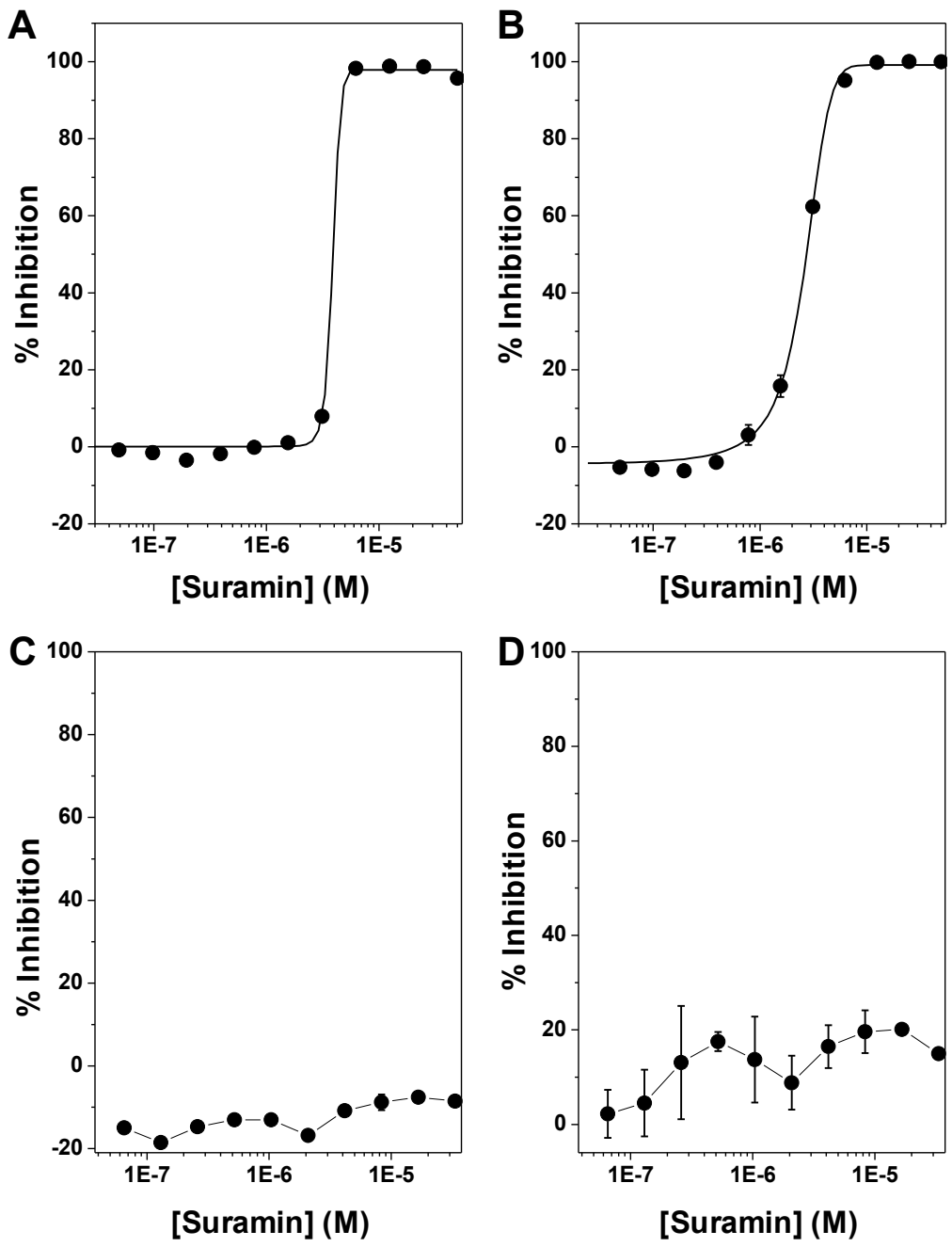


Figure 2.3. The discovery of suramin as a potent inhibitor of HMGA-DNA interactions. **(A)** The AlphaScreen Primay assay with IC_{50} of 2.6 μM. **(B)** The TR-FRET LANCE secondary screen. **(C)** The couterscreen assay using the BRD4 AlphaScreen Assay. **(D)** The cytotoxicity assay. All assays were described in "Methods". The standard deviation was calculated according to three independent experiments. The curve represents the best fit of a four parameter logistic that was determined by nonlinear regression. Data points represent mean ± SEM.

2.4.3 Suramin and analogues strongly bind to HMGA2 and ATHP3.

We next sought to determine the mechanism by which suramin inhibits HMGA2-DNA interactions. Since suramin carries 6 negative charges (Figure S2.4), it should not bind to DNA due to the fact that DNA is highly negatively charged. Instead, it should bind to HMGA2 because HMGA2 is positively charged [176]. ITC studies revealed that suramin physically interacts with HMGA2 (Figure 2.4 A, Table 2.1). These studies revealed that there are two types and a total of five suramin binding sites on HMGA2. The first type of three suramin binding sites has a binding constant of $4.08 \pm 0.92 \times 10^6$ M^{-1} with the following thermodynamic parameters: ΔG, − 9.02 kcal/mol; ΔH, − 14.58 kcal/mol; and – TΔS, 5.56 kcal/ mol. The second type of two suramin binding sites has a binding constant of $4.34 \pm 2.16 \times 10^4$ M^{-1} with the following thermodynamic parameters: ΔG, − 6.33 kcal/mol; ΔH, − 6.57 kcal/mol; and − TΔS, 0.24 kcal/mol. These five suramin binding sites of HMGA2 were confirmed by our mass spectrometric experiments at high suramin to HMGA2 ratios (Figure S2.5). Because HMGA2 contains three highly positively charged "AT-hook" DNA binding motifs, it is reasonable to assume that suramin strongly binds to these highly positively charged motifs through charge-charge interactions. Our ITC experiments of suramin titrating into an "AT-hook" peptide 3 (ATHP3) solution confirmed this hypothesis: suramin binding to ATHPs has a binding constant of $4.06 \pm 0.32 \times 10^7$ M^{-1} and 1:1 binding molar ratio (Figure 2.4 B, Table 2.1). The following are its binding thermodynamic parameters: ΔG, − 10.38 kcal/mol; ΔH, − 17.13 kcal/mol; and − TΔS, 6.75 kcal/mol.

Next, we examined the structure activity relationship (SAR) of suramin using several suramin analogues (Figures. S2.4, S2.6, S2.7). As shown in Table 2.1, all of these

analogues except sodium 1-naphthalenesulfonate potently inhibit HMGA2-DNA interactions with IC_{50} values ranging from 0.43 to 10.63 μM (the results of sodium 1-naphthalenesulfonate are not included in Table 2.1). Likewise, all these chemical compounds except sodium 1-naphthalenesulfonate strongly bind to both HMGA2 and ATHP3. Intriguingly, although the negative charge is important, it is not the only parameter that determines their inhibition potency. For example, NF110 only carries 4 negative charges. However, it strongly inhibits HMGA2-DNA interactions with an IC_{50} of 0.87 ± 0.04 μM. In contrast, although NF023 has 6 negative charges, its inhibition IC_{50} was determined to be 10.63 ± 0.46 μM, tenfold higher than that of NF110. These results suggest that both the charge and structure are very important for the inhibition of HMGA2-DNA interactions by these related compounds.

Table 2.1. Tightly binding of suramin and analogues to HMGA2 and ATHP3.

Suramin and analogous	IC_{50} (μM)	HMGA2		ATHP3	
		K_a (M^{-1})	Binding site	K_a (M^{-1})	Binding site
Suramin	2.78 ± 0.10	$4.08 \pm 0.92 \times 10^6$	2.81	$4.06 \pm 0.32 \times 10^7$	0.98
		$4.34 \pm 2.16 \times 10^4$	2		
NF449	0.43 ± 0.18	$5.53 \pm 6.14 \times 10^7$	0.89	$1.10 \pm 7.35 \times 10^{11}$	0.86
		$8.17 \pm 1.33 \times 10^5$	1.63		
NF110	0.87 ± 0.039	$2.04 \pm 0.48 \times 10^8$	2.06	$2.17 \pm 0.17 \times 10^7$	1.01
		$7.16 \pm 0.67 \times 10^5$	1.77		
NF546	5.49 ± 2.81	$1.48 \pm 0.47 \times 10^7$	1.83	$4.28 \pm 0.47 \times 10^6$	0.81
		$3.37 \pm 8.81 \times 10^5$	1.98		
NF340	6.95 ± 9.36	$7.17 \pm 4.68 \times 10^6$	2.12	$1.89 \pm 0.16 \times 10^7$	0.8
		$2.27 \pm 0.57 \times 10^5$	0.3	$1.77 \pm 0.06 \times 10^5$	1.06
NF023	10.63 ± 0.46	$2.47 \pm 4.53 \times 10^6$	1.74	$9.75 \pm 3.47 \times 10^6$	1

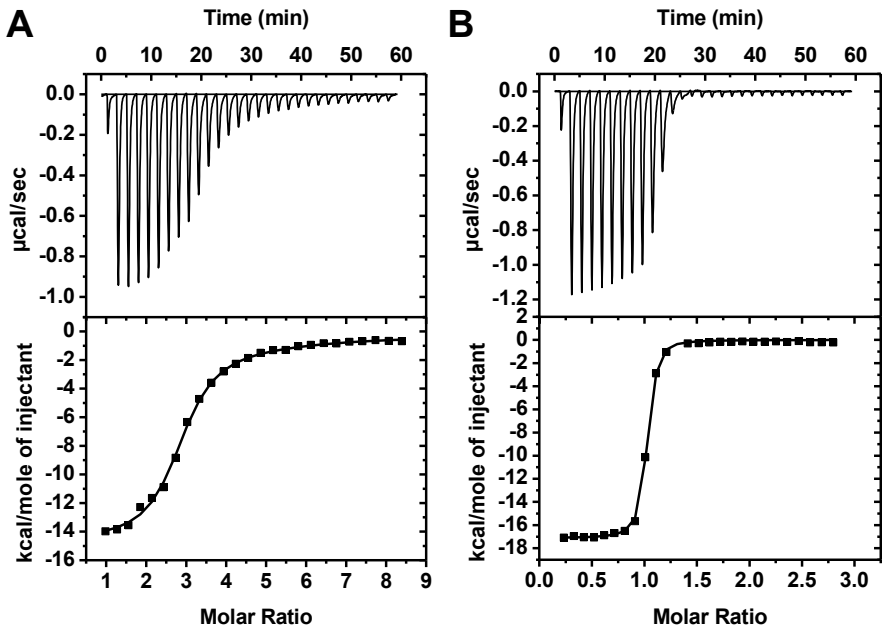


Figure 2.4. Sample raw data from isothermal titration calorimetry (ITC) experiments for the titration of suramin to HMGA2 **(A)** and ATHP3 **(B)**. ITC experiments were performed according to conditions as described in "Methods". The ITC data were fit using the software supplied by the manufacturer to yield thermodynamic parameters.

2.4.4 Effects of suramin on tumor cells.

Previous studies showed that HMGA2 plays a critical role in the epithelial-mesenchymal transition (EMT) induced by transforming growth factor-β (TGF-β) [156]. Further studies showed that HMGA2 directly binds to the promoter region of the transcriptional factor SNAIL and coregulates SNAIL expression with intracellular transducers, Smads during the EMT induction [157]. Since HMGA proteins including HMGA2 are highly expressed in glioblastomas and glioblastoma-derived brain tumor stem cells (BTSCs) [166], where they play pivotal roles in regulating self-renewal, differentiation and symmetric/asymmetric division [166, 177], BTSCs are a good cell assay system to evaluate the effect of suramin on the expression of SNAIL and several differentiation and stemness markers, such as MSI1 (Musashi RNA Binding Protein 1),

ID2 (inhibitor of DNA binding 2), and OLIG2 (oligodendrocyte transcription factor) [166]. Here, we exposed two BTSC lines (BTSC#83 and BTSC#30p) to different concentrations of suramin. Consistent with the cytotoxicity studies (Table S2.3), up to 200 μM of suramin did not significantly affect the cell growth of these two cell lines (Figure S2.8 A). Intriguingly, in the presence of 100 μM and 200 μM (Figure 2.5 A) of suramin, cells were stimulated to adhere to the bottom of the well and extend neurite-like structures (as in BTSC#83; Figure 2.5 A, top panel) or acquire at epithelial morphology (as in BTSC#30p; Figure 2.5 A, bottom panel), suggesting induction of differentiation in both BTSC cell lines. Western blotting and qRT-PCR experiments showed that suramin slightly reduced the expression of HMGA2 and HMGA1 in both cell lines (Figure 2.5 B, C; data not shown). Nevertheless, suramin significantly reduced the expressions of SNAIL, MSI1, and ID2 (Figure 2.5 C, Figure S2.8 B, D, E) and, in contrast, increased the expression of OLIG2, an oligodendrocyte differentiation marker [166] (Figure 2.5 C). Our previous results showed that silencing of another AT-hook protein HMGA1 in BTSC#83 and BTSC#30p downregulated MSI1 expression and induced cells to adhere and extend neurite-like structures or acquire at epithelial morphology [166]. Figure S2.8 also shows that silencing of HMGA1 induces reduction in ID2 expression in BTSC#30p but not in BTSC#83 (Figure S2.8 C). Similarly, treatment with 100 μM Suramin for 48 h induces reduction in ID2 expression in BTSC#30p but not in BTSC#83 (Figure S2.8 D, E), suggesting that HMGA silencing produces similar effects to suramin treatment. Although these results support a hypothesis by which suramin targets HMGA-DNA interactions in these two BTSC cell lines and therefore induces their differentiation, suramin may also work with other cellular

components/enzymes, such as protein-tyrosine phosphatases [178] and various growth factors [179], to achieve its biological functions.

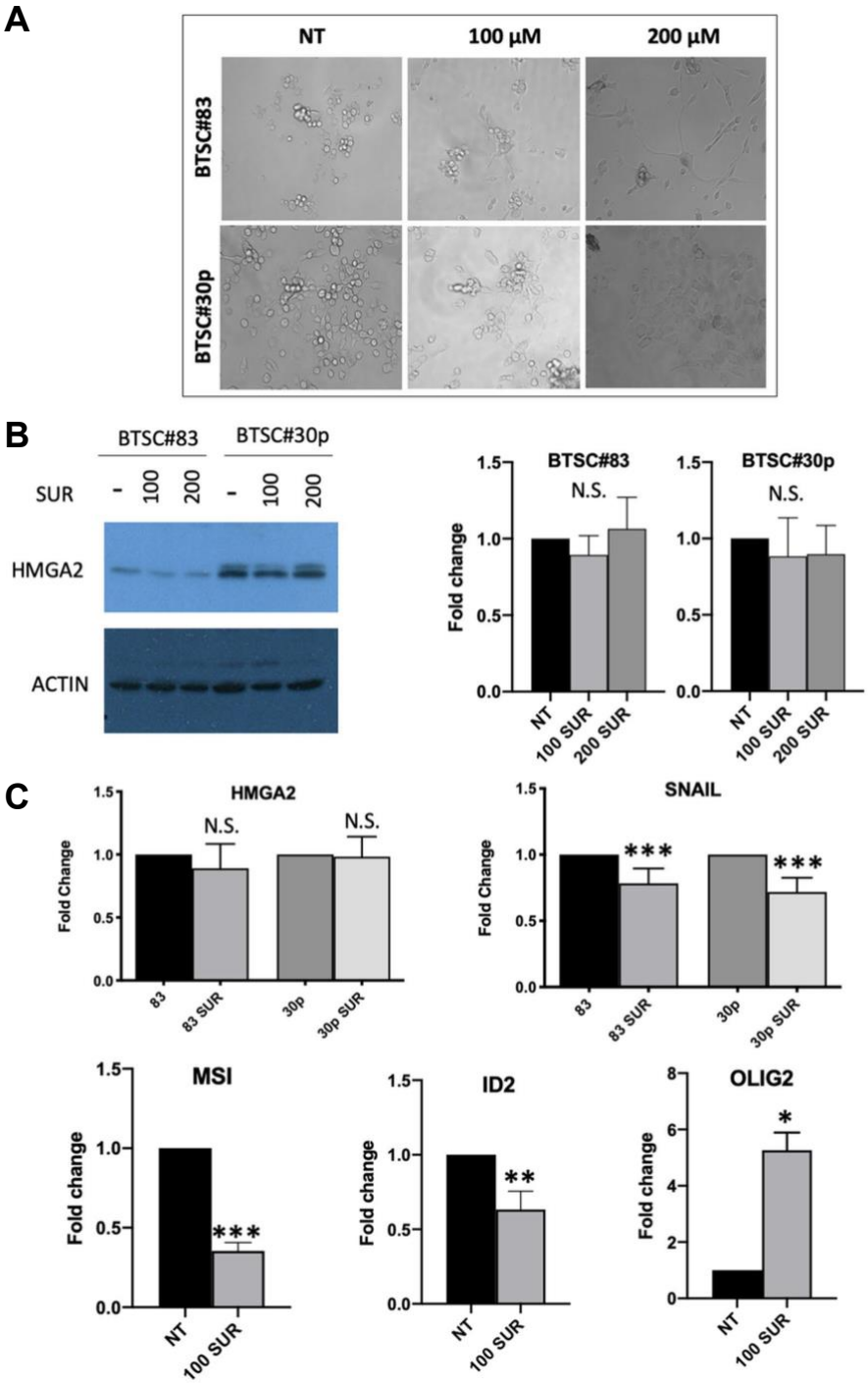


Figure 2.5. Suramin induces the differentiation of brain tumor stem cells (BTSCs) #83 and #30p. **(A)** Light Microscopy of BTSC#83 and BTSC#30p treated with various

concentrations of suramin for 6 days, compared to non-treated (NT) cells. **(B)** Western blot analysis (left panel) and densitometry (right panels) of HMGA2 expression after the 6-day treatment of suramin. This image was generated using two different blots of the same western blot by antibodies against HMGA2 (the top panel) and actin (the bottom panel), respectively. The original images are provided in the supplemental information Figure S2.11. **(C)** Expression of HMGA2, SNAIL, MSI1, ID2, and OLIG2 in suramin-treated BTSCs. qRT-PCR analyses of HMGA2 and SNAIL in BTSC#83 and BTSC#30p treated with suramin 100 μM for 48 h, compared to non-treated cells. qRT-PCR analyses for stemness (MSI1 and ID2) and differentiation markers (OLIG2) in BTSC#30p, treated with suramin 100 μM for 9 days. Fold changes are normalized for actin expression. Data represent the mean value +/− SD of two or three independent experiments performed in duplicate. (*$p \leq 0.05$; **$p \leq 0.01$; ***$p \leq 0.001$; Student's t test).

Next, we examined whether suramin is able to dissociate HMGA2 from its target DNA sequences inside cells. Since the ID2 gene promoter region contains multiple AT-rich DNA sequences and can be regulated by HMGA2 [156] presumably through binding to these AT-rich DNA sequences, Chromatin Immunoprecipitation Analysis (ChIP) was performed to determine whether suramin could dissociate HMGA2 from the ID2 gene promoter region. The role of HMGA1 and HMGA2 in regulating ID2 transcription was confirmed by analyzing its expression in mouse embryonic fibroblasts (MEFs) knock-out for HMGA1, HMGA2, and the A1/A2 double mutant, respectively [62]. These three knock-out MEFs showed a significant down-regulation of ID2 expression (Figure S2.9), indicating that both HMGA1 and HMGA2 can regulate the ID2 expression. We identified two AT-rich regions, hereafter indicated as "region A" and "region B", in the human ID2 promoter, containing the HMGA2 consensus sequence [46] (Table S2.1) for our ChIP experiments. Due to availability, the high proliferation rate, and the expression of HMGA proteins, the anaplastic thyroid carcinoma cell line 8505c was used. Briefly, after 8505c cells were treated with 100 and 200 μM suramin for 6 days, the ChIP experiments were performed. Figure S2.10 shows our results. In the absence of suramin, the sample immunoprecipitated

with the anti-HMGA2 antibody is enriched both in region A and B (Figure S2.10 A), indicating that HMGA2 is able to bind to both regions. Unexpectedly, we did not detect significant HMGA2 binding reduction in the presence of 100 μM of suramin (Figure S2.10). Nevertheless, 200 μM of suramin caused a drastic reduction in the enrichment of the anti-HMGA2 precipitated sample with respect to the sample immunoprecipitated with IgG only, indicating the loss of the binding of HMGA2 to regions A and B. These results demonstrate that suramin can inhibit HMGA2-DNA binding in 8505c cells. Similarly, the mRNA expression level of ID2 was significantly decreased in the presence of 200 and 400 μM of suramin (Figure S2.10 B).

2.5 Discussion

In this article, we report the development of a miniaturized, automated AlphaScreen uHTS assay to identify small molecule inhibitors targeting HMGA2-DNA interactions. This uHTS assay has excellent screening parameters with a Z' and S/B ratio at 0.83 and 438, respectively (Table S2.2). After screening the LOPAC1280 compound library, we found that suramin, a highly negatively charge anti-parasitic drug that is used to treat African sleeping sickness and river blindness [180, 181], is a potent inhibitor of HMGA2-DNA interactions. Further, our results show that suramin's inhibition of HMGA2-DNA interactions stems from its binding to HMGA2 with high affinity, particularly to ATHPs. Our molecular modeling studies confirmed this hypothesis (Figure 2.6). The suramin-ATHP3 complex shows that the suramin sulfonated groups are close to the Arg and Lys residues, providing ionic/hydrogen-bond interactions. The ring structures at the two ends of suramin are folded to embrace the central proline ring from either side of the residue.

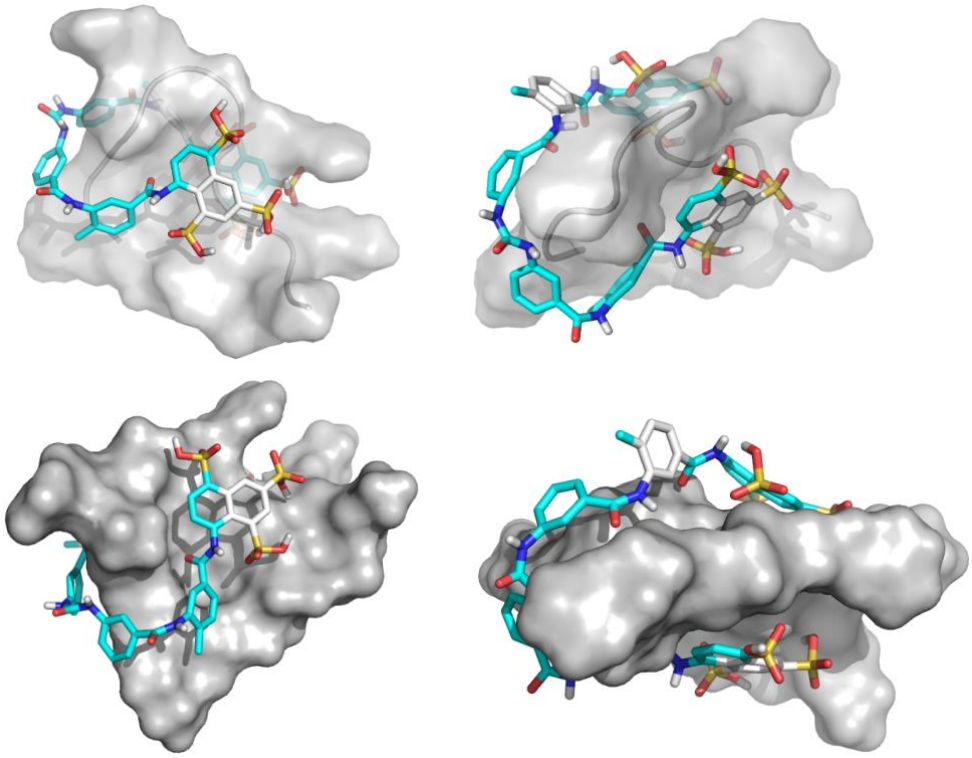


Figure 2.6. The simulated structure of the suramin-ATHP3 complex obtained by molecular docking and simulations.

DNA-binding proteins, such as transcriptional factors are excellent targets for anticancer therapy [182]. Indeed, several clinically important anticancer drugs, such as doxorubicin and cisplatin [183, 184] bind to DNA, disrupt protein-DNA interactions, and as a result prevent DNA-binding proteins including transcriptional factors from binding to their target DNA sites [185]. However, a major disadvantage of these anticancer drugs is that they bind to DNA nonspecifically. Consequently, they kill cancer cells as well as normal cells causing serious side effects. Therapeutically active compounds that specifically block oncogenic transcriptional factors from binding to their DNA recognition sites would be superior medicines. Such medicines would be expected to have fewer, less severe side effects, thereby potentially increasing the dose and duration of chemotherapy and perhaps improving treatment outcomes as well as quality of life. Transcription factors

are widely considered to be “undruggable” because they usually do not have enzymatic activities, and lack deep pockets to accommodate drug-like molecules [150]. Further, the lack of HTS assays to identify inhibitors from small molecule repositories also contributes to the perceived undruggable nature of transcriptional factors [90]. Our results refute these notions, at least with respect to the potential druggability of HMGA2. Previously, we developed protein-DNA or protein-RNA interaction enzyme-linked immunosorbent assays (PDI-ELISA or PRI-ELISA) to identify inhibitors targeting specific protein-nucleic acids interactions [90]. These methods are versatile and can be applied to any nucleic acid binding proteins as long as an antibody is available. In fact, one may use tagged proteins, such as His-tagged proteins so that antibodies against His-tag can be used in PDI-ELISA or PRI-ELISA. Regardless, as we pointed out above, because streptavidin-coated 1536-well plates are expensive, PDI-ELISA or PRI-ELISA cannot be configured into a miniaturized, automated uHTS assay in a 1536 well plate format. Additionally, too many washing steps were used, which make PDI-ELISA or PRI-ELISA a “lengthy” procedure and unsuitable for automated uHTS studies. Here we showed that AlphaScreen assays could be utilized to screen for inhibitors against HMGA-DNA interactions by using His-tagged HMGA2 linked to nickel chelate (Ni-NTA) acceptor beads. The use of AlphaScreen uHTS assays has several advantages. Because His-tag can be added to most of DNA-binding proteins and usually does not interfere with their DNA binding activities, AlphaScreen assays can be used to screen and identify inhibitors targeting most DNA-binding proteins. Another advantage is that AlphaScreen uHTS assays are cost-effective. For instance, in this study, only one 1536-well plate was used in the screening of the LOPAC1280 compound library. It is anticipated that AlphaScreen uHTS assays using His-

tagged DNA binding proteins will be utilized to identify inhibitors for other DNA-binding proteins in the future.

The most intriguing result of this uHTS study is the identification of suramin as a potent inhibitor for HMGA2-DNA interactions. The inhibition is likely through a mechanism by which suramin tightly binds to highly positively charged ATHPs. This is not surprising because suramin, a polysulphonated naphylurea, carries 6 negative charges at physiological conditions (Figure S2.4). As shown in Figure 2.6, charge-charge interactions and hydrogen bonding between the suramin sulfonated groups and Arg/Lys residues play critical roles in the binding of suramin to ATHPs. This inhibition mechanism is different from that of netropsin and other DNA minor groove binders that prevent HMGA2 from binding to the minor groove of AT-rich DNA sequences [90, 142]. In fact, suramin is the first chemical compound that was found to tightly bind to the intrinsically disordered protein (IDP) HMGA2 and the "unstructured" DNA-binding motif ATHPs. Since HMGA1 proteins also carry multiple "AT-hook" DNA-binding motifs [186, 187], suramin is expected to strongly inhibit HMGA1-DNA interactions as well. Furthermore, it is reasonable to predict that suramin tightly binds to many positively charged motifs on protein surface that contain lysine and/or arginine residues.

Suramin is a century-old drug synthesized in the 1920s by Bayer to treat human African trypanosomiasis (HAT) [188]. However, the mechanism or mode of action is still unknown [189]. Suramin is usually administered by intravenous injection due to the poor intestinal absorption [175] and binds to serum proteins, such as albumin and low-density lipoprotein (LDL), immediately after administration [175]. It is believed that the parasite takes up the drug through receptor-mediated endocytosis of the protein-bound drug [190].

Suramin is highly effective against blood-stream forms of the parasite but not very active against procyclic trypanosomes [191]. Since the bloodstream forms of *Trypanosoma brucei* lacks a functional mitochondrion and entirely depend on glycolysis for their energy needs, this led to the hypothesis that the glycolytic pathway of the parasite is the target of the drug [192]. However, so far there is no direct evidence to support this hypothesis [175]. Another interesting feature of *T. brucei* is that its mitochondrion contains the so-called kinetoplast DNA (kDNA), comprising of ~ 73% AT base pairs [193]. Recent studies showed that certain DNA minor groove-binding compounds were able to enter the mitochondrion of *T. brucei*, bind to AT sequences of kDNA and, as a result, cause cell death [194]. It was suggested that these minor groove-binding compounds might disrupt the functions of kDNA binding proteins, such as the TbKAP6 protein that is essential for kDNA replication and maintenance and also for cell viability [195]. TbKAP6 contains two HMG boxes and binds to DNA minor groove [195]. Interestingly, this protein carries several highly positively charged motifs similar to ATHP [195]. One possibility is that suramin tightly binds to these positively charged motifs in TbKAP6 and prevents its binding to kDNA. Further studies are needed to support this hypothetic mechanism.

Another interesting function of suramin is its well-studied anticancer and anti-metastasis activities [179, 196, 197] although the mechanism is still obscure. Possible mechanisms include inhibiting important enzymes, such as cullin-RING E3 ubiquitin ligases [198], protein-tyrosine phosphatases [178], and blocking various growth factors binding to their receptors [179]. In fact, suramin has “so many” targets that its cellular and physiological activities are expected to stem from not just one single target but multiple targets [199].

Here we reveal another potential target, HMGA2 and/or HMGA proteins for this intriguing small molecule. It is possible that binding of suramin to HMGA2 prevents the protein from binding to the target DNA sequence of the SNAIL promoter [157] and decreases the expression of SNAIL (Figure 2.5 C) [156, 157]. The differentiation of BTSCs induced by suramin is consistent with this hypothesis (Figure 2.5A). Our ChIP analysis showing that high concentrations of suramin can dissociate HMGA2 from its target DNA sequences inside cells is also consistent with this hypothesis (Figure S2.10). Nevertheless, it will be difficult to design experiments demonstrating that HMGA2 and/or HMGA proteins are the suramin's only target or main target in cells. What we learned so far is that suramin likely prefers binding to positively charge motifs on protein surface similar to the way it binds to the "AT-hook" DNA-binding motifs (Figure 2.6). It is also possible to identify a specific inhibitor for HMGA2 and HMGA proteins with the established uHTS in the future. Please note that suramin has several disadvantages for further development. First, suramin binds to many proteins in vitro and in vivo. It may not be very specific for HMGA2 targeting. Secondly, suramin is not orally bioavailable and must be given intravenously. Third, it has limited bioavailability through tissue due to the fact that 98% of the drug are protein bound in blood serum. Furthermore, suramin causes a fair number of side effects and unpredictable pharmacokinetics. Based on the study results of structure activity relationship (SAR) of suramin by using several suramin analogs, NF110 which carries 4 negative charges inhibits HMGA2-DNA interactions with a lower IC_{50} compared with suramin. We proposed NF110 compound to be a better drug to target HMGA2 and used to treat HMGA2 related disease based on the Lipinski's rule of five [200]. Such as, NF110 meets some of the criteria of drug like four simple physicochemical parameter ranges

(MWT $\leq$ 500, log P $\leq$ 5, H-bond donors $\leq$ 5, H- bond acceptors $\leq$ 10). More detailed studies will be need in the future to confirm our hypothesis.

CHAPTER 3: Inhibition of HMGA2 binding to AT-rich DNA by its negatively charged C-terminus

3.1 Abstract

The mammalian high mobility group protein AT-hook 2 (HMGA2) is a small DNA-binding protein that specifically binds to and bends the AT-rich DNA sequences. It works as a general transcriptional factor and plays an important role in tumorigenesis, adipogenesis, and stem cell youth. HMGA2 is an intrinsically disordered protein (IDP) consisting of three positively charged “AT-hook” DNA binding motifs and a highly negatively charged C-terminus. Previous results showed that the HMGA2 could self-associate into homodimers in aqueous buffer solution through the electrostatic interactions between the positively charged “AT-hooks" and the negatively charged C-terminus, suggesting that the negatively charged C-terminus should significantly modulate the “AT-hooks” and HMGA2 binding to AT-rich DNA sequences. Indeed, new results in this dissertation demonstrate that the negatively charged C-terminus greatly affected the DNA-binding properties of HMGA2. For instance, the C-terminal deletion mutant HMGA2Δ95-108 binds much more tightly to an AT-rich DNA oligomer FL814 compared with the wildtype HMGA2. A synthetic peptide derived from the C-terminus of HMGA2 (the C-terminus peptide or CTP) strongly inhibits HMGA2 binding to FL814. This inhibition is most likely through its binding to the positively charged “AT-hook” motifs as demonstrated by different biochemical and biophysical methods, such as PDI-ELISA, fluorescence anisotropy, and AKTA FPLC size exclusion chromatography. Furthermore, molecular modeling studies show that electrostatic interactions and hydrogen bonding are the major forces for the CTP binding to the “AT-hooks.” This study suggests that the CTP

or CTP mimics of HMGA2 may be used as a potent inhibitor to block HMGA2 binding to AT-rich DNA sequences.

3.2 Introduction

The mammalian high mobility group protein AT-hook 2 (HMGA2) is a nuclear protein that plays a critical role in the epithelial-mesenchymal transition (EMT) during cell development and differentiation [15, 156]. It is expressed at the embryonic stage and usually not expressed in the adult and differentiated tissues. HMGA2's chromosome location is at 12q14.3 [16]. It consists of three highly positively charged "AT-hooks" motifs which specifically bind to the minor groove of AT-rich DNA sequences and a negatively charged C-terminus that plays an important role in protein-protein interactions [4, 46]. Previous studies have shown that HMGA2 was related to fat cell proliferation and obesity [61, 89, 159]. HMGA2 knockout mice were severely deficient in fat cells and developed pygmy phenotype [59]. Disruption of HMGA2 caused a reduction in the obesity induced by leptin deficiency (Lep^{ob}/Lep^{ob}) in a gene-dose-dependent manner in mice [61]. These results indicate that HMGA2 may be a potential target for the treatment of obesity. Along with being a potential target for the treatment of obesity, HMGA2 is related to oncogenesis. Aberrant and/or over expression of HMGA2 was directly associated with the formation of a variety of tumors. It includes benign tumors of mesenchymal origin, such as lipomas [160], uterine leiomyomas [161], and fibroadenomas of the breast [201], and malignant tumors, such as lung cancer [163, 202], breast cancer [134], prostate cancer [136, 203], and melanoma [139, 140]. The expression level of HMGA2 often correlates with the degree of malignancy and the existence of metastasis [163, 204, 205]. These results indicate that HMGA2 is a potential therapeutic target of anti-cancer drugs [118, 120]. Additionally,

some studies showed that HMGA2 is also related to human height [68], human intelligence [77], and stem cell youth [101].

