UNIVERSITY OF CALIFORNIA, SAN DIEGO

Selectivity in the Interactions Between Positively Charged Small Molecules and Negatively Charged Biopolymers

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Chemistry

by

Lev Elson-Schwab

Committee in charge:

Professor Yitzhak Tor, Chair
Professor Michael D. Burkart
Professor Thomas Hermann
Professor Simpson Joseph
Professor James T. Kadonaga

2006

The dissertation of Lev Elson-Schwab is approved, and it is acceptable in quality for publication on microfilm:

Chair

University of California, San Diego

2006

To my wife, parents and brother

Table of Contents

List of Figures

List of Abbreviations

AIDS	Acquired Immunodeficiency Syndrome
2-AP	2-Aminopurine
$Arg_{8/9}$	$Arginine_{8/9}$
Boc	*Tert*-Butoxycarbonyl
BODIPY	*N*-(4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-*s*-indacene)
BSA	Bovine Serum Albumen
Cbz	Benzyl Chloroformate
CD	Circular Dichorism
CHO	Chinese Hamster Ovarian
DIS	Dimerization Initiation Site
DMF	*N*,*N*-Dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic Acid
2-DOS	2-Deoxystreptamine
EDTA	Ethylenediaminetetraacetic Acid
FACS	Fluorescent Activated Cell Sorting
Fl	Fluorescein
HAART	Highly Active Antiretroviral Therapy
HPLC	High Pressure Liquid Chromatography
IdoA	L-iduronic Acid
FGF	Fibroblast Growth Factor
GAG	Glycosaminoglycan
GlcA	Uronic Acid
GlcN	D-Glucosamine
Gly	Glycine
HIV	Human Immunodeficiency Virus
kDa	Kilodalton
LTR	Long Terminal Repeat
M	Molar
mM	Milimolar
μM	Micromolar
NaCl	Sodium Chloride
NES	Nuclear Export Signal
NCI	National Cancer Institute
NLS	Nuclear Localization Signal
nm	nanometer
nM	Nanomolar
NMR	Nuclear Magnetic Resonance
nt	Nucleotide
PBS	Primer Binding Site/Phosphate Buffered Saline

pM	Picomolar
PTD	Protein Transduction Domain
RNA	Ribonucleic Acid
rpm	Revolutions per Minute
rRNA	Ribosomal RNA
RRE	Rev Response Element
Ru^{II}	Ruthenium II
SL1	Stem Loop 1
TAR	Trans-Activation Response
tRNA	Transfer RNA
Ser	Serine

Acknowledgements

I have been privileged to work as part of an excellent group of scientists in the Tor Lab, they have all helped to make my graduate school experience entertaining as well as educational. I would like to thank my mentor, Professor Yitzhak Tor for all that he has given me. I do not think that I could have asked for a better boss. He has been a patient, generous and inspirational person to work for, and I do not think that I will ever forget his lessons, both life and scientific.

As with any scientific project, many people have contributed to this work. Groundwork done by Dr. Nathan Luedtke opened up many avenues of previously unstudied work, a few of which I have followed. I was very lucky to work with two very talented biologists, Omai Garner and Manuela Schuksz, who helped to make the work that we did together both profitable and a highly enjoyable experience.

I would especially like to thank my family for the support which they have given me throughout my life. My parents, Michael Schwab and Margret Elson have given me more than any person can expect. Their love and guidance have enabled me to accomplish more than I thought I could. Many people have inspired me in one way or another throughout my life, but none more than my mom and dad.

And, last but not least there are two more people without whom I would not be who or where I am. My brother, Ilan Elson-Schwab, is one of the most amazing people that I know. My wife, Lea-Ann Torres is a continuous source of inspiration for me. Her passion and her love help fuel our life together. Without her, none of this work would have been possible.

Vita

Lev Elson-Schwab
Tor Lab
Department of Chemistry and Biochemistry
University of California, San Diego

Education	1999	*B.S., Biochemistry and Molecular Biology* University of California, Santa Cruz
	2006	*Ph.D., Chemistry* University of California, San Diego, Department of Chemistry and Biochemistry
Research Experience	1999-present	*Yitzhak Tor Laboratory* University of California, San Diego
	1999	*Monique Cosman Laboratory* Lawrence Livermore National Laboratory
	1997-1999	*Harry Noller Laboratory* University of California, Santa Cruz
	1995 – 1997	*Joseph D. Puglisi Laboratory* University of California, Santa Cruz and Stanford University
Awards and Grants	2001-2003	*Molecular and Genetics Training Grant* National Institute of Health and University of California, San Diego
Teaching Experience	2004	*Chemistry 143A* *Organic Chemistry Laboratory* University of California, San Diego

Evaluation Score: 3.3/4.0

2004 *Chemistry 12*
Molecules and Reactions
University of California, San Diego
Evaluation Score: 3.5/4.0

2001 *Chemistry 13*
Chemistry of Life
University of California, San Diego
Evaluation Score: 3.3/4.0

2000 and 2003 *Chemistry 6A*
General Chemistry
University of California, San Diego
Evaluation Score: 3.5/4.0

2000 *Chemistry 112B*
Molecular Biochemistry Laboratory
University of California, San Deigo
Evaluation Score: 3.2/4.0

2000 *Chemistry 6B*
General Chemistry
University of California, San Diego
Evaluation Score: 3.0/4.0

Journal Articles

L. Elson-Schwab, O. Garner, M. Schuksz, J. Esko and Y. Tor, Binding and delivery of guanidinoglycosides depends on heparan sulfate (in preparation)

J. Boer, K. F. Blount, N. W. Luedtke, L. Elson-Schwab, Y. Tor, RNA-selective Modification by a Pt(II) Complex Conjugated to Amino- and Guanidino-glycosides. *Angew. Chem. Int. Ed.* **2005**, *44*, 927–932.

J. Aldrich-Wright, C. Brodie, E. C. Glazer, N. W. Luedtke, L. Elson-Schwab, Y. Tor, Symmetrical Dinuclear Complexes with High DNA

J. Aldrich-Wright, C. Brodie, E. C. Glazer, N. W. Luedtke, L. Elson-

Book Chapters

L. Elson-Schwab and Y. Tor, "Targeting HIV-1 RNA with Aminoglycoside Antibiotics and their Derivatives" in *Aminoglycoside Antibiotics: From Chemical Biology to Drug Discovery*, Ed. Dev Arya, *Wiley-VCH* (in press)

Acknowledgements

N.W. Luedtke and Y. Tor, Fluorescence-based methods for evaluating the RNA affinity and specificity of HIV-1 Rev-RRE inhibitors. *Biopolymers*, **2003**, 70, 103-119

E.V. Puglisi and J.D. Puglisi, HIV-1 A-rich loop mimics the tRNA anticodon structure. *Nature Struct. Biol*., **1998**, 5, 1033-1036

ABSTRACT OF THE DISSERTATION

Selectivity in the Interactions Between Positively Charged Small Molecules and Negatively Charged Biopolymers

by

Lev Elson-Schwab

Doctor of Philosophy in Chemistry

University of California, San Diego, 2006

Professor Yitzhak Tor, Chair

There are multiple negative charged biopolymers that exist within the context of the cell including cell surface proteoglycans, the cell membrane, RNA and DNA. Guanidinoglycosides, a family of small, positively charged molecules bind with surprising selectivity to cell surface heparan sulfate proteoglycans. This binding event allows these compounds to internalize into the cell via receptor mediated endocytosis, and be released into the cytoplasm. Guanidinoglycosides also have a high affinity for RNA and show selectivity for the RRE over other RNA and DNA.

Guanidinoglycosides have significant potential as both cellular delivery agents and as scaffolds for the design of anti-HIV therapeutics. Through conjugation to therapeutic agents, guanidinoglycosides can increase the efficacy of these compounds by allowing them to cross cellular membranes and localize within the cytoplasm, and through modifications which increase their selectivity for specific RNA targets guanidinoglycosides can achieve more potent anti-viral activity.

Chapter 1

1.1 Introduction

According to figures generated by the World Health Organization, approximately 40 million people worldwide are currently living with HIV/AIDS. While HIV infection rates have been decreasing in several countries, the global number of people living with HIV continues to rise, with the majority of new cases being diagnosed in young people (15-24 years old).[1] The AIDS epidemic continues to outpace the global response. Since the first AIDS diagnosis in 1981, progress has been made in understanding the biology of the virus, preventing new infections and improving the quality and length of life for people living with HIV.[2] However, there is a pressing need for the development of new anti-retroviral therapies.