HMGA2 is a member of the high-mobility group family of the DNA-binding architectural transcription factors, which enables it to bind to DNA to participate in the conformational regulation of active chromatin on its specific downstream target genes. It is thought to modulate the assembly and the function of transcriptional complexes. Previous studies identified that phosphorylation of HMGA proteins by casein kinase 2 (CK2) altered its confirmation and modulated its DNA binding properties by attenuating binding affinity and reducing the extent of contacts between the DNA and protein [206, 207]. HMGA2 has a highly negatively charged C-terminus, a 14 amino acid residue which contains 3 serine residues and 2 threonine residues that could be phosphorylated by CK2. After fully phosphorylated by CK2, there will be totally 18 negative charges on the C-terminus at physiological conditions, 10 negative charges from phosphate groups and 8 negative charges from 7 glutamic acid residues, and 1 aspartic acid residue respectively. Phosphorylated C-terminus is a super negatively charged motif and it may play a major role in reducing the binding affinity of DNA and protein as the electrostatic interaction is an important force for their binding [50]. HMGA2 is an intrinsically disordered protein (IDP) with a unique feature of asymmetric charge distribution over its primary structure [176]. Previously, our lab identified that HMGA2 self-associated to homodimer through electrostatic interactions between the highly positively charged “AT-hooks” and the highly negatively charged C-terminus in aqueous solutions [21]. All the information mentioned above indicates that the C-terminus of HMGA2 could regulate the DNA-binding affinity

of HMGA2, the interference of HMGA2-DNA interactions and involve HMGA-HMGA partner protein interactions [55].

Here, in this work the truncated HMGA2 (HMGA2Δ95-108) binds much more tightly to AT-rich DNA sequences compared with the wildtype HMGA2 that was demonstrated. The C-terminal motif binding to "AT-hooks" of HMGA2, resulted in inhibiting HMGA2 binding to AT rich DNA sequences. The C-terminus of HMGA2 worked as a good inhibitor to target HMGA2 to block its binding to AT rich sequences for therapeutic purposes.

3.3 Materials and methods

3.3.1 Proteins and other reagents.

HMGA2 and HMGA2Δ95-108 were purified as described previously [50, 176]. An extinction coefficient of 5810 M^{-1} cm^{-1} at 280 nm was used to determine their concentration [176]. All peptides including the CTP (the C-terminal peptide, CETEETSSQESAEED), the positively charged AT-hook 2 peptide (CEPSPKRPRGRPK), and the positively charged AT-hook 3 peptide (KRPRGRPRKW) were custom-synthesized by Advanced ChemTech, Inc. Biotin-labeled DNA oligomer FL814 carrying a specific SELEX binding site of HMGA2 [46, 90] was purchased from Eurofins MWG Operon, Inc. The tetramethylrhodamine-5-maleimide (TMR) and tris-(2-carboxyethyl) phosphine hydrochloride (TCEP) were obtained from Molecular Probes, Inc. Betaine and the N-Acetyl-L-tryptophan was purchased from the Sigma Aldrich Chemical Co., Inc. Glycerol and Thermo Scientific™ Pierce™ High Binding Streptavidin Coated Plates used were bought from Thermo Fisher Scientific, Inc. Antibodies HMGA2 (D1A7) Rabbit mAb and Anti-rabbit IgG, HRP-linked Antibody were purchased from Cell Signaling Technology, Inc. All chemicals were used without further purification.

3.3.2 Protein-DNA interaction ELISA (PDI-ELISA) assay.

The PDI-ELSA assays were described previously [90] and used to determine the apparent DNA dissociation constant (K_d) by nonlinear-least-squares fitting. The following equation was used: $y = \frac{(a+x+K_d)-\sqrt{(a+x+K_d)^2-4ax}}{2a}$, where y, a, and x represented the DNA-binding ratio, the total DNA concentration and the total protein concentration. The PDI-ELISA assay was also used to determine the inhibition IC_{50} against HMGA2-DNA interactions. The apparent inhibitory IC_{50} values were obtained using the following equation: $y = \min + (\frac{(\max-\min)}{1+10^{((x-\log IC_{50})\times \text{Hill slope}}})$, where x, y, max, and min represented the inhibitor's concentration, the inhibition level, the maximum inhibition value, and the minimum inhibition value, respectively.

3.3.3 Isothermal titration calorimetry (ITC).

ITC experiments were conducted using a VP-ITC titration calorimeter (MicroCal, Inc., Northampton, MA) interfaced to a computer with Origin 7.0, supplied by the manufacturer and used for data acquisition. The samples were extensively dialyzed against 1×BPE buffer 6 mM Na_2HPO_4, 2 mM NaH_2PO_4, and 1 mM Na_2EDTA) plus 184 mM NaCl or 1×BPE buffer plus 184 mM NaCl and one osmolyte, such as betaine. Typically, 15 µL of HMGA2 (50 µM) or HMGA2Δ95-108 (30 µM) was injected every 200 seconds, up to a total of 18 injections, into a FL814 sample (1.44 mL of 2.5 µM and 2 µM, respectively) in 1×BPE buffer plus 184 mM NaCl in the sample cell. The heat liberated or absorbed was observed as a peak corresponding to the power required to keep the sample and reference cells at identical temperatures. The peak produced during the injection was converted to heat output by integration and corrected for the cells volume and sample concentration. Control

experiments were also carried out to determine the contribution of the heat of dilution arising from (1) HMGA2 or HMGA2Δ95-108 into buffer and (2) buffer into FL814. The net enthalpy for the titration reaction was determined by subtraction of the component heats of dilution. The ITC data were fitted using the built-in curve fitting model for a single set of identical binding sites to obtain the thermodynamic parameters.

3.3.4 Labeling CTP with TMR.

Since CTP contained a cysteine residue, it can be labeled with the fluorescence reagent TMR through the thiol group. Specifically, 100 μM of CTP was incubated with 200 μM of TMR in the presence of 200 μM TCEP in 10 mM sodium phosphate buffer (pH 7.2) containing 20 mM NaCl for 2 hours at room temperature to yield TMR labeled CTP (CTP-TMR). To purify CTP-TMR, the labeling mixtures were loaded onto a pre-equilibrated Q-sepharose column (200 μL) and CTP-TMR was eluted using 10 mM sodium phosphate buffer (pH 7.2) containing 500 mM of NaCl. An extinction coefficient of 95000 M^{-1} cm^{-1} at 541 nm in methanol was used to determine the CTP-TMR concentration. The purity of CTP-TMR was verified by using size exclusion chromatography (Superdex Peptide 10/300 GL) of AKTA FPLC.

3.3.5 Fluorescence anisotropy experiments.

Fluorescence anisotropy experiments were used to determine the binding constants of CTP to HMGA2, HMGA2Δ95-108, or ATHPs using the Varian Cary Eclipse fluorescence spectrophotometer with the excitation wavelength at 543 nm and the emission wavelength at 575 nm. Excitation and emission slits are 5 and 10 nm, respectively. For the CTP-TMR binding experiments, increasing concentrations of HMGA2, HMGA2Δ95-108, and ATHPs were titrated into a fixed concentration of CTP-TMR in a buffer solution containing 10

mM Tris-HCl (pH 8.0) and 20 mM NaCl at room temperature. The anisotropy values of CTP-TMR were determined and calculated by the equation [208]: $r = \frac{I_{VV} - GI_{VH}}{I_{VV} + 2GI_{VH}}$, where r is the calculated anisotropy, I_{VV} is the observed polarized intensity corresponds to vertically polarized excitation and vertically polarized emission, I_{VH} is the observed polarized intensity corresponds to vertically polarized excitation and horizontally polarized emission, G factor is the ratio of the sensitivities of the detection system for vertically and horizontally polarized light. The binding data were fit with the following equation [209]: $y = \frac{(a+x+K_d) - \sqrt{(a+x+K_d)^2 - 4ax}}{2a}$, y is the binding ratio of A/A_{max}, where A and A_{max} are the relative and maximum anisotropy, respectively; x is the titrate concentration, a is the CTP-TMR concentration, and K_d the dissociation constant.

3.3.6 AKTA FPLC size exclusion chromatography.

Size exclusion chromatography in an AKTA FPLC system (Amersham Biosciences, US) was used to confirm that the CTP-TMR binds to HMGA2 and ATHP3. Specifically, 10 µM of CTP-TMR was mixed with 10 µM of ATHP3 in 10 mM Tris-HCl (pH 8.0) buffer plus 20 mM NaCl and incubated for 30 min at room temperature. The CTP-TMR and ATHP3 mixture was loaded onto to a Superdex Peptide 10/300 GL column and eluted with 10 mM Tris-HCl (pH 8.0) plus 20 mM NaCl at at a flow rate of 0.5 mL/min. HMGA2 (MW of 11,820 Da) and N-Acetyl-L-tryptophan (MW of 246.3 Da) were used to calibrate the column.

3.3.7 Molecular dynamics simulation.

The molecular dynamics (MD) simulation of CTP and ATHP3 was performed using the Charmm-Gui web interface [169]. The peptide was solvated in a cubic water box with

TIP3 water and the system was neutralized by adding 36 Cl^- ions. The final system contained ~23,000 atoms. The NAMD 2.14_CUDA (Compute Unified Device Architecture) [39] was used to perform all-atom molecular dynamics with CHARMM36M force field [37]. The particle mesh Ewald (PME) method [171] was used for the calculation of the long-range ionic interactions. The system was minimized for 100,000 steps, followed by a 250 ps equilibration at 310 K with 1 fs time step. A 100-ns production simulation with 2-fs time step was performed at constant pressure of 1 atm and T = 310 K. The Nose-Hoover Langevin-piston method [172] was used for pressure coupling, with a piston period of 50 fs and a decay of 25 fs. The Langevin temperature coupling with a friction coefficient of 1 ps^{-1} was used to maintain the temperature.

3.4 Results

3.4.1 The apparent DNA binding constants of HMGA2 and the truncated mutant HMGA2Δ95-108 binding to AT-rich DNA oligomer FL814 determined with PDI-ELISA and ITC.

The mouse HMGA2 is a 108 amino acid residue protein, which carries three positively charged “AT-hooks" and a highly negatively charged C-terminus (Figure 3.1 A) [176]. The 94 amino acid truncated mutant HMGA2Δ95-108 also contains the 3 positively charged “AT-hooks” and, however, lacks the negatively charged C-terminus [21]. The apparent DNA binding constants of HMGA2 and HMGA2Δ95-108 binding to AT-rich DNA oligomer FL814 were determined to be $1.18 \pm 0.16 \times 10^6$ and $5.04 \pm 0.60 \times 10^7$ M^{-1}, respectively, by using HMGA2-DNA PDI-ELISA titration experiments (Figure 3.2 A). Apparently, HMGA2Δ95-108 binds more tightly to AT-rich FL814 than HMGA2 does,

suggesting that the negatively charged C-terminus significantly affects HMGA2 binding to AT-rich DNA.

Since HMGA2 and HMGA2Δ95-108 are highly positively charged proteins, we next examined the salt dependence of HMGA2 and HMGA2Δ95-108 binding to FL814 using PDI-ELISA titration experiments in 2×SSCT buffer in the presence of a various concentrations of NaCl (0.390-0.572 M) at room temperature. Figure 3.2B shows our results. It is apparent that the DNA binding constant decreases with increasing concentrations of NaCl for both HMGA2 and HMGA2Δ95-108. According to the polyelectrolyte theory of Record *et al*. [210], the slopes of the linear lines in Figure 3.2 B can be used to calculate the counter ions associated with the binding reaction with the following equation: $SK = \frac{\partial \ln K}{\partial \ln[Na^+]} = -Z\psi$, where Z is the charge on the ligands and ψ is the fraction of counter ions associated with each DNA phosphate group (ψ = 0.88 for the double-stranded B-form of DNA). The quantity SK is equivalent to the number of the counter ions released upon binding of a ligand with net charge *Z*. It was found that 8.4 ± 0.5 and 9.6 ± 1.0 counter ions, respectively, were released from DNA upon HMGA2 and HMGA2Δ95-108 binding to FL814. According to these results, we calculated the charges of $+9.6 \pm 0.6$ for the HMGA2 binding to FL814 and $+10.9 \pm 1.2$ for HMGA2Δ95-108 binding to FL814 (Table 3.1). These results are consistent with our previously published results [50].

The binding properties of HMGA2 and HMGA2Δ95-108 binding to AT-rich DNA oligomer FL814 were also investigated using ITC. Due to technical difficulties, 2×SSCT buffer could not be used. 1×BPE containing 184 mM NaCl were used here. Specifically,

50 μM of HMGA2 or 30 μM of HMGA2Δ95-108 was titrated into a 2.5 or 2 μM FL814 sample, respectively, in 1×BPE containing 184 mM NaCl in an Microcal calorimeter at 25 ℃. Figure 3.2 C and D show the results. DNA binding enthalpies of HMGA2 and HMGA2Δ95-108 binding to FL814 were estimated to be -14.04 and -9.95 kcal/mol, respectively. The ITC experiments cannot be used to obtain/estimate the DNA binding constants for HMGA2 and HMGA2Δ95-108 binding to FL814 due to the fact that very high concentrations of proteins and DNA were used for a detectible heat signal. With the salt dependent binding results, we calculated the DNA binding constants of HMGA2 and HMGA2Δ95-108 to be $3.29 \pm 0.45 \times 10^{8}$ and $3.13 \pm 0.32 \times 10^{10}$ M^{-1}, respectively. The thermodynamic parameters of HMGA2 binding to FL814 are: ΔG, -11.62 kcal/mol; ΔH, -14.04 kcal/mol; and -TΔS, 2.42 kcal/mol (Figure 3.2 E). The thermodynamic parameters of HMGA2Δ95-108 binding to FL814 are: ΔG, -14.32 kcal/mol; ΔH, -9.95 kcal/mol; and -TΔS, -4.37 kcal/mol (Figure 3.2 E).

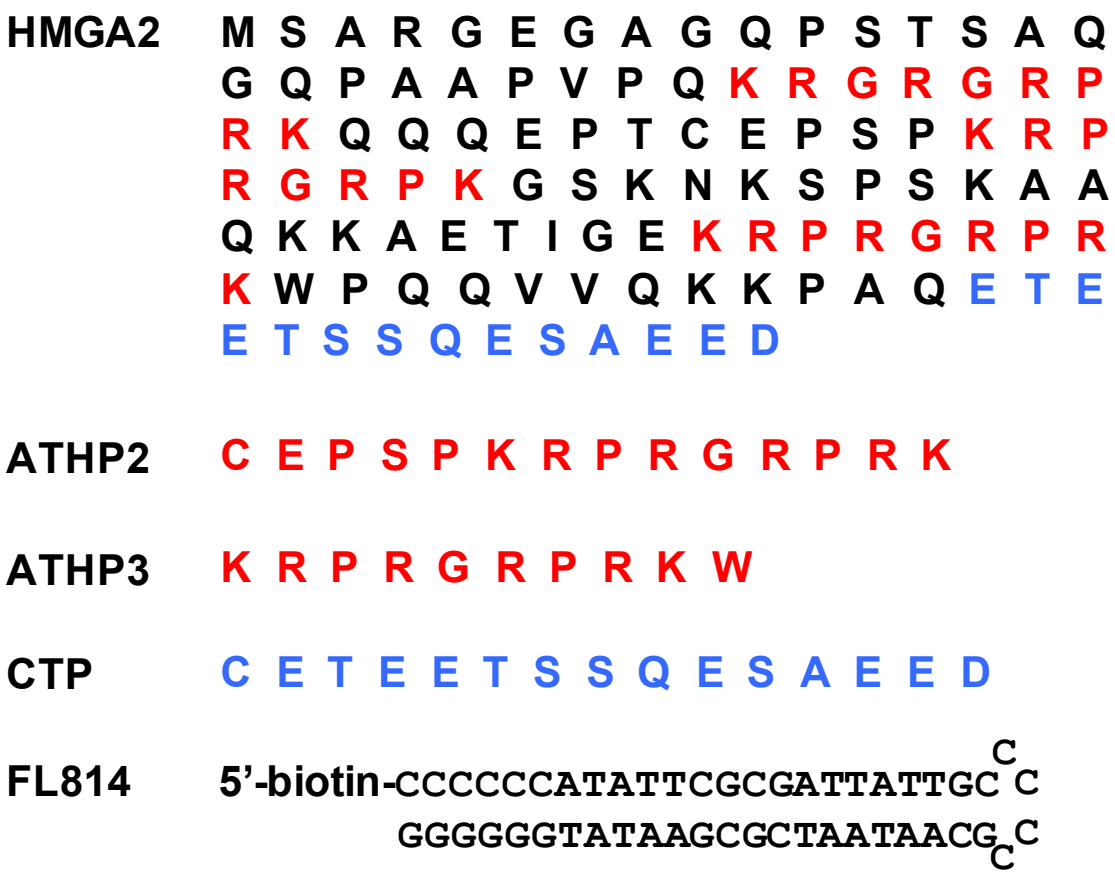


Figure 3.1. Amino acid sequences of HMGA2, ATHP2, ATHP3, and CTP. The DNA sequence of AT-rich DNA oligomer FL814.

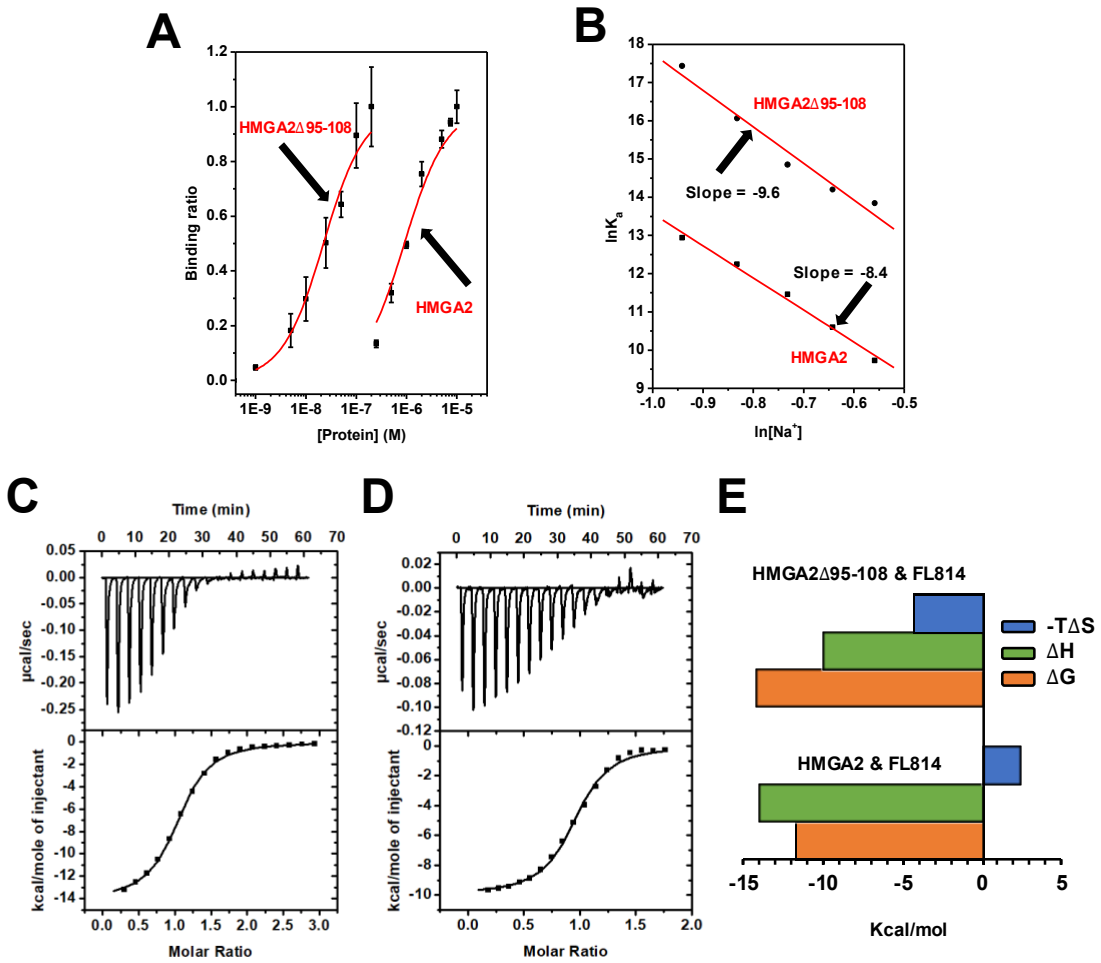


Figure 3.2. HMGA2Δ95-108 binds more tightly to the AT-rich DNA oligomer FL814 than HMGA2 does. **(A)** The PDI-ELISA titration experiments were used to determine the DNA binding constants of HMGA2 and HMGA2Δ95-108 binding to FL814 in 2×SSCT buffer at room temperature. **(B)** The DNA binding constants of HMGA2 (squares) and HMGA2Δ95-108 (circles) binding to FL814 are dependent on the Na^+ concentration at room temperature in 2×SSCT buffer. **(C)** and **(D)** Isothermal titration calorimetry experiments were used to study HMGA2 **(C)** and HMGA2Δ95-108 **(D)** binding to FL814 in 1×BPE containing 184 mM NaCl. **(E)** Comparison of the thermodynamic profiles for HMGA2 and HMGA2Δ95-108 binding to FL814.

Table 3.1. Summary of counterions and hydration changes for HMGA2 and HMGA2Δ95-108 binding to DNA.

	Counterions	Ligand charges	$\frac{\partial \ln(K_s/K_0)}{\partial [Osm]}$	Δn_w
HMGA2	-8.4 ± 0.5	+9.6 ± 0.6	-0.002 ± 0.001	+0.12 ± 0.08
HMGA2Δ95-108	-9.6 ± 1.0	+10.9 ± 1.2	+0.006 ± 0.006	-0.31 ± 0.33

3.4.2 Hydration changes for HMGA2 and HMGA2Δ95-108 binding to AT-rich DNA oligomer FL814.

Since hydration plays important roles in the protein-DNA interactions [211], we decided to investigate the hydration changes for HMGA2 and HMGA2Δ95-108 binding to FL814 using PDI-ELISA experiments in the presence of an osmolytes, such as betaine or glycerol, that perturbs the water activity. Figure 3.3 shows our results. To our surprise, high concentrations of osmolytes do not significantly affect HMGA2 or HMGA2Δ95-108 binding to FL814. These results were confirmed by an ITC experiment in which different concentrations of HMGA2 were titrated into FL814 solution in the presence of 1 Osm of betaine (Figure 3.4). The ITC profiles and thermodynamic paraments derived from this ITC experiment is almost identical to those calculated from the ITC experiment in Figure 3.2 C, suggesting that hydration does not greatly affect the binding of HMGA2 to FL814. Assuming that the osmolytes do not directly interact with FL814 or HMGA2, the change in hydration can be estimated by using the equation [212, 213]: $\frac{\partial \ln(K_s/K_0)}{\partial [Osm]} = -\Delta n_w/55.5$, where K_s and K_0 are the binding constant in the presence or absence of an osmolyte, respectively. “Osm" is the osmolality of the solution. Δn_w is the difference of the bound water molecules between the protein-DNA complex and the free reactants, HMGA2 or

HMGA2Δ95-108 and FL814. The hydration results are summarized in Table 3.1. According to our calculation, the binding reaction between HMGA2 or HMGA2Δ95-108 and FL814 does not add or release the water molecules.

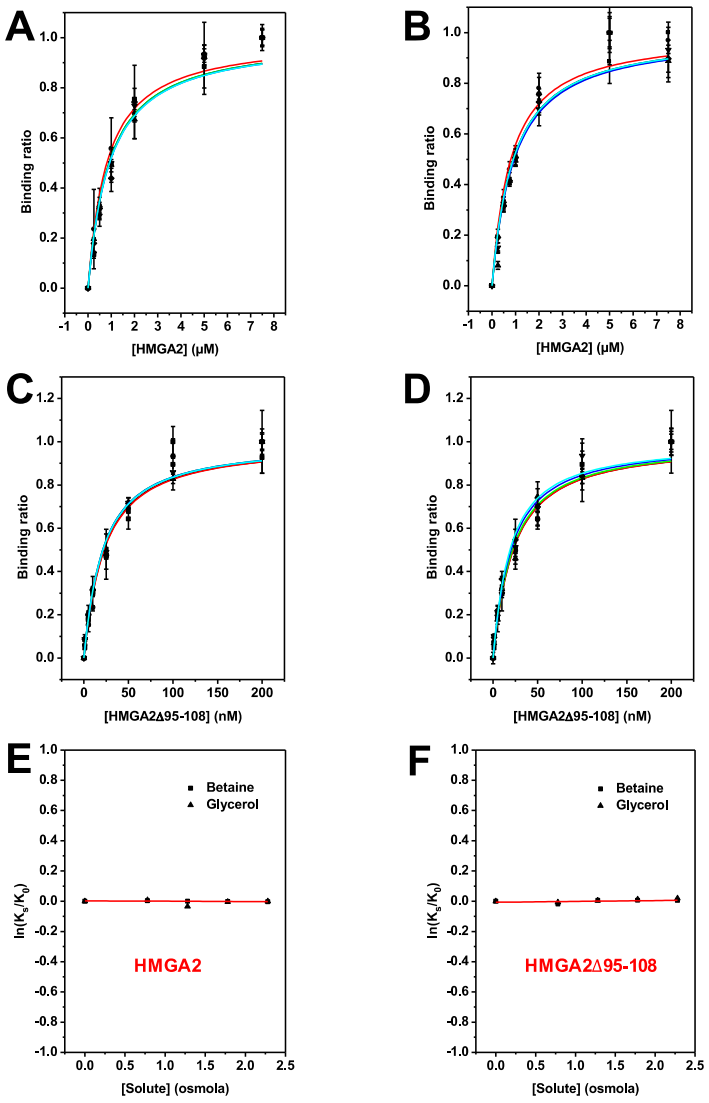


Figure 3.3. The DNA-binding experiments of HMGA2 and HMGA2Δ95-108 binding to FL814 in 2×SSCT buffer containing an osmolyte. **(A)** Betaine, **(B)** Glycerol, **(C)** Betaine, **(D)** Glycerol. Symbols: 0 Osm, black squares; 0.78 Osm, red circles; 1.28 Osm, green triangles; 1.78 Osm, blue inverted triangles; and 2.28 Osm, cyan rhombus. The standard deviation was calculated according to three independent titration experiments. **(E)** and **(F)** The DNA binding constants of HMGA2 **(E)** and HMGA2Δ95-108 **(F)** are independent of the osmolyte concentrations.

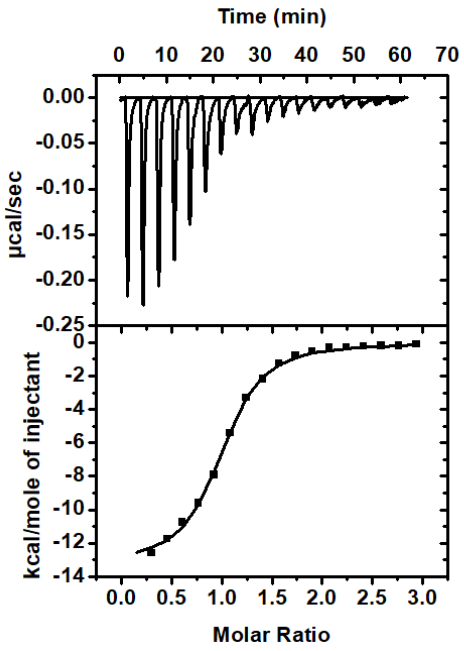


Figure 3.4. An ITC experiment of HMGA2 titrated to FL814 in 1×BPE plus 184 mM NaCl in the presence of 1 Osm of betaine.

3.4.3. Binding of the C-terminal peptide (CTP) to HMGA2, HMGA2Δ95-108, and ATHPs.

HMGA2Δ95-108 binding more tightly to FL814 suggests that the negatively charged C-terminal motif of HMGA2 inhibits HMGA2 binding to AT-rich FL814. This result is consistent with our previously published results showing that HMGA2 self-associates into homodimers in aqueous buffer solution and the self-association stems from the negatively charged C-terminus interacting with the positively charged AT-hooks of HMGA2 through electrostatic interactions [21]. In this study, we designed and purchased a 15 aa synthetic peptide the C-terminal peptide (CTP) containing the 14 aa C-terminus (Figure 3.1). A cysteine residue was added to the N-terminus for the labeling of the CTP with tetramethylrhodamine-5-maleimide and fluoroecien-5-maleimide. If the hypothesis by which the negatively charged C-terminal motif of HMGA2 inhibits HMGA2 binding to AT-rich DNA is correct, the CTP should inhibit HMGA2 and HMGA2Δ95-108 binding to

FL814. Figure 3.5 shows the results of PDI-ELISA experiments in which the CTP was titrated into the HMGA2-FL814 or HMGA2Δ95-108-FL814 complexes in 2×SSCT. The CTP indeed inhibited HMGA2 or HMGA2Δ95-108 binding to AT-rich FL814. These two PDI-ELISA experiments yield an IC_{50} of 52.36 ± 8.42 and 58.36 ± 6.53 μM, respectively, for the CTP inhibiting HMGA2 or HMGA2Δ95-108 binding to AT-rich FL814. A likely mechanism for the inhibition is that the CTP directly binds to the "AT-hooks" and prevent the "AT-hooks" binding to DNA.

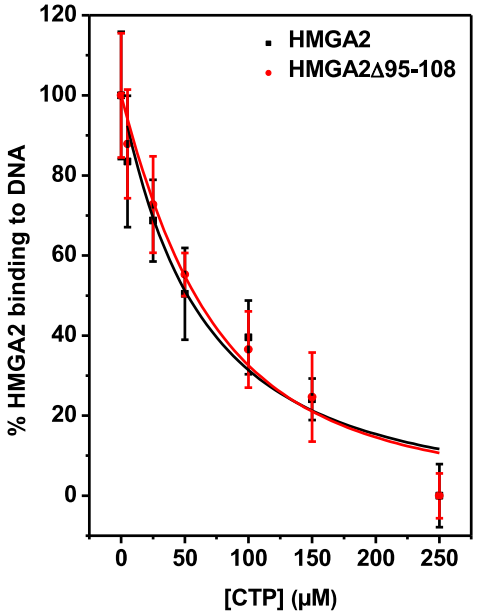


Figure 3.5. The C-terminal peptide (CTP) potently inhibits HMGA2 (black line and squares) and HMGA2Δ95-108 (red line and circles) binding to AT-rich DNA oligomer FL814 in 2×SSCT.

We next labeled the CTP with tetramethylrhodamine-5-maleimide (TMR) to produce a fluorescent CTP-TMR and then used fluorescence anisotropy to investigate how the CTP interacts with HMGA2, HMGA2Δ95-108, ATHP2, and ATHP3. Figure 3.6 shows our results. Consistent with our hypothesis, the CTP binds to HMGA2, HMGA2Δ95-108, ATHP2, and ATHP3. These fluorescence anisotropy titration experiments also yield

binding constants of $2.59 \pm 0.20 \times 10^6$, $4.49 \pm 0.52 \times 10^7$, $5.03 \pm 0.63 \times 10^5$, $6.06 \pm 1.54 \times 10^5\ M^{-1}$, respectively, for the CTP binding to HMGA2, HMGA2Δ95-108, ATHP2, and ATHP3 in 10 mM Tris-HCl (pH 8.0) containing 20 mM NaCl.

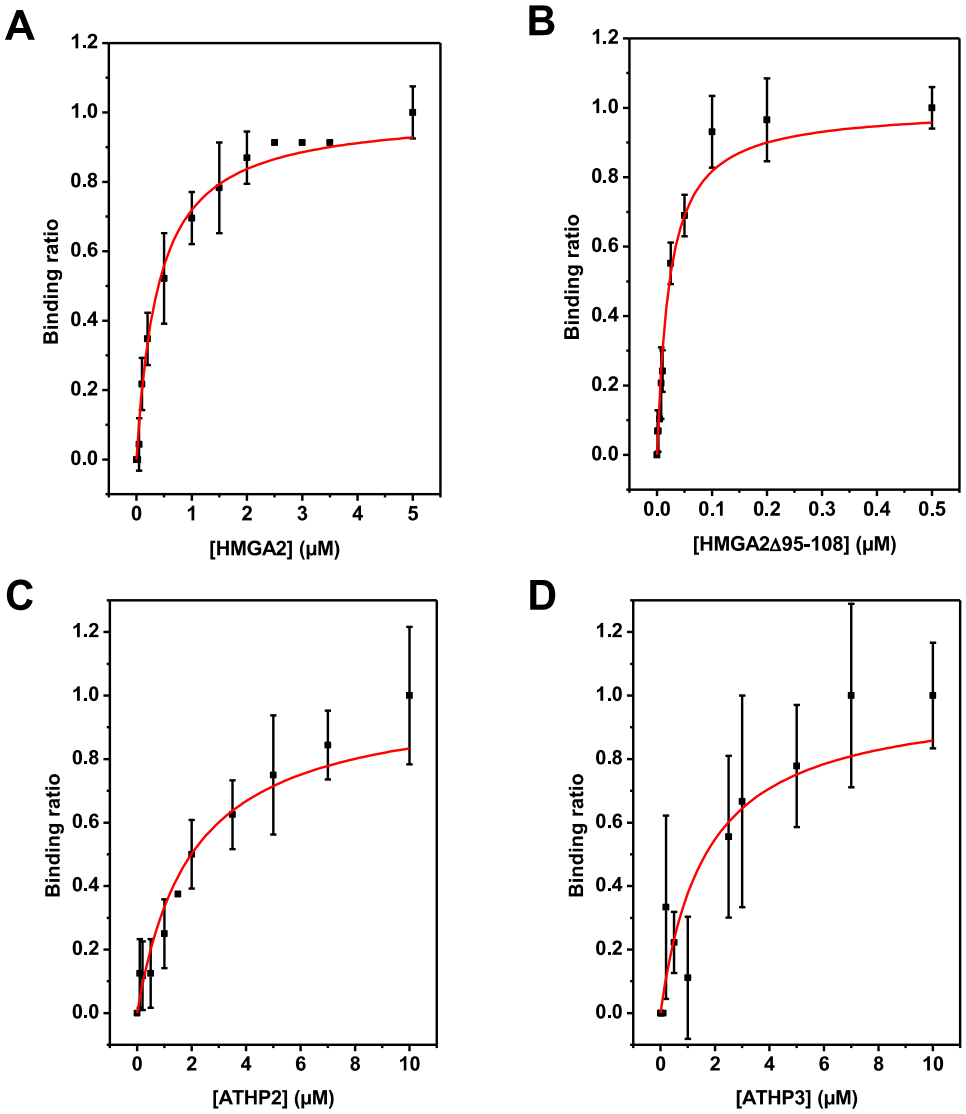


Figure 3.6. Fluorescence anisotropy titration experiments show that the CTP tightly binds to HMGA2 with $K_a = 2.59 \pm 0.20 \times 10^6\ M^{-1}$ (A), HMGA2Δ95-108 with $K_a = 4.49 \pm 0.52 \times 10^7\ M^{-1}$ **(B)**, ATHP2 with $K_a = 5.03 \pm 0.63 \times 10^5\ M^{-1}$ **(C)** and ATHP3 with $K_a = 6.06 \pm 1.54 \times 10^5\ M^{-1}$ **(D)** in 10 mM Tris-HCl (pH 8.0) buffer plus 20 mM NaCl at room temperature.

We also used a size exclusion chromatography experiment with a Superdex Peptide 10/300 GL column in an AKTA FPLC system to confirm that the CTP binds to ATHP3. Figure 3.7 shows our experimental results. HMGA2 (MW of 11,820 Da) and N-Acetyl-L-tryptophan (MW of 246.3 Da) were used to calibrate this column at different buffer conditions. Due to technical difficulties, we had to use two different buffer solutions/conditions for the exclusion chromatography experiments. Figure 3.7 A shows the exclusion chromatography experimental results in 10 mM Tris-HCl (pH 8.0) containing 200 mM NaCl. In this buffer condition, CTP-TMR and ATHP3 did not strongly interact with the column and eluted as unique peaks as the mobile phase went through the column. Two peaks of the CTP-TMR chromatography profile suggest that CTP-TMR interacts with the resin of the Superdex Peptide 10/300 GL column under this experimental condition. Figure 3.7 B shows the exclusion chromatography experimental results in 10 mM Tris-HCl (pH 8.0) containing 20 mM NaCl. Under this buffer condition, the mixture with equal amount of CTP-TMR and ATHP3 was applied to the column and eluted as a single peak indicating that CTP-TMR tightly binds to ATHP3. Please note that the single peak was eluted at 10.96 mL, earlier than the peaks of the CTP-TMR and ATHP3 individually. We also found that the CTP-TMR and ATHP3 alone strongly interacted with the resin of the Superdex Peptide 10/300 GL column in 10 mM Tris-HCl (pH 8.0) containing 20 mM NaCl.

We performed a molecular dynamic study of the CTP interacting with ATHP3. Figure 3.8 shows the simulation results. It looks likely that charge-charge interactions and hydrogen bonding play a critical role in their interactions.

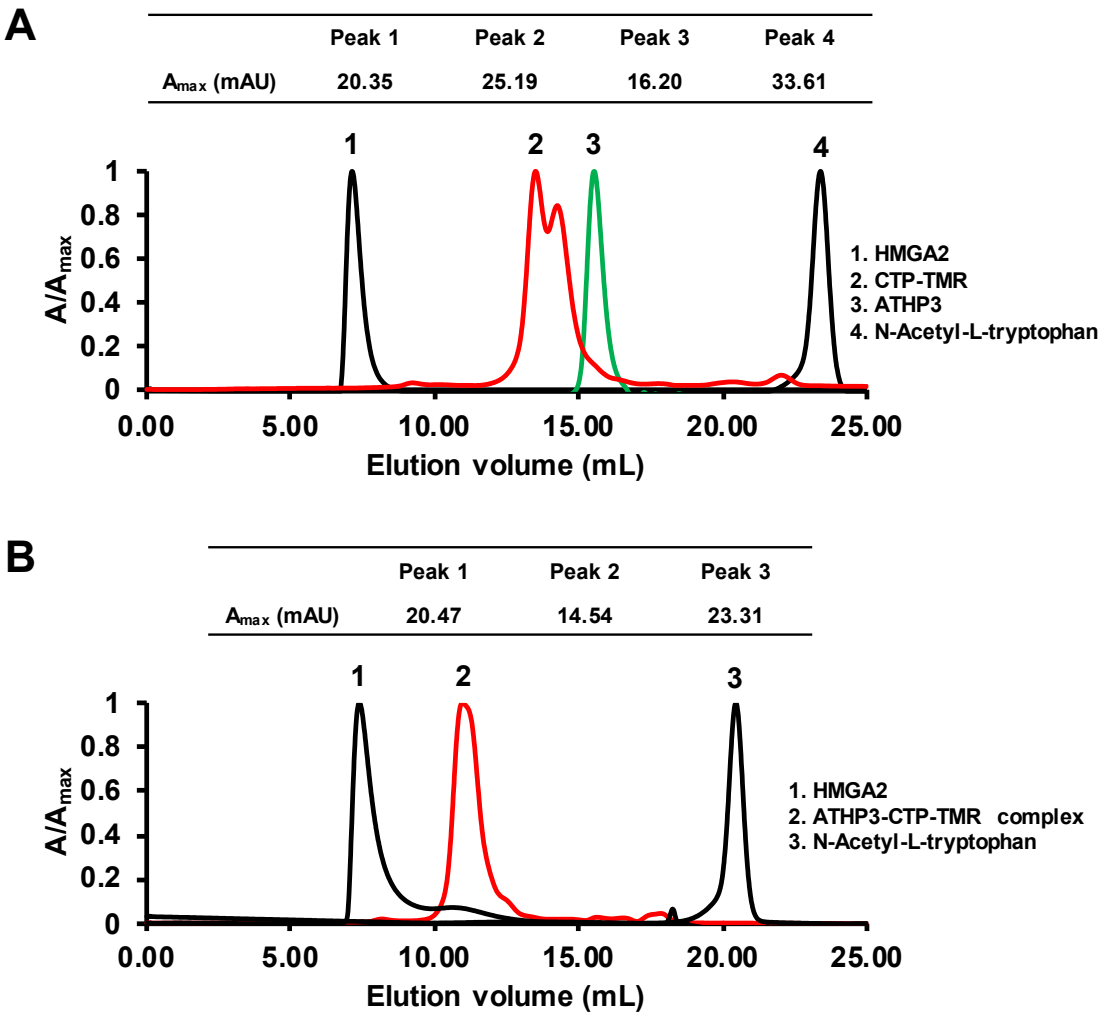


Figure 3.7. The formation of an ATHP3-CTP complex demonstrated by using FPLC size exclusion chromatography. **(A)** AKTA FPLC size exclusion chromatography experiments in 10 mM Tris-HCl (pH 8.0) plus 200 mM NaCl. **(B)** AKTA FPLC size exclusion chromatography experiments in 10 mM Tris-HCl (pH 8.0) plus 20 mM NaCl. The Superdex Peptide 10/300 GL column was calibrated using HMGA2 (11820 Da) and N-Acetyl-L-tryptophan (246.3 Da).

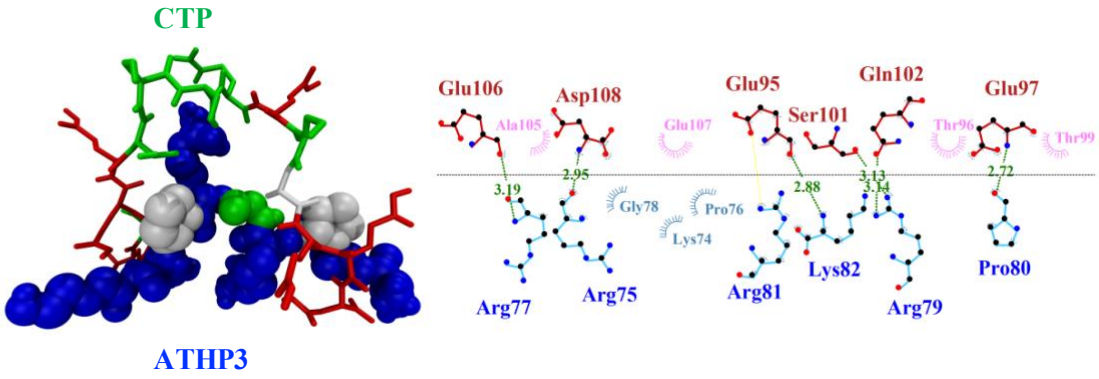


Figure 3.8. A simulated structure of ATHP3-CTP complex using docking and molecular dynamics. In this simulated structure, the CTP wraps around the ATHP3 (left panel). The right panel shows possible charge-charge interactions and hydrogen bonding in the complex.

3.5 Discussion

HMGA2 has an intriguing feature of the asymmetric charge distribution over its primary structure [176]. As shown in Figure 3.1, it consists of three highly positively charged "AT-hooks" motifs mainly concentrated in the center of the sequence and the negatively charged residues at the C-terminus. Previously, HMGA2 was identified to self-associate to homodimer through electrostatic interactions between the highly charged "AT-hooks" and the highly negatively charged C-terminus in aqueous solutions [21]. In our study, the higher apparent association constant that was identified by the PDI-ELISA assay and isothermal titration calorimetry experiments by the truncated HMGA2Δ95-108 binding to AT-rich DNA sequences compared to the wildtype HMGA2. HMGA2 binds to FL814 at 25 °C with a ΔG of -11.62 kcal/mol, a ΔH of -14.04 kcal/mol, and an unfavorable ΔS of -8.12 cal/(mol·K) (-TΔS, 2.42 kcal/mol) (Figure 3.2 E). It is the enthalpy change contributing the most energy to drive the binding reaction of HMGA2 and FL814. The HMGA2Δ95-108 binding to FL814 at 25 °C enthalpically drove the favorable ΔH of -9.95

kcal/mol, a ΔG of -14.32 kcal/mol, and a favorable ΔS of +14.66 cal/(mol·K) (-TΔS, -4.37 kcal/mol) (Figure 3.2E). The binding of both HMGA2 and HMGA2Δ95-108 to FL814 is associated with a favorable enthalpy contribution. The favorable enthalpy may come from the formation of non-covalent bonds between the protein and DNA such as electrostatic interactions, hydrogen bonds, and hydrophobic interactions. The C-terminus of HMGA2 regulated the DNA-binding affinity of HMGA2 that involved partner proteins interactions. The methods used such as the protein-DNA interactions ELISA assay, isothermal titration calorimetry, fluorescence anisotropy assay, and the peptide-peptide interactions confirmed that the CTP could regulate the DNA-binding affinity of HMGA2. Furthermore, the experimentation confirmed that the HMGA2Δ95-108 binds more tightly to DNA than the HMGA2 due to the lack of the negatively charged C-terminus which could interact with positively charged "AT-hooks" and then interfere with "AT-hooks" binding to DNA.

According to the HMGA2 and truncated HMGA2Δ95-108 are both cations, therefore their binding to FL814 is thermodynamically linked to Na^+ binding to DNA, and their binding constants depend on the total Na^+ concentration in the buffer system. From the PDI-ELISA results, we saw the binding constants decreased with increasing salt concentration. Based on the polyelectrolyte theory of Record et al. [210], our calculations indicated that there are 8.4 ± 0.5 and 9.6 ± 1.0 counter ions released from the DNA respectively by HMGA2 and HMGA2Δ95-108 binding to FL814 (Table 3.1). Water is an internal part of DNA structure [212, 214]. Previous studies have shown that there are at least two hydration layers surrounding duplex DNA, the first layer consists of about 20 water molecules per nucleotide. Hydration plays an important role in the binding affinity and specificity of protein-DNA interactions [211, 215]. Hydration contributions to binding

arise from differences in molecular interactions available to water in the bulk medium relative to those surrounding the macromolecules. When a protein binds to DNA, the reaction may uptake or release water. Interestingly, our PDI-ELISA assay experiments tested the role of water in the binding of HMGA2 to FL814. From the results, the binding constants of HMGA2 and HMGA2Δ95-108 to FL814 were almost unchanged in the presence of increasing concentrations of osmolytes (Figure 3.3). Based on the $\frac{\partial \ln(K_s/K_0)}{\partial [Osm]} = -\Delta n_w/55.5$ equation , our calculated values Δn_w of the difference in the number of bound water molecules between the complex and the free reactants indicate that within experimental error no water was released or taken up upon complex formation (Figure 3.3 E, F and Table 3.1).

Here we are also very interested in what kind of properties were involved in the HMGA2-DNA interactions and how the C-terminus of HMGA2 worked in interfering the binding of HMGA2 to DNA. Our results showed that CTP inhibits HMGA2 and HMGA2Δ95-108 binding to AT-rich DNA through tightly binding to the AT-hooks of HMGA2 and HMGA2Δ95-108. From the PDI-ELISA results, CTP can inhibit HMGA2-FL814 interactions with an IC_{50} of 52.36 ± 8.42 μM, relatively a little lower than the IC_{50} of CTP in HMGA2Δ95-108 and FL814 interaction which was 58.36 ± 6.53 μM (Figure 3.5). These results were consistent with our first experiment results that HMGA2Δ95-108 binds tightly to FL814 because it is lacking the negatively charged C-terminus compared with HMGA2. From our fluorescence anisotropy results, CTP tightly binds to HMGA2 and HMGA2Δ95-108 with binding constants of $2.59 \pm 0.20 \times 10^6\ M^{-1}$, $4.49 \pm 0.52 \times 10^7\ M^{-1}$ respectively (Figure 3.6 A, B). The CTP tightly binds to positively charged ATHP2

and ATHP3 with binding constants of $5.03 \pm 0.63 \times 10^5\ M^{-1}$, $6.06 \pm 1.54 \times 10^5\ M^{-1}$ respectively (Figure 3.6 C, D). When the CTP binding constant is compared to the ATHP2 and ATHP3, they all have similar binding affinities. However, the CTP binds more tightly to HMGA2Δ95-108 than HMGA2 when comparing the binding constants. This indicated that the C-terminus of HMGA2 may interfere with CTP peptide binding to "AT-hook" motifs in HMGA2 especially the ATHP3 motif which is next to the C-terminus. We hypothesized the charge-charge interaction played a vital role in the CTP binding to the AT-hooks of HMGA2 and the hypothesis was confirmed by the AKTA FPLC size exclusion chromatography experiments. The CTP-TMR and ATHP3 can bind and form a stable complex at a low salt concentration buffer system of 10 mM Tris-HCl (pH 8.0) plus 20 mM NaCl (Figure 3.7). When incubating the CTP-TMR with ATHP3 in 10 mM Tris-HCl (pH 8.0) plus 200 mM NaCl which had a higher salt concentration, there was no complex elution peak of ATHP3-CTP-TMR due to the increasing total Na^+ concentration makes electrostatic interaction decrease. To characterize the CTP and AT-hooks interactions we tried many other methods, such as fluorescence resonance energy transfer (FRET) assay, DNA UV- melting studies, and ITC studies. As the working conditions of peptide-peptide interaction in these methods were very tricky, we did not get valuable information. Alternatively, these results of CTP binding to positively charged “AT-hook” DNA binding domains of HMGA2 were confirmed by our molecular dynamics simulation studies. As shown in the simulated structure of ATHP3-CTP complex (Figure 3.7), the ATHP3 was surrounded by CTP. The CTP peptide bound to ATHP3 mainly through hydrogen bonding and charge-charge interaction, such as the interaction of Glu-95 (CTP) and Lys-82 (ATHP3) by hydrogen bonding and Glu-95 (CTP) and Arg-82 (ATHP3) by

electrostatic interactions. Our results show that the inhibition is through CTP peptide binding to the AT-hooks of HMGA2, and therefore blocking its AT-rich DNA binding capacity.

3.6 Conclusions

In this study, we characterized that C-terminal deletion mutant HMGA2Δ95-108 binding considerably more tightly to the AT-rich DNA sequences compared with the wildtype HMGA2 by the biochemical and biophysical studies, such as PDI-ELISA and ITC. Our results showed that the binding affinity of HMGA2 and DNA is dependent on salt concentration in the reaction buffer. It was found that 8.4 ± 0.5 and 9.6 ± 1.0 counterions released from the DNA respectively by HMGA2 and HMGA2Δ95-108 binding to AT-rich DNA sequences oligomer FL814. However, there were no hydration changes upon complex formation as HMGA2 and HMGA2Δ95-108 binding to DNA within experimental error. Additionally, we presented the C-terminus peptide of HMGA2 binding to positively charged "AT-hooks" motifs mainly by hydrogen bond and electrostatic interactions, and as a result inhibiting HMGA2 binding to DNA. Our findings will promote the development of the new therapeutic strategy against HMGA2-associated cancers and obesities by using the highly negatively charged C-terminus peptide of HMGA2 or synthetic CTP mimics as an inhibitor.

CHAPTER 4: Synthesis and characterization of daunorubicin-GTP, doxorubicin-GTP, daunorubicin-dGTP, and doxorubicin-dGTP conjugates

4.1 Abstract

The anthracycline antibiotics daunorubicin and doxorubicin are among the most potent anti-cancer drugs. The mechanism of action (MOA) of these two drugs stems from the poisoning of DNA topoisomerase II. However, their therapeutic efficacy is compromised by the multidrug resistance (MDR) of cancer cells and a cumulative, irreversible cardiotoxicity. Additionally, because these two drugs are not able to penetrate the blood-brain barrier (BBB), they cannot be used to treat brain tumors. Here, we show that formaldehyde can efficiently and specifically conjugate daunorubicin (DNR) and doxorubicin (DOX) to guanosine-5'-triphosphate (GTP) and 2'-deoxyguanosine-5'-triphosphate (dGTP) to yield daunorubicin-GTP (DNR-GTP), doxorubicin-GTP (DOX-GTP), daunorubicin-dGTP (DNR-dGTP), and doxorubicin-dGTP (DOX-dGTP) conjugates. The linkage is between the 2'-NH_2 of guanine and the 3'-NH_2 of daunosamine. The specific conjugation of DNR/DOX with GTP/dGTP likely results from the fact that GTP and dGTP can interact with DNR and DOX, and form noncovalent DNR/DOX-GTP/dGTP complexes at a one-to-one molar ratio. The DNR/DOX-GTP/dGTP conjugates were characterized by different methods including UV-Vis, fluorescence, CD, and mass spectroscopy. Our results also show that DNR/DOX-GTP/dGTP conjugates bind to DNA by intercalation. Furthermore, we found that the DNR/DOX-GTP/dGTP conjugates rapidly accumulated in human cancer cells and were cytotoxic to both doxorubicin-sensitive SK-OV-3 and doxorubicin-resistant NCI/ADR-RES cells. Interestingly, our results suggest that these DNR/DOX-GTP/dGTP conjugates may overcome doxorubicin resistance for

certain cancer cells. Our western blotting results showed that these DNR/DOX-GTP/dGTP conjugates change the topoisomerase I and IIα's expression patterns.

4.2 Introduction

Daunorubicin (daunomycin) and doxorubicin (adriamycin) are two anthracycline antibiotics which were found and developed in the 1960s [216, 217]. They are among the most potent anti-cancer drugs against a wide spectrum of human cancers, such as, leukaemias, breast cancer, lung cancer, and ovarian cancer [218-220]. A number of models have been proposed for the mechanism of action of the two drugs causing DNA damage and cell death, including DNA intercalation, DNA topoisomerase II inhibition, DNA adduct formation, and oxidative stress [221-225]. Although daunorubicin (DNR) and doxorubicin (DOX) are widely used chemotherapy drugs, there are some problems that limit their anticancer potential. Their therapeutic effects are limited by a cumulative and irreversible cardiotoxicity [226-228], and by the multidrug resistance (MDR) of the cancer cells developed in the process of treatment [229, 230].

In the process of solving the problems, scientists made many efforts to overcome these limits, such as, conjugating polymers, including peptides and poly(ethylene glycol)s (PEGs) to DNR and DOX to increase their treatment efficacy and reduce the toxicity [231, 232]. Another approach was developed as crosslinking daunorubicin/doxorubicin (DNR/DOX) to double strand DNA (dsDNA) by formaldehyde to form DNR/DOX-DNA adducts [233]. The DNR/DOX molecules (Figure 4.1) contain an amino sugar group called daunosamine which can react with formaldehyde to form a Schiff base and then react with DNA to form the adducts which offer the potential for overcoming formerly mentioned limitations [220]. Having these findings, we hypothesized that DNR/DOX might be crosslinked to a single

nucleoside triphosphate (NTP) or deoxyribonucleotide triphosphate (dNTP) to form a conjugate via Schiff base chemistry.

In this paper, DNR and DOX are found specifically covalently crosslinked to GTP/dGTP (Figure 4.1) by formaldehyde to yield DNR/DOX-GTPs and DNR/DOX-dGTPs. Herein, we report the synthesis, characterization, and cytotoxicity of these new derivatives which result from the reaction of formaldehyde at the 3'-NH_2 group of DNR or DOX and the 2'-NH_2 group of guanine. These derivatives, denoted DNR-GTP, DNR-dGTP, DOX-GTP, and DOX-dGTP bind to DNA by intercalation as the parent drugs DNR/DOX. These derivatives show enhanced cytotoxicity toward tumor cells, especially, the DOX-GTP improved cellular uptake and accumulation of DOX in the cells and induced potent toxicity to the DOX-resistant NCI/ADR-RES.

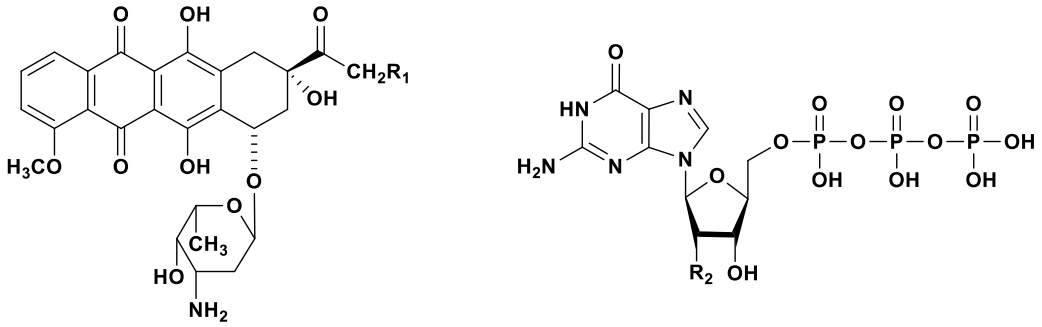


Figure 4.1. Chemical structures of daunorubicin (DNR, R_1 = H), doxorubicin (DOX, R_1 = OH), Guanosine-5'-triphosphate (GTP, R_2 = OH), and 2'-deoxyguanosine-5'-triphosphate (dGTP, R_2 = H).

4.3 Materials and Methods

4.3.1 Materials.

Daunorubicin (DNR) and Doxorubicin (DOX) was obtained from Waterstone Technology, LLC. A molar extinction coefficient of 11,500 M^{-1} cm^{-1} at 480 nm was used to determine the concentrations of DNR and DOX. Adenosine-5'-triphosphate disodium (ATP) and guanosine-5'-triphosphate sodium salt (GTP) were purchased from Amersham Pharmacia Biotech, Inc. The molar extinction coefficient of 15,400 M^{-1} cm^{-1} at 259 nm and 13,700 M^{-1} cm^{-1} at 253 nm, respectively was used to determine the concentrations of ATP and GTP. Cytidine-5'-triphosphate (CTP) and uridine 5'-triphosphate (UTP) were purchased from NEW ENGLAND BioLabs, Inc. 2'-deoxyadenosine-5'-triphosphate (dATP), 2'-deoxyguanosine-5'-triphosphate (dGTP), 2'-deoxycytidine-5'-triphosphate (dCTP), thymidine-5'-triphosphate (dTTP), Tris buffer (pH 8.0) saturated phenol, 2-propanol, HCl, acetic acid, Gibco™ RPMI 1640 Medium (11-875-093), Corning™ Cellgro™ Cell Culture Phosphate Buffered Saline, Corning™ Penicillin-Streptomycin Solution, Corning™ 0.25% Trypsin and 0.1% EDTA in HBSS without Calcium, Magnesium and Sodium Bicarbonate (MT25053CI), and BioLife 35mm Tissue Culture Dish were purchased from Thermo Fisher Scientific, Inc. Sodium borate decahydrate ($Na_2B_4O_7 \cdot 10\ H_2O$), formaldehyde solution (36.5%, for molecular biology), 1-butanol, and chloroform (99.9+%, A.C.S. HPLC grade) were purchased from Sigma Aldrich Chemical Co., Inc. Plasmocin™ prophylactic (ant-mpp) was purchased from InvivoGen. HyClone™ standard fetal bovine serum (FBS, SH30088.03) was purchased from GE Healthcare Life Sciences. 96-Well CytoOne®plate, TC-treated (CC7682-7596) was purchased from USA Scientific, Inc. 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) was

purchased from Tocris Bioscience. SK-OV-3 human ovarian cancer cells were obtained from American Type Culture Collection (Manassas, VA). NCI/ADR-RES adriamycin resistant ovarian cancer cells were a gift from Dr. Daping Fan at the University of South Carolina. HMGA1 (D6A4) XP® Rabbit mAb (#7777), HMGA2 (D1A7) Rabbit mAb (#8179), TOP1 (E4Z1Q) Rabbit mAb (#38650), Topoisomerase IIα (D10G9) XP® Rabbit mAb (#12286), β-Actin (D6A8) Rabbit mAb (#8457), and Anti-rabbit IgG (H+L) (DyLightTM 800 4× PEG Conjugate) (#5151) were purchased from Cell Signaling Technology, Inc.

4.3.2 Synthesis and purification of DNR-GTP, DNR-dGTP, DOX-GTP, and DOX-dGTP.

Formaldehyde crosslinking experiments were conducted at 37 ℃ in 10 mM sodium borate buffer pH 8.2 containing DNR or DOX, GTP or dGTP, and 2% (v/v) HCHO (unless specified otherwise). After 24 hours of incubation, reactions were stopped by two times of phenol extraction. The products were precipitated with 10 volumes of 1-butanol, dried using a CENTRIVAP concentrator, and stored at a -80 ℃ freezer.

4.3.3 Thin Layer Chromatography (TLC).

DNR, DOX, and their derivatives were spotted on TLC Silica Gel 60 F_{254} and developed with a solvent consisting of chloroform: methanol: acetic acid = 16:4:1 (v/v) at room temperature. After the development, the TLC Silica Gels were air dried, and visualized and photographed under visible and UV light.

4.3.4 Mass spectrometry.

Solid DNR-GTP samples were dissolved in water to make a 93 µM (~100 ppm) solution. Working solutions of 20 ppm DNR-GTP were prepared in 1:1 Optima grade Methanol/Water. The samples were directly infused into the ESI source of a Bruker Solarix

FT-ICR, operated at 256 KWord in negative ion mode. Scan acquisition of 50 coadded scans was collected. Mass spectrometry was calibrated utilizing Tuning Mix calibration standard in the range between 112 to 1033 with 5 calibration points obtaining a standard deviation of 0.613. A solvent blank was also analyzed as reference. Solvents, methanol utilized in this study were analytical grade or better and purchased from Fisher Scientific (Pittsburgh, PA). A Tuning Mix calibration standard (G24221A) was obtained from Agilent Technologies (Santa Clara, CA) and used as received. Mass spectra were processed using Bruker Compass Data Analysis version 5.1 (Bruker Daltonik GmbH).

4.3.5 Visible absorbance, fluorescence, and circular dichroism (CD) spectroscopy.

Visible absorbance spectra of DNR-GTP, DNR-dGTP, DOX-GTP, and DOX-dGTP in 1×BPE (6 mM Na_2HPO_4, 2 mM NaH_2PO_4, and 1 mM Na_2EDTA, pH 7.0) were recorded on a Varian Cary Bio-50 UV-Vis spectrometer at room temperature. The fluorescence emission spectra were recorded on a Varian Cary Eclipse fluorescence spectrophotometer at room temperature. CD spectra were recorded on a Jasco Model J-815 CD spectrometer at room temperature. The molar ellipticity $[\theta]$ was calculated from the equation $[\theta] = 100\theta/cl$, where θ is the measured ellipticity in degrees, c is the concentration of DNR or DNR-GTP in molar, and l is the path length in centimeters.

4.3.6 Absorbance titration experiments.

Absorbance titration experiments using a Varian Cary Bio-50 UV-Vis spectrophotometer were used to study GTP binding to DNR, and to determine their binding constants. Specifically, the visible absorbance of DNR or DOX were measured while increasing concentrations of GTP were titrated into 10 μM of DNR in 10 mM sodium borate buffer pH 8.2 at room temperature. The concentrations of free (C_f) and bound DNR

(C_b) were calculated using these two equations: $C_B = (A^0 - A) / (\varepsilon_f - \varepsilon_b)$ and $C_f = C_T - C_b$, where A^0 and A is the absorbance of DNR in the absence and presence of GTP, respectively. C_T is the total DNR concentration. ε_f and ε_b are extinction coefficients of DNA in in the absence and presence of GTP, respectively. ε_b was determined in the presence of a large excess of GTP with no further change in absorbance at 480 nm.

4.3.7 Fluorescence titration experiments.

Fluorescence titration experiments using a Varian Cary Eclipse fluorescence spectrophotometer were employed to study nucleotides (ATP, GTP, CTP, and UTP) binding to DNR at room temperature with λ_{ex}= 480 nm and λ_{em}= 590 nm. A slit width of 10 nm was used for both excitation and emission light slits. Specifically, fluorescence spectra were recorded by titrating one nucleotide into 1 µM of DNR in 10 mM sodium borate buffer pH 8.2 and used to calculate the concentrations of free (C_f) and bound DNR (C_b): $C_f = C_T (I - I_\infty) / (I_0 - I_\infty)$ and $C_b = C_T - C_f$, where C_T is the total DNR concentration; I_0 is the fluorescence intensity of free DNR; and I_∞ is the fluorescence intensity of the DNR in the presence of a large excess of a nucleotide.

For the DNA binding studies, fluorescence spectra were recorded while increasing concentrations of DNA were titrated into 1 µM of DNR or DNR-GTP in 1×BPES (6 mM Na_2HPO_4, 2 mM NaH_2PO_4, 1 mM Na_2EDTA, and 185 mM NaCl, pH 7.0).

4.3.8 Fluorescence anisotropy titration.

Fluorescence anisotropy titration experiments were performed using a Varian Cary Eclipse fluorescence spectrophotometer with λ_{ex}= 480 nm and λ_{em}= 590 nm. A slit width of 10 nm was used for both excitation and emission light slits. To measure the fluorescence

anisotropy, increasing concentrations of DNA were titrated into 1 μM solutions of DNR or DNR-GTP in 1×BPES at room temperature.

4.3.9 Continuous variation analysis.

Binding stoichiometry of GTP to DNR was obtained using the method of continuous analysis [234, 235]. 100 μM of DNR and GTP in the 10 mM sodium borate buffer pH 8.2 were prepared. A series of solution mixtures were made by varying volumes of these two equally concentrated solutions to keep the sum of the concentrations of DNR and GTP constant at 100 μM. The absorbance of DNR in the presence or absence of GTP was measured at 480 nm at room temperature. Similarly, fluorescence intensity at λ_{ex} = 480 nm with λ_{em} = 590 nm was also measured. The difference in absorbance (ΔA) or fluorescence intensity (ΔF) was plotted against the molar fraction of DNR to determine the binding stoichiometry of GTP to DNR.

4.3.10 Cell culture.

SK-OV-3 (DOX-sensitive) and NCI/ADR-RES (DOX-resistant) human ovarian cancer cells were cultured in RPMI 1640 medium supplemented with 10% FBS, 100 IU/mL of penicillin, 100 μg/mL of streptomycin, and 2.5 μg/mL of plasmocin in 25 cm^2 tissue culture flasks at 37 ℃ in a humid atmosphere with 5% CO_2 and 95% air. The cells were subcultured once they reached 80% confluence and cell density, determined with a hemocytometer prior to each experiment.

4.3.11 Fluorescence microscopic imaging.

SK-OV-3 (4×10^5) and NCI/ADR-RES (4×10^5) cells were seeded in 35 mm tissue culture dishes with 2 mL RPMI 1640 medium supplemented with 10% FBS, 100 IU/mL of penicillin, 100 μg/mL of streptomycin, and 2.5 μg/mL of plasmocin and incubated at

37 ℃ in a humid atmosphere with 5% CO_2 and 95% air for 48 hours. Cells incubated for additional 4 hours with fresh medium containing 5 μM of DNR, DNR-GTP, DOX, or DOX-GTP. After the incubation, cells were washed thrice with phosphate buffered saline (PBS) prior to the fluorescence imaging. A Home-built inverted fluorescence microscope with an LED light source (X-cite 120 LED Boost, USA) for sample illumination was used to acquire all fluorescence images. The fluorescence intensity was collected by using a 20× objective lens (Nikon, CFI S Plan Fluor ELWD, NA ~0.45) and a sCMOS camera (ThorLabs, USA). ImageJ was used for data analysis.

4.3.12 MTT cell viability assay.

Cytotoxicity of DNR, DNR-GTP, DOX, and DOX-GTP against SK-OV-3 and NCI/ADR-RES cells was evaluated by MTT cell viability assays. Briefly, cells were seeded in 96-well plates (7500 cells/well) in 100 μL of RPMI 1640 medium containing 10% FBS and grew at 37 ℃ for 24 hours in a humid atmosphere with 5% CO_2 and 95% air. Subsequently, the cells grew in fresh medium (RPMI 1640 medium containing 10% FBS) containing different concentrations of DNR, DNR-GTP, DOX, or DOX-GTP for additional 72 hours. After the 72 hours of incubation, 10 μL of 3-[4,5-dimethylthiazole-2-yl]-2,5-diphenyltetrazolium bromide (MTT; 5 mg/mL) was added to each well and incubated for 4 hours at 37 °C. The supernatant was then removed. 100 μL of acidified isopropanol was added to solubilize the MTT-formazan products for 15 min on an orbital shaker. Absorbance at 590 nm was measured with a microplate reader (BioTek, Synergy). Cell survival rate (r) was calculated with the following equation [236]: $\mathrm{r} = \frac{\mathrm{A}-\mathrm{A}_{bk}}{A^0-\mathrm{A}_{bk}} \times 100$, where A^0 and A represent the absorbance of wells with cells in the absence and presence of a compound, respectively. A_{bk} is the absorbance of wells without cells. The half maximal

inhibitory concentration (IC_{50} value), the concentration of a compound that inhibits 50% cell growth, was calculated according to the cell survival rate.