Development of antiretroviral therapy targeted at HIV has been a major challenge since the discovery of the virus.[3,4] Early antiretroviral drugs targeted the viral reverse transcriptase and protease enzymes. More recently, drug cocktails or HAART (highly active antiretroviral therapy), where multiple treatments are administered simultaneously, have shown increased efficacy.[5,6] However, current therapeutics can cause adverse effects including abdominal pain and vomiting, bone disorders, liver and kidney toxicity, sexual dysfunction and psychiatric disorders.[7,8] Additionally, HIV has an extremely high rate of mutation. This causes viral strains which are resistant to current therapeutics to evolve and appear extremely quickly.[9,10] This highlights the need for development of new, independent classes of anti-HIV agents.[11,12]

Host-virus interactions, of which there are many during the HIV lifecycle, present appealing targets for therapeutic intervention. Drugs can interact with viral receptors, enzymes, structural components, genes or transcripts. Recently, as the understanding of the importance of RNA in the viral lifecycle has emerged, viral RNA has become increasingly attractive as a potential therapeutic target. Small molecules which can target viral-specific RNA sites and prevent the formation of key RNA–protein and RNA–RNA interfaces are becoming increasingly attractive as potential drugs.[13,14]

It is important to have a detailed understanding of the multiple functions of RNA in the viral life cycle when designing and analyzing novel potential RNA targeted small molecules. As a retrovirus, HIV utilizes RNA as its genetic material. Viral replication relies on reverse-transcribing the viral RNA genome into DNA, and later transcribing it back into RNA. However, RNA also serves many other key functions during the viral life cycle. These functions rely on specific RNA–RNA or RNA–protein interactions, and are essential for viral replication. The elaborate life cycle of HIV requires an ordered pattern of viral gene expression, and can be divided into three major phases (Figure 1.1).[15,16]

The first stage involves recognition of the host cell, fusion/uncoating, reverse transcription and integration into the host cell chromosome. Integration within the host genome establishes a permanent presence of the viral genome within the infected cells. The next major stage includes transcription of viral RNA, translation of viral proteins and processing of the RNA and proteins into their mature forms, generating the

necessary building blocks for new viral particles. In the final step, new viral particles are assembled and released from the cell to infect new cells (Figure 1.1).

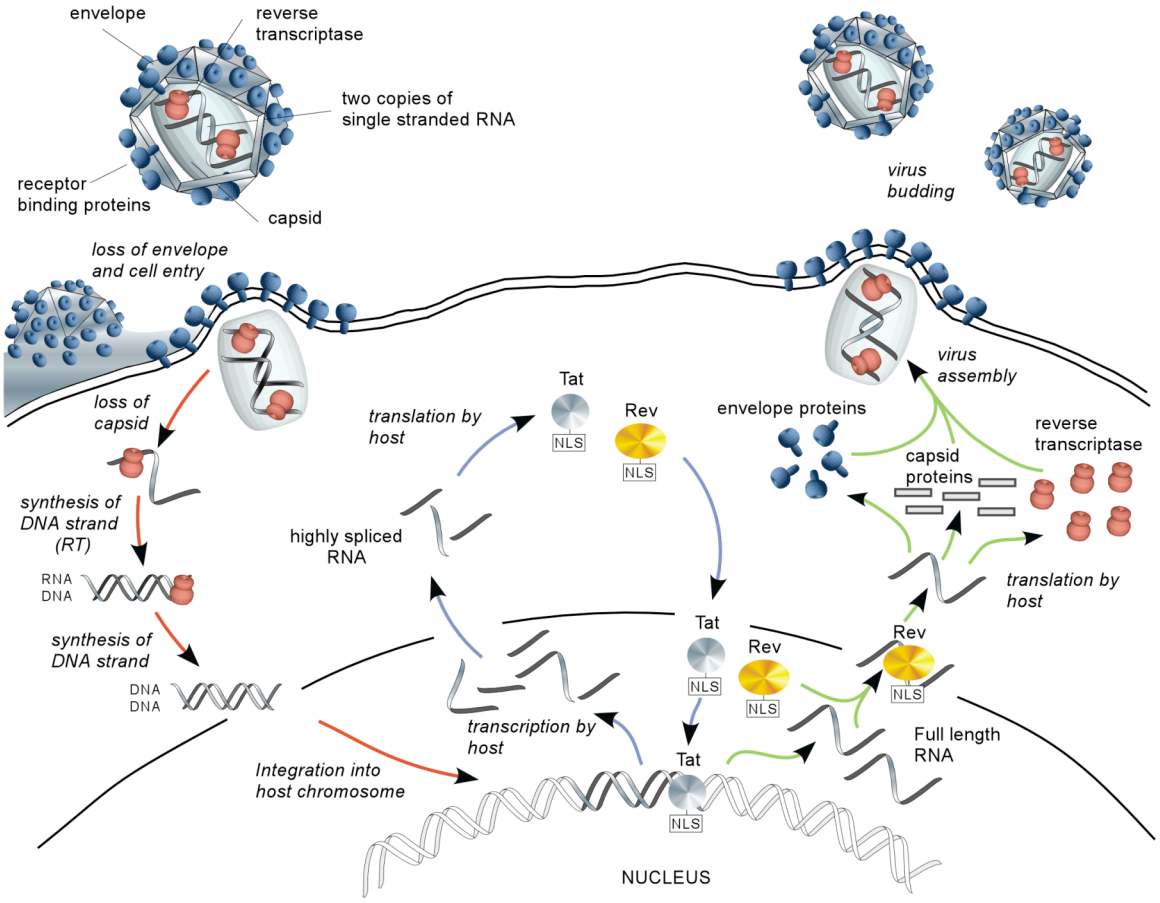


Figure 1.1. Schematic illustration of the HIV-1 life cycle. Following receptor binding, membrane fusion and partial loss of capsid, the single stranded viral RNA genome is released into the host cytoplasm where is it reverse transcribed to double stranded DNA which is then integrated into the host genome. Short, viral RNA transcripts are produced, and the Tat and Rev regulatory proteins are expressed. Tat and Rev both re-enter the nucleus with Tat facilitating the formation of full length viral transcripts and Rev mediating export of fully unspliced and partially spliced viral RNA. The unspliced RNA allows for the biosynthesis of the remaining viral proteins, and new viral particles with the complete HIV genome are formed. Reproduced with permission from Wiley-VCH.

Each stage involves at least one critical RNA–RNA or RNA–protein interaction. Initiation of reverse transcription involves the binding of a specific tRNA, which is packaged in the viral particle, to a short sequence of viral RNA. Transcription and accumulation of viral RNA is dependent on sequence specific interactions of the Tat and Rev proteins with their respective RNA binding sites on the viral transcript, specifically the trans-activating region (TAR) and Rev response element (RRE).[17,18] The Tat protein acts as a transcriptional activator[19,20] whereas Rev acts post-transcriptionally to export fully unspliced and partially spliced viral RNA transcripts from the nucleus into the cytoplasm.[21,22] Viral assembly is initiated by the formation of an RNA–RNA dimer. The full genomic viral RNA dimerizes and is selectively packaged into new viral particles.[23]

RNA–RNA and RNA–protein interactions in the life cycle of the HIV virus can be highly selective, making them attractive targets for potential therapeutic intervention. Small molecules which can specifically disrupt these viral-specific molecular recognition events therefore may show potential as future antiviral agents. Aminoglycoside antibiotics are a family of naturally occurring amino sugars which are known to bind to a variety of RNA sequences, and inhibit RNA–protein interactions. This RNA binding activity gives these molecules potential as scaffolds for future RNA targeted anti-HIV therapeutics.

1.2 Aminoglycoside Antibiotics

Aminoglycoside antibiotics represent a family of highly charged naturally occurring pseudo-oligosaccharides.[24] The common core of most aminoglycosides is 2-deoxystreptamine (2-DOS) (Figure 1.2).[25] The 2-DOS core is glycosidated with mono- or di-saccharides, typically at the 4 and 5 or 4 and 6 positions, generating most aminoglycosides. The antibacterial activity of aminoglycoside antibiotics is derived

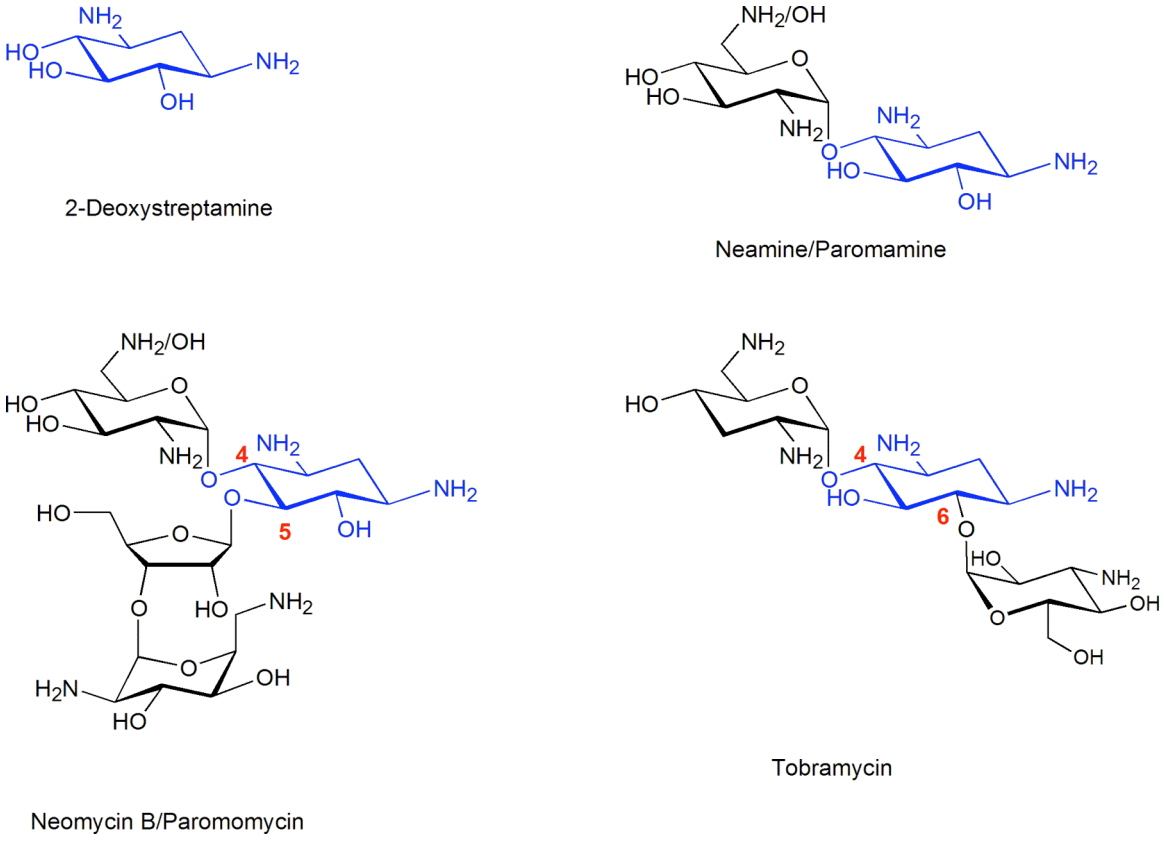


Figure 1.2. Structures of several aminoglycosides and their common cores, highlighting the 2-deoxystreptamine (2-DOS) moiety which is a structural feature found among many aminoglycoside class antibiotics.

from their ability to bind to the bacterial ribosomal decoding site (A-site) which causes misreading of the genetic code and translational inhibition, resulting in bacteriacidal activity.[26]

The discovery that aminoglycosides interfere with protein biosynthesis through specific binding to the decoding region on prokaryotic 16S rRNA[27] helped to define these antibiotics as RNA binders. This research triggered an extensive search for other potential RNA targets, and many have since been discovered. Aminoglycosides have been shown to bind to (and inhibit) self-splicing group I introns,[28,29] the hammerhead ribozyme, [30,31] the hepatitis delta ribozyme,[32,33] RNase P,[34] and even transfer RNA.[35] Multiple other aminoglycoside binding sites on both small and large RNA molecules have been discovered.[36]

Aminoglycosides bind to a wide range of RNA targets, signifying that they are rather promiscuous binders. Electrostatic interactions have been demonstrated to play a critical role in RNA–aminoglycoside interactions,[30] prompting a model suggesting "structural electrostatic complementarity".[37,38] However, structural motifs within folded RNA may provide aminoglycoside recognition sites. Small changes in RNA or aminoglycoside structure can cause large differences in the binding interaction,[39,40] suggesting that these interactions can potentially be highly specific. It is doubtful that these naturally occurring antibiotics will display the necessary selectivity to become therapeutically relevant,[41] but they may serve as scaffolds for the development of higher affinity and more selective RNA binders.

However, aminoglycosides have been shown to bind to viral RNA, blocking the formation of RNA-peptide and RNA-protein complexes. This is shown in both the HIV

Tat–TAR[42] and Rev–RRE[43] systems, and has highlighted aminoglycosides as potential viral RNA binding molecules. There are multiple RNA–RNA and RNA–protein interactions in the HIV life cycle, some of which may serve as potential targets for aminoglycoside based therapeutics based on aminoglycoside scaffolds.

1.3 RNA Interactions in The HIV life cycle

For a viral particle to infect the host cell, receptors on the host cell must recognize the viral particle to allow for fusion and uncoating. The viral coat protein Gp120 binds to CD4 receptors on the cell surface. This interaction does not lead to fusion or active entry into cells, but rather it promotes viral association with the host entry receptors, chemokine co-receptors CCR5 and CXCR4 (Figure 1.3).[44,45] Natural ligands to these receptors prevent HIV entry through direct competition for co-receptor binding and down-regulation of the co-receptors. Individuals lacking CCR5 expression are highly resistant to HIV infection.[46] After Gp120 has bound to CD4 and CCR5 receptors, the viral protein Gp41 induces fusion with the host cell membrane.[47]

Uncoating involves viral proteins $p17^{MA}$, Nef, Vif and a variety of host factors. $p17^{MA}$ is phosphorylated by host kinases. Upon phosphorylation, $p17^{MA}$ promotes dissociation of the matrix protein from the plasma membrane, and is also responsible for connecting the reverse transcription complex to the host actin filaments to allow translocation to the nucleus.[48] After uncoating, the reverse transcription complex is formed. It is composed of viral RNA, reverse transcriptase, integrase, $tRNA^{Lys}_3$, $p17^{MA}$, nucleocapsid p7, Vpr (viral protein R) and multiple host proteins.

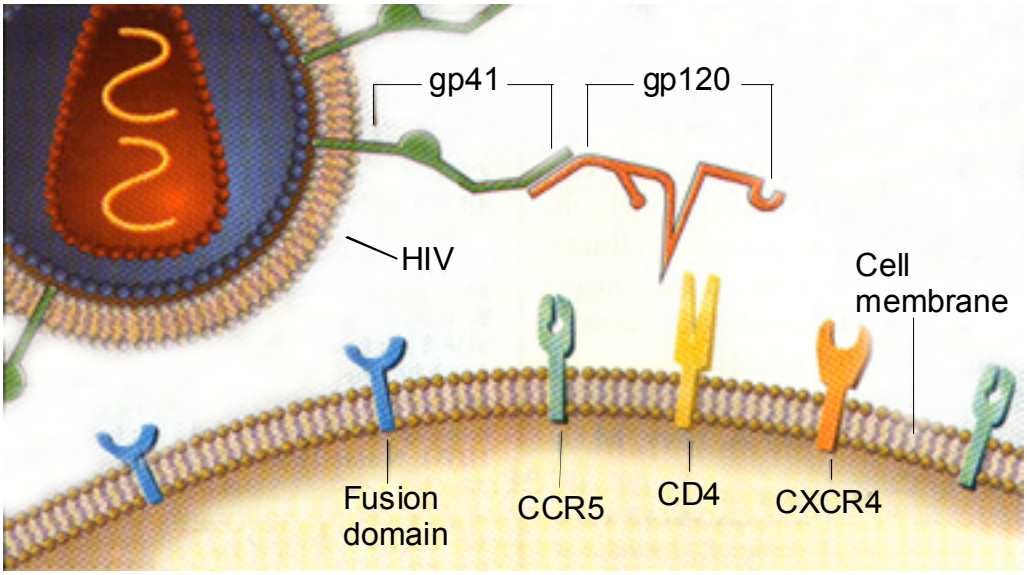


Figure 1.3. Gp120 binds sequentially to CD4 receptors then to CCR5 entry receptors. This second attachment brings the viral envelops closer to the cell surface, allowing interaction between gp41 on the viral envelope and a fusion domain on the cell surface.

1.4 Initiation of Reverse Transcription and the Primer Binding Site

Initiation of viral reverse transcription in cells infected with HIV relies on a critical RNA-RNA interaction between $tRNA^{Lys}_3$, a tRNA preferentially packaged into the viral particle[49], and a specific sequence of viral RNA. The 3′-terminal 18 nucleotides of $tRNA^{Lys}_3$ are complementary to the primer binding site (PBS) sequence located in the 5′-long terminal repeat (LTR) of the viral RNA genome. The UUU anticodon of the tRNA is complementary to and binds to an adenosine rich loop located 8 nucleotides upstream (5′) of the PBS. This RNA-RNA duplex, which is formed when $tRNA^{Lys}_3$ binds to the PBS, fits within the active site of HIV-1 reverse transcriptase,[50,51]