4.3.13 Western blotting experiment.

SK-OV-3 and NCI/ADR-RES cells were seeded at a density of 4×10^5 cells in 35 mm tissue culture dishes with 2 mL of RPMI 1640 medium containing 10% FBS and incubated at 37 ℃ for 48 hours. Subsequently, the cells were treated with fresh medium containing 10 μM of DNR, DNR-GTP, DOX, or DOX-GTP and incubated for additional 24 hours. To prepare whole cell extracts, cells were washed twice with ice-cold 1×PBS and lysed on ice for 5 min in 1×RIPA (radioimmunoprecipitation assay) buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 5 mM EDTA, 1 mM EGTA, 1% NP-40, 0.1% SDS, and 0.5% sodium deoxycholate) supplemented with 1×Halt Protease inhibitor cocktails, 1×Halt Phosphatase inhibitor cocktails, and 1 mM of phenylmethanesulfonyl fluoride (PMSF). The cell extracts were transferred to a sterile microcentrifuge tube, sonicated for 10 seconds with a Fisherbrand™ Model 120 Sonic Dismembrator at 30% amplitude, and centrifuged at 15,000 g at 4 °C for 15 min to remove the cell debris. The protein concentrations of the cell extracts (supernatant) were determined using the Pierce BCA Assay Kit (Thermo Fisher Scientific) according to the manufactory's instruction.

To determine the expression level of TOP2α, TOP1, HMGA2, and HMGA1 in NCI/ADR-RES and SK-OV-3 cells, the proteins of cell extracts were separated using an 8 or 15% SDS-PAGE gel. After the electrophoresis, the proteins were transferred to nitrocellulose membranes in a transfer buffer (48 mM Tris-HCl (pH 9.2), 39 mM glycine, 20% (v/v) methanol) at 100 V for 1 hour. The membranes were then blocked with 5% nonfat dry milk in Tris buffer saline with 0.1% Tween®20 detergent (TBST: 20 mM Tris-

HCl (pH 7.5), 150 mM NaCl, 0.1% Tween 20) at room temperature for 30 min and incubated with a primary antibody against human DNA topoisomerase I, DNA topoisomerase IIα, HMGA1, or HMGA2 in TBST overnight at 4 °C. Subsequently, the membrane was washed thrice with TBST and incubated at room temperature with the anti-rabbit IgG (H+L) (DyLight™ 800 4× PEG Conjugate) (#5151, 1:10,000 v/v, Cell Signaling Technology) secondary antibody in TBST for 1 h. After the membrane was washed thrice with TBST, images were taken using an Odyssey CLx near infrared scanner (Licor Biosciences) to determine the protein expression levels. The Western Blotting results were corrected for dilution, normalized to the β-actin signal, and analyzed with the Image Studio 2.0 software (Licor Biosciences).

4.4 Results

4.4.1 Synthesizing DNR-GTP, DNR-dGTP, DOX-GTP, and DOX-dGTP conjugates.

Previous studies showed that formaldehyde can efficiently crosslink DNR and DOX to the guanine of DNA through a methyl bridge/linkage between the 2-amino group of guanine and the 3'-amino group of daunosamine [237, 238]. In this study, we decided to investigate whether formaldehyde can also crosslink DNR or DOX to GTP or dGTP and produce DNR-GTP, DNR-dGTP, DOX-GTP, and DOX-dGTP conjugates. To our surprise, at 37 °C in 10 mM sodium borate buffer pH 8.2, formaldehyde efficiently crosslink DNR or DOX to GTP or dGTP and yield DNR-GTP, DNR-dGTP, DOX-GTP, and DOX-dGTP conjugates. Figure 4.2 shows the procedure we used for the synthesis of the DNR-GTP, DNR-dGTP, DOX-GTP, and DOX-dGTP conjugates. Phenol can efficiently extract unreacted DNR and DOX from the aqueous buffer solutions. A large excess of 1-butanol efficiently precipitated the DNR-GTP, DNR-dGTP, DOX-GTP, and DOX-dGTP

conjugates. TLC experiments showed the successful of the DNR-GTP, DNR-dGTP, DOX-GTP, and DOX-dGTP conjugates. Since these conjugates carry 4 negative charges, they did not migrate under this experimental condition. As expected, DNR in the presence or absence of formaldehyde migrated rapidly in the TLC experiments and was near the front of the mobile phase (Figure 4.2 B). Interestingly, the DNR-GTP conjugate did not significantly fluorescent under the UV light.

We also examined whether formaldehyde could crosslink DNR or DOX to other nucleotides. Figure 4.3 A and B show our results. Formaldehyde was only able to crosslink DNR and DOX to GTP and dGTP although ATP, dATP, CTP, and dCTP also carry an amino group. Temperature dependent studies show that 55 ℃ is the optimal temperature for the crosslinking reactions (Figure 4.3 C). The synthesis yields are dependent on the GTP concentration (Figure 4.3 D), GNR concentration (Figure 4.3 E), and reaction time (Figure 4.3 F).

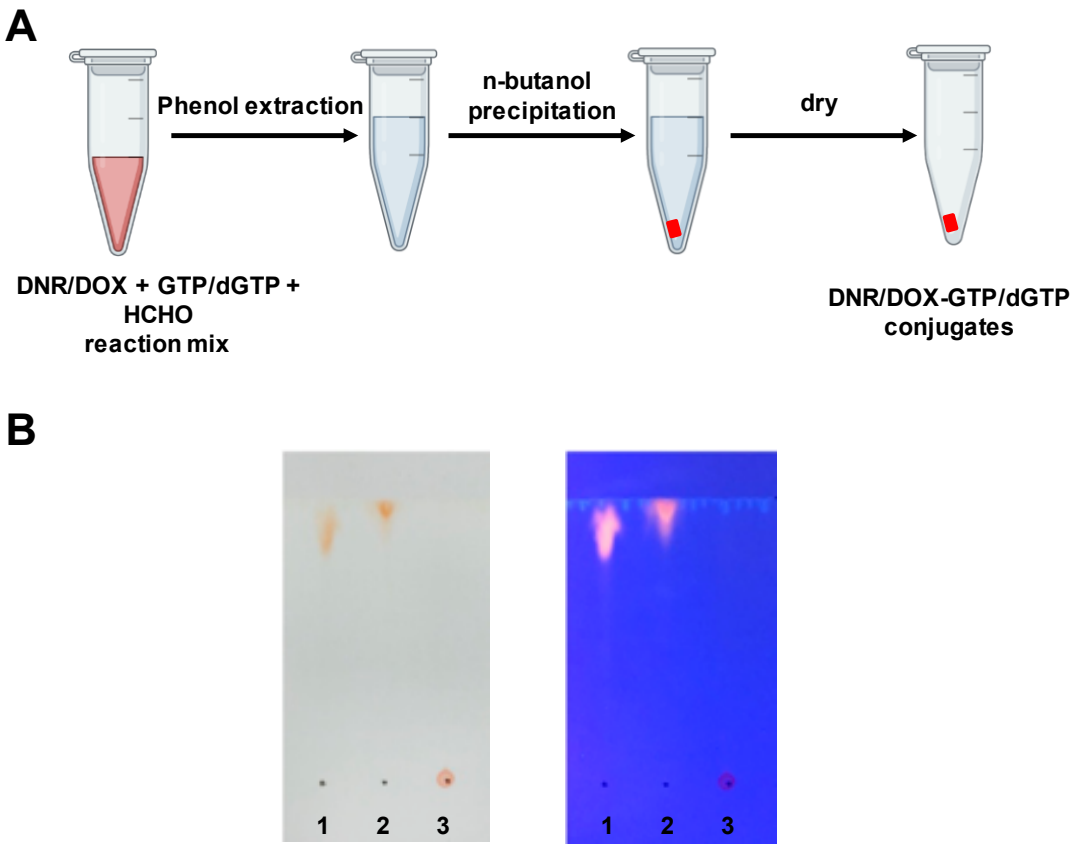


Figure 4.2. (A) An experimental procedure to synthesize and purify DNR-GTP, DNR-dGTP, DOX-GTP, and DOX-dGTP conjugates in 10 mM sodium borate buffer pH 8.2. **(B)** Thin layer chromatography (TLC) of DNR (lane 1), DNR in the presence of 2% HCHO (lane 2), the DNR-GTP conjugate (lane 3). Samples were spotted on Silica Gel 60 F_{254} and developed with a mobile phase of chloroform: methanol: acetic acid = 16:4:1 (v/v). After the development, the TLC plates were photographed under visible (left panel) and UV light (right panel).

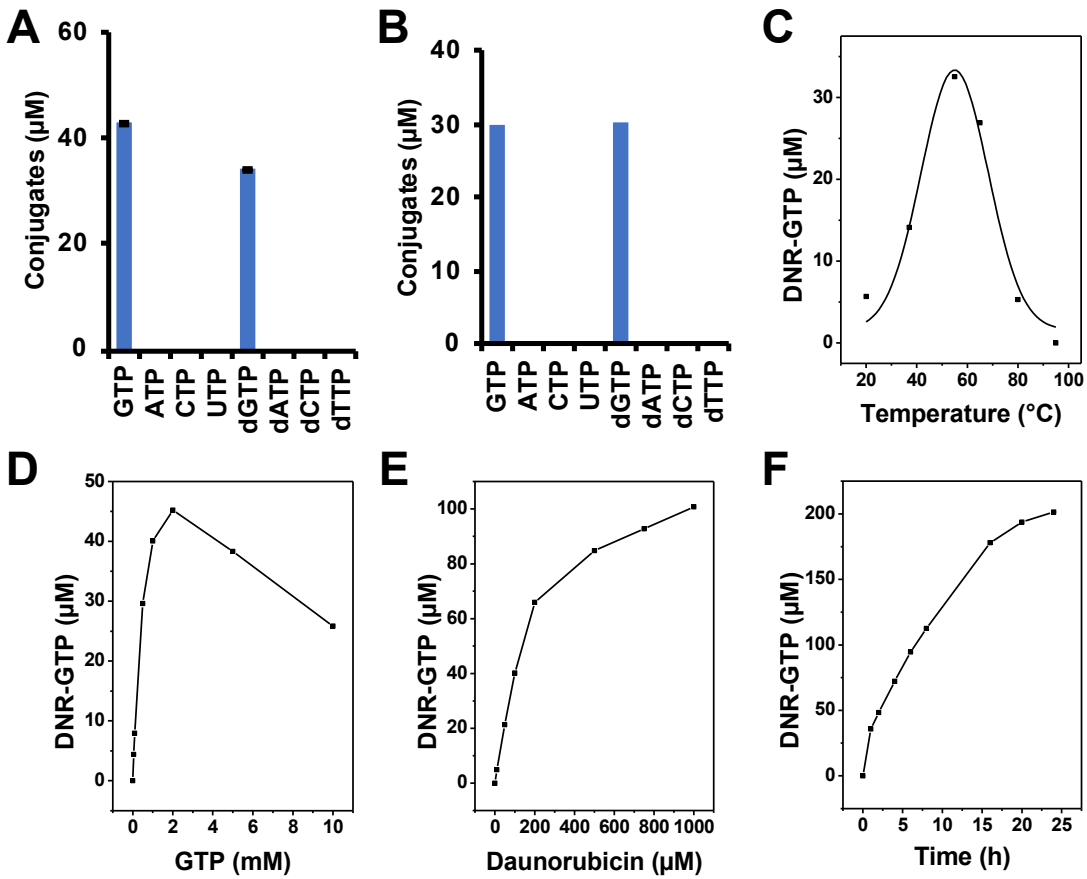


Figure 4.3. Crosslinking reactions of DNR/DOX to GTP and dGTP by HCHO. Specific crosslinking of DNR/DOX to GTP and dGTP by HCHO. The crosslinking reactions were conducted in 10 mM sodium borate buffer pH 8.2, containing 100 μM DNR/DOX, 1 mM one of NTPs/dNTPs, 2% (v/v) HCHO at 37 ℃ and RT, respectively overnight **(A**, **B)**. Temperature effects on the crosslinking of DNR to GTP by HCHO. The crosslinking reactions were conducted in 10 mM sodium borate buffer pH 8.2, containing 100 μM DNR, 1 mM GTP, 2% (v/v) HCHO at 20 ℃, 37 ℃, 55 ℃, 65 ℃, 80 ℃, 95 ℃, respectively for 2 h **(C)**. Formaldehyde crosslinking titration experiments were conducted at 37 °C in 10 mM sodium borate buffer pH 8.2, and 0.125% (v/v) HCHO for 2 h. Fixed concentration of DNR (100 μM) were crosslinked with increasing concentrations of GTP by HCHO **(D)** and fixed concentration of GTP (1 mM) were crosslinked with increasing concentrations of DNR by HCHO **(E)**. Time course of DNR-GTP synthesis in the conditions of 2 mM DNR, 1 mM GTP, 10 mM sodium borate buffer pH 8.2, 0.5% (v/v) HCHO at 37 ℃ for 24 h **(F)**.

The formation of the DNR-GTP conjugate was confirmed using electrospray ionization mass spectrometry (ESI-MS) experiments in negative ion mode that produced a doubly charged signal $[M-2H]^{2-}$ at m/z 536.07798 and δ = 0.65 ppm (Figure 4.4 A, left

panel). This spectrum was deconvoluted and produces a singly charged diprotonated ion $[M-H]^-$ at m/z 1073.16307 with chemical formula $C_{39}H_{45}N_6O_{24}P_3$ and δ = 0.50 ppm. Figure 4.4 B shows a proposed chemical structure of the DNR-GTP conjugate according to these MS results. A methyl bridge covalently links the 2-amino group of guanine and the 3'-amino group of daunosamine. An additional formaldehyde molecule forms an oxazolidine ring between the 3'-NH_2 and 4'-OH of daunosamine.

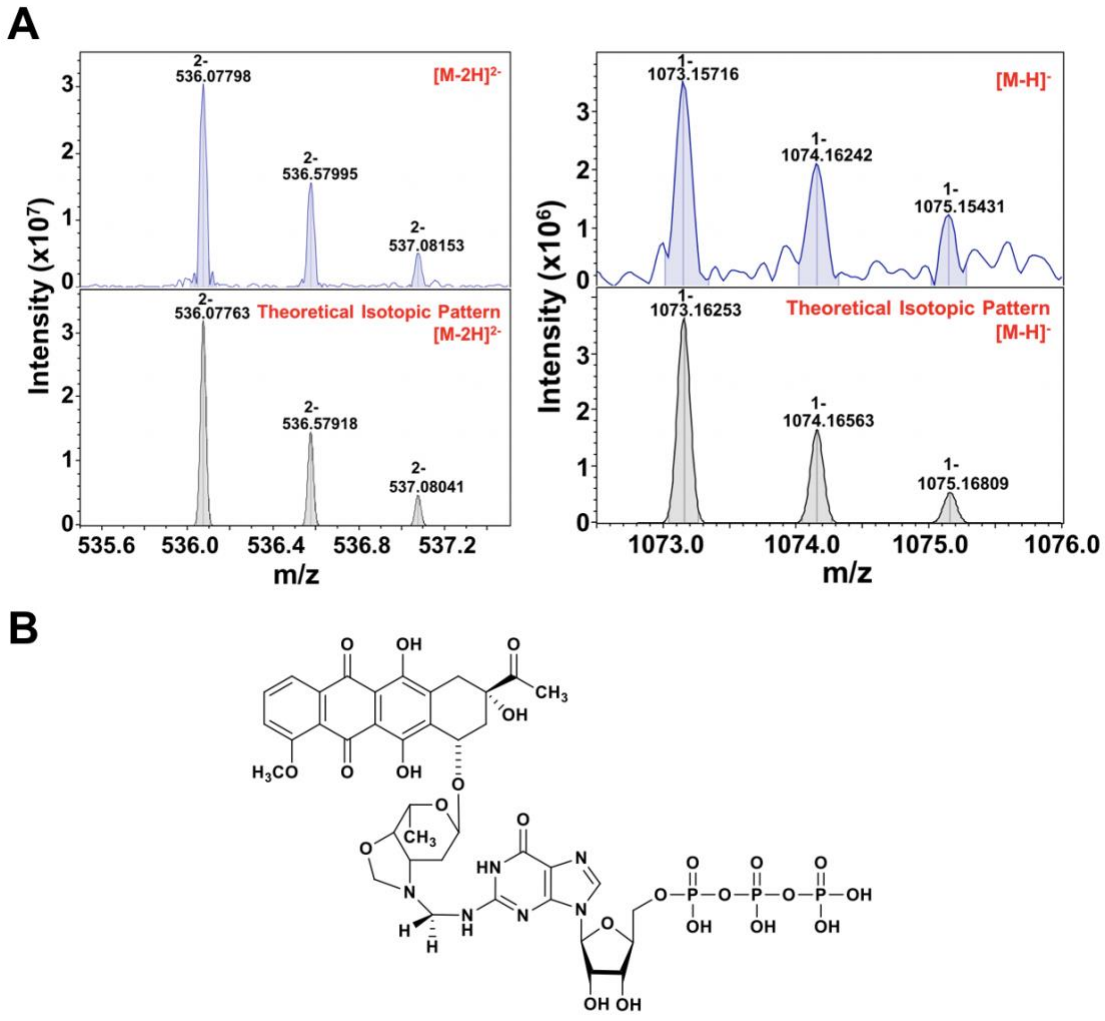


Figure 4.4. (A) Electrospray ionized mass spectrum of DNR-GTP negative ion mode. **(B)** A proposed chemical structure for the DNR-GTP conjugate.

4.4.2 Stability and optical properties of the DNR-GTP conjugate.

The stability of the DNR-GTP conjugate was examined using TLC experiments and fluorescence spectroscopy. Figure 4.5 shows the TLC experiments. The DNR-GTP was quite stable at room temperature and 4 °C. Very little of the DNR-GTP conjugate was decomposed, especially at 4 °C. Figure 4.6 shows a time course of the DNR-GTP conjugate in 1×BPE buffer at room temperature. The fluorescence intensity of the DNR-GTP conjugate was significantly enhanced for the diluted DNR-GTP conjugate solutions. Since the fluorescence of the DNR-GTP conjugate is quenched, these results suggest that the diluted DNR-GTP conjugate solutions may not be stable and decomposed into individual components. Nevertheless, light may play a significant role in the instability of the diluted DNR-GTP conjugate solutions since it is not possible to block all light when measured the fluorescence spectra of the diluted DNR-GTP conjugate solutions.

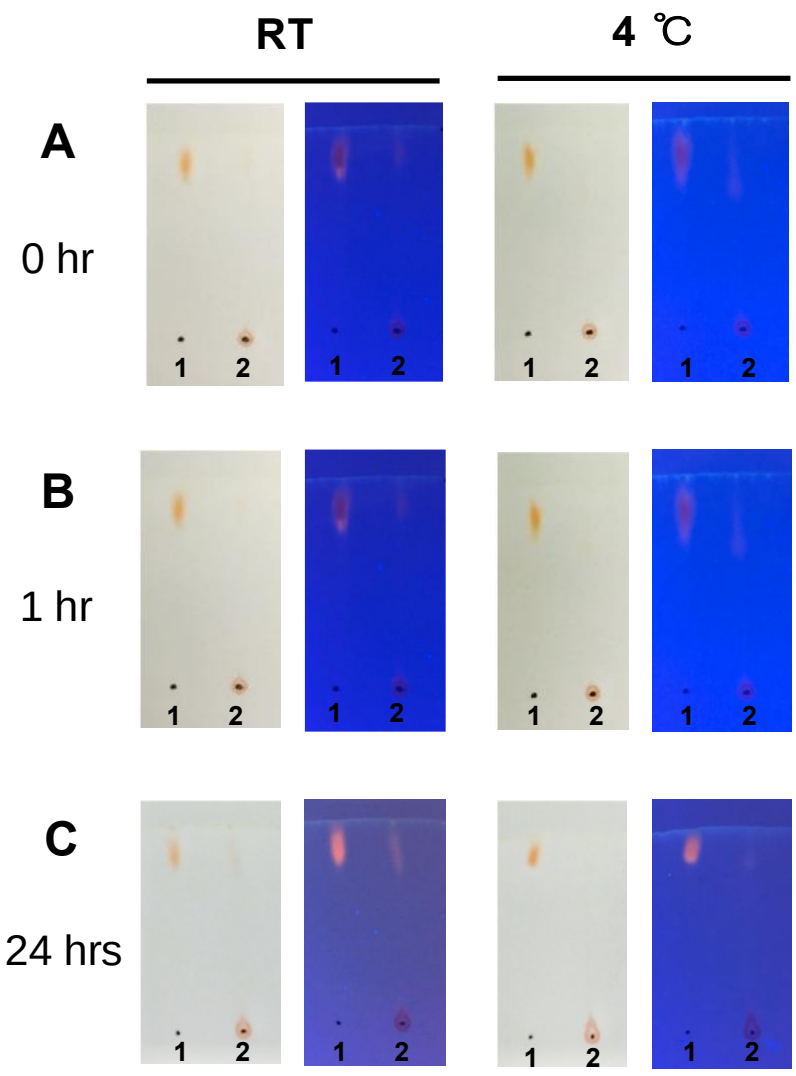


Figure 4.5. Stability of DNR-GTP in 1×BPE at RT and 4 °C by thin layer chromatography. TLC profile of 2 mM (1) DNR, (2) DNR-GTP. Samples were prepared at RT and 4 ℃ respectively and spotted on Silica Gel 60 F_{254} and developed with a solvent system of chloroform : methanol : acetic acid = 16:4:1 (v/v) and then was visualized under visible and UV light in the time course of **(A)** 0 hour (hr), **(B)** 1 hour (hr), and **(C)** 24 hours (hrs).

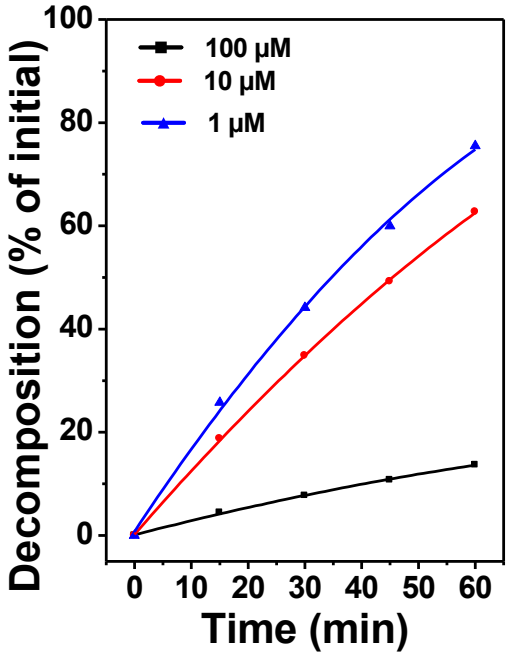


Figure 4.6. Stability of DNR-GTP in 1×BPE buffer at room temperature. Decomposition (% of initial DNR-GTP) at 100 µM (black line), 10 µM (red line), and 1 µM (blue line) for the time course of 0, 15, 30, 45, 60 min.

Figure 4.7 A shows the visible spectra of the DNR-GTP conjugate. Crosslinking DNR to GTP causes a red shift of absorbance maximum from 480 to 486 nm and a decrease of the molar extinction coefficient. The spectrum of the DNR-GTP conjugate is similar to that of DNR binding to DNA. Figure 4.7 B shows the fluorescence emission spectra of DNR and the DNR-GTP conjugate in the absence or presence of DNA. The emission spectra of DNR in the absence or presence of DNA are identical to the previously published results [239]. Crosslinking DNR to GTP significantly quenched the fluorescence intensity of DNR, similar to DNR binding to DNA. The fluorescence emission of the DNR-GTP was further quenched in the presence of a large excess of DNA.

Figure 4.7 C shows the CD spectra of DNR and the DNR-GTP conjugate in the absence or presence of DNA. The CD spectra of DNR in the absence or presence of DNA

are similar to the published results [222]. However, the CD spectra of the DNR-GTP are dramatically different from that of DNR. The DNR-GTP conjugate has three bands with two negative bands and one positive band. The first band is a negative band around 300 and possibly corresponds to the electronic transition around the short axis of DNR chromophore. The second band is a positive band around 460 nm. The final band is a negative band around 520 nm. In the presence of excess DNA, the CD spectrum of the DNR-GTP conjugate did not change significantly. Excess DNA only slightly enhanced the magnitude of the negative CD band around 300 nm, indicating that the DNR-GTP conjugate also binds to and intercalates into DNA base pairs. Certain optical properties of DNR and the DNR-GTP conjugate are summarized in Table 4.1.

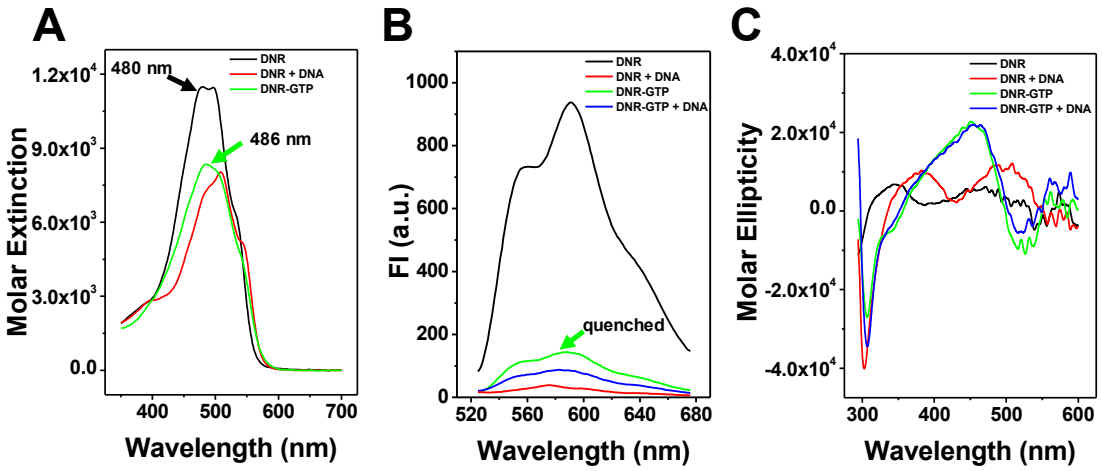


Figure 4.7. (A) Visible absorbance spectra of DNR in the absence (black line) or presence (red line) of 1 mM stDNA and the DNR-GTP conjugate (green line) in 1×BPE buffer. **(B)** Fluorescence emission spectra of DNR and DNR-GTP in the absence and presence of 1 mM stDNA in 1×BPE buffer. **(C)** CD spectra of DNR and DNR-GTP in the absence and presence of 1 mM stDNA in 1×BPE buffer.

Table 4.1. Optical properties of DNR and DNR-GTP.

	λ_{max} (nm)	ε (M^{-1} cm^{-1})	Relative fluorescence [a]
DNR	480	11500	1
DNR-GTP	486	8387	0.15

[a] At λ_{em} = 590 nm with λ_{ex} = 480 nm.

4.4.3 DNR binding to GTP in aqueous buffer solutions.

The DNR and GTP binding studies were performed by using visible absorbance titration and fluorescence titration. As shown in Figure 4.8, GTP binds to DNR with a relative high binding constant of $K_a = 9.0 \pm 0.2 \times 10^3\ M^{-1}$ at low concentration range and a low binding constant of $K_a = 0.25 \pm 0.05 \times 10^3\ M^{-1}$ (nonspecific binding) at high concentration range. The binding stoichiometry of DNR and GTP was determined to be 1:1 by using absorbance titration Job plot in 10 mM sodium borate buffer pH 8.2. Our fluorescence titration results showed that all the NTPs (ATP, GTP, CTP, UTP) bound to DNR with similar moderate binding affinity (Figure 4.9). GTP bound to DNR with 1:1 stoichiometry was also confirmed by the fluorescence titration Job plot in 10 mM sodium borate buffer pH 8.2.

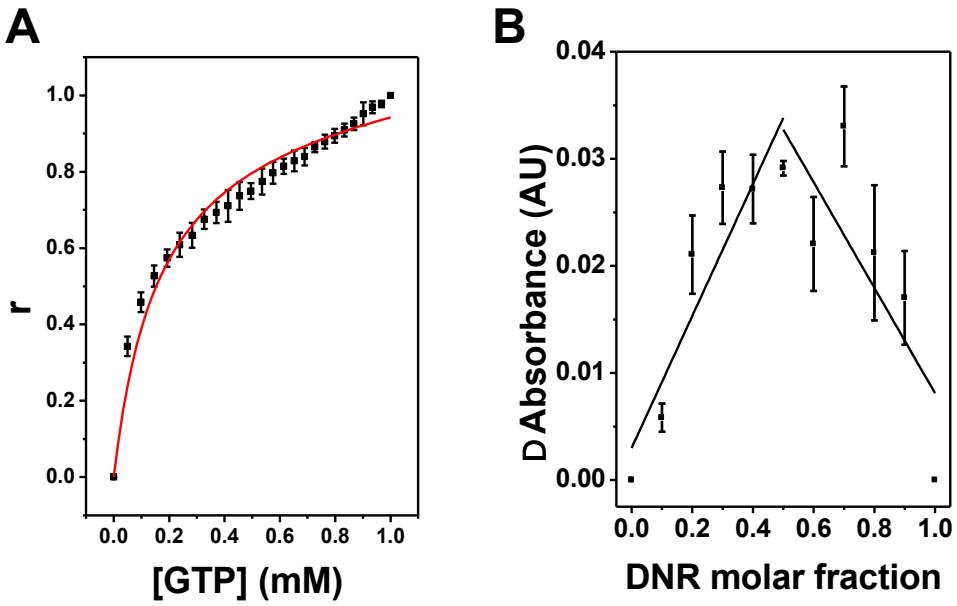


Figure 4.8. Binding studies of GTP and DNR in 10 mM sodium borate buffer pH 8.2 by absorbance titration. **(A)** The absorbance titration of GTP to DNR in 10 mM sodium borate buffer pH 8.2. GTP binds to DNR with $K_a = 6.20 \pm 0.73 \times 10^3$ M^{-1}. **(B)** The binding stoichiometry of DNR and GTP was determined to be 1:1 by using absorbance titration Job plot in 10 mM sodium borate buffer pH 8.2. Error bars represent the standard deviation from three independent experiments.

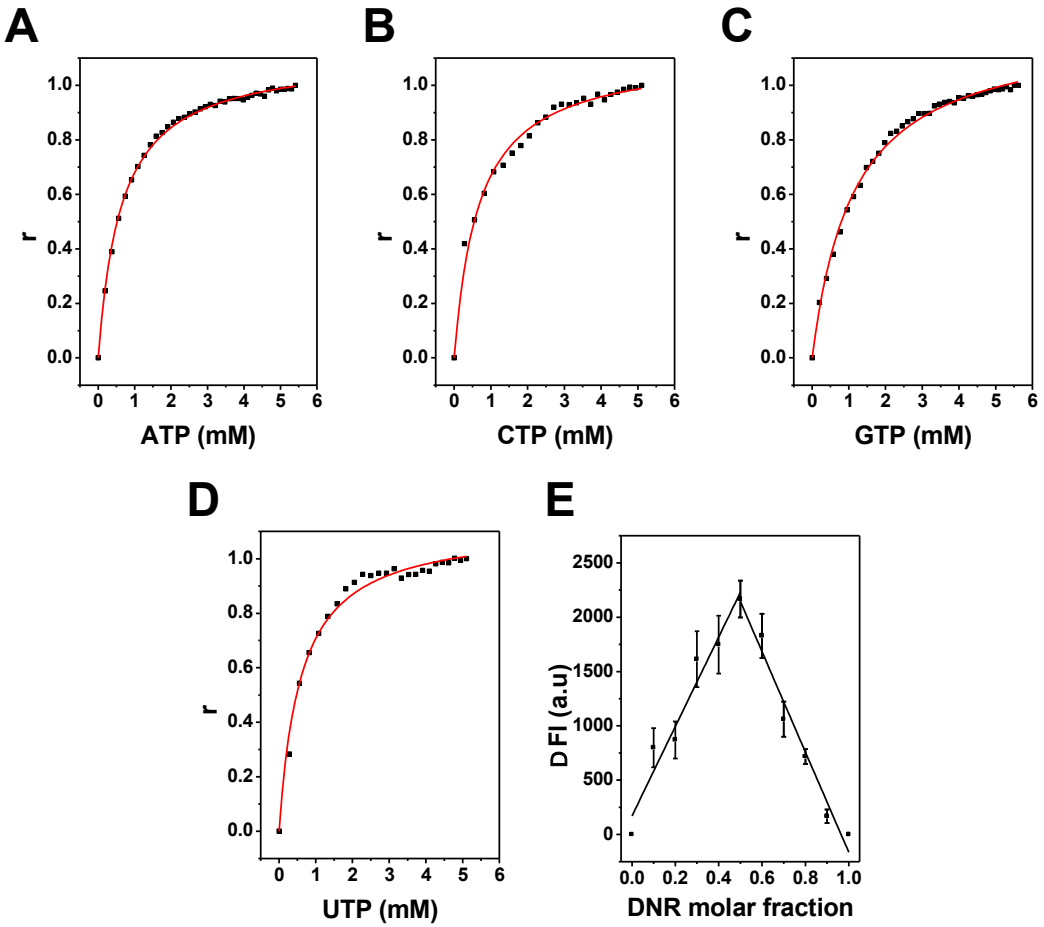


Figure 4.9. Binding studies of NTPs and DNR in 10 mM sodium borate buffer pH 8.2 by fluorescence titration. The fluorescence titration of ATP, CTP, GTP, and UTP to DNR in 10 mM sodium borate buffer pH 8.2. **(A)** ATP binds to DNR with $K_a = 1.56 \pm 0.03 \times 10^3$ M^{-1}. **(B)** CTP binds to DNR with $K_a = 1.48 \pm 0.09 \times 10^3$ M^{-1}. **(C)** GTP binds to DNR with $K_a = 0.88 \pm 0.02 \times 10^3$ M^{-1}. **(D)** UTP binds to DNR with $K_a = 1.69 \pm 0.12 \times 10^3$ M^{-1}. **(E)** The binding stoichiometry of DNR and GTP was determined to be 1:1 by using fluorescence titration Job plot in 10 mM sodium borate buffer pH 8.2.

4.4.4. Binding of the DNR-GTP conjugate to DNA.

The fluorescence and fluorescence anisotropy titration experiments were used to study how DNR and DNR-GTP bind to stDNA in 1×BPES buffer (6 mM Na_2HPO_4, 2 mM NaH_2PO_4, 1 mM Na_2EDTA, and 185 mM NaCl, pH 7.0). Figure 4.10 shows our results. Consistent with previously published results, binding of DNR to DNA greatly quenched the fluorescence intensity (Figure 4.10 A) and increased the anisotropy values of DNR (Figure 4.10 B). Intriguingly, titrating DNA to a fixed concentration of the DNR-GTP conjugate also quenched the fluorescence intensity and increased the anisotropy values of the DNR-GTP conjugate, indicating that the DNR-GTP conjugate binds to DNA under this experimental condition. Nevertheless, the magnitudes of fluorescence quenching and anisotropy increase is much smaller than that of DNR, suggesting that the DNR-GTP conjugate binds to DNA much weaker compared with DNR binding to DNA.

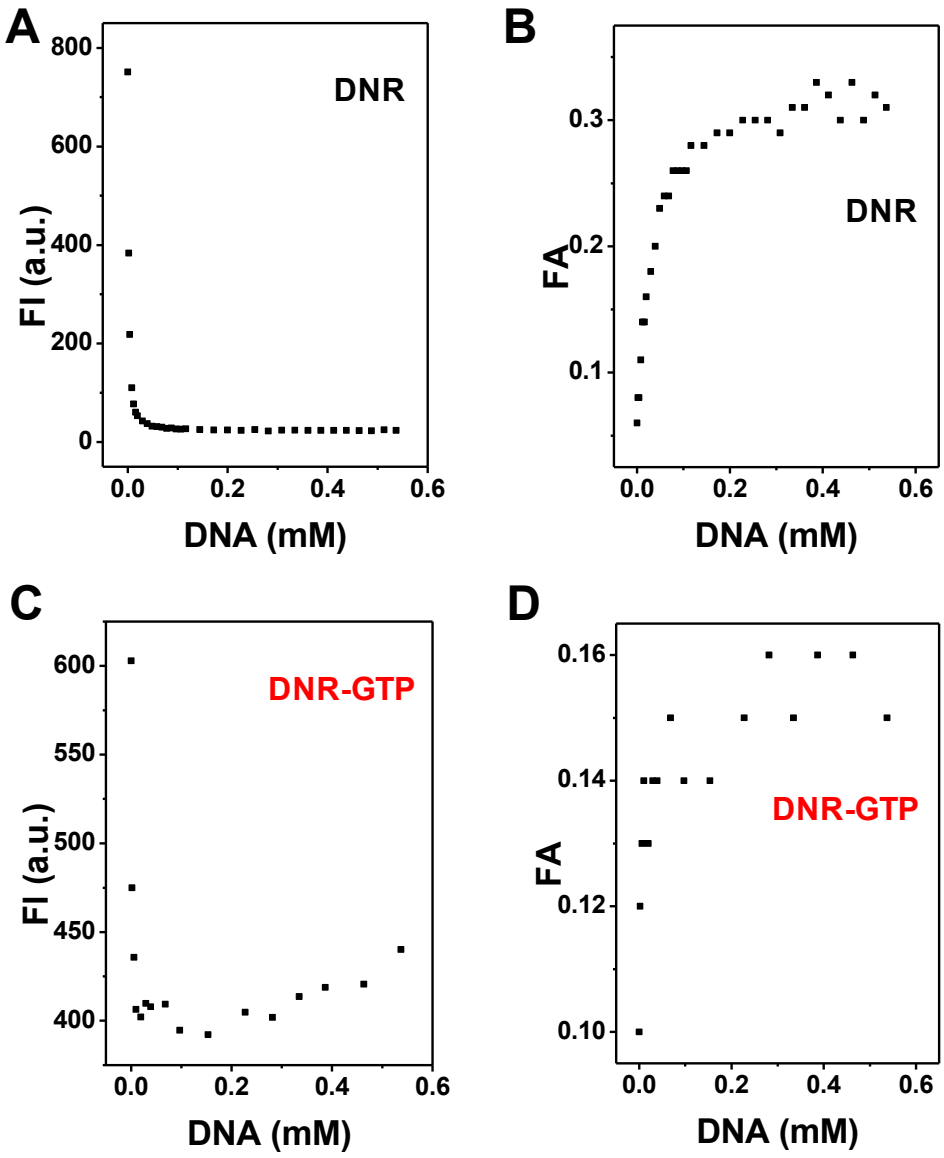


Figure 4.10. Fluorescence intensity (A and C) and anisotropy titration (B and D) experiments in 1×BPES buffer (6 mM Na_2HPO_4, 2 mM NaH_2PO_4, 1 mM Na_2EDTA, and 185 mM NaCl, pH 7.0). Increasing concentrations of DNA were titrated into a fixed concentration of DNR or the DNR-GTP conjugate in 1×BPES. The fluorescence intensity at 590 nm with λ_{em} = 480 nm were measured. The fluorescence anisotropy was determined with λ_{ex}= 480 nm and λ_{em}= 590 nm. **(A)** and **(B)** DNR. **(C)** and **(D)** the DNR-GTP conjugate.

4.4.5 In vitro cytotoxicity and other effects against cancer cells.

In this study, we used the MTT cell viability assays to study the in vitro cytotoxicity of the DOX-GTP and DNR-GTP conjugates and to determine their IC_{50} against two different human ovarian cell lines: the DOX-sensitive SK-OV-3 and the DOX-resistant NCI/ADR-RES cell lines. Figure 4.11 shows our results. The IC_{50} values of DOX against SK-OV-3 and NCI/ADR-RES cells were determined to be 0.48 ± 0.08 and 40 ± 0.6 μM, respectively. These results are consistent with the previously published results [240-243]. The IC_{50} values of the DOX-GTP conjugate against SK-OV-3 and NCI/ADR-RES cells were estimated to be 0.83 ± 0.08 and 10 ± 0.1 μM, respectively. These results suggest that the DOX-GTP is less toxic to the DOX-sensitive SK-OV-3 cells and more toxic to the DOX-resistant NCI/ADR-RES cells. This is more pronounced for the IC_{90} values of DOX and the DOX-GTP conjugate. The IC_{90} values of DOX and the DOX-GTP conjugate against the DOX-resistant NCI/ADR-RES cells are 131 ± 0.74 and 20 ± 0.24 μM, respectively. Similar results are obtained for DNR and DNR-GTP conjugates.

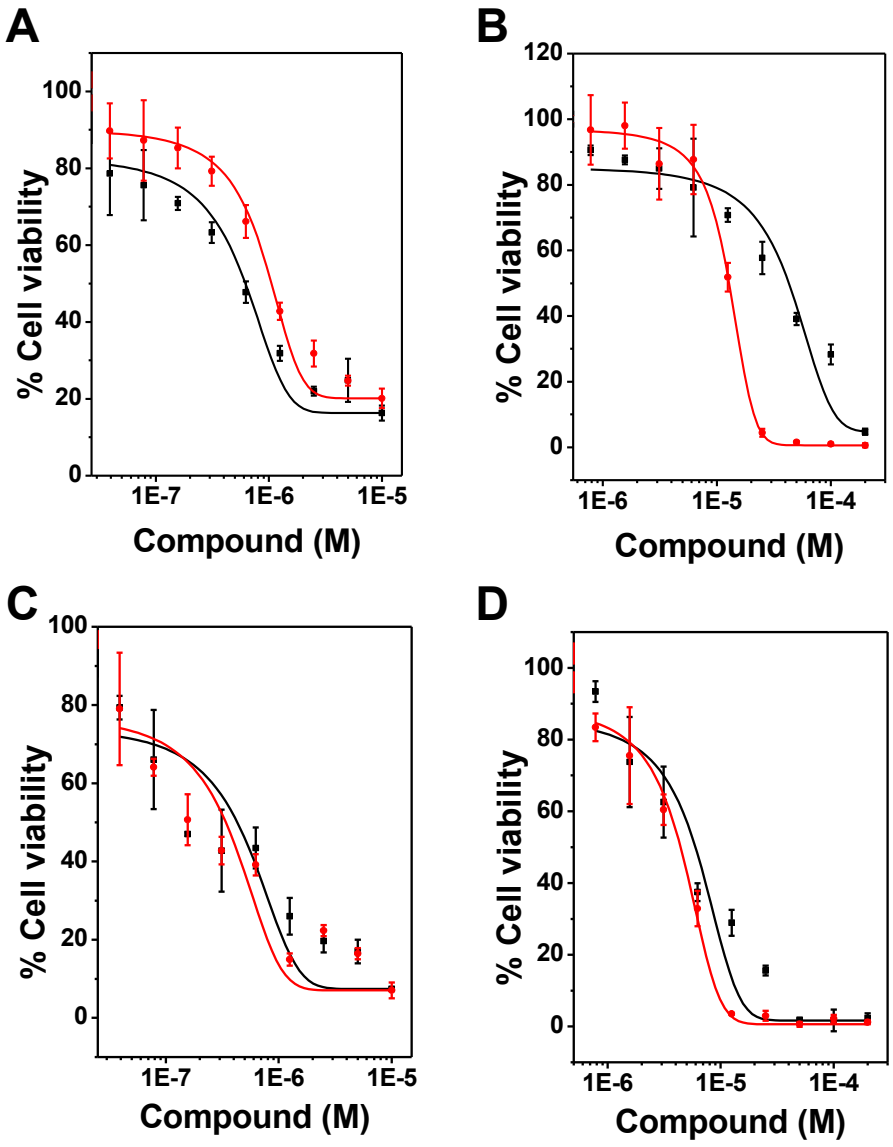


Figure 4.11. Cytotoxicity of DOX, DOX-GTP, DNR and DNR-GTP against human ovarian cell line SK-OV-3 and NCI/ADR-RES. Cells were treated with indicated concentrations of DOX, DOX-GTP, DNR, and DNR-GTP for 72 h and cytotoxicity was then evaluated by MTT assay. **(A)** IC_{50} was determined to be 0.48 ± 0.08 µM and 0.83 ± 0.08 µM for DOX (black squares and line) and DOX-GTP (red circles and line) respectively in SK-OV-3 cell. **(B)** IC_{50} was determined to be 40 ± 0.6 µM and 10 ± 0.1 µM for DOX (black squares and line) and DOX-GTP (red circles and line) respectively in NCI/ADR-RES cell. **(C)** IC_{50} was determined to be 0.37 ± 0.12 µM and 0.30 ± 0.08 µM for DNR (black squares and line) and DNR-GTP (red circles and line) respectively in SK-OV-3 cell. **(D)** IC_{50} was determined to be 5.87 ± 1.00 µM and 4.39 ± 0.29 µM for DNR (black squares and line) and DNR-GTP (red circles and line) respectively in NCI/ADR-RES cell. All values are expressed as the mean ± SD for triplicated samples.

Fluorescence microscopy was used to investigate the intracellular accumulation and localization of DOX, the DOX-GTP conjugate, DNR, and the DNR-GTP conjugate in SK-OV-3 and NCI/ADR-RES cells. As expected, all four compounds entered cells rapidly and accumulated in cell nuclei (Figure 4.12 A). In contrast, much less DOX and DNR accumulated in NCI/ADR-RES cells (Figure 4.12 B). These results are consistent with previously published results [241, 244, 245]. Interestingly, more of the DOX-GTP conjugate and the DNR-GTP conjugate entered and accumulated in NCI/ADR-RES cells. The cellular distribution is different from that in SK-OV-3 cells. For instance, most cells with the DOX-GTP conjugate and the DNR-GTP conjugate have a very bright point. Since the DOX-GTP conjugate and the DNR-GTP conjugate contain GTP and may be used by RNA polymerase, the bright point may be nucleoli. More studies are needed to confirm this hypothesis. We also estimated the fluorescence intensity of DOX, the DOX-GTP conjugate, DNR, and the DNR-GTP conjugate in SK-OV-3 and NCI/ADR-RES cells. Figure 4.13 Shows our results. The fluorescence intensity of these compounds is consistent with the IC_{50} determined by the MTT cell viability assays.

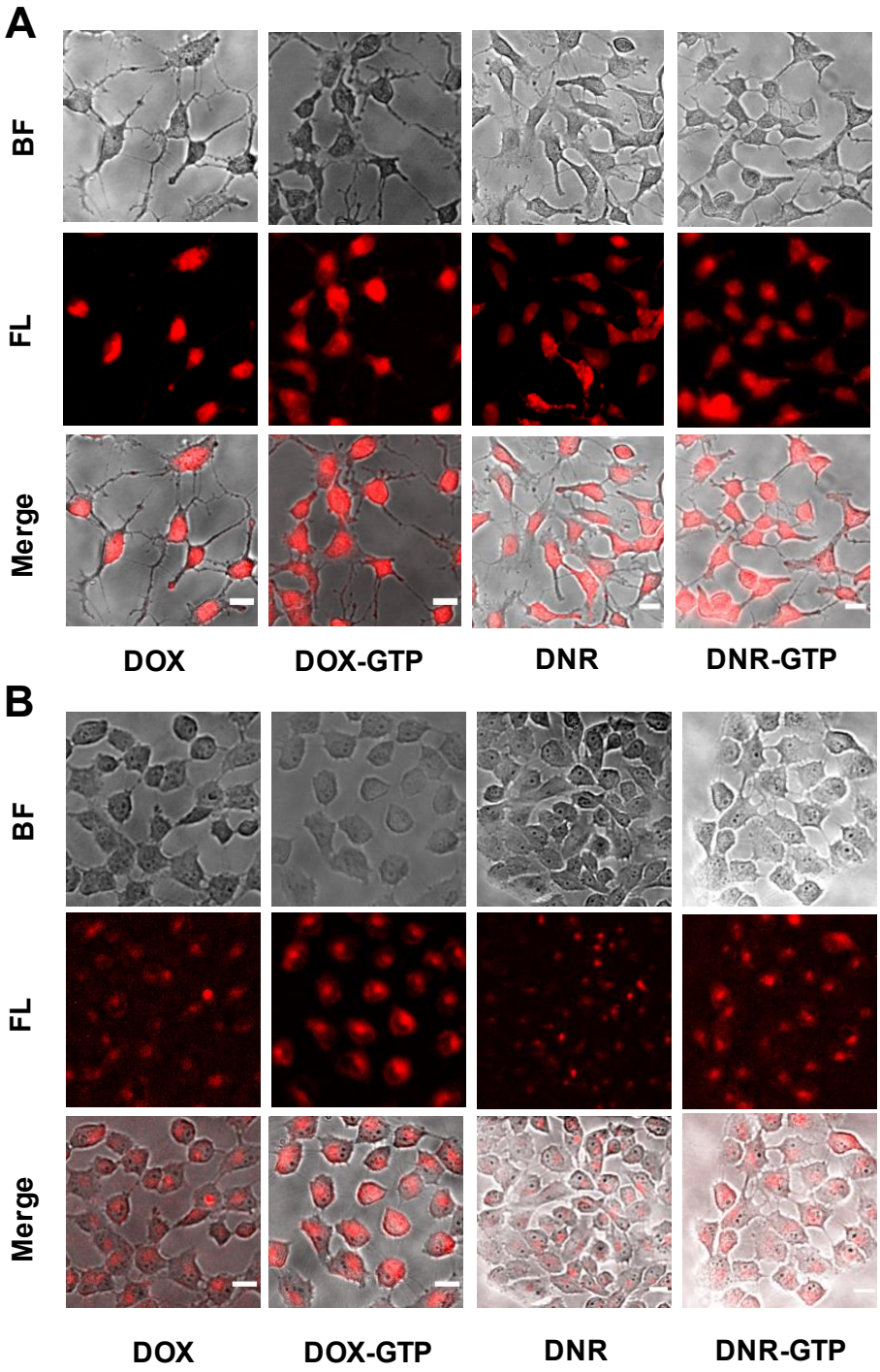


Figure 4.12. Fluorescence microscopy images of SK-OV-3 **(A)** and NCI/ADR-RES **(B)** cells after incubation with 5 μM of DOX, DOX-GTP, DNR and DNR-GTP respectively for 4 h. The images were taken under 20×. Scale bars 20 μm. (BF: bright field image; FL: fluorescence image; Merge: merged image)

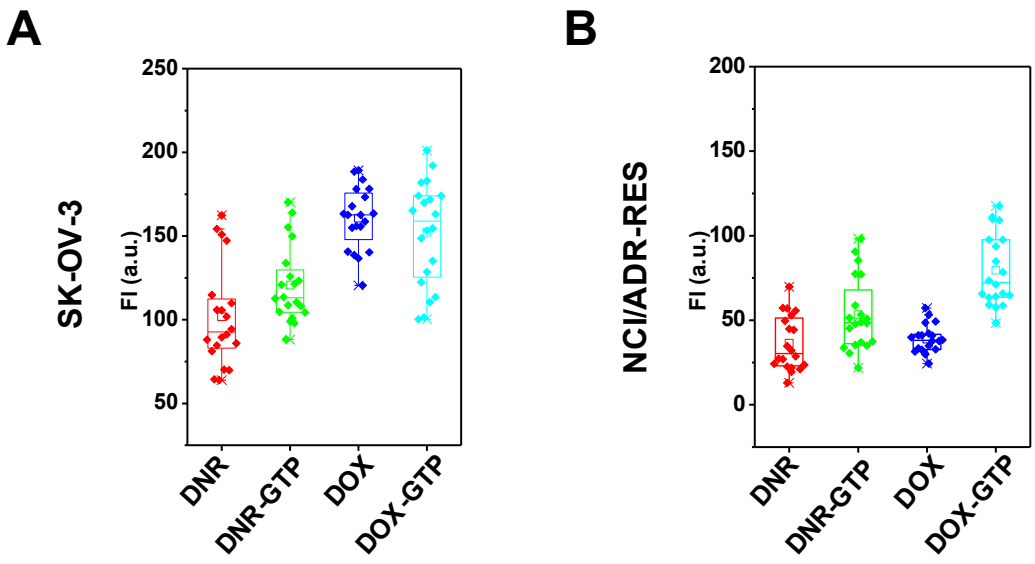


Figure 4.13. Fluorescence intensity quantification of the accumulation of DNR, DNR-GTP, DOX and DOX-GTP in SK-OV-3 **(A)** and NCI/ADR-RES cells **(B)**.

Doxorubicin and daunorubicin are human DNA topoisomerase II poisons [223, 224]. Previous studies showed that HMGA1 and HMGA2 are related to human DNA topoisomerase II poisoning [126, 127]. We therefore decided to examine the expression levels of human DNA topoisomerase I, topoisomerase IIα, HMGA1, and HMGA2 using Western blotting experiments. Figure 4.14 Shows our results. Interestingly, DOX-resistant NCI/ADR-RES cells expressed much more HMGA1 than that of the DOX-sensitive SK-OV-3 cells. Likewise, DOX-resistant NCI/ADR-RES cells also expressed more human DNA topoisomerase I than that of the DOX-sensitive SK-OV-3 cells. Treatment with 10 μM of DNR, DNR-GTP, DOX, or DOX-GTP for 48 hours did not caused significant changes to the expression levels of human DNA topoisomerase I, topoisomerase IIα, HMGA1, and HMGA2.

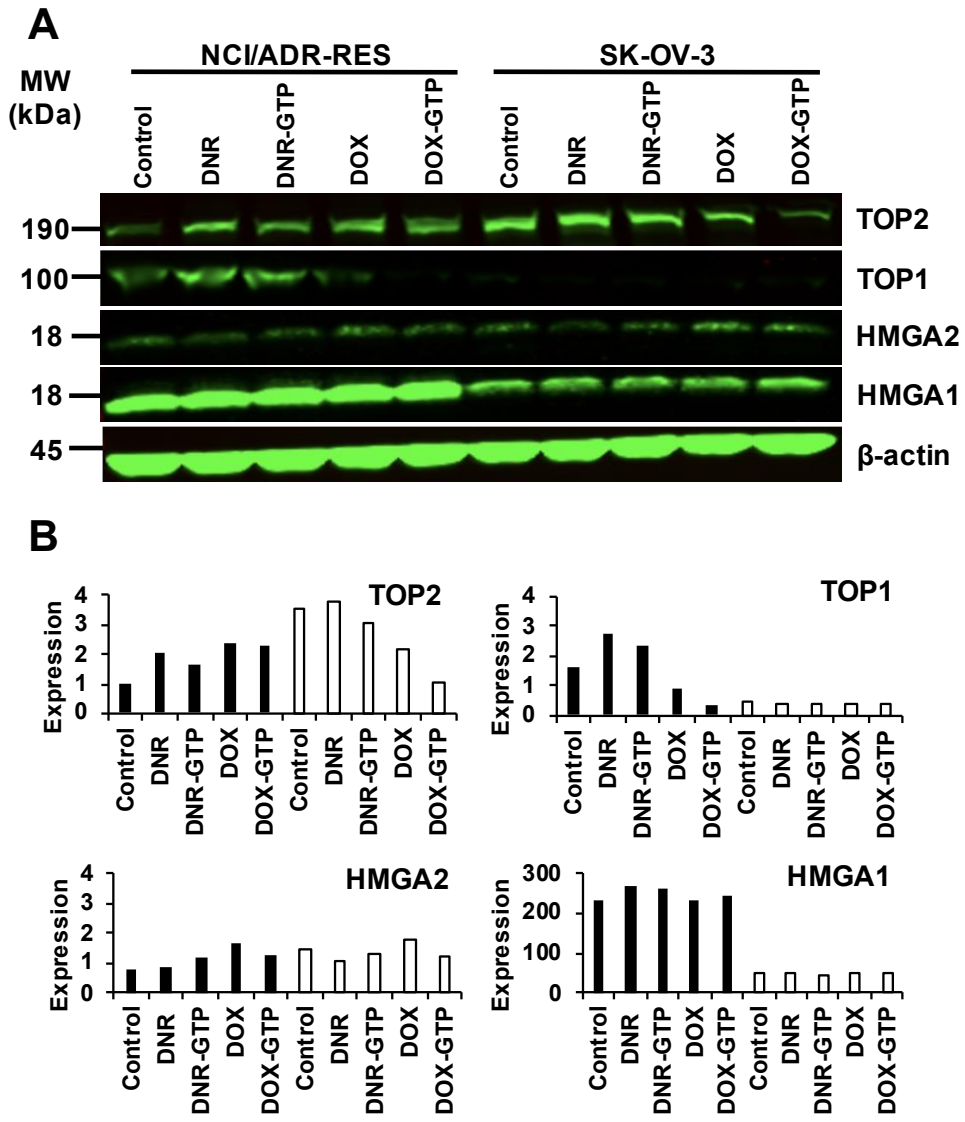


Figure 4.14. The western blot analysis of TOP2α, TOP1, HMGA2, and HMGA1 expression in NCI/ADR-RES and SK-OV-3 cells with treatment of DNR, DNR-GTP, DOX and DOX-GTP, respectively. **(A)** The western blot analysis of TOP2α, TOP1, HMGA2, and HMGA1 expression of NCI/ADR-RES and SK-OV-3 cells in the absence (-) and presence (+) of 10 μM DNR, DNR-GTP, DOX and DOX-GTP, respectively. The loading amount of whole cell extracts is 20 μg, and β-actin was used as a loading control. The original images are provided in the supplemental information Figure S4.1 **(B)** Normalized signal intensity of TOP2α, TOP1, HMGA2, and HMGA1 in the western blot of NCI/ADR-RES (black fill) and SK-OV-3 (black line) cells in the absence and presence of 10 μM DNR, DNR-GTP, DOX and DOX-GTP. The western blot protein signal intensity for NCI/ADR-RES and SK-OV-3 cell whole cell extracts were first normalized to β-actin signal to correct for loading, then divided by the TOP2α signal intensity in NCI/ADR-RES in the absence of drug treatment for the comparation of relative protein levels.

4.5 Discussion

4.5.1 Specific Crosslinking of DNR/DOX to GTP or dGTP by HCHO to Form DNR/DOX-GTPs and DNR/DOX-dGTPs.

In our results, NTPs (ATP, GTP, CTP, UTP) and dNTPs (dATP, dGTP, dCTP, dTTP) were used in the reaction for conjugate formation. We found that DNR/DOX was specifically covalently crosslinked to GTP/dGTP by HCHO to yield DNR/DOX-GTP/dGTPs products (Figure 4.3 A and B). The specific conjugation of DNR/DOX with GTP/dGTP likely results from the fact that GTP and dGTP can interact with DNR and DOX, and form noncovalent DNR/DOX-GTP/dGTP complexes at a one-to-one molar ratio in the absence of HCHO. From the results of DNR and GTP binding studies by absorbance titration and fluorescence titration, GTP was determined to bind to DNR with a binding constant of $6.20 \pm 0.73 \times 10^3$ M^{-1}. Further, the binding stoichiometry of GTP and DNR was determined to be 1:1 in molar by using continuous variation analysis in 10 mM sodium borate buffer pH 8.2. From our ESI-MS studies, the chemical formula of DNR-GTP was determined to be $C_{39}H_{45}N_6O_{24}P_3$. Two molecules of HCHO are used in the formation of the conjugate, the 3'-NH_2 group of DNR crosslinked with 2'-NH_2 of guanine by one molecule of formaldehyde via Schiff base chemistry and one additional molecule of formaldehyde was used to form the oxazolidine ring in the group of daunosamine (Figure 4.4 B). The formation of the bis-oxazolidinylmethane is due to the interaction of the amino alcohol functionality of the DNR with formaldehyde [246]. This finding is consistent to previous studies that the crosslinking of anthracycline drugs to dsDNA was through the methylene group generated from formaldehyde linked the 3'-NH_2 group of DNR/DOX on one side to the 2'-NH_2 group of guanine on the other side [246, 247].

4.5.2 Optimize the Synthesis Reaction Conditions.

To improve the yield of DNR-GTP, we optimized the experimental conditions. We found that the DNR-GTP crosslink reaction is thermolabile by conducting the crosslinking reactions at 20 ℃, 37 ℃, 55 ℃, 65 ℃, 80 ℃, and 95 ℃, respectively for 2 hours shown in Figure 4.3 C. The yield of DNR-GTP decreased dramatically at high temperatures over the optimal temperature of 55 ℃, especially at 95 ℃ where there were almost no products left. The heat labile property of DNR-GTP was also confirmed by the stability studies. First, we tested the stability of DNR-GTP samples at the concentration of 2 mM at 4 °C and RT respectively by TLC. Our results show the DNR-GTP is more stable at a relative low temperature 4 °C compared with that at RT in Figure 4.5. More importantly, we found that DNR-GTPs are more susceptible to light by fluorescence spectrometry at relatively low concentrations (Figure 4.6).

From the formaldehyde crosslinking titration experiments (Figure 4.3 D), it was found that excess starting materials of GTP result in the decreasing of DNR-GTP products. The GTPs may form dimers by HCHO through linking the 2'-NH_2 group of guanine and prevent the crosslinking to DNR. This could also happen in the case of excess DNR, which dimerizes to form daunoform, consisting of two drug molecules bound together with three methylene groups as illustrated in reference [246]. In order to synthesize and purify the DNR-GTPs on a large scale and as well as reduce the amount of free GTP contamination left in the products, the following reaction conditions were used: 2 mM DNR, 1 mM GTP, 10 mM sodium borate buffer pH 8.2, 0.5% (v/v) HCHO at 37 °C for 24 hours (Figure 4.3 F).

4.5.3 Determination of DNR/DOX and DNR/DOX-GTPs' Intracellular Distribution in SK-OV-3 and NCI/ADR-RES Cells.

The intracellular drug accumulation is a complex process including drug uptake into the cell, retention and distribution in the cell, and sometimes efflux from the drug resistant cell which has over expressed plasma membrane drug efflux pumps, such as the P-gp and the MRP [248-250]. The intracellular accumulation of anthracycline drugs is an important parameter for evaluation of their cytotoxic action. From the results in Figure 4.12 A, the intracellular localization and accumulation of drugs were examined in SK-OV-3 with the fluorescence patterns demonstrating that there was a homogeneous distribution of DOX, DOX-GTP, DNR, and DNR-GTP, which are predominantly localized in nuclei with similar noticeable fluorescence intensity. We also examined the drug's distribution and accumulation in the DOX-resistant NCI/ADR-RES cells treated with DOX, DOX-GTP, DNR, and DNR-GTP (Figure 4.12 B) showing that the overall fluorescence signals were significantly weaker compared to the data from SK-OV-3 cells with most of the fluorescence signals of DNR and DOX presenting as diffuse or dot-like patterns localized in both the cytoplasm and the nuclei. More importantly, DNR-GTP and DOX-GTP improved their cellular uptake to the cells and increased the accumulation both in cytoplasm and nuclei compared to free DNR and DOX in drug resistant cells likely because of the conjugate's increased structural polarity due to GTP being a highly negatively charged molecule (Figure 4.4 B) which may interfere with partial functions of the plasma membrane drug efflux pumps. The high fluorescence intensity of DNR-GTP and DOX-GTP in NCI/ADR-RES cells were located in nuclei as discrete regions or formed local clusters in cytoplasm may be related to certain active transport of the drugs into the cells

in addition to their diffusion. In summary, our preliminary results have demonstrated that the accumulation and intracellular distribution of DNR/DOX differs in DOX-sensitive SK-OV-3 and DOX-resistant NCI/ADR-RES cells from the DNR/DOX-GTP formulations. It has been shown that DNR/DOX-GTP formulations accumulate at higher degree in DOX-resistant NCI/ADR-RES cells compared to free DNR/DOX form.

4.5.4 Improved Cytotoxicity Sensitivity of DOX-resistant NCI/ADR-RES Cells to DNR/DOX-GTPs.

From the cytotoxicity studies of DOX-GTP and DNR-GTP against the DOX-resistant NCI/ADR-RES cells, it demonstrated that the DNR-GTP was slightly more cytotoxic compared to free DNR and more importantly, DOX-GTP showed a 4-fold higher cytotoxicity relative to free DOX (Figure 4.11). We are curious about how the conjugates improve the drug sensitivity of NCI/ADR-RES. Previous studies showed that DNR and DOX catalyzed the production of superoxide and hydrogen peroxide by the redox machinery of the quinone functionality and finally to produce some formaldehydes [251-254]. DNR and DOX react with DNA in presence of formaldehyde to form virtual crosslinks which induce programmed cell death [255]. As a consequence, DNR-GTP and DOX-GTP, which carry their additional own formaldehyde into the cells, are more effective against the NCI/ADR-RES cells. An alternative explanation for increased drug resistance in NCI/ADR-RES cells would be the presence of an overexpressed plasma membrane drug efflux pump, P-glycoprotein, which reduces the intracellular accumulation of DOX. The increased intracellular accumulation of DOX-GTP may be a direct result of it being a poor substrate for the efflux pump compared to free DOX. The DNR/DOX-GTP

formulations accumulate at higher degree in cells, consistent with the higher cytotoxicity observed in cells.

Rapidly dividing cells maintain chromatin supercoiling homeostasis via TOP1 and TOP2α. Several important anticancer drugs perturb this homeostasis by targeting TOP1 and TOP2α. The anthracycline antibiotics DNR and DOX are among the most potent anticancer drugs. The mechanism of action of these two drugs stems from the DNA intercalation and the inhibition of TOP2α. From our CD results that DNR-GTP binds to DNA by intercalation and DNR-GTP treatment on NCI/ADR-RES cells induced TOP2α expression from western blot studies which indicate that it may exert the cytotoxic activity by the same mechanism as free DNR. Previous studies showed that the oncofetal chromatin structuring HMGA2 protein attenuates genotoxic damage induced by TOP2α target compounds through the regulation of local DNA topology [127]. From our western blot data, there were no significant changes detected at protein levels of HMGA2 after 48 hours treatment with 10 μM of DNR, DNR-GTP, DOX, and DOX-GTP relative to the blank controls without drugs treatments in both SK-OV-3 and NCI/ADR-RES indicating that the drug sensitivity of both cell lines is not related to the expression of HMGA2. In summary, DNR-GTP and DOX-GTP, which can skip formaldehyde production by oxidation of intracellular components and bypass the resistant mechanism from drug efflux pump, improved the cytotoxicity against DOX-resistant NCI/ADR-RES cells.

4.6 Conclusions

In these studies, we successfully crosslinked DNR and DOX to GTP and dGTP by HCHO in solution to synthesize the conjugates. The reaction is nucleobase specific to guanine and with a requirement of the 3'-NH_2 group of DNR or DOX, 2'-NH_2 group of

guanine and two molecules of formaldehyde. The synthesized DNR/DOX-GTPs are efficiently purified by phenol extraction and butanol precipitation. The DNR/DOX-GTPs were characterized by UV-Vis, fluorescence, CD, and mass spectrometry, and showing that they bind to DNA by intercalation. More importantly, from the cell studies, the prodrug of DOX-GTP improved the cellular uptake and accumulation of DOX in the cell and further induced potent toxicity in the DOX-resistant NCI/ADR-RES. Our studies provide a new strategy to conjugate DNR/DOX to a single nucleotide and provide a new therapeutic approach to treat DOX-resistant cancer cell lines.

CHAPTER 5: Summary and Future Work

5.1 Summary

The dissertation research focuses on the study of HMGA2-DNA interactions, and synthesis, characterization of new daunorubicin/doxorubicin derivatives (DNR/DOX-GTP/dGTP conjugates). There are three main objectives in this dissertation. The first one is to develop a miniaturized, automated high-throughput screening assay to identify inhibitors targeting HMGA2-DNA interactions. The second one is to study the interactions of the negatively charged C-terminus peptide (CTP) and positively charged AT-hook peptides (ATHPs) and find out how the C-terminus affects the DNA-binding properties of HMGA2. The last objective is to improve daunorubicin/doxorubicin therapeutic efficacy by formaldehyde crosslinking to GTP/dGTP to form DNR/DOX-GTP/dGTP conjugates to overcome the multidrug resistance (MDR) of certain cancer cells.

In chapter 1, the biochemical and biophysical properties of HMGA2 and its association with adipogenesis were reviewed. As discussed above, HMGA2 plays an important role in adipogenesis and is an excellent target for the treatment of obesity. Since the overexpression and/or aberrant-expression of HMGA2 is directly linked to the formation of a variety of malignant tumors, including lung cancer, breast cancer, prostate cancer, leukemia, and melanoma, HMGA2 appears to be an attractive target for anticancer drugs. Several strategies may be used to target HMGA2 for therapeutic purposes were discussed. Such as, target the AT-rich DNA-binding sequences that HMGA2 recognizes, identify compounds that bind to HMGA2 and prevent it from binding to AT-rich DNA sequences, use the negatively charged C-terminus to inhibit HMGA2 binding to AT-rich DNA sequences, and target HMGA2's mRNA.

In chapter 2, a miniaturized, automated AlphaScreen ultra-high-throughput screening assay to identify inhibitors targeting HMGA2-DNA interactions was developed. After screening the LOPAC1280 compound library, several compounds were identified that strongly inhibit HMGA2-DNA interactions including suramin, a century-old, negatively charged antiparasitic drug. The results show that the inhibition is likely through suramin binding to the "AT-hook" DNA-binding motifs and therefore preventing HMGA2 from binding to the minor groove of AT-rich DNA sequences. Biochemical and biophysical studies show that charge-charge interactions and hydrogen bonding between the suramin sulfonated groups and Arg/Lys residues play critical roles in the binding of suramin to the "AT-hook" DNA-binding motifs. Furthermore, the results suggest that HMGA2 may be one of suramin's cellular targets.