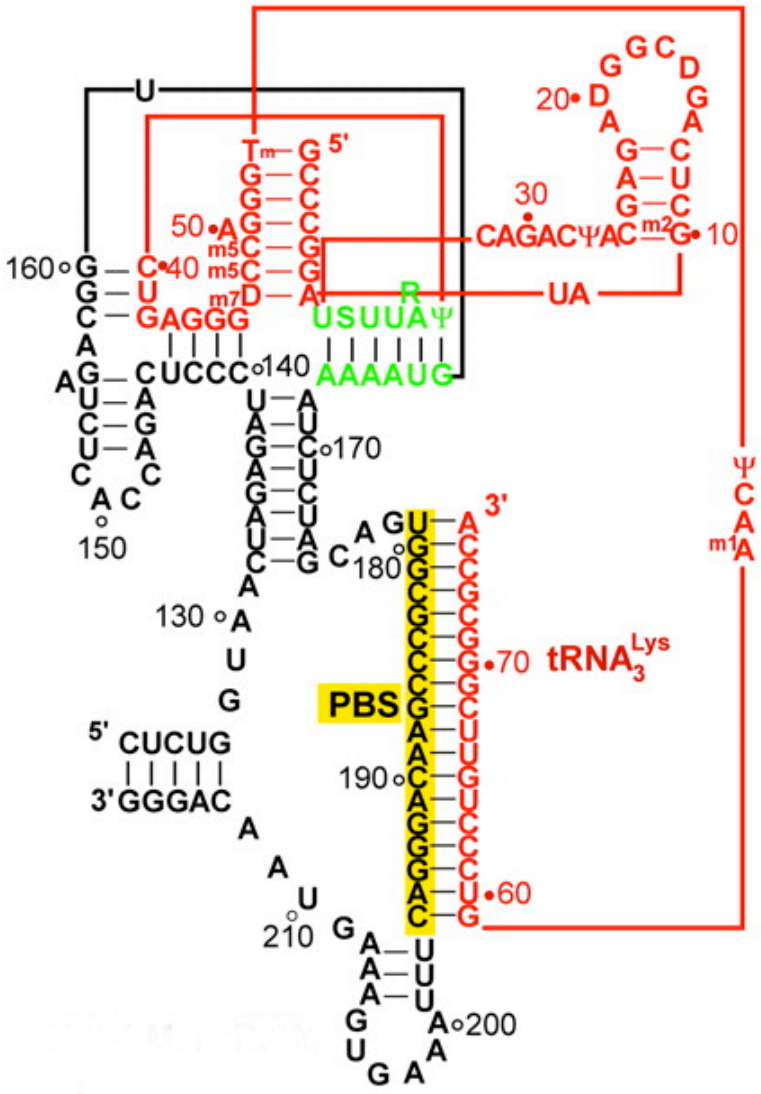


Figure 1.4. Secondary structure of $tRNA^{Lys}_3$ (red) bound to viral RNA. Reverse transcription is initiated by the binding of $tRNA^{Lys}_3$ to the primer binding site (PBS). The viral A-rich loop and complementary anticodon loop of the tRNA (green) form a kissing loop which is essential to initiation. Part of the anticodon arm and of the variable loop of $tRNA^{Lys}_3$ interact with viral sequences upstream in the PBS (yellow highlight) to facilitate initiation of reverse transcription. (figure modified from reference 55)

but multiple interactions between the viral RNA and $tRNA^{Lys}_3$ are necessary for efficient initiation of reverse transcription.[49] This interaction nucleates the formation of the reverse transcription complex which contains viral RNA, reverse transciptase, $tRNA^{Lys}_3$, $p17^{MA}$, nucleocapsid p7 and Vpr (Viral protein R), as well as multiple host factors.[48]

The A-rich loop within the PBS plays a key role in retroviral replication and likely pairs with the complementary $tRNA^{Lys}_3$ anticodon (Figure 1.4).[50-52] Structural analysis of this A-rich loop shows a U-turn motif, an RNA structure which was first seen in $tRNA^{Phe53}$, but which is a common structural motif in many tRNAs.[54] If, as evidence suggests, the four adenosines in the A-rich loop of the PBS pairing with the four uridines in the anticodon stem loop of $tRNA^{Lys}_3$ is a critical event in the HIV life cycle, this could provide a potential target for therapeutic intervention.

The therapeutic potential of this RNA target is questionable due to the structural similarity of the A-rich loop and the anticodon stem loop of tRNA. Any potential drug must have not only reasonable affinity for its target but must also display a high degree of selectivity. A high level of discrimination may be difficult to achieve due to the prevalence of transfer RNA in cells.

Additionally, the therapeutic potential of aminoglycosides may by limited as they are known to bind double stranded RNA regions,[55,56] making it unlikely that they would provide a good framework for synthetic ligands targeting the A-rich loop. A molecule which preferentially binds the complex formed between the A-rich loop and $tRNA^{Lys}_3$ could potentially stabilize this interaction, promoting viral replication rather

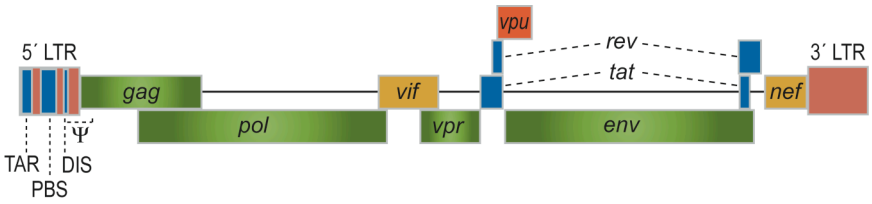


Figure 1.5. Schematic view of the HIV-1 genome, showing gene structure and splicing patterns, as well as important RNA sites located within the 5′-long terminal repeat (LTR). Many RNA sequences known to be involved in key RNA-RNA and RNA-protein interactions are located at the far 5′ end of the genome.

than hindering it. Based on these considerations, it is unlikely that the PBS would be classified as a good target for modified aminoglycosides designed to interfere with reverse transcription.

Once the viral RNA genome (Figure 1.5) has been transcribed into DNA through reverse transcription, it must enter the nucleus and be integrated into the host genome. The pre-integration complex (viral DNA, Reverse transcriptase, $p17^{MA}$, Vpr and various host factors) forms and connects to host microtubules, which allow it to translocate into the nucleus.[57] Integration is promoted by integrase, a viral enzyme, as well as several host factors, and takes place at transcriptional "hot spots".[48,58-61] This process is for the most part RNA independent, making it unlikely that aminoglycosides or related analogs, which preferentially bind to RNA, would be selective inhibitors.

Once integrated into the host chromosome, viral transcripts and proteins must be produced for progeny viral particles to be formed. Initiation of viral transcription is also an RNA independent process where host transcription promoter and enhancer elements such as NF-κB, a transcription factor that regulates immune response, bind to the 5′-LTR. This causes the host transcriptional complex to also bind to the 5′-LTR, and serves to initiate transcription.[58,62-64] Again, with no specific RNA–RNA or RNA–

protein interactions, there is no obvious target for inhibition with aminoglycosides or aminoglycoside based molecules.

After transcription has been initiated, RNA–protein and RNA–RNA interactions play a critical role in the creation of full length viral transcripts. The large, multi-protein transcription complex has a recognition factor for non-host DNA and will quickly release from viral DNA, creating short, abortive transcripts. Processing and nuclear export of these transcripts leads to the translation of the HIV Tat protein,[64,65] a small early phase viral protein that plays a key role in the formation of full length viral RNA transcripts.

1.5 Regulation of viral transcription and the TAR RNA

HIV Tat can cross nuclear membranes and localize in the nucleus[66,67] where it binds to the TAR (trans-activating region) RNA, a stem-loop RNA found on the 5′-LTR (Figure 1.6). The formation of this RNA–protein complex stimulates the transcription of the full length viral genome by preventing the release of the transcription complex from the non-host viral DNA,[63,68] and is essential for transcription of full length viral RNA. Tat is involved in a positive feedback mechanism that helps to ensure high levels of HIV transcription. Being a key step in the life cycle of the virus, the Tat-TAR interaction has been the subject of structural studies,[69] discovery assay development, [70,71] and has been widely studied as a potential therapeutic target.[19,20,72-74]

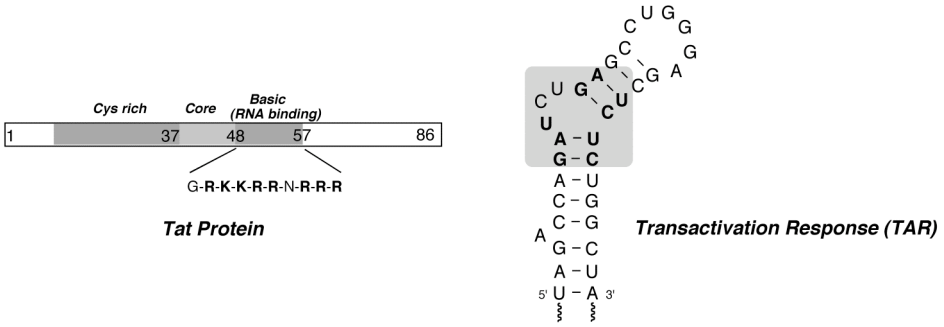


Figure 1.6. Schematic representation of the Tat protein and its functional regions. The arginine-rich basic RNA binding domain is highlighted. The secondary structure of its RNA target, the TAR RNA is shown with residues critical for Tat binding highlighted.