In chapter 3, it was demonstrated that the negatively charged C-terminus greatly affected the DNA-binding properties of HMGA2. For instance, the C-terminal deletion mutant HMGA2Δ95-108 binds much more tightly to the AT-rich DNA sequences compared with the wildtype HMGA2. The studies showed that the negatively charged CTP strongly inhibited HMGA2 binding to AT-rich DNA sequences through its binding to the positively charged AT-hooks motifs by PDI-ELISA, fluorescence anisotropy, and AKTA FPLC size exclusion chromatography. A structure of ATHP3-CTP complex based on molecular dynamic simulation showed that charge-charge interactions and hydrogen bonds were the major forces to form the complex. This study suggests that the CTP be used as a potent inhibitor to target HMGA2. These findings will help to develop a new therapeutic strategy against HMGA2-associated cancers and obesities by using the highly negatively charged CTP or synthetic CTP mimics as an inhibitor.

In chapter 4, it showed that formaldehyde can efficiently and specifically conjugate daunorubicin (DNR) and doxorubicin (DOX) to guanosine-5'-triphosphate (GTP) and 2'-deoxyguanosine-5'-triphosphate (dGTP) to yield daunorubicin-GTP (DNR-GTP), doxorubicin-GTP (DOX-GTP), daunorubicin-dGTP (DNR-dGTP), and doxorubicin-dGTP (DOX-dGTP) conjugates. The linkage is between the 2'-NH_2 of guanine and the 3'-NH_2 of daunosamine. The specific conjugation of DNR/DOX with GTP/dGTP likely results from the fact that GTP and dGTP can interact with DNR and DOX, and form noncovalent DNR/DOX-GTP/dGTP complexes at a one-to-one molar ratio. The DNR/DOX-GTP/dGTP conjugates were characterized by different methods including UV-Vis, fluorescence, CD, and mass spectroscopy. The results also show that DNR/DOX-GTP/dGTP conjugates bind to DNA by intercalation. Furthermore, it was found that the DNR/DOX-GTP/dGTP conjugates rapidly accumulated in human cancer cells and were cytotoxic to both doxorubicin-sensitive SK-OV-3 and doxorubicin-resistant NCI/ADR-RES cells. Interestingly, the results suggest that these DNR/DOX-GTP/dGTP conjugates may overcome doxorubicin resistance for certain cancer cells.

5.2 Future work

The first work is inspired by the phenomena as below. The preliminary fluorescence microscopic results demonstrate that DNR/DOX-GTPs improved their cellular uptake to the cells and increased the accumulation in both cytoplasm and nuclei. The fluorescence microscopic observation was consistent with the enhanced cell cytotoxicity data and suggested that DNR-GTP and DOX-GTP could be effectively taken up by cells and exert their cytotoxic activity against the DOX-resistant NCI/ADR-RES cells. Significantly, the dot-like patterns of DNR/DOX-GTPs in the nuclei indicates that DNR/DOX-GTPs may be

used by the process of DNA transcription to make RNA and concentrate at this specific area. This process may also exert additional toxicity to kill the cells. This hypothesis will need to be verified by Hoechst33258 staining of the nucleus to find the specific location of the DNR/DOX-GTPs. Another hypothesis is that DNR/DOX-dGTPs may take part in the DNA replication process to exert additional toxicity to cancer cells. The fluorescence microscopic studies of accumulation and distribution patterns of DNR/DOX-dGTPs, and their cytotoxic activity against the DOX-sensitive SK-OV-3 and DOX-resistant NCI/ADR-RES cells will be tested in the future work.

The second work is inspired by a new phenomenon of mixing doxorubicin and GTP in 10 mM sodium borate buffer resulting in the formation of a covalently bound DOX-GTP conjugate. The structure of DOX-GTP conjugate was identified by our mass spectrometry. The linkage is between the 2'-NH_2 of guanine and the 14-CH_2OH of doxorubicin which means this new product is different from above discussed DOX-GTP which was crosslinked by formaldehyde. This may be the mechanism of doxorubicin killing cells. In the future work of this project may need focus on these areas, such as doxorubicin and GTP/dGTP binding studies, doxorubicin-GTP/dGTP conjugates synthesis and characterizations on NMR, optical properties, cell cytotoxicity, etc.

REFERENCES

1. Su, L., Z. Deng, and F. Leng, The mammalian high mobility group protein AT-Hook 2 (HMGA2): biochemical and biophysical properties, and its association with adipogenesis. International Journal of Molecular Sciences, 2020. 21(10): p. 3710.

2. Bianchi, M.E. and M. Beltrame, Upwardly mobile proteins. EMBO Reports, 2000. 1(2): p. 109-114.

3. Friedmann, M., et al., Organization, inducible-expression and chromosome localization of the human HMG-I (Y) nonhistone protein gene. Nucleic Acids Research, 1993. 21(18): p. 4259-4267.

4. Manfioletti, G., et al., cDNA cloning of the HML-C phosphoprotein, a nclear protein associated with neoplastic and undifferentiated phenotypes. Nucleic Acids Research, 1991. 19(24): p. 6793-6797.

5. Zhou, X., et al., Genomic structure and expression of the murine Hmgi-c gene. Nucleic Acids Research, 1996. 24(20): p. 4071-4077.

6. Hauke, S., et al., Sequencing of intron 3 of HMGA2 uncovers the existence of a novel exon. Genes, Chromosomes and Cancer, 2002. 34(1): p. 17-23.

7. Goodwin, G.H. and E.W. Johns, Isolation and characterisation of two calf-thymus chromatin non-histone proteins with high contents of acidic and basic amino acids. European Journal of Biochemistry, 1973. 40(1): p. 215-219.

8. Goodwin, G.H., C. Sanders, and E.W. Johns, A new group of chromatin-associated proteins with a high content of acidic and basic amino acids. European Journal of Biochemistry, 1973. 38(1): p. 14-19.

9. Bustin, M., Revised nomenclature for high mobility group (HMG) chromosomal proteins. Trends in Biochemical Sciences, 2001. 26(3): p. 152-153.

10. Johns, E., The Chromosomal Proteins. 2012: Elsevier.

11. Lund, T., et al., On the presence of two new high mobility group-like proteins in HeLa S3 cells. FEBS Letters, 1983. 152(2): p. 163-167.

12. Giancotti, V., et al., Changes in nuclear proteins on transformation of rat epithelial thyroid cells by a murine sarcoma retrovirus. Cancer Research, 1985. 45(12 Part 1): p. 6051-6057.

13. Goodwin, G.H., et al., Fractionation by high-performance liquid chromatography of the low-molecular-mass high-mobility-group (HMG) chromosomal proteins present in proliferating rat cells and an investigation of the HMG proteins present

in virus transformed cells. European Journal of Biochemistry, 1985. 149(1): p. 47-51.

14. Patel, U.A., et al., Expression and cDNA cloning of human HMGI-C phosphoprotein. Biochemical and Biophysical Research Communications, 1994. 201(1): p. 63-70.

15. Asher, H., et al., Disruption of the architectural factor HMGI-C: DNA-binding AT hook motifs fused in lipomas to distinct transcriptional regulatory domains. Cell, 1995. 82(1): p. 57-65.

16. Schoenmakers, E.F., et al., Recurrent rearrangements in the high mobility group protein gene, HMGI-C, in benign mesenchymal tumours. Nature Genetics, 1995. 10(4): p. 436-444.

17. Ashar, H., et al., Genomic characterization of human HMGIC, a member of the accessory transcription factor family found at translocation breakpoints in lipomas. Genomics, 1996. 31(2): p. 207-214.

18. Lee, Y.S. and A. Dutta, The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. Genes & Development, 2007. 21(9): p. 1025-1030.

19. Yu, F., et al., let-7 regulates self renewal and tumorigenicity of breast cancer cells. Cell, 2007. 131(6): p. 1109-1123.

20. Park, S.-M., et al., Let-7 prevents early cancer progression by suppressing expression of the embryonic gene HMGA2. Cell Cycle, 2007. 6(21): p. 2585-2590.

21. Frost, L., et al., The dimerization state of the mammalian high mobility group protein AT-hook 2 (HMGA2). PLoS One, 2015. 10(6): p. e0130478.

22. Giancotti, V., et al., Elevated levels of a specific class of nuclear phosphoproteins in cells transformed with v-ras and v-mos oncogenes and by cotransfection with c-myc and polyoma middle T genes. The EMBO Journal, 1987. 6(7): p. 1981-1987.

23. Yie, J., et al., The role of HMG I (Y) in the assembly and function of the IFN-β enhanceosome. The EMBO Journal, 1999. 18(11): p. 3074-3089.

24. Edwards, L.K., Biochemical characterization of mammalian high mobility group protein A2. Master of Science Thesis, Florida International University, Miami, FL, USA, March 2006.

25. Danielsson, J., et al., The intrinsically disordered RNR inhibitor Sml1 is a dynamic dimer. Biochemistry, 2008. 47(50): p. 13428-13437.

26. Sigalov, A.B., Structural biology of intrinsically disordered proteins: Revisiting unsolved mysteries. Biochimie, 2016. 125: p. 112-118.

27. Sigalov, A., D. Aivazian, and L. Stern, Homooligomerization of the cytoplasmic domain of the T cell receptor ζ chain and of other proteins containing the immunoreceptor tyrosine-based activation motif. Biochemistry, 2004. 43(7): p. 2049-2061.

28. Lanza, D.C., et al., Human FEZ1 has characteristics of a natively unfolded protein and dimerizes in solution. Proteins: Structure, Function, and Bioinformatics, 2009. 74(1): p. 104-121.

29. Simon, S., et al., Regulation of Escherichia coli SOS mutagenesis by dimeric intrinsically disordered umuD gene products. Proceedings of the National Academy of Sciences, 2008. 105(4): p. 1152-1157.

30. Pieprzyk, J., et al., Homodimerization propensity of the intrinsically disordered N-terminal domain of Ultraspiracle from Aedes aegypti. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 2014. 1844(6): p. 1153-1166.

31. Reeves, R. and M.S. Nissen, The AT-DNA-binding domain of mammalian high mobility group I chromosomal proteins. A novel peptide motif for recognizing DNA structure. Journal of Biological Chemistry, 1990. 265(15): p. 8573-8582.

32. Geierstanger, B.H., et al., Short peptide fragments derived from HMG-I/Y proteins bind specifically to the minor groove of DNA. Biochemistry, 1994. 33(17): p. 5347-5355.

33. Huth, J.R., et al., The solution structure of an HMG-I (Y)–DNA complex defines a new architectural minor groove binding motif. Nature Structural Biology, 1997. 4(8): p. 657-665.

34. Fonfría-Subirós, E., et al., Crystal structure of a complex of DNA with one AT-hook of HMGA1. PloS One, 2012. 7(5): p. e37120.

35. Chen, B., J. Young, and F. Leng, DNA bending by the mammalian high-mobility group protein AT hook 2. Biochemistry, 2010. 49(8): p. 1590-1595.

36. Schenk, E.R., et al., Isomerization kinetics of AT hook decapeptide solution structures. Analytical Chemistry, 2014. 86(2): p. 1210-1214.

37. Huang, J., et al., CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nature Methods, 2017. 14(1): p. 71-73.

38. Ho, B.K. and K.A. Dill, Folding very short peptides using molecular dynamics. PLoS Computational Biology, 2006. 2(4): p. e27.

39. Phillips, J.C., et al., Scalable molecular dynamics with NAMD. Journal of Computational Chemistry, 2005. 26(16): p. 1781-1802.

40. Solomon, M.J., F. Strauss, and A. Varshavsky, A mammalian high mobility group protein recognizes any stretch of six AT base pairs in duplex DNA. Proceedings of the National Academy of Sciences, 1986. 83(5): p. 1276-1280.

41. Maher, J.F. and D. Nathans, Multivalent DNA-binding properties of the HMG-1 proteins. Proceedings of the National Academy of Sciences, 1996. 93(13): p. 6716-6720.

42. Thanos, D. and T. Maniatis, The high mobility group protein HMG I (Y) is required for NF-κB-dependent virus induction of the human IFN-β gene. Cell, 1992. 71(5): p. 777-789.

43. Du, W., D. Thanos, and T. Maniatis, Mechanisms of transcriptional synergism between distinct virus-inducible enhancer elements. Cell, 1993. 74(5): p. 887-898.

44. Baldassarre, G., et al., Onset of natural killer cell lymphomas in transgenic mice carrying a truncated HMGI-C gene by the chronic stimulation of the IL-2 and IL-15 pathway. Proceedings of the National Academy of Sciences, 2001. 98(14): p. 7970-7975.

45. John, S., et al., Regulation of cell-type-specific interleukin-2 receptor alpha-chain gene expression: potential role of physical interactions between Elf-1, HMG-I (Y), and NF-kappa B family proteins. Molecular and Cellular Biology, 1995. 15(3): p. 1786-1796.

46. Cui, T. and F. Leng, Specific recognition of AT-rich DNA sequences by the mammalian high mobility group protein AT-hook 2: a SELEX study. Biochemistry, 2007. 46(45): p. 13059-13066.

47. Winter, N., et al., Chromatin immunoprecipitation to analyze DNA binding sites of HMGA2. PloS One, 2011. 6(4): p. e18837.

48. Schwanbeck, R., G. Manfioletti, and J.R. Wisniewski, Architecture of high mobility group protein IC· DNA complex and its perturbation upon phosphorylation by Cdc2 kinase. Journal of Biological Chemistry, 2000. 275(3): p. 1793-1801.

49. Piekielko, A., et al., GENES: STRUCTURE AND REGULATION-Distinct organization of DNA complexes of various HMGI/Y family proteins and their modulation upon mitotic phosphorylation. Journal of Biological Chemistry, 2001. 276(3): p. 1984-1992.

50. Cui, T., et al., Energetics of binding the mammalian high mobility group protein HMGA2 to poly (dA-dT) 2 and poly (dA)-poly (dT). Journal of Molecular Biology, 2005. 352(3): p. 629-645.

51. Nissen, M.S. and R. Reeves, Changes in Superhelicity Are Introduced into Closed Circular DNA by Binding of High Mobility Group Protein I/Y (∗). Journal of Biological Chemistry, 1995. 270(9): p. 4355-4360.

52. Schwanbeck, R. and J.R. Wisniewski, Cdc2 and Mitogen-activated Protein Kinases Modulate DNA Binding Properties of the Putative Transcriptional RegulatorChironomus High Mobility Group Protein I. Journal of Biological Chemistry, 1997. 272(43): p. 27476-27483.

53. Noro, B., et al., Molecular dissection of the architectural transcription factor HMGA2. Biochemistry, 2003. 42(15): p. 4569-4577.

54. Sgarra, R., et al., HMGA molecular network: from transcriptional regulation to chromatin remodeling. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 2010. 1799(1-2): p. 37-47.

55. Maurizio, E., et al., Conformational role for the C-terminal tail of the intrinsically disordered high mobility group A (HMGA) chromatin factors. Journal of Proteome Research, 2011. 10(7): p. 3283-3291.

56. Battista, S., et al., The expression of a truncated HMGI-C gene induces gigantism associated with lipomatosis. Cancer Research, 1999. 59(19): p. 4793-4797.

57. Xiang, X., K.F. Benson, and K. Chada, Mini-mouse: disruption of the pygmy locus in a transgenic insertional mutant. Science, 1990. 247(4945): p. 967-969.

58. Benson, K.F. and K. Chada, Mini-mouse: phenotypic characterization of a transgenic insertional mutant allelic to pygmy. Genetics Research, 1994. 64(1): p. 27-33.

59. Zhou, X., et al., Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. Nature, 1995. 376(6543): p. 771-774.

60. Ashar, H.R., et al., In vivo modulation of HMGA2 expression. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 2010. 1799(1-2): p. 55-61.

61. Anand, A. and K. Chada, In vivo modulation of Hmgic reduces obesity. Nature Genetics, 2000. 24(4): p. 377-380.

62. Federico, A., et al., Hmga1/Hmga2 double knock-out mice display a “superpygmy” phenotype. Biology Open, 2014. 3(5): p. 372-378.

63. Zaidi, M.R., Y. Okada, and K.K. Chada, Misexpression of full-length HMGA2 induces benign mesenchymal tumors in mice. Cancer Research, 2006. 66(15): p. 7453-7459.

64. Arlotta, P., et al., Transgenic mice expressing a truncated form of the high mobility group IC protein develop adiposity and an abnormally high prevalence of lipomas. Journal of Biological Chemistry, 2000. 275(19): p. 14394-14400.

65. Broberg, K., et al., Fusion of RDC1 with HMGA2 in lipomas as the result of chromosome aberrations involving 2q35-37 and 12q13-15. International Journal of Oncology, 2002. 21(2): p. 321-326.

66. Ashar, H.R., et al., HMGA2 is expressed in an allele-specific manner in human lipomas. Cancer Genetics and Cytogenetics, 2003. 143(2): p. 160-168.

67. Ligon, A.H., et al., Constitutional rearrangement of the architectural factor HMGA2: a novel human phenotype including overgrowth and lipomas. The American Journal of Human Genetics, 2005. 76(2): p. 340-348.

68. Weedon, M.N., et al., A common variant of HMGA2 is associated with adult and childhood height in the general population. Nature Genetics, 2007. 39(10): p. 1245-1250.

69. Sanna, S., et al., Common variants in the GDF5-UQCC region are associated with variation in human height. Nature Genetics, 2008. 40(2): p. 198-203.

70. Weedon, M.N., et al., Genome-wide association analysis identifies 20 loci that influence adult height. Nature Genetics, 2008. 40(5): p. 575-583.

71. Yang, T.L., et al., HMGA2 is confirmed to be associated with human adult height. Annals of Human Genetics, 2010. 74(1): p. 11-16.

72. Liu, J.Z., et al., Genome-wide association study of height and body mass index in Australian twin families. Twin Research and Human Genetics, 2010. 13(2): p. 179-193.

73. Takeshita, H., et al., Confirmation that SNPs in the high mobility group-A2 gene (HMGA2) are associated with adult height in the Japanese population; wide-ranging population survey of height-related SNPs in HMGA2. Electrophoresis, 2011. 32(14): p. 1844-1851.

74. Carty, C.L., et al., Genome-wide association study of body height in African Americans: The women's health initiative SNP health association resource (share). Human Molecular Genetics, 2012. 21(3): p. 711-720.

75. Horikoshi, M., et al., New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. Nature Genetics, 2013. 45(1): p. 76-82.

76. Allen, H.L., et al., Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature, 2010. 467(7317): p. 832-838.

77. Stein, J.L., et al., Identification of common variants associated with human hippocampal and intracranial volumes. Nature genetics, 2012. 44(5): p. 552-561.

78. Buysse, K., et al., The 12q14 microdeletion syndrome: additional patients and further evidence that HMGA2 is an important genetic determinant for human height. European Journal of Medical Genetics, 2009. 52(2-3): p. 101-107.

79. Lynch, S.A., et al., The 12q14 microdeletion syndrome: six new cases confirming the role of HMGA2 in growth. European Journal of Human Genetics, 2011. 19(5): p. 534-539.

80. Rosen, E.D. and O.A. MacDougald, Adipocyte differentiation from the inside out. Nature Reviews Molecular Cell Biology, 2006. 7(12): p. 885-896.

81. Sarjeant, K. and J.M. Stephens, Adipogenesis. Cold Spring Harbor Perspectives in Biology, 2012. 4(9): p. a008417.

82. Tang, Q.Q. and M.D. Lane, Adipogenesis: from stem cell to adipocyte. Annual Review of Biochemistry, 2012. 81: p. 715-736.

83. Ghaben, A.L. and P.E. Scherer, Adipogenesis and metabolic health. Nature Reviews Molecular Cell Biology, 2019. 20(4): p. 242-258.

84. Lee, J.-E., et al., Transcriptional and epigenomic regulation of adipogenesis. Molecular and Cellular Biology, 2019. 39(11): p. e00601-18.

85. Green, H. and M. Meuth, An established pre-adipose cell line and its differentiation in culture. Cell, 1974. 3(2): p. 127-133.

86. Yuan, Y., et al., STAT3 stimulates adipogenic stem cell proliferation and cooperates with HMGA2 during the early stage of differentiation to promote adipogenesis. Biochemical and Biophysical Research Communications, 2017. 482(4): p. 1360-1366.

87. Ayoubi, T.A., et al., Regulation of HMGIC expression: an architectural transcription factor involved in growth control and development. Oncogene, 1999. 18(36): p. 5076-5087.

88. Vernochet, C., et al., PPARγ-dependent and PPARγ-independent effects on the development of adipose cells from embryonic stem cells. FEBS Letters, 2002. 510(1-2): p. 94-98.

89. Sun, T., et al., MicroRNA let-7 regulates 3T3-L1 adipogenesis. Molecular Endocrinology, 2009. 23(6): p. 925-931.

90. Alonso, N., et al., A rapid and sensitive high-throughput screening method to identify compounds targeting protein–nucleic acids interactions. Nucleic Acids Research, 2015. 43(8): p. e52-e52.

91. Xi, Y., et al., HMGA2 promotes adipogenesis by activating C/EBPβ-mediated expression of PPARγ. Biochemical and Biophysical Research Communications, 2016. 472(4): p. 617-623.

92. Price, N.L., et al., SREBP-1c/MicroRNA 33b genomic loci control adipocyte differentiation. Molecular and Cellular Biology, 2016. 36(7): p. 1180-1193.

93. Dröge, P. and C.A. Davey, Do cells let-7 determine stemness? Cell Stem Cell, 2008. 2(1): p. 8-9.

94. Hammond, S.M. and N.E. Sharpless, HMGA2, microRNAs, and stem cell aging. Cell, 2008. 135(6): p. 1013-1016.

95. Pietras, E.M. and E. Passegué, Lin king HSCs to their youth. Nature Cell Biology, 2013. 15(8): p. 885-887.

96. Parisi, S., et al., HMGA proteins in stemness and differentiation of embryonic and adult stem cells. International Journal of Molecular Sciences, 2020. 21(1): p. 362.

97. Rogalla, P., et al., HMGI-C expression patterns in human tissues. Implications for the genesis of frequent mesenchymal tumors. The American Journal of Pathology, 1996. 149(3): p. 775.

98. Gattas, G.J., et al., HMGIC expression in human adult and fetal tissues and in uterine leiomyomata. Genes, Chromosomes and Cancer, 1999. 25(4): p. 316-322.

99. Rommel, B., et al., HMGI-C, a member of the high mobility group family of proteins, is expressed in hematopoietic stem cells and in leukemic cells. Leukemia & Lymphoma, 1997. 26(5-6): p. 603-607.

100. Li, O., et al., High-level expression of DNA architectural factor HMGA2 and its association with nucleosomes in human embryonic stem cells. Genesis, 2006. 44(11): p. 523-529.

101. Nishino, J., et al., Hmga2 promotes neural stem cell self-renewal in young but not old mice by reducing p16Ink4a and p19Arf Expression. Cell, 2008. 135(2): p. 227-239.

102. Caron, L., et al., A new role for the oncogenic high-mobility group A2 transcription factor in myogenesis of embryonic stem cells. Oncogene, 2005. 24(41): p. 6281-6291.

103. Markowski, D.N., et al., BMP4 increases expression of HMGA2 in mesenchymal stem cells. Cytokine, 2011. 56(3): p. 811-816.

104. Copley, M.R., et al., The Lin28b–let-7–Hmga2 axis determines the higher self-renewal potential of fetal haematopoietic stem cells. Nature Cell Biology, 2013. 15(8): p. 916-925.

105. Wei, J., et al., let-7 enhances osteogenesis and bone formation while repressing adipogenesis of human stromal/mesenchymal stem cells by regulating HMGA2. Stem Cells and Development, 2014. 23(13): p. 1452-1463.

106. Singh, I., et al., Hmga2 is required for canonical WNT signaling during lung development. BMC Biology, 2014. 12(1): p. 1-16.

107. Parameswaran, S., et al., Hmga2 regulates self-renewal of retinal progenitors. Development, 2014. 141(21): p. 4087-4097.

108. Yu, K.-R., et al., Rapid and efficient direct conversion of human adult somatic cells into neural stem cells by HMGA2/let-7b. Cell Reports, 2015. 10(3): p. 441-452.

109. Kalomoiris, S., et al., Fibroblast Growth Factor 2 Regulates High Mobility Group A2 Expression in Human Bone Marrow-Derived Mesenchymal Stem Cells. Journal of Cellular Biochemistry, 2016. 117(9): p. 2128-2137.

110. Rowe, R.G., et al., Developmental regulation of myeloerythroid progenitor function by the Lin28b–let-7–Hmga2 axis. Journal of Experimental Medicine, 2016. 213(8): p. 1497-1512.

111. Yu, K.-R., et al., HMGA2 regulates the in vitro aging and proliferation of human umbilical cord blood-derived stromal cells through the mTOR/p70S6K signaling pathway. Stem Cell Research, 2013. 10(2): p. 156-165.

112. Cavazzana-Calvo, M., et al., Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia. Nature, 2010. 467(7313): p. 318-322.

113. Zhou, X. and K. Chada, HMGI family proteins: architectural transcription factors in mammalian development and cancer. The Keio Journal of Medicine, 1998. 47(2): p. 73-77.

114. Tallini, G. and P. Dal Cin, HMGI (Y) and HMGI-C dysregulation: a common occurrence in human tumors. Advances in Anatomic Pathology, 1999. 6(5): p. 237-246.

115. Reeves, R. and L.M. Beckerbauer, HMGA proteins as therapeutic drug targets. Progress in Cell Cycle Research, 2003. 5: p. 279-286.

116. Reeves, R., HMGA proteins: flexibility finds a nuclear niche? Biochemistry and Cell Biology, 2003. 81(3): p. 185-195.

117. Young, A.R. and M. Narita, Oncogenic HMGA2: short or small? Genes & Development, 2007. 21(9): p. 1005-1009.

118. Fusco, A. and M. Fedele, Roles of HMGA proteins in cancer. Nature Reviews Cancer, 2007. 7(12): p. 899-910.

119. Cleynen, I. and W.J. Van de Ven, The HMGA proteins: a myriad of functions. International Journal of Oncology, 2008. 32(2): p. 289-305.

120. Fedele, M. and A. Fusco, HMGA and cancer. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 2010. 1799(1-2): p. 48-54.

121. Sgarra, R., et al., High Mobility Group A (HMGA) proteins: Molecular instigators of breast cancer onset and progression. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2018. 1869(2): p. 216-229.

122. Zhang, S., Q. Mo, and X. Wang, Oncological role of HMGA2. International Journal of Oncology, 2019. 55(4): p. 775-788.

123. Ozturk, N., et al., HMGA proteins as modulators of chromatin structure during transcriptional activation. Frontiers in Cell and Developmental Biology, 2014. 2: p. 5.

124. Giancotti, V., et al., Epigenetic contribution of High-Mobility Group A proteins to stem cell properties. International Journal of Cell Biology, 2018. 2018.

125. Yu, H., et al., Chaperoning HMGA2 protein protects stalled replication forks in stem and cancer cells. Cell Reports, 2014. 6(4): p. 684-697.

126. Ahmed, S.M., et al., The chromatin structuring protein HMGA2 influences human subtelomere stability and cancer chemosensitivity. PloS One, 2019. 14(5): p. e0215696.

127. Ahmed, S.M. and P. Dröge, Oncofetal HMGA2 attenuates genotoxic damage induced by topoisomerase II target compounds through the regulation of local DNA topology. Molecular Oncology, 2019. 13(10): p. 2062-2078.

128. Goodwin, G.H., et al., Analysis of the high mobility group proteins associated with salt-soluble nucleosomes. Nucleic Acids Research, 1979. 7(7): p. 1815-1835.

129. Cao, X., et al., SUMOylation of HMGA2: selective destabilization of promyelocytic leukemia protein via proteasome. Molecular Cancer Therapeutics, 2008. 7(4): p. 923-934.

130. Wu, Y., et al., The regulation of acetylation and stability of HMGA2 via the HBXIP-activated Akt–PCAF pathway in promotion of esophageal squamous cell carcinoma growth. Nucleic Acids Research, 2020. 48(9): p. 4858-4876.

131. Bustin, M. and R. Reeves, High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. Progress in Nucleic Acid Research and Molecular Biology, 1996. 54: p. 35-100b.

132. Rogalla, P., et al., HMGIC expression patterns in non-small lung cancer and surrounding tissue. Anticancer Research, 1998. 18(5A): p. 3327-3330.

133. Kumar, M.S., et al., HMGA2 functions as a competing endogenous RNA to promote lung cancer progression. Nature, 2014. 505(7482): p. 212-217.

134. Sun, M., et al., HMGA2/TET1/HOXA9 signaling pathway regulates breast cancer growth and metastasis. Proceedings of the National Academy of Sciences, 2013. 110(24): p. 9920-9925.

135. El Ayachi, I., et al., The WNT10B network is associated with survival and metastases in chemoresistant triple-negative breast cancer. Cancer Research, 2019. 79(5): p. 982-993.

136. Mueller, M.H., et al., HMGA2 expression in the PC-3 prostate cancer cell line is autonomous of growth factor stimulation. Anticancer Research, 2013. 33(8): p. 3069-3078.

137. Marquis, M., et al., High expression of HMGA2 independently predicts poor clinical outcomes in acute myeloid leukemia. Blood Cancer Journal, 2018. 8(8): p. 1-12.

138. Murakami, M., et al., Regulatory expression of genes related to metastasis by TGF-β and activin A in B16 murine melanoma cells. Molecular Biology Reports, 2010. 37(3): p. 1279-1286.

139. Raskin, L., et al., Transcriptome profiling identifies HMGA2 as a biomarker of melanoma progression and prognosis. Journal of Investigative Dermatology, 2013. 133(11): p. 2585-2592.

140. Hou, X., et al., Let-7a inhibits migration of melanoma cells via down-regulation of HMGA2 expression. American Journal of Translational Research, 2016. 8(9): p. 3656.

141. Moon, H., et al., Melanocyte stem cell activation and translocation initiate cutaneous melanoma in response to UV exposure. Cell Stem Cell, 2017. 21(5): p. 665-678. e6.

142. Miao, Y., et al., Inhibition of high-mobility-group A2 protein binding to DNA by netropsin: a biosensor-surface plasmon resonance assay. Analytical Biochemistry, 2008. 374(1): p. 7-15.

143. Neidle, S., DNA minor-groove recognition by small molecules. Natural Product Reports, 2001. 18(3): p. 291-309.

144. Zimmer, C. and U. Wähnert, Nonintercalating DNA-binding ligands: specificity of the interaction and their use as tools in biophysical, biochemical and biological investigations of the genetic material. Progress in Biophysics and Molecular Biology, 1986. 47(1): p. 31-112.

145. Wemmer, D.E., Designed sequence-specific minor groove ligands. Annual Review of Biophysics and Biomolecular Structure, 2000. 29(1): p. 439-461.

146. Baraldi, P.G., et al., DNA minor groove binders as potential antitumor and antimicrobial agents. Medicinal Research Reviews, 2004. 24(4): p. 475-528.

147. Dervan, P.B., Design of sequence-specific DNA-binding molecules. Science, 1986. 232(4749): p. 464-471.

148. Wemmer, D.E. and P.B. Dervan, Targeting the minor groove of DNA. Current Opinion in Structural Biology, 1997. 7(3): p. 355-361.

149. Dervan, P.B. and R.W. Bürli, Sequence-specific DNA recognition by polyamides. Current Opinion in Chemical Biology, 1999. 3(6): p. 688-693.

150. Yan, C. and P.J. Higgins, Drugging the undruggable: transcription therapy for cancer. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2013. 1835(1): p. 76-85.

151. Lambert, M., et al., Targeting transcription factors for cancer treatment. Molecules, 2018. 23(6): p. 1479.

152. Helmer, D. and K. Schmitz, Peptides and peptide analogs to inhibit protein-protein interactions. Protein Targeting Compounds, 2016: p. 147-183.

153. Rai, J., Peptide and protein mimetics by retro and retroinverso analogs. Chemical Biology & Drug Design, 2019. 93(5): p. 724-736.

154. French, S.W., M.C. Schmidt, and J.C. Glorioso, Involvement of a high-mobility-group protein in the transcriptional activity of herpes simplex virus latency-active promoter 2. Molecular and Cellular Biology, 1996. 16(10): p. 5393-5399.

155. Berlingieri, M.T., et al., Inhibition of HMGI-C protein synthesis suppresses retrovirally induced neoplastic transformation of rat thyroid cells. Molecular and Cellular Biology, 1995. 15(3): p. 1545-1553.

156. Thuault, S., et al., Transforming growth factor-β employs HMGA2 to elicit epithelial–mesenchymal transition. The Journal of Cell Biology, 2006. 174(2): p. 175-183.

157. Thuault, S., et al., HMGA2 and Smads co-regulate SNAIL1 expression during induction of epithelial-to-mesenchymal transition. Journal of Biological Chemistry, 2008. 283(48): p. 33437-33446.

158. Su, L., et al., Identification of HMGA2 inhibitors by AlphaScreen-based ultra-high-throughput screening assays. Scientific Reports, 2020. 10(1): p. 1-14.

159. Chung, J., et al., High mobility group A2 (HMGA2) deficiency in pigs leads to dwarfism, abnormal fetal resource allocation, and cryptorchidism. Proceedings of the National Academy of Sciences, 2018. 115(21): p. 5420-5425.

160. Fedele, M., et al., Role of the high mobility group A proteins in human lipomas. Carcinogenesis, 2001. 22(10): p. 1583-1591.

161. Ligon, A.H. and C.C. Morton, Genetics of uterine leiomyomata. Genes, Chromosomes and Cancer, 2000. 28(3): p. 235-245.

162. Abe, N., et al., Determination of high mobility group I (Y) expression level in colorectal neoplasias: a potential diagnostic marker. Cancer Research, 1999. 59(6): p. 1169-1174.

163. Meyer, B., et al., HMGA2 overexpression in non-small cell lung cancer. Molecular Carcinogenesis: Published in Cooperation with the University of Texas MD Anderson Cancer Center, 2007. 46(7): p. 503-511.

164. Reeves, R., Nuclear functions of the HMG proteins. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 2010. 1799(1-2): p. 3-14.

165. Lehn, D., et al., A conformational study of the sequence specific binding of HMG-I (Y) with the bovine interleukin-2 cDNA. Biochemistry International, 1988. 16(5): p. 963-971.