The interactions between aminoglycoside antibiotics and the HIV-1 TAR RNA have been extensively studied. It was shown that aminoglycosides bind TAR and inhibit the formation of the Tat-TAR complex.[34,71] Aminoglycoside derivatives as well as other small molecules have been studied as TAR binders.[75-80] It was found that the conjugation of arginine to aminoglycosides can generate potent TAR binders[81-83] which show antiviral activity in infected cells.[84] Several other neomycin conjugates containing chloramphenicol and linezolid have been synthesized and found to display approximately 10-fold higher affinity to TAR compared to the parent aminoglycoside.[85]

Other families of small molecules have also been found to effectively compete with the Tat peptide for TAR RNA binding.[75-80] Footprinting studies with diphenylfuran derivates suggest that the bulge region is critical for RNA binding, and that TAR recognition appears to be shape-specific rather than sequence specific.[86] Many small molecules including polyamine-acridine conjugates,[87] oligomeric peptides, peptidomimetics and oligoureas have been reported to bind to TAR and inhibit the Tat-TAR interaction.[88]

The binding of a series of natural aminoglycosides has been measured with the TAR RNA[89], and as observed in other aminoglycoside-RNA interactions,[90-92] the affinities correlate with the number of amines on the aminoglycoside, implying that electrostatics dominate the binding interaction. However, the number of amines in the aminoglycoside is not the only factor which influences binding. Paromomycin, kanamycin B and tobramycin all have five amines but display a range of affinities to the pyrene-labeled TAR RNA.[89] Aminoglycosides are promiscuous RNA binders capable of binding numerous RNA sequences with comparable affinity. This lack of selectivity likely results from the electrostatic driven binding mode of aminoglycosides and their conformational adaptability. The inherent flexibility of the glycosidic bonds allows for a large degree of rotational freedom, giving these molecules the ability to assume a variety of structural conformations, allowing for structural rearrangement and adaptation to a diverse array of RNA targets.[38]

This structural plasticity leads to the formation of aminoglycoside complexes with diverse RNA targets in which the aminoglycosides assume distinct and, sometimes, markedly different conformations. Structural analysis of neomycin B and paromomycin bound to the HIV-1 TAR RNA and the prokaryotic ribosomal A-site show markedly different conformations. When bound to the A-site, paromomycin is compactly arranged with ring III in close proximity to the 2-DOS ring.[93-95] This is in contrast to the structure of neomycin B bound to the HIV-1 TAR RNA where the aminoglycoside is arranged in a much more extended conformation.[96]

These data leads to the hypothesis that if the conformational flexibility of aminoglycosides were reduced, their target selectivity could potentially be altered. To

study the effects of aminoglycoside conformation on selectivity, conformationally constrained neomycin and paromomycin analogs were prepared.[97] By locking the aminoglycoside in a conformation similar to that when bound to the A-site RNA, but different from that when bound to the TAR RNA, the derivatives were designed to mimic the A-site bound aminoglycoside structure and disfavor the TAR-bound conformation. Not unexpectedly, locking the aminoglycosides into a relatively compact structure had a minimal impact of A-site affinity.[98] Surprisingly, fixing these neomycin-class aminoglycosides into what was considered a TAR-disfavored structure had almost no deleterious effect on binding to this HIV-1 TAR RNA sequence.

These observations can be rationalized by suggesting that the A-site and TAR RNAs possess inherently different selectivities towards aminoglycoside antibiotics.[97] The inherent flexibility of the TAR RNA, coupled with any remaining conformational flexibility within the constrained aminoglycoside analogs, may make this RNA site a very accommodating target for this class of polycationic ligands. It is possible that design of a larger, more rigid molecule will allow enhanced selectivity for flexible RNA structures like the TAR RNA. In contrast, the deeply encapsulating A-site is a more discriminating RNA target, meaning that small changes in ligand structure or conformation may potentially show greater difference in A-site binding.

This suggests that future design of RNA-directed therapeutics must consider the inherent structural selectivity of the RNA target as well as the selectivity of the ligand towards the specific target RNA. A discriminating RNA sequence such as the A-site may prove to be a much better RNA target than the HIV-1 TAR because *in vivo* selectivity for the TAR may prove difficult to attain.

1.6 Regulation of viral replication: RNA export and the RRE RNA

Full length viral RNA transcripts are processed by the host machinery prior to nuclear export. However, if only processed viral RNA transcripts are exported from the nucleus, viral replication would halt. Both partially spliced and fully unspliced viral RNA transcripts must be exported from the nucleus to serve as open reading frames for proteins (Gag, Pol and Env) which are essential components to the viral life cycle.[99] Additionally, each new viral particle must contain two copies of the fully unspliced viral RNA, which serves as the primary genome for progeny virions.

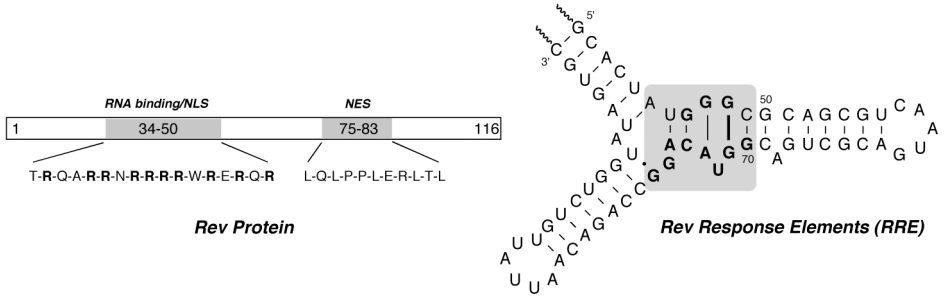


Figure 1.7. Schematic representation of the Rev protein. The RNA binding and nuclear localization domain (amino acids 34-50) and the nuclear export signal (amino acids 75-83) serve as its two key functional domains. The secondary structure of the Rev Response Element (RRE) RNA is shown with the high affinity Rev binding site highlighted, and essential residues shown in bold. The two non-Watson-Crick base pairs (G48:G71 and G47:A73), and the bulged U72 are highly conserved and necessary for the HIV-1 life cycle.[100,101]

The Rev protein serves to facilitate the nuclear export of viral RNA which has not been fully processed. Rev is a 116 amino acid polypeptide and, like Tat, it is produced early in the viral life cycle (Figure 1.7).[21,102] Rev is comprised of three functional domains: (a) an arginine-rich motif (amino acids 34–50) that functions as

both an RNA binding domain and as a nuclear localization signal (NLS), (b) a leucine-rich nuclear export signal (NES, amino acids 75–83), that associates with the cellular export receptor protein CRM1, and (c) a multimerization domain which flanks the RNA binding domain (Figure 1.7).[21,102]

The Rev protein binds to a unique high affinity site on the RRE RNA with low nM affinity[103-106] and high selectivity. After this binding event, the Rev protein multimerizes on the entire 350 nt long RRE RNA. This process is facilitated with both protein-protein and RNA-protein interactions. Once bound to the RRE of a partially or fully unspliced viral RNA transcript, the complex is exported from the nucleus. The fact that the arginine rich motif of the Rev protein serves as both the RRE binding domain and the nuclear localization signal allows the protein to continuously shuttle back and forth between the cytoplasm and the nucleus in a defined process. After translation, the Rev protein enters into the nucleus, binds to the RRE RNA, multimerizes and exits the nucleus with RRE inclusive transcripts. Once in the cytoplasm, the Rev-RRE complex dissociates and the Rev, whose NLS is now exposed is free to re-enter the nucleus and export additional viral RNA transcripts.

The RRE RNA sequence presents an intriguing target for therapeutic intervention as this viral RNA sequence serves as both the high-affinity Rev binding site and as part of the open reading frame for the gp41 “fusion domain” of the *env* protein.[107] This dual role is most likely the cause for the unusually high rate of sequence conservation found at the primary Rev binding site on the RRE.[108] The nucleotides comprising this RNA sequence are almost invariant between genetically diverse groups of HIV isolates, making the primary Rev binding site on the RRE a

highly attractive target for small molecule therapeutics for two reasons. Small molecules that bind to, and inactivate, this site should possess consistent activities among diverse HIV groups. Additionally, because of the RRE's dual function, evolution of resistant variants may be prevented or impeded.

Aminoglycoside antibiotics, particularly neomycin and tobramycin have been shown to competitively inhibit the binding of the Rev protein to the RRE and interfere with the function of Rev in infected cells.[43] These data established a precedent for the use of aminoglycosides as an attractive antiviral strategy. Additionally, it has been shown that RRE variants that are highly neomycin B-resistant generally show decreased affinity for Rev.[109] This suggests that HIV mutants harboring a neomycin B-resistant RRE variant are also likely to display compromised replication.

To further evaluate the binding of aminoglycosides to the RRE, truncated RRE constructs have been developed based on the RRE IIB stem. 2-aminopurine (2-AP), a fluorescent adenosine isoster has replaced the A at position 68 or the U at position 72.[110] These substitutions allow the RRE RNA to retain its ability to bind Rev while also acting as a reporter for ligand binding. By measuring fluorescence changes upon neomycin titrations, a high affinity neomycin binding site (K_d ~250 nM) and one or more low affinity (high μM to low mM) sites were found. Competition experiments where a complex of neomycin and RRE-72(2-AP) was titrated with the Rev peptide suggested that neomycin occupation of the high affinity site is not disruptive to Rev binding.[110] However, studies with the larger, naturally-occurring RRE construct suggest that for this RNA construct, binding of neomycin to its highest affinity binding site does competitively inhibit Rev binding.

Truncations made to the native RRE sequence likely alter the neomycin binding site. In the larger RRE66, footprinting experiments indicate that neomycin binds near the three-way junction of stem IIA, IIB and IIC.[43] When the RRE is truncated by the removal of stems IIA and IIC, this region is dramatically altered. This minimization appears to affect the mechanism by which neomycin displaces Rev from the RRE.