166. Colamaio, M., et al., HMGA1 silencing reduces stemness and temozolomide resistance in glioblastoma stem cells. Expert Opinion on Therapeutic Targets, 2016. 20(10): p. 1169-1179.

167. Ito, T., et al., In vitro irradiation is able to cause RET oncogene rearrangement. Cancer Research, 1993. 53(13): p. 2940-2943.

168. Puca, F., et al., HMGA1 silencing restores normal stem cell characteristics in colon cancer stem cells by increasing p53 levels. Oncotarget, 2014. 5(10): p. 3234.

169. Qi, Y., et al., CHARMM-GUI martini maker for coarse-grained simulations with the martini force field. Journal of Chemical Theory and Computation, 2015. 11(9): p. 4486-4494.

170. Huang, J. and A.D. MacKerell Jr, CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. Journal of Computational Chemistry, 2013. 34(25): p. 2135-2145.

171. Essmann, U., et al., A smooth particle mesh Ewald method. The Journal of Chemical Physics, 1995. 103(19): p. 8577-8593.

172. Brooks, M.M., A. Hallstrom, and M. Peckova, A simulation study used to design the sequential monitoring plan for a clinical trial. Statistics in Medicine, 1995. 14(20): p. 2227-2237.

173. Humphrey, W., A. Dalke, and K. Schulten, VMD: visual molecular dynamics. Journal of Molecular Graphics, 1996. 14(1): p. 33-38.

174. Trott, O. and A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of Computational Chemistry, 2010. 31(2): p. 455-461.

175. Barrett, M.P. and I.H. Gilbert, Targeting of toxic compounds to the trypanosome's interior. Advances in Parasitology, 2006. 63: p. 125-183.

176. Cui, T., et al., Large scale preparation of the mammalian high mobility group protein A2 for biophysical studies. Protein and Peptide Letters, 2007. 14(1): p. 87-91.

177. Puca, F., et al., HMGA1 negatively regulates NUMB expression at transcriptional and post transcriptional level in glioblastoma stem cells. Cell Cycle, 2019. 18(13): p. 1446-1457.

178. Zhang, Y.-L., et al., Suramin is an active site-directed, reversible, and tight-binding inhibitor of protein-tyrosine phosphatases. Journal of Biological Chemistry, 1998. 273(20): p. 12281-12287.

179. McGeary, R.P., et al., Suramin: clinical uses and structure-activity relationships. Mini Reviews in Medicinal Chemistry, 2008. 8(13): p. 1384-1394.

180. Büscher, P., et al., Human african trypanosomiasis. The Lancet, 2017. 390(10110): p. 2397-2409.

181. Wainwright, M., Dyes, trypanosomiasis and DNA: a historical and critical review. Biotechnic & Histochemistry, 2010. 85(6): p. 341-354.

182. Darnell, J.E., Transcription factors as targets for cancer therapy. Nature Reviews Cancer, 2002. 2(10): p. 740-749.

183. Cagel, M., et al., Doxorubicin: nanotechnological overviews from bench to bedside. Drug Discovery Today, 2017. 22(2): p. 270-281.

184. Riddell, I.A. and S.J. Lippard, 1. Cisplatin and oxaliplatin: our current understanding of their actions. Metallo-Drugs: Development and Action of Anticancer Agents, 2018: p. 1-42.

185. Meredith, A.-M. and C.R. Dass, Increasing role of the cancer chemotherapeutic doxorubicin in cellular metabolism. Journal of Pharmacy and Pharmacology, 2016. 68(6): p. 729-741.

186. Reeves, R., Molecular biology of HMGA proteins: hubs of nuclear function. Gene, 2001. 277(1-2): p. 63-81.

187. Reeves, R. and L. Beckerbauer, HMGI/Y proteins: flexible regulators of transcription and chromatin structure. Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression, 2001. 1519(1-2): p. 13-29.

188. Dressel, J., The discovery of germanin by oskar dressel and richard kothe. Journal of Chemical Education, 1961. 38(12): p. 620.

189. Barrett, M.P., et al., Human African trypanosomiasis: pharmacological re-engagement with a neglected disease. British Journal of Pharmacology, 2007. 152(8): p. 1155-1171.

190. Fairlamb, A. and I. Bowman, Uptake of the trypanocidal drug suramin by bloodstream forms of Trypanosoma brucei and its effect on respiration and growth rate in vivo. Molecular and Biochemical Parasitology, 1980. 1(6): p. 315-333.

191. Scott, A.G., et al., Characterisation of clones lines of Trypanosoma brucei expressing stable resistance to MelCy and suramin. Acta Tropica, 1996. 60(4): p. 251-262.

192. Opperdoes, F.R., Compartmentation of carbohydrate metabolism in trypanosomes. Annual Reviews in Microbiology, 1987. 41(1): p. 127-151.

193. Jensen, R.E. and P.T. Englund, Network news: the replication of kinetoplast DNA. Annual Review of Microbiology, 2012. 66: p. 473-491.

194. Wilson, W.D., et al., Antiparasitic compounds that target DNA. Biochimie, 2008. 90(7): p. 999-1014.

195. Wang, J., et al., TbKAP6, a mitochondrial HMG box-containing protein in Trypanosoma brucei, is the first trypanosomatid kinetoplast-associated protein

essential for kinetoplast DNA replication and maintenance. Eukaryotic Cell, 2014. 13(7): p. 919-932.

196. Kaur, M., et al., Suramin's development: what did we learn? Investigational New Drugs, 2002. 20(2): p. 209-219.

197. La Rocca, R.V., et al., Suramin, a novel antitumor compound. The Journal of Steroid Biochemistry and Molecular Biology, 1990. 37(6): p. 893-898.

198. Wu, K., et al., Suramin inhibits cullin-RING E3 ubiquitin ligases. Proceedings of the National Academy of Sciences, 2016. 113(14): p. E2011-E2018.

199. Wiedemar, N., D.A. Hauser, and P. Mäser, 100 years of suramin. Antimicrobial Agents and Chemotherapy, 2020. 64(3): p. e01168-19.

200. Lipinski, C.A., Lead-and drug-like compounds: the rule-of-five revolution. Drug Discovery Today: Technologies, 2004. 1(4): p. 337-341.

201. Staats, B., et al., A fibroadenoma with at (4; 12)(q27; q15) affecting the HMGI-C gene, a member of the high mobility group protein gene family. Breast Cancer Research and Treatment, 1996. 38(3): p. 299-303.

202. Di Cello, F., et al., HMGA2 participates in transformation in human lung cancer. Molecular Cancer Research, 2008. 6(5): p. 743-750.

203. Shi, Z., et al., Silencing of HMGA2 promotes apoptosis and inhibits migration and invasion of prostate cancer cells. Journal of Biosciences, 2016. 41(2): p. 229-236.

204. Wu, J., et al., Elevated HMGA2 expression is associated with cancer aggressiveness and predicts poor outcome in breast cancer. Cancer Letters, 2016. 376(2): p. 284-292.

205. Rizzi, C., et al., The expression of the high-mobility group A2 protein in colorectal cancer and surrounding fibroblasts is linked to tumor invasiveness. Human Pathology, 2013. 44(1): p. 122-132.

206. Piekiełko, A., et al., Distinct organization of DNA complexes of various HMGI/Y family proteins and their modulation upon mitotic phosphorylation. Journal of Biological Chemistry, 2001. 276(3): p. 1984-1992.

207. Schwanbeck, R., et al., Consecutive steps of phosphorylation affect conformation and DNA binding of the chironomus high mobility group A protein. Journal of Biological Chemistry, 2001. 276(28): p. 26012-26021.

208. Fluorescence Anisotropy, in Principles of Fluorescence Spectroscopy, J.R. Lakowicz, Editor. 2006, Springer US: Boston, MA. p. 353-382.

209. Narula, G., et al., The strictly conserved Arg-321 residue in the active site of Escherichia coli topoisomerase I plays a critical role in DNA rejoining. Journal of Biological Chemistry, 2011. 286(21): p. 18673-18680.

210. Record, M.T., C.F. Anderson, and T.M. Lohman, Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. Quarterly Reviews of Biophysics, 1978. 11(2): p. 103-178.

211. Schwabe, J.W., The role of water in protein—DNA interactions. Current Opinion in Structural Biology, 1997. 7(1): p. 126-134.

212. Qu, X. and J.B. Chaires, Hydration changes for DNA intercalation reactions. Journal of the American Chemical Society, 2001. 123(1): p. 1-7.

213. Parsegian, V.A., R. Rand, and D. Rau, Energetics of biological macromolecules. Methods Enzymol, 1995. 259: p. 43-94.

214. Schneider, B., D. Cohen, and H.M. Berman, Hydration of DNA bases: analysis of crystallographic data. Biopolymers: Original Research on Biomolecules, 1992. 32(7): p. 725-750.

215. Garner, M. and D. Rau, Water release associated with specific binding of gal repressor. The EMBO Journal, 1995. 14(6): p. 1257-1263.

216. Dimarco, A., et al., Daunomycin: a new antibiotic with antitumor activity. Cancer Chemotherapy Reports, 1964. 38: p. 31-38.

217. Di Marco, A., Adriamycin (NSC-123,127): a new antibiotic with antitumor activity. Cancer Chemotherapy Reports, 1969. 53: p. 33-37.

218. Young, R.C., R.F. Ozols, and C.E. Myers, The anthracycline antineoplastic drugs. New England Journal of Medicine, 1981. 305(3): p. 139-153.

219. Zhu, H., et al., Doxorubicin redox biology: redox cycling, topoisomerase inhibition, and oxidative stress. Reactive Oxygen Species (Apex, NC), 2016. 1(3): p. 189.

220. M Cutts, S., et al., Potential therapeutic advantages of doxorubicin when activated by formaldehyde to function as a DNA adduct-forming agent. Current Topics in Medicinal Chemistry, 2015. 15(14): p. 1409-1422.

221. Chaires, J.B., et al., Parsing the free energy of anthracycline antibiotic binding to DNA. Biochemistry, 1996. 35(7): p. 2047-2053.

222. Leng, F., et al., Base specific and regioselective chemical cross-linking of daunorubicin to DNA. Journal of the American Chemical Society, 1996. 118(20): p. 4731-4738.

223. Thorn, C.F., et al., Doxorubicin pathways: pharmacodynamics and adverse effects. Pharmacogenetics and Genomics, 2011. 21(7): p. 440.

224. Gewirtz, D., A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin. Biochemical Pharmacology, 1999. 57(7): p. 727-741.

225. Coldwell, K., et al., Detection of adriamycin-DNA adducts by accelerator mass spectrometry, in Drug-DNA Interaction Protocols. 2010, Springer. p. 103-118.

226. Zhao, L. and B. Zhang, Doxorubicin induces cardiotoxicity through upregulation of death receptors mediated apoptosis in cardiomyocytes. Scientific Reports, 2017. 7(1): p. 1-11.

227. Zhou, S., et al., Cumulative and irreversible cardiac mitochondrial dysfunction induced by doxorubicin. Cancer Research, 2001. 61(2): p. 771-777.

228. Koleini, N., et al., Oxidized phospholipids in Doxorubicin-induced cardiotoxicity. Chemico-Biological Interactions, 2019. 303: p. 35-39.

229. Tada, Y., et al., Increased expression of multidrug resistance-associated proteins in bladder cancer during clinical course and drug resistance to doxorubicin. International Journal of Cancer, 2002. 98(4): p. 630-635.

230. Shen, F., et al., Quantitation of doxorubicin uptake, efflux, and modulation of multidrug resistance (MDR) in MDR human cancer cells. Journal of Pharmacology and Experimental Therapeutics, 2008. 324(1): p. 95-102.

231. Moon, C., et al., In vitro assessment of a novel polyrotaxane-based drug delivery system integrated with a cell-penetrating peptide. Journal of Controlled Release, 2007. 124(1-2): p. 43-50.

232. Veronese, F.M., et al., PEG− doxorubicin conjugates: influence of polymer structure on drug release, in vitro cytotoxicity, biodistribution, and antitumor activity. Bioconjugate Chemistry, 2005. 16(4): p. 775-784.

233. Cutts, S.M., et al., The power and potential of doxorubicin-DNA adducts. IUBMB Life, 2005. 57(2): p. 73-81.

234. Job, P., Formation and stability of inorganic complexes in solution. Annali Di Chimica, 1928. 9(113).

235. Tooney, N.M., Biophysical chemistry - part I: The conformation of biological, macromolecules, Cantor and Schimmel, W. H. Freeman and Company, San Francisco, 1980, 341 pp. Journal of Polymer Science: Polymer Letters Edition, 1980. 18(9): p. 643-644.

236. van Meerloo, J., G.J.L. Kaspers, and J. Cloos, Cell Sensitivity Assays: The MTT Assay, in Cancer Cell Culture: Methods and Protocols, I.A. Cree, Editor. 2011, Humana Press: Totowa, NJ. p. 237-245.

237. Gao, Y.-G., et al., Facile formation of a crosslinked adduct between DNA and the daunorubicin derivative MAR70 mediated by formaldehyde: molecular structure of the MAR70-d (CGTnACG) covalent adduct. Proceedings of the National Academy of Sciences, 1991. 88(11): p. 4845-4849.

238. Zhang, H., et al., Simultaneous incorporations of two anticancer drugs into DNA. The structures of formaldehyde-cross-linked adducts of daunorubicin-d (CG (araC) GCG) and doxorubicin-d (CA (araC) GTG) complexes at high resolution. Journal of Biological Chemistry, 1993. 268(14): p. 10095-10101.

239. Chaires, J.B., N. Dattagupta, and D.M. Crothers, Studies on interaction of anthracycline antibiotics and deoxyribonucleic acid: equilibrium binding studies on the interaction of daunomycin with deoxyribonucleic acid. Biochemistry, 1982. 21(17): p. 3933-3940.

240. Sangthong, S., et al., Overcoming doxorubicin-resistance in the NCI/ADR-RES model cancer cell line by novel anthracene-9, 10-dione derivatives. Bioorganic & Medicinal Chemistry Letters, 2013. 23(22): p. 6156-6160.

241. Lu, J., et al., Targeted delivery of doxorubicin by folic acid-decorated dual functional nanocarrier. Molecular Pharmaceutics, 2014. 11(11): p. 4164-4178.

242. Patankar, N.A., et al., Topotecan and doxorubicin combination to treat recurrent ovarian cancer: the influence of drug exposure time and delivery systems to achieve optimum therapeutic activity. Clinical Cancer Research, 2013. 19(4): p. 865-877.

243. Ittiudomrak, T., et al., α-Mangostin and apigenin induced cell cycle arrest and programmed cell death in SKOV-3 ovarian cancer cells. Toxicological Research, 2019. 35(2): p. 167-179.

244. Rusetskaya, N., et al., Distribution and accumulation of liposomal form of doxorubicin in breast cancer cells of MCF-7 line. Experimental Oncology, 2011.

245. Zhou, Y., et al., Loperamide, an FDA-approved antidiarrhea drug, effectively reverses the resistance of multidrug resistant MCF-7/MDR1 human breast cancer cells to doxorubicin-induced cytotoxicity. Cancer Investigation, 2012. 30(2): p. 119-125.

246. Fenick, D.J., D.J. Taatjes, and T.H. Koch, Doxoform and daunoform: anthracycline− formaldehyde conjugates toxic to resistant tumor cells. Journal of Medicinal Chemistry, 1997. 40(16): p. 2452-2461.

247. Cutts, S.M. and D.R. Phillips, Use of oligonucletides to define the site of interstand cross-links induced by Adriamycin. Nucleic Acids Research, 1995. 23(13): p. 2450-2456.

248. Belhoussine, R., et al., Confocal scanning microspectrofluorometry reveals specific anthracyline accumulation in cytoplasmic organelles of multidrug-resistant cancer cells. Journal of Histochemistry & Cytochemistry, 1998. 46(12): p. 1369-1376.

249. Larsen, A.K., A.E. Escargueil, and A. Skladanowski, Resistance mechanisms associated with altered intracellular distribution of anticancer agents. Pharmacology & Therapeutics, 2000. 85(3): p. 217-229.

250. Khariton, N., et al., Distribution and accumulation of liposomal form of doxorubicin in breast cancer cells of MCF-7 line. Experimental Oncology, 2011. 33: p. 2.

251. Cullinane, C., A. van Rosmalen, and D.R. Phillips, Does adriamycin induce interstrand cross-links in DNA? Biochemistry, 1994. 33(15): p. 4632-4638.

252. Taatjes, D.J., et al., Redox pathway leading to the alkylation of DNA by the anthracycline, antitumor drugs adriamycin and daunomycin. Journal of Medicinal Chemistry, 1997. 40(8): p. 1276-1286.

253. Taatjes, D.J., et al., Alkylation of DNA by the anthracycline, antitumor drugs adriamycin and daunomycin. Journal of Medicinal Chemistry, 1996. 39(21): p. 4135-4138.

254. Skladanowski, A. and J. Konopa, Relevance of interstrand DNA crosslinking induced by anthracyclines for their biological activity. Biochemical Pharmacology, 1994. 47(12): p. 2279-2287.

255. Skladanowski, A. and J. Konopa, Adriamycin and daunomycin induce programmed cell death (apoptosis) in tumour cells. Biochemical Pharmacology, 1993. 46(3): p. 375-382.

APPENDICES

Supplementary information (figures and tables)

Chapter 2.

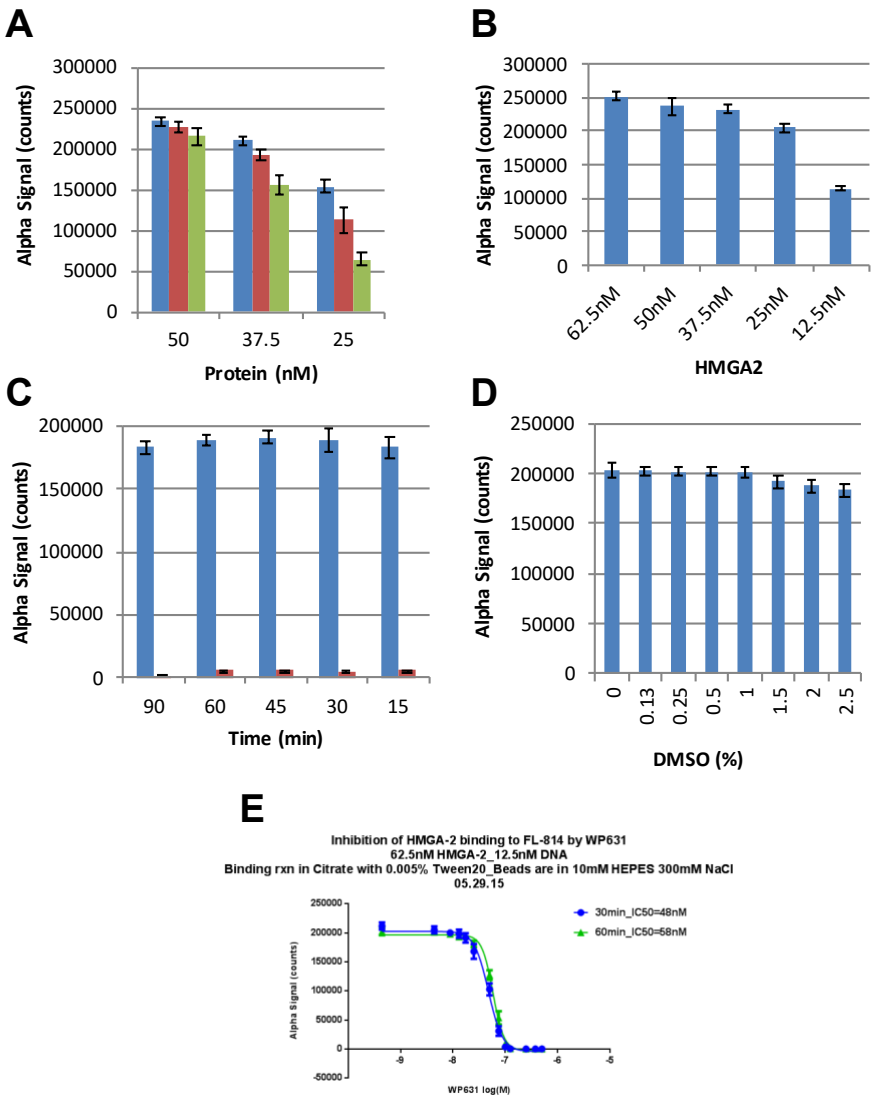


Figure S2.1. DNA-binding studies to determine the optimal conditions for HMGA2-FL814 interactions in 1×assay buffer (30 mM Citrate buffer, pH 7.0, 300 mM NaCl, 0.005% Tween 20) for the AlphaScreenprimary assay. **(A)** Varying HMGA2 and FL814 concentrations. Blue, red, and green columns represent 12.5, 6.25, and 3.125 nM of FL814. **(B)** Different concentrations of HMGA2 titrating into 12.5 nM FL814. **(C)** Time course of HMGA2-FL814 interactions. Blue and red columns represent the reactions in the presence and absence of HMGA2, respectively. **(D)** DMSO tolerant assays. **(E)** WP631 strongly inhibits HMGA2-FL814 interactions.

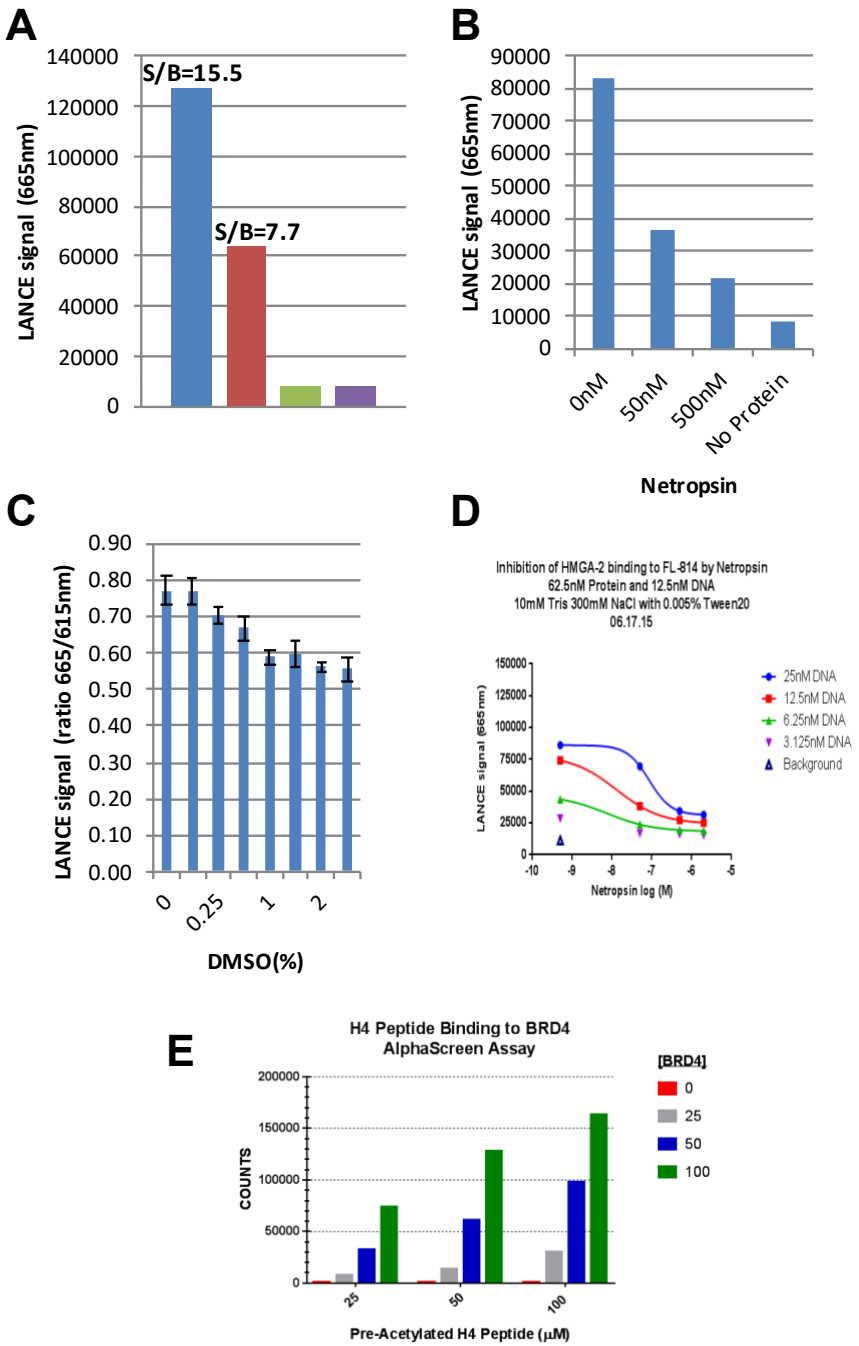


Figure S2.2. (A-D) DNA-binding studies to determine the optimal conditions for HMGA2-FL814 interactions for the LANCE time-resolved fluorescence energy transfer (TR-FRET) assay. **(A)** Varying HMGA2 and FL814 concentrations. Blue column: 62.5 nM HMGA2 and 12.5 nM FL814. Red column: 31.25 nM HMGA2 and 6.25 nM FL814. Green colun: 12.5 nM FL814 only. Purple column: 6.25 nM FL814 only. **(C)** DMSO tolerant assays. **(B)** and **(D)** Netropsin strongly inhibits HMGA2-FL814 interactions. **(E)** The AlphaScreen assay for H4 peptide binding to BRD4.

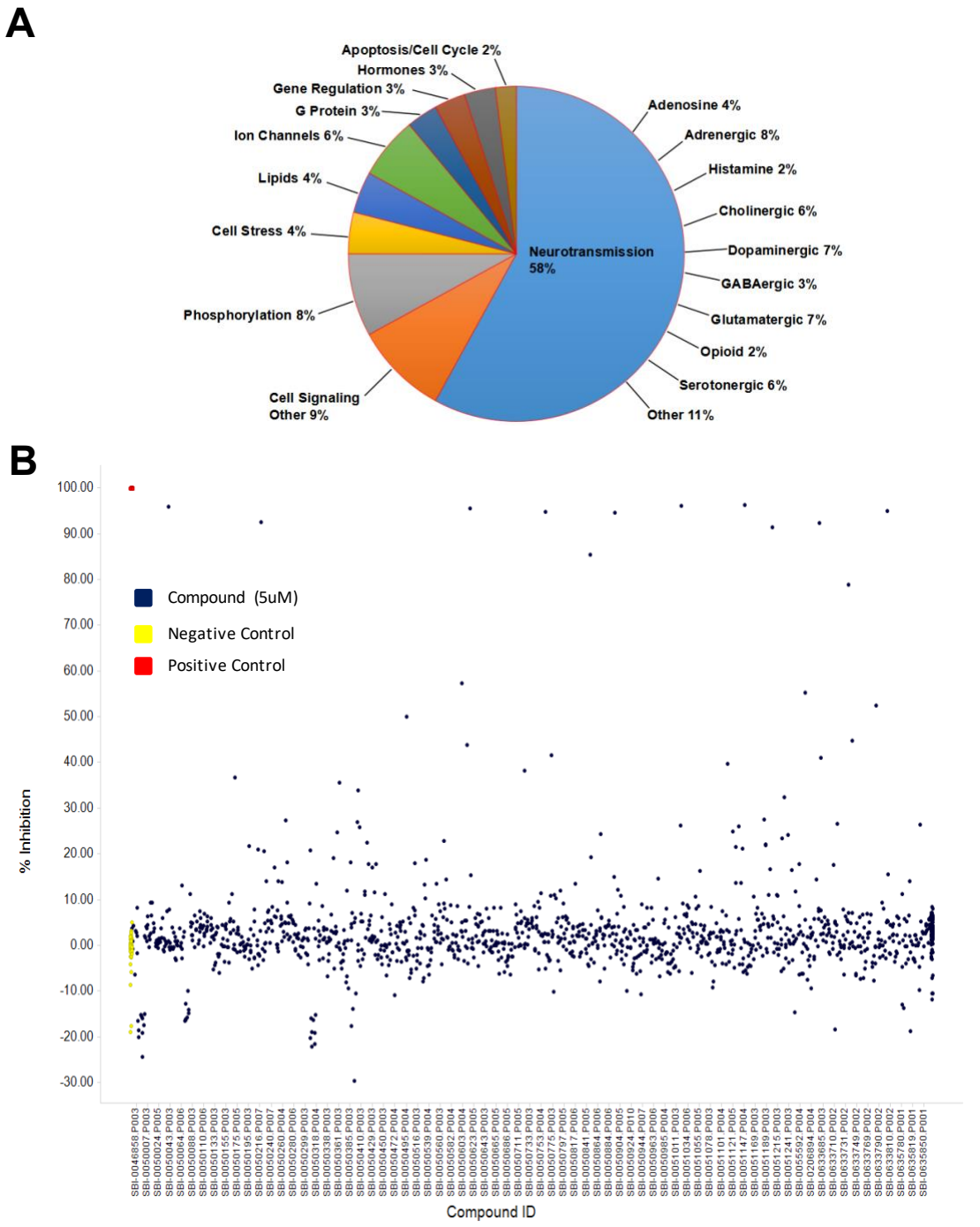
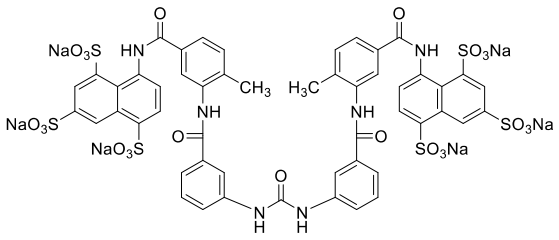
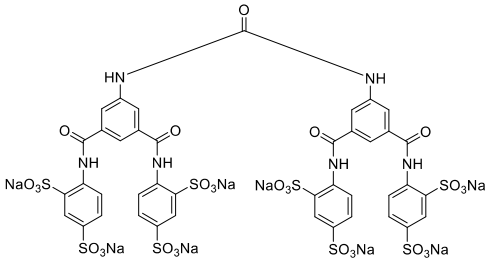


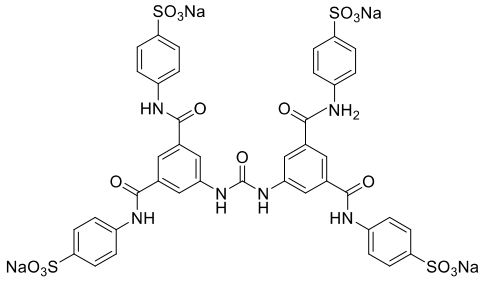
Figure S2.3. HMGA2-DNA pilot screens using the Sigma LOPAC1280 compound library. **(A)** The biologically annotated collection of LOPAC1280 compound library that contains 1280 pharmacologically active compounds. **(B)** HMGA2-DNA pilot screens using the Sigma LOPAC1280 compound library. Netropsin was used as positive controls.



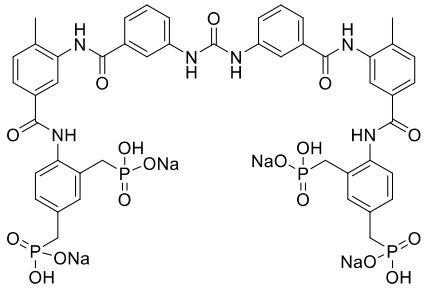
Suramin



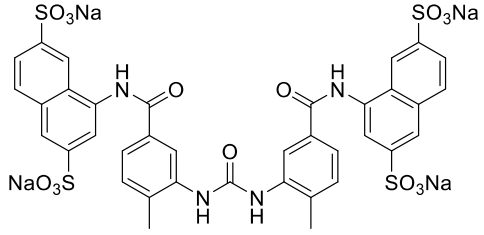
NF449



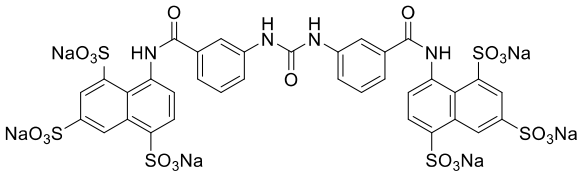
NF110



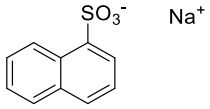
NF546



NF340



NF023



N0015

Figure S2.4. Chemical structures of suramin and analogues.

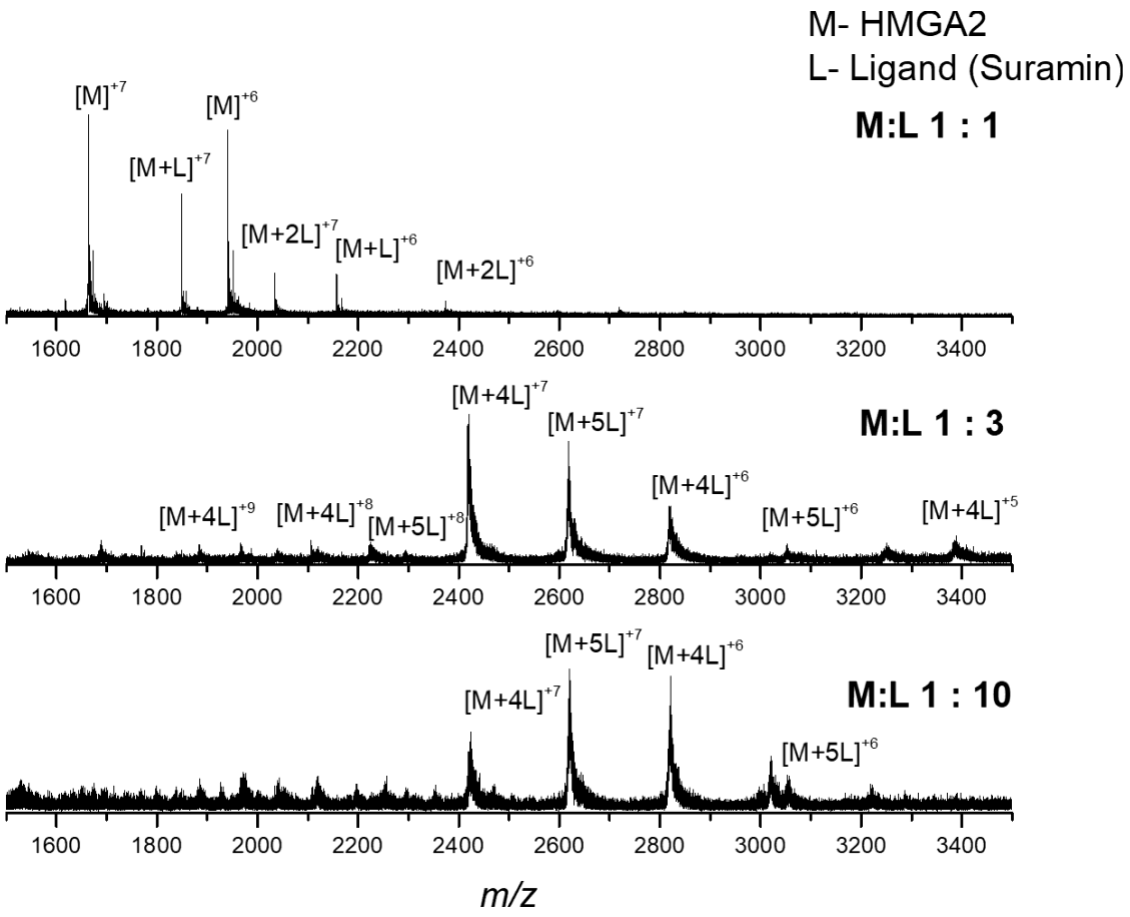


Figure S2.5. Typical native mass spectrometry spectra of a mixture of HMGA2: Ligand (Suramin) at 1:1, 1:3 and 1:10 ratio. Notice the increase in the numbers of ligands up to [M+4L] and [M+5L] per HMGA2 molecule observed in the MS from 1:1 to 1:10 ratios.