To facilitate the discovery of potent RRE binding small molecules, compounds which show both improved affinity and selectivity, a novel solid-phase assay has been developed.[111] The components of this assay include insoluble agarose beads covalently modified with streptavidin, a biotinylated RRE fragment and a fluorescein-labeled Rev peptide (RevFl). Assembly of the three components generates an immobilized ternary complex where the biotinylated RRE binds to the beaded streptavidin and the RevFl binds to the RRE with the same affinity as measured in solution by fluorescence anisotropy ($K_d \cong 3$ nM).[112] To determine the inhibitory potency of RRE binders, the immobilized Rev-RRE complex is challenged with increasing ligand concentrations. Measuring the displacement of RevFl generates information regarding ligand affinity. This assay can be performed in the presence of large amounts of competing nucleic acids, allowing assessment of the selectivity traits of RRE binders.[111]

Additionally, RRE-RevFl association and dissociation can be followed in solution.[112] Use of these assays has lead to the discovery and study of several new families of RRE binders including aminoglycoside dimers and conjugates,[113] cyclized peptides,[114] and octahedral Ru^{II} complexes.[115]

Many approaches have been explored in attempts to increase the affinity and selectivity of aminoglycoside-based ligands to the RRE. Some neamine derivatives

with numerous peptide conjugates bound to the 5 position of the 2-DOS ring show enhanced affinity compared to the core aminoglycoside.[116] Neamine has also been modified at the 6′ primary aminomethyl group with various aromatic acids.[117,118] More potent and selective RRE binders have been found by conjugating chloramphenicol and linezolid to neomycin via a hydroxyl residue on the D-ribose moiety.[85] Some neomycin-dipeptide conjugates have also been shown to have both improved RRE affinity and selectivity.[119]

Analysis of the NMR structure of the arginine-rich RNA binding domain or Rev complexed to a short RRE RNA construct reveals purine-purine base pairing, as well as a bulged out pyrimidine residue.[120] Non-canonical motifs with extended base pairing regions may constitute favored intercalation sites.[121] A neomycin-acridine conjugate (neo-S-acridine) shows a marked increase in affinity. Its K_i of 3 nM is approximately the same affinity as the Rev peptide.[122] Footprinting experiments have established that neo-S-acridine competitively inhibits Rev-RRE complex formation by binding the RRE in the same region as the arginine-rich Rev_{34-50} peptide.[122]

Neo-S-acridine displays mediocre RRE selectivity, binding to competing nucleic acids including tRNA and DNA.[113] However, the selectivity of the aminoglycoside-intercalator conjugates can be accomplished by varying the linker length. Neo-C-acridine, which has a longer linker displays lower RRE selectivity while neo-N-acridine, the ligand with the shortest of the three linkers, exhibits improved RRE selectivity. Compared to the neomycin-acridine conjugates, tobramycin-acridine and kanamycin A-acridine conjugates have slightly lower RRE affinity but improved specificity for the RRE over other RNA targets.[113]

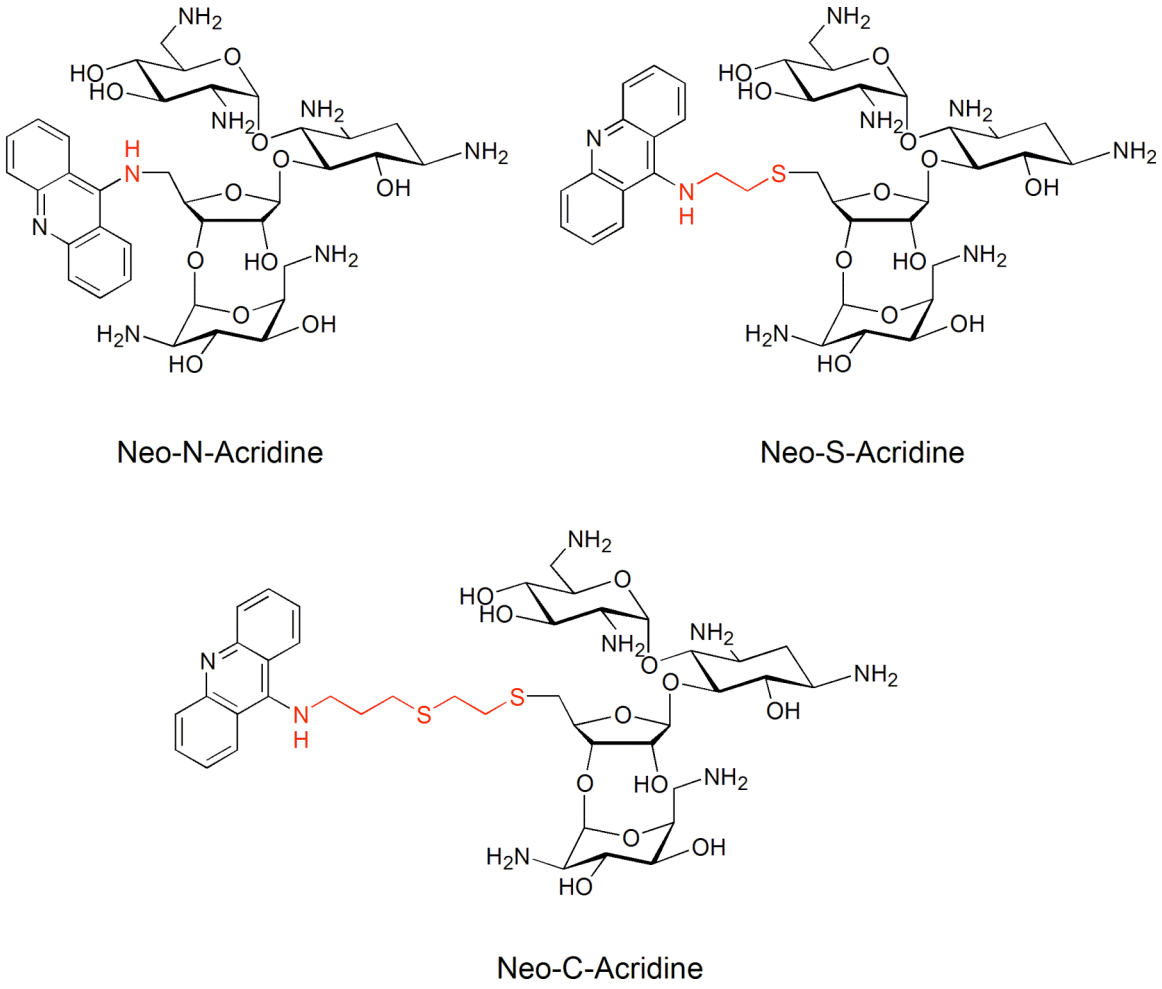


Figure 1.8. Structures of the three Neo-acridine conjugates. All three compounds show similar affinity for the RRE RNA. However, linker length plays a large role in RRE selectivity with the neo-N-acridine having displaying the highest selectivity and neo-C-acridine the lowest.

Structural modifications and conjugation of functionalities can improve RRE selectivity to a degree, but additional RNA-selective scaffolds are essential for developing RRE selective ligands. Numerous proteins utilize guanidinium groups, typically within arginine rich domains, for RNA recognition.[123,124] Guanidinium groups are highly basic and planar, capable of both stacking and directed H-bonding interactions. A new family of RRE ligands, where all of the ammonium groups of the

natural aminoglycosides have been converted into guanidinium groups has been synthesized.[125,126] This new family, termed "guanidinoglycosides" display a substantial increase in RRE affinity when compared to their parent aminoglycosides.[125] RRE selectivity could also be improved, although this was determined by the total number of amine/guanidine residues on the molecule. Guanidino-paromomycin which has 5 basic groups increases its RRE selectivity while guanidino-neomycin, which has 6 basic groups, shows diminished selectivity.

While a good balance between ligand affinity and selectivity to the RRE RNA may be difficult to achieve, it remains an important target for future antiretroviral agents. Additionally, the strategy of using guanidinoglycosides scaffolds may prove invaluable as they display increased RNA selectivity. By mimicking the binding site of the natural ligand for the RRE, the arginine (guanidine)-rich region of the Rev peptide and adding other functionalities, it may be possible to design ligands which display both high affinity and selectivity to the RRE RNA.

1.7 Viral packaging and RNA dimerization

Once the viral RNA has been translated and the large polyprotein has been processed to form individual viral proteins, new viral particles begin to form. Packaged into each virion are two copies of fully unspliced viral RNA. To be packaged into the virion, the RNA must dimerize in a highly orchestrated process involving a self-complementary RNA stem loop interaction. HIV specific RNA events such as the

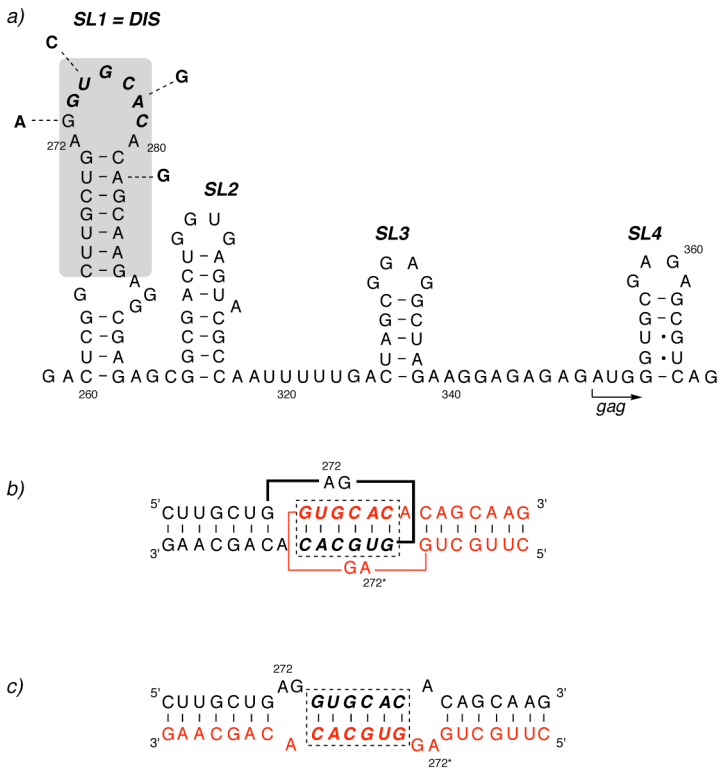


Figure 1.9. a) secondary structure of the HIV-1 encapsidation region (ψ-site) including the four stem loops. The SL1 or DIS dimerizes in first a kissing, then extended loop structure, initiating full viral genomic dimerization. b) secondary structure of the 23-nt long kissing loop complex which extends into c) the rearranged extended duplex.

initiation of viral dimerization may provide potential targets for future therapeutic intervention.

In mature viral particles, the viral genome is non-covalently dimerized near its 5′-end.[127] The initial events leading to dimerization have been confined to stem loop 1 (SL1) which is also known as the dimerization initiation site (DIS). The DIS is the first of four stem loops found on the *ca.* 120 nt long HIV ψ-packaging site.[23] This stem loop RNA structure initiates dimerization by forming a "kissing loop" structure where a self complementary sequence of six nucleotides from each loop associates via base pairing.

Once formed, this kissing loop can rearrange itself into a more stable extended duplex,[128] in a process which is facilitated by the nucleocapsid protein (NCp).[129]

RNA dimerization and packaging take place in the late stages of the viral particle, but the ability to disrupt the assembly of mature viral particles should constitute a viable antiretroviral therapeutic approach. Changes in the DIS have been shown to dramatically diminish viral infectivity,[130,131] giving small molecules that modulate this process potential to exhibit activity across diverse strains and show therapeutic potential.

HIV viral dimerization has been extensively studies by structural, biophysical and biochemical approaches.[132-134] However, much less has been reported regarding the binding of low molecular weight ligands to the viral RNA sequence. It has been shown that aminoglycosides bind to the packaging region of HIV-1.[135-137] Footprinting experiments suggest that the primary binding site for neomycin is found on SL1. Binding at this, and additional secondary sites, induces conformational changes in the RNA. The DIS, found on SL1, warrants more attention for aminoglycoside based ligands. A structural analysis[138] of this RNA shows that it bears a remarkable structural resemblance to the prokaryotic ribosomal A-site, the cognate target of aminoglycoside antibiotics. This similarity extends beyond the primary, secondary and tertiary structures, but also in the ability of the DIS to bind certain aminoglycosides,[139] leading to the possibility that the design and synthesis of modified aminoglycosides which can distinguish between the A-site and the DIS may show therapeutic potential.

1.8 RNA Binding Molecules Within the Cell

With the exception of viral particle fusion with the host cell and budding of new virions, every step of the HIV life cycle takes place within the context of the cell. This fact can not be overlooked when evaluating molecules designed to target specific RNA interactions.

For an RNA targeted molecule to be therapeutically relevant, it must meet certain criteria. Molecules must show *in vitro* activity, binding to the RNA target and inhibiting the RNA–RNA or RNA–protein complex formation targeted with reasonable affinity. Additionally, molecules must show selectivity for a specific RNA target. Compounds which bind to multiple different RNA (or DNA) molecules with comparable affinity to the target RNA are not attractive potential therapeutics. One other issue which must be addressed when evaluating any potential therapeutic agent is that of delivery. A molecule may show excellent *in vitro* activity in terms of affinity for its target RNA, activity in disrupting the RNA–RNA or RNA–protein complex formation and selectivity between different RNA targets, but not be active *in vivo* because it is never able to enter the cell to reach its target.

Delivery is an issue which can not be ignored when evaluating the therapeutic potential of any anti-HIV targeted molecule. A compound must be able to cross the cell membrane and enter into the cell, but additionally this compound must be delivered in such a way that the target RNA is accessible. Molecules which are sequestered in lysosomal vesicles may be in the cell, yet show little to no biological activity due to the fact that they can not get close enough to their target to act.

A compound can enter the cell in one of two main ways. It may possess its own inherent uptake activity, crossing cell membranes and localizing within the cytoplasm of the cell, or it can be linked to a delivery agent which will facilitate cellular uptake and cytoplasmic accumulation. One such class of delivery agents includes arginine-rich proteins/peptides and other oligomers with a high density of guanidinium groups.

1.9 Cellular Uptake and the Guanidinum Group

The HIV-1 Tat is an 86-amino acid protein which binds to the trans-activating response element (TAR), a segment of viral RNA, enhancing viral transcription.[16] Integrated viral DNA is transcribed and processed in the nucleus and then exported into the cytoplasm where it is translated and the Tat protein is formed. To bind to the TAR RNA, Tat must cross the nuclear membrane and enter the nucleus. In 1988 it was discovered that the HIV-1 Tat protein can cross lipid bi-layers [66,67] and enter the nucleus. This provided one of the first examples of a protein having the ability to cross a cellular membrane.

Since then, it has been shown that many other poly-cationic proteins, from multiple organisms, have the ability to translocate through cell membranes and accumulate in the nucleus. Proteins with multiple arginine groups are most efficiently internalized, but there appears to be no specific receptor since the proteins have little or no sequence homology other than the presence of 5-11 arginine residues.[140] CD spectra is also suggestive that there is a lack of secondary structure homology as well. The

efficiency of incorporation appears to correspond only to the number of arginine residues.[141]

Arginine rich peptides can also serve as carrier molecules. When complexed to large molecules, such as carbonic anhydrase (29 kDa), they are able to deliver these molecules to the nucleus. Carbonic anhydrase without a carrier protein is only taken into the cell by endocytosis in a limited manner.[141] Arginine rich peptides have also been used as carriers for other large molecules including β-galactosidase (120 kDa), DNA oligonucleotides, magnet beads (>50 nm diameter), chelators with radioisotopes, liposomes (>200 nm diameter), other proteins, and fluorescent tags.[142-147]

In fact, peptides containing multiple arginine residues need not even be linear. Branched chain arginine peptides also show robust translocation through cell membranes.[148] Cellular uptake of branched chain peptides is dependent only on the total number of arginine residues. These peptides are internalized as fast as the HIV-1 Tat protein, and just like other polyarginine peptides, they are able to carry relatively large molecules across the cell membrane.

A peptide backbone is not necessary for efficient cellular uptake. Polyarginines with oligocarbamate backbones have been shown to have enhanced uptake efficiency when compared with the corresponding peptides (Figure 1.10). These oligocarbamates are also able to carry biotin not only across cellular membranes but across multiple epidermal layers where it then shows a high degree of nuclear localization.[149]

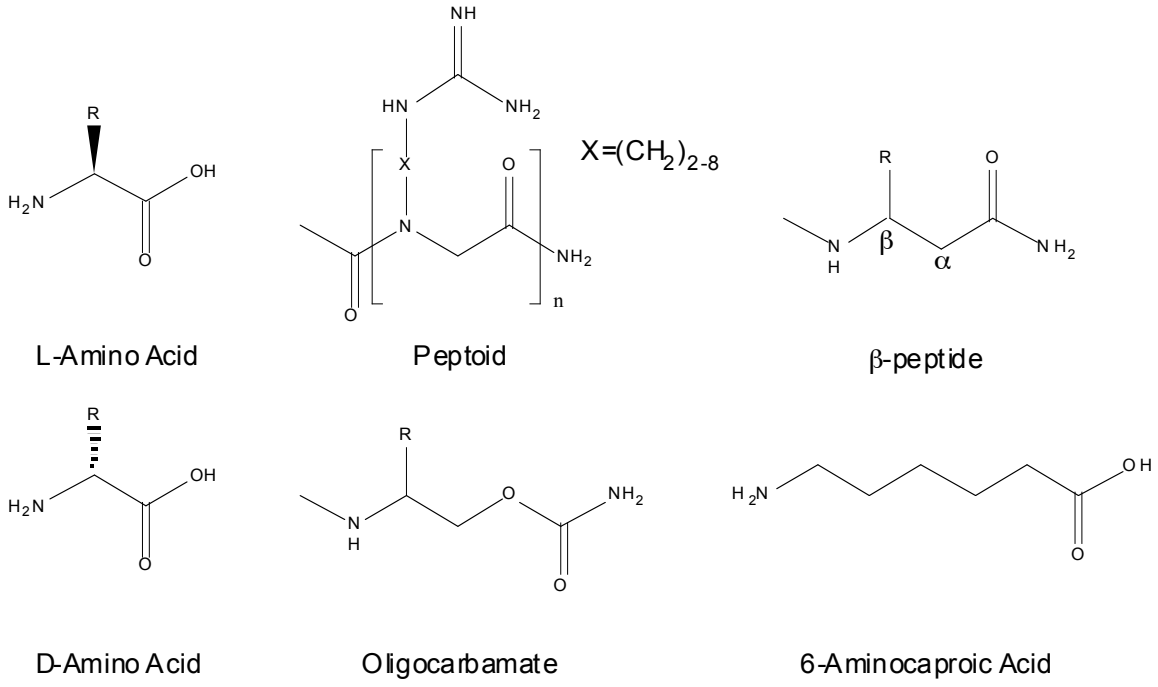


Figure 1.10. Examples of different backbones used with multiple guanidinium group-containing polymers. Guanidinium group polymers with multiple different linear and branched scaffolds can efficiently enter bring large cargos across membranes and into the cell.