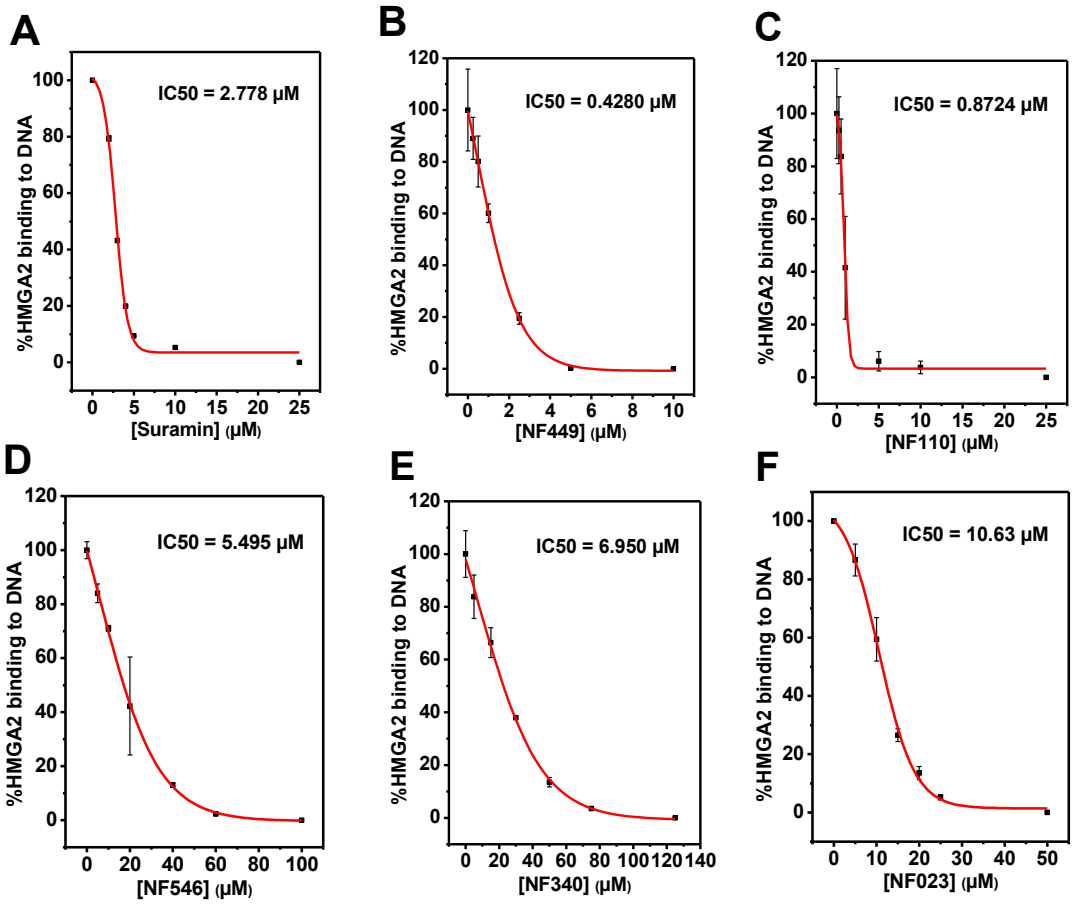


Figure S2.6. Inhibition of HMGA2 binding to FL814 by suramin **(A)** and analogues **(B-F)**. The apparent inhibitory IC_{50} was determined by using PDI-ELISA.

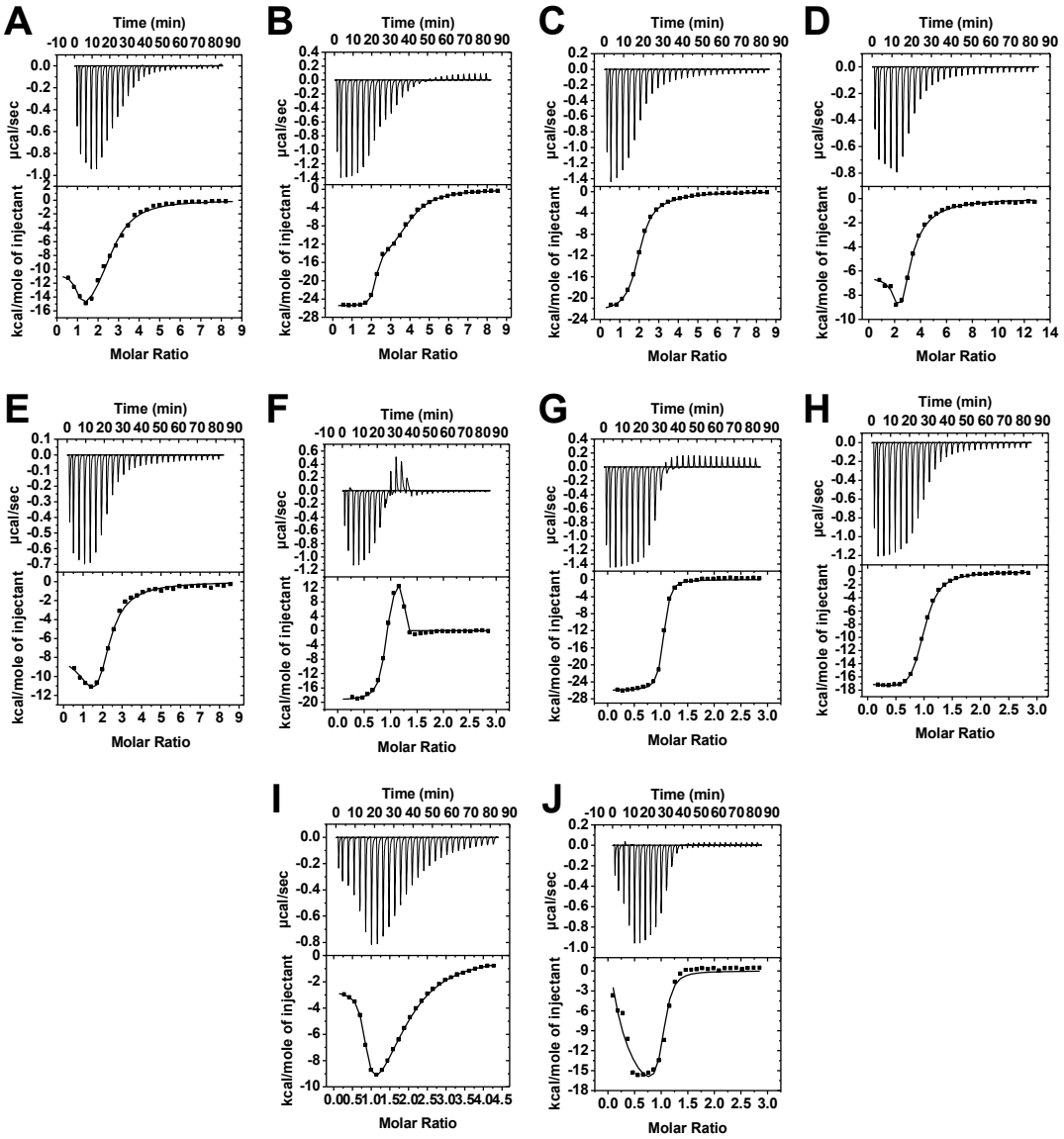


Figure S2.7. Sample raw data from isothermal titration calorimetry (ITC) experiments for the titration of suramin analogues to HMGA2 **(A-E)** and ATHP3 **(F-J).**
ITC experiments were performed according to conditions as described in Materials and Methods. The ITC data were fit using the software supplied by the manufacturer to yield thermodynamic parameters. **(A)** and **(F)** NF449; **(B)** and **(G)** NF110; **(C)** and **(H)** NF546; **(D)** and **(I)** NF 340; **(E)** and **(J)** NF023.

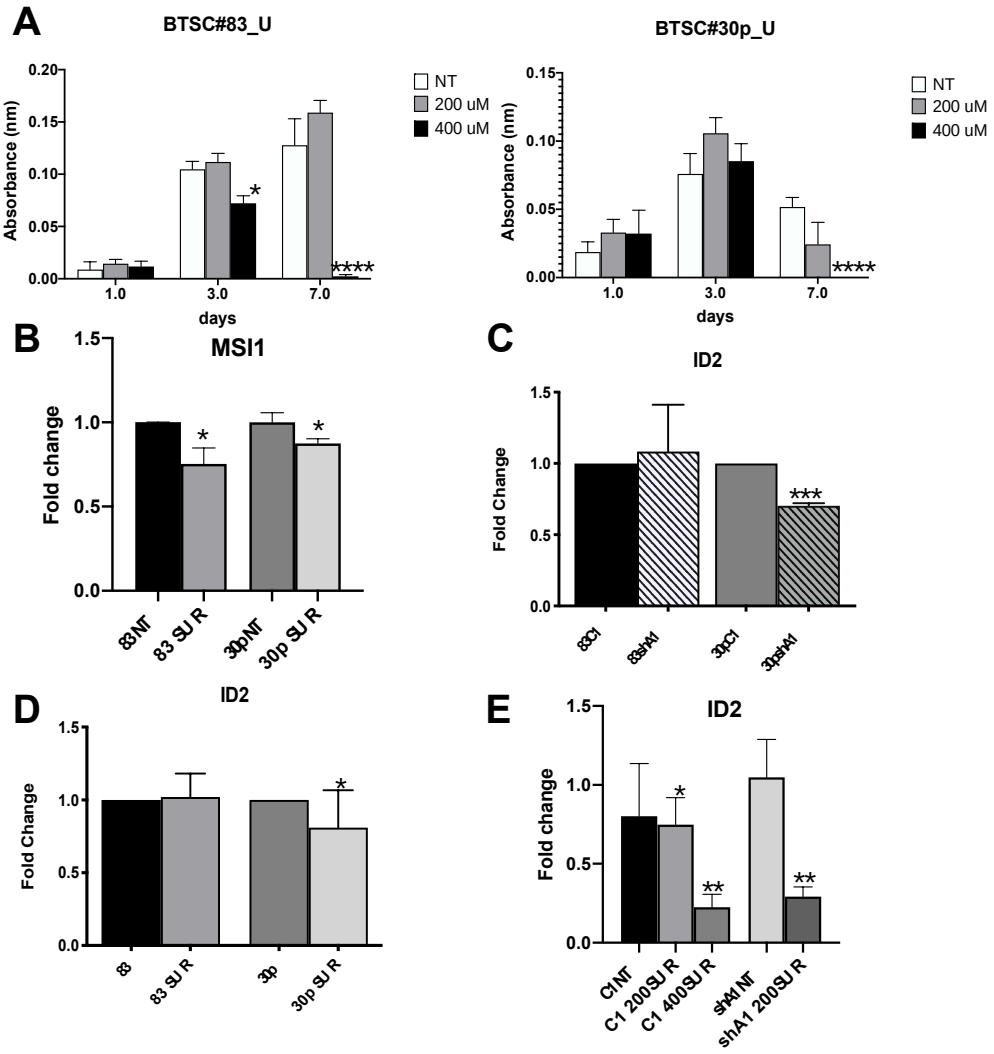


Figure S2.8. Suramin mimics HMGA1 silencing effects on gene expression. **(A)** Cell proliferation of BTSC#83 an #30p cells in U-bottom wells treated with suramin for 24, 72, and 168 hrs. The mean values ± SD in triplicate are shown**. (B)** qRT-PCR results of MSI1 in BTSC#83 and #30p treated with 100 µM suramin for 48 hours. **(C)** qRT-PCR of ID2 in BTSC#83 and #30p, silenced for HMGA1 (83sha1 and 30pshA1), normalized for actin expression and compared to scramble-transfected cells (83C1 and 30pC1). **(D)** qRT-PCR analyses of ID2 in BTSC#83 and #30p treated with 100 µM suramin for 48 hrs, compared to non-treated cells. Fold changes are normalized for actin expression. **(B-D)** Data represent the mean values ± SD of 3 independent experiments performed in duplicate. (* $p \leq 0.05$; *** $p \leq 0.001$; Student's t test). **(E)** qRT-PCR analyses of ID2 in scramble-transfected (C1) and HMGA1-knockdown (shA1) BTSC#83 treated with suramin 200 and 400 µM for 48 hrs, compared to non-treated cells (NT). Fold changes are normalized for actin expression. Data represent the mean value ± SD of 3 independent experiments performed in duplicate. (* $p \leq 0.05$; ** $p \leq 0.01$, Two-way ANOVA, followed by Dunnett's multiple comparisons test).

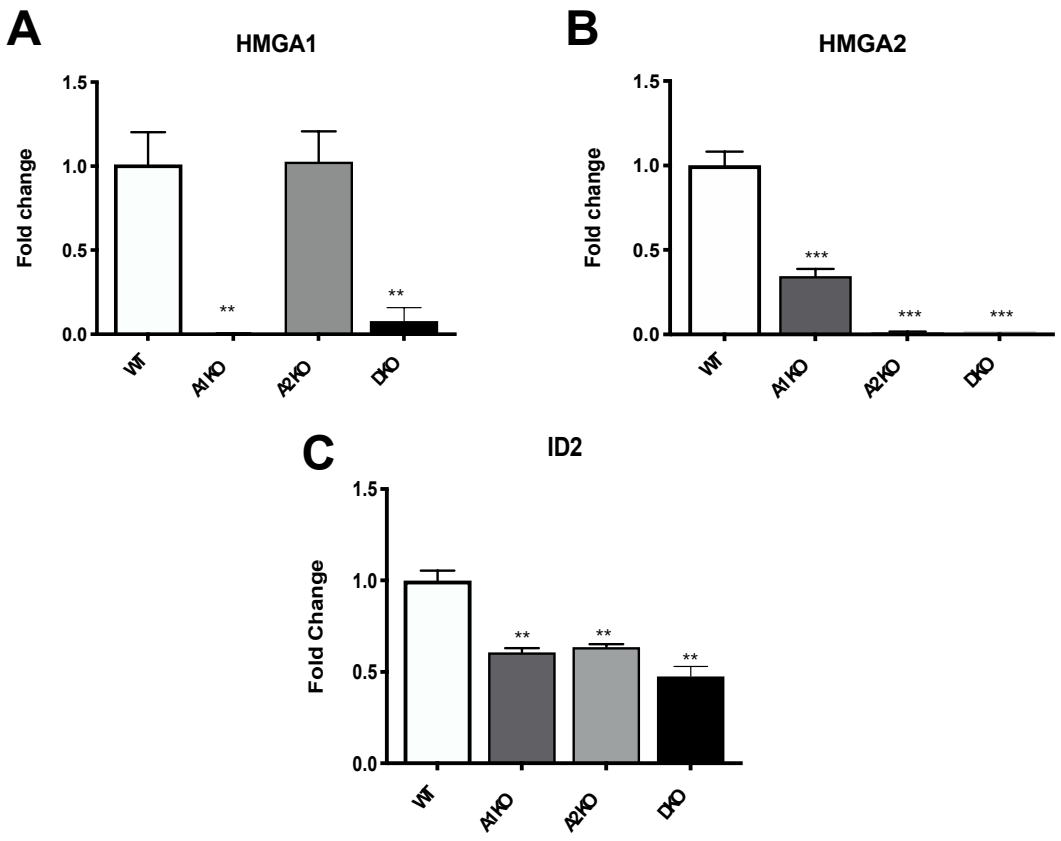


Figure S2.9. HMGA1 and HMGA2 affect ID2 expression in knock-out cells. qRT-PCR for **(A)** HMGA1, **(B)** HMGA2 and **(C)** ID2 expression in wild-type (WT), HMGA1 knock-out (A1 KO), HMGA2 knock-out (A2 KO), and HMGA1 and HMGA2 double knock-out (DKO) mouse embryonic fibroblasts (MEFs). Data represent the mean value ± SD of one representative experiment, performed in duplicate. (** $p \leq 0.01$, Two-way ANOVA, followed by Dunnett's multiple comparisons test). Primers for mouse HMGA1 and HMGA2 and for mouse ID2 have been described (Federico et al., 2014). Primers for mouse ID2 are: ID2 forward: 5'-ggaccacagcttgggcat; ID2 reverse, 5'-cgttcatgttgtagagcagactcat.

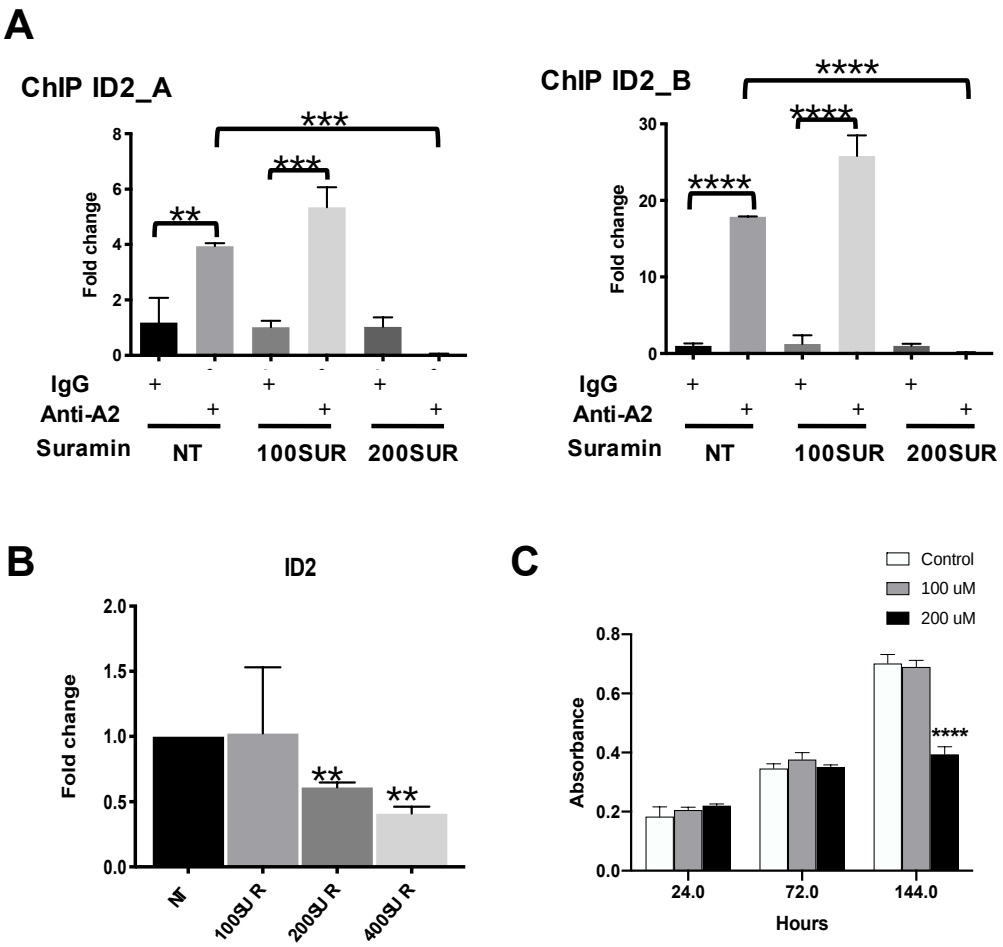


Figure S2.10. Effects of Suramin on HMGA2 in 8505c cells. **(A)** ChIP assay, revealed by qPCR, detecting the *in vivo* binding of HMGA2 to the sub-region A and B of the ID2 promoter in 8505c chromatin extracts. The relative fold enrichment of the ID2 promoter region A by anti-HMGA2 is indicated as vertical bars. The $2^{-\Delta\Delta Ct}$ formula was used to calculate the relative fold enrichment, normalized to its IgG reference and relative to the calibrator (input). (** $p \leq 0.01$, ***, $p < 0.001$****, $p < 0.0001$). **(B)** qRT-PCR analyses for ID2 expression in 8505c cells treated with Suramin 100, 200 and 400 µM for 48 hrs. Data represent the mean value +/- SD of 3 independent experiments performed in duplicate. (** $p \leq 0.01$). **(C)** Cell proliferation assay of 8505c cells treated with 0, 100 and 200 µM Suramin for 24, 72 and 144 hours. Columns represent the mean values +/- SD of a representative experiment performed in triplicate. Asterisks indicate the statistical significance of the absorbance of 200 µM suramin treated cells compared to not treated (control) cells. (****, $p < 0.0001$).

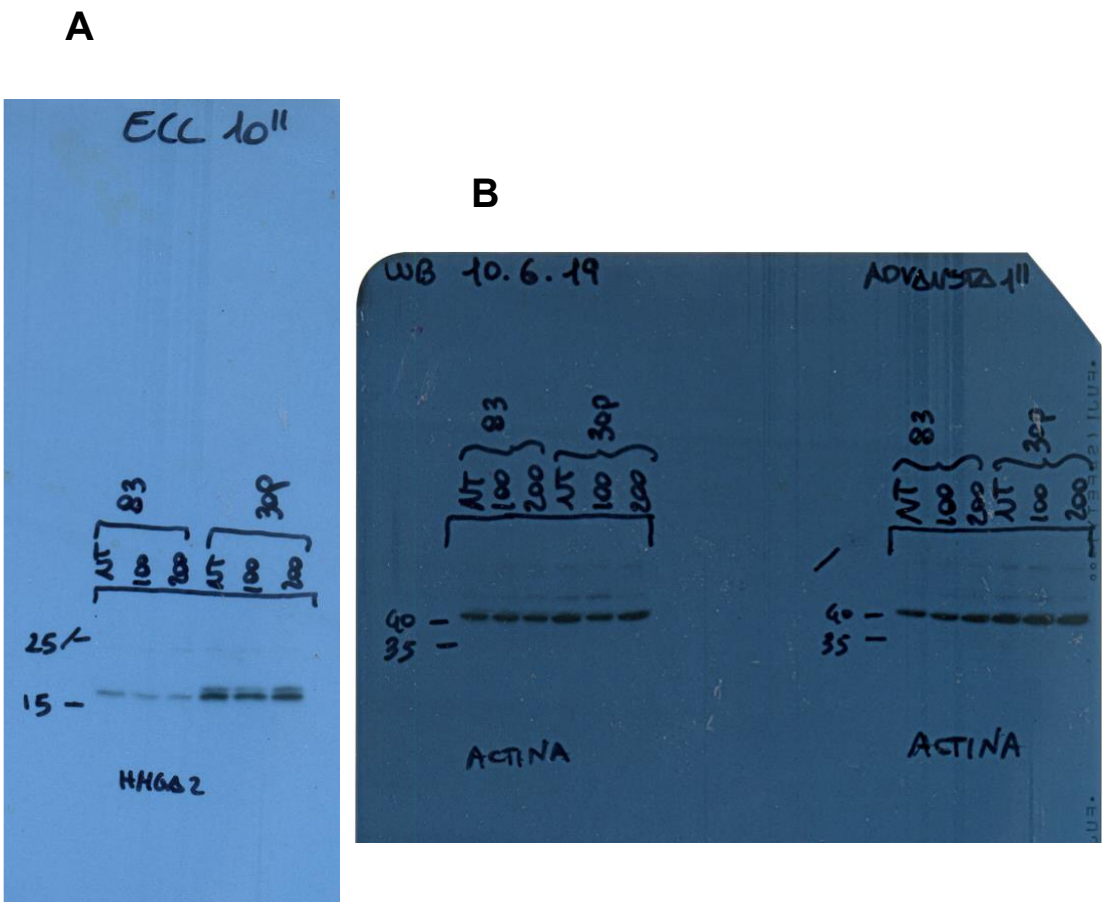


Figure S2.11. The western blotting images that were used to generate Figure 2.5B. **(A)** The western blotting results of HMGA2. **(B)** The western blotting results of actin.

Table S2.1. DNA oligonucelotides used for primers of the RT-PCR experiments.

Oligo names	Sequence (5-3')	Gene
HMGA2_F[a]	TCCCTCTAAAGCAGCTCAAAA	HMGA2
HMGA2_R[a]	ACTTGTTGTGGCCATTTCCT	HMGA2
SNAIL_F	AGTGGTTCTTCTGCGCTACT	SNAIL
SNAIL_R	GGGCTGCTGGAAGGTAAACT	SNAIL
ID2_F	TGTCAAATGACAGCAAAGCAC	ID2
ID2_R	GTTGTTGTTGTGCAAAGAATAAAAG	ID2
MSI-1_F	CCAATGGGTACCACTGAAGC	MSI-1
MSI-1_R	CACTCGTGGTCCTCAGTCAG	MSI-1
OLIG2_F	GACAAGCTAGGAGGCAGTGG	OLIG2
OLIG2_R	CGGCTCTGTCATTTGCTTCT	OLIG2

[a] F and R represent forward and reverse orientations, respectively.

Table S2.2. Parameters for the primary HTS assay of HMGA2-DNA interactions using LOPAC library.

# of compounds	1280
Tested concentration	5 μM
Z' value	0.83
RZ' value	0.9
[a]S/B	438
# of compounds with inhibition>50%	17
Hit rate	1.25%

[a] S/B represents signal versus background ratio.

Table S2.3. Inhibitors identified by the HTS assays.

Compounds	IC_{50} (μM)			
	Alpha Screen	Lance Screen	BRD4	Cytotoxicity
WP631	0.044	0.17	2.05	0.17
Cisplatin	> 100	9.00	> 100	> 100
cDPCP[a]	> 100	15.41	> 100	> 100
Ro 90-7501[a]	1.32	11.58	16.90	> 100
ATA[a]	0.32	0.21	5.00	> 100
Mitoxantrone	0.18	1.77	0.99	1.65
Suramin	4.06	2.58	> 100	> 100
Carboplatin	> 100	> 100	> 100	> 100

[a] cDPCP, cis-Diammine(pyridine)chloroplatinum(II) chloride; Ro 90-7501, 2′-(4-Aminophenyl)-[2,5′-bi-1H-benzimidazol]-5-amine; ATA, aurintricarboxylic acid.

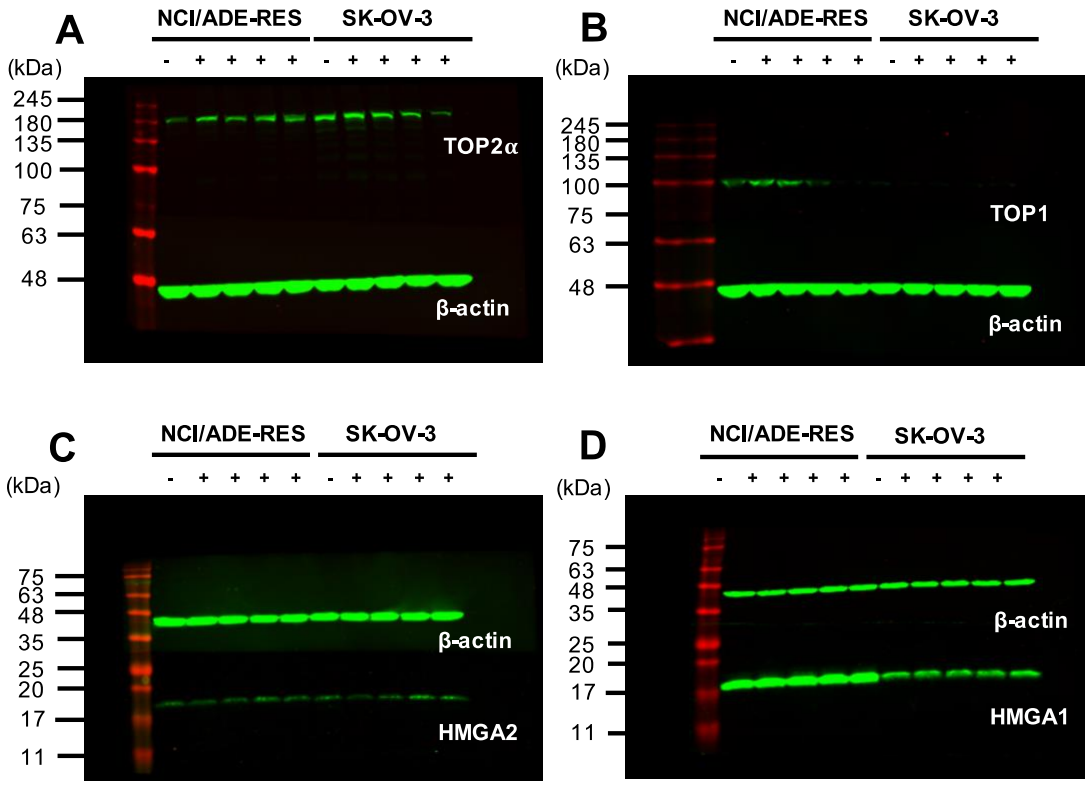


Figure S4.1. The western blotting images that were used to generate Figure 4.15A. The western blotting results of TOP2α **(A)**, TOP1 **(B)**, HMGA2 **(C)**, and HMGA1 **(D)** expression in NCI/ADR-RES and SK-OV-3 cells in the absence (-) and presence (+) of 10 μM DNR, DNR-GTP, DOX, and DOX-GTP, respectively. The loading amount of whole cell extracts is 20 μg, and β-actin was used as a loading control.

VITA

LINJIA SU

Born, Shandong, China

2009-2013	B.S., Applied Chemistry Yantai University Shandong, China
2013-2016	M.S., Physical Chemistry Capital Normal University Beijing, China
2016-2021	Ph.D., Chemistry Florida International University Miami, Florida, USA

PUBLICATIONS, PATENTS AND PRESENTATIONS

Pandey, P.; Su, L.; Leng, F.; He, J. Quantitative single-cell delivery of protein molecules via a multifunctional nanopipette. In preparation.

Pandey, P.; Bhattarai, N.; Su, L.; Wang, X.; Leng, F.; Gerstman, B.; Chapagain, P. P.; He, J. Detecting individual proteins and their surface charge variations in solution by potentiometric nanoimpact method. Submitted, 2021.

Su, L.; Leng, F. Synthesis and characterization of daunorubicin-GTP, doxorubicin-GTP, daunorubicin-dGTP, and doxorubicin-dGTP conjugates. In preparation.

Su, L.; Fernandez-Lima, F.; Chapagain, P. P.; Leng, F. Inhibition of HMGA2 binding to AT-rich DNA by its negatively charged C-terminus. In preparation.

Su, L.; Bryan, N.; Battista, S.; Freitas, J.; Garabedian, A.; D'Alessio, F.; Romano, M.; Falanga, F.; Fusco, A.; Kos, L.; Chambers, J.; Fernandez-Lima, F.; Chapagain, P. P.; Vasile, S.; Smith, L.; Leng, F. Identification of HMGA2 inhibitors by AlphaScreen-based ultra-high-throughput screening assays. Scientific Reports 2020, 10 (1), 18850.

Su, L.; Deng, Z.; Leng, F. The mammalian high mobility group protein AT-hook 2 (HMGA2): biochemical and biophysical properties, and its association with adipogenesis. International Journal of Molecular Sciences 2020, 21 (10), 3710.

Su, L.; Hu, S.; Zhang, L.; Wang, Z.; Gao, W.; Yuan, J.; Liu, M. A fast and efficient replacement of CTAB with MUA on the surface of gold nanorods assisted by a water-immiscible ionic liquid. Small 2017, 13 (11), 1602809.

Su, L.; Yuan, J. A method for the surface modification of gold nanoparticles in ionic liquids, China Patent No. 201510868027.7

Su, L.; Leng, F. Synthesize and characterization of GTP-daunorubicin, GTP-doxorubicin, dGTP-daunorubicin, and dGTP-doxorubicin conjugates. The 2021 Annual BSI Research Symposium, Florida International University, 2021, Miami, FL, US.

Su, L.; Deng, Z.; Chambers, J.; Leng, F. Inhibition of HMGA2 binding to AT-rich DNA by its own negatively charged C-terminus. The 2020 Annual BSI Research Symposium. Florida International University, 2020, Miami, FL, US.

Su, L.; Deng, Z.; Leng, F. Potent inhibition of HMGA2 binding to AT-rich DNA by its own negatively charged C-terminus. The 64th Annual Meeting of Biophysical Society, 2020, San Diego, CA, US.

Su, L.; Chambers, J.; Leng, F. Potent inhibition of the mammalian high mobility group protein AT-hook 2 binding to DNA by natural polyphenols. The 2019 Southeastern Regional Meeting (SERMACS), 2019, Savannah, GA, US.

Su, L.; Chambers, J.; Leng, F. Tight binding of natural polyphenols to the intrinsically disordered mammalian high mobility group protein AT-hook 2. The 63rd Annual Meeting of Biophysical Society, 2019, Baltimore, MA, US.

Su, L.; Chambers, J.; Leng, F. Tight binding of natural polyphenols to the intrinsically disordered mammalian high mobility group protein AT-hook 2. The 2019 Graduate Student Appreciation Week (GSAW), Florida International University, 2019, Miami, FL, US.

Su, L.; Chambers, J.; Leng, F. Tight binding of natural polyphenols to the intrinsically disordered mammalian high mobility group protein AT-hook 2. The 2019 Annual BSI Research Symposium, Florida International University, 2019, Miami, FL, US.

Su, L.; Garabedian, A.; Fernandez-Lima, F.; Chapagain, P. P.; Vasile, S.; Smith, L.; Leng, F. Identification of suramin as a potent and specific inhibitor of the mammalian high mobility group protein AT-hook 2 (HMGA2)-DNA interactions. The 94th Florida Annual Meeting and Exposition (FAME) conference, 2018, Tampa, FL, US.

Su, L.; Vasile, S.; Smith, L.; Leng, F. Identification of suramin as a potent and specific inhibitor of the mammalian high mobility group protein AT-hook 2 (HMGA2)-DNA interactions." The 63rd Annual Meeting of Biophysical Society, 2018, San Francisco, CA, US.