Many examples of different backbones for arginine-based polymers have been used including L and D-amino acids, oligocarbamates, peptoids, β-peptides and 6-aminocaproic acids (Figure 1.10).[149-154] All of these backbones were sufficient to allow the molecules to effectively translocate across cellular membranes. The only factor that these compounds have in common is the presence of multiple guanidinium groups.

Further evidence suggesting that guanidium density is all that is needed for efficient cellular uptake was demonstrated when a fluorescent tag (BODIPY) was conjugated to the amino and guanidino forms of both tobramycin and neomycin (Figure

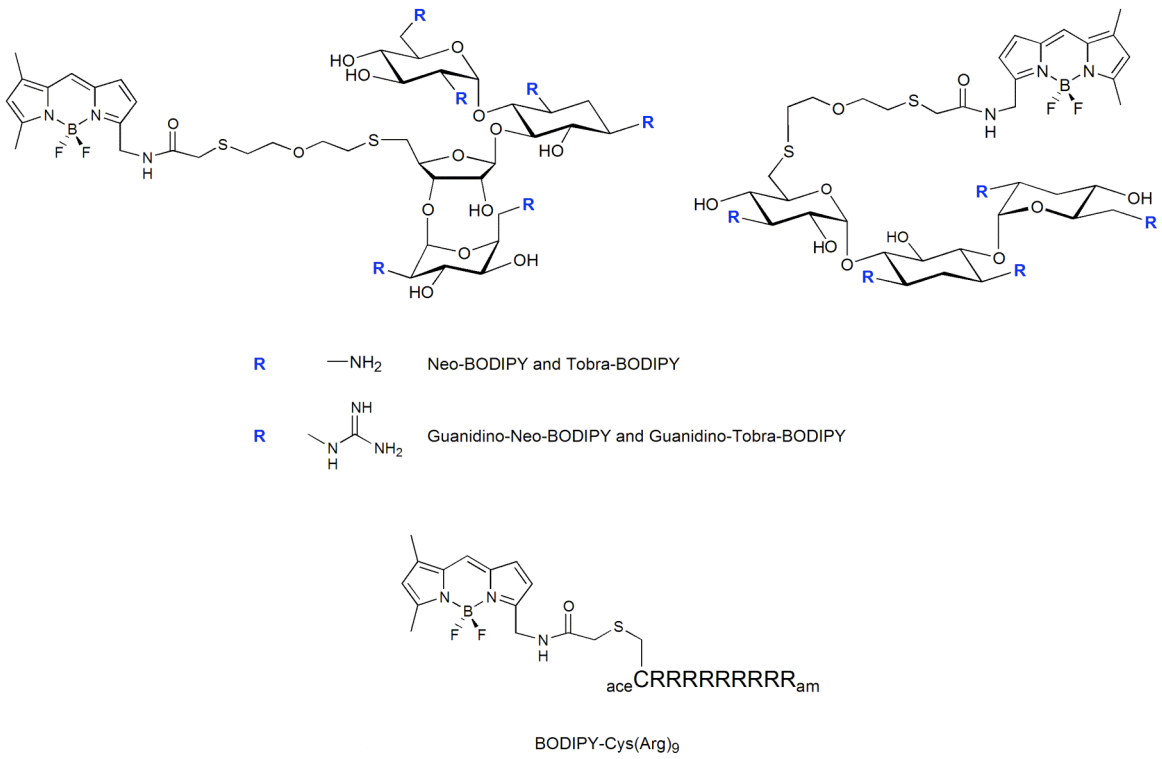


Figure 1.11. Neomycin and Tobramycin in both amino and guanidino forms have been conjugated to a BODIPY fluorophore. The guanidinylated derivates show enhanced cellular uptake when compared to the amino forms of each molecule. An Arg_9 peptide was also conjugated to the fluorophore. This peptide shows efficient uptake which can be inhibited in the presence of guanidino-neomycin, suggesting a common mechanism of action between arginine-rich peptides and guanidinoglycosides.

1.11).[155] Guanidinylated neomycin-BODIPY and guanidinylated tobramycin-BODIPY show a marked increase in cellular uptake as compared to neomycin-BODIPY and tobramycin-BODIPY. In addition, uptake of BODIPY-Arg_9 (Figure 1.11) is inhibited by guanidinylated-neomycin, showing evidence that guanidinoglycosides and arginine rich peptides are likely to cross the cell membrane through a common mechanism. While elucidation of the uptake mechanism of poly-arginine peptides is not complete, studies have begun to shed some light on the mechanism of action.

The positive charge on the guanidinium group is not sufficient for uptake. When the guanidinium group on the arginines is methylated, there is a marked decrease in uptake, and when it is doubly methylated, uptake is completely inhibited.[156] These mono and di-methylated compounds retain their charge, but lose their hydrogen-bonding capabilities and ability to form salt bridges.

When a fluorescein labeled Arg_8 is dissolved in water and put into a water-octanol bi-layer, it remains in the aqueous layer indefinitely, but when sodium laurate is added, forming a lipid bi-layer between the aqueous and organic layers, the Arg_8 relocalizes to the octanol. This indicates that the interaction between the peptide and the lipid bi-layer is required for translocation. Additionally, when the K^+ Nernst potential, the electrical potential/voltage difference across a cell membrane, is lowered with the addition of K^+ ions, there is a linear decrease in the uptake of both Tat and Arg_8. Cellular uptake of Arg_8 is greatly reduced when cells are pre-treated with gramicidin A. Gramicidin A forms pores in cell membranes and is known to reduce membrane potential. Conversely, Arg_8 uptake is enhanced when cells are pre-treated with valinomycin, an antibiotic that selectively shuttles K^+ ions across membranes, increasing membrane potential.

These observations lead to an uptake model where water-soluble, positively charged guanidinium groups form hydrogen bonds with the polar phospholipids head groups on the cellular surface. The resultant ion pair partitions into the lipid bi-layer, and then migrates across the membrane at a rate proportional to membrane potential. Once across the membrane, the complex dissociates, releasing the guanidinium containing compound into the cytosol.[156]

Other models have arginine rich peptides binding to cell surface heparan sulfate proteoglycans and being taken into the cell through heparan sulfate mediated endocytosis.[157] Once in a vesicle, inside of the cell, heparan sulfate is degraded by heparanases, increasing the effective peptide concentration. Peptides not bound to heparan sulfate then disrupt the lipid bi-layer of the vesicle promoting leakage.[157]

The uptake of Arg_9 was examined in both living and fixed CHO-K1 (Chinese Hamster Ovarian wild type) cells to determine if this has an effect on intracellular localization, as was shown with the HIV Tat peptide.[158] In living cells, Arg_9 was found almost exclusively in vesicles within the cell, as evidenced by co-staining with FM-143 (a marker for endocytotic vesicles). However, when the cells were fixed with paraformaldehyde (4% w/v) or acetone/methanol, the labeled Arg_9 was found to be diffused throughout the cytoplasm. Fixed cells do not retain any biological activity, and the process of fixation can create holes in the cell membrane, making any internalization seen when compounds were added to fixed cells likely a non-specific process with little biological relevance.

Polyarginine peptides have a high affinity for heparin. Heparin is a natural glycosaminoglycan similar to heparan sulfate, consisting of the same repeating disaccharides, but with a higher level of sulfation.[159] Cells lacking heparan sulfate have shown a remarkably reduced internalization of Arg_9 when compared with wild type cells. An overlay of the fluorescence of tetramethylrhodamine labeled Arg_9 in cells, FM-143 and Hoescht 33342 (a cell permeable nuclear stain) shows the peptide in vesicles and not in the nucleus.[157]

CHO cells lacking heparan sulfate (CHO-pgsD-677) or heparan sulfate and chondroitin sulfate (CHO-pgsA-745) showed remarkably decreased uptake of Arg_9 when compared to wild type CHO-K1 cells. The fact that there is virtually no uptake of Arg_9 by GAG-deficient cell lines indicates that interaction of Arg_9 and cell surface glycosaminoglycans is essential for internalization. Arg_9 shows a very high affinity for heparin, eluting from immobilized heparin at approximately 1.5 M NaCl.[157]

As previously mentioned, guanidinoglycosides show a high degree of cellular uptake in multiple different cell lines. Guanidino-tobramycin which has only five guanidinium groups has a very similar transport efficiency to Arg_9 even though it has four less guanidinium groups. Guanidino-neomycin, with six guanidinium groups, shows slightly better cellular uptake than Arg_9.[155] This indicates that the guanidine-density rather than the number of guanidinium groups may be responsible for efficient uptake of guanidine-rich molecules. Additionally, because uptake of Arg_9 is significantly inhibited in the presence of guanidino-neomycin, there may be a possible common mechanism of uptake.

Bioavailability of pharmaceutical agents presents a major problem for drug development. Molecules which show high activity *in vitro* may not be able to cross cell membranes and thus could potentially have very low *in vivo* activity. This highlights a need for relatively small carrier molecules which can efficiently cross cell membranes and accumulate in living cells.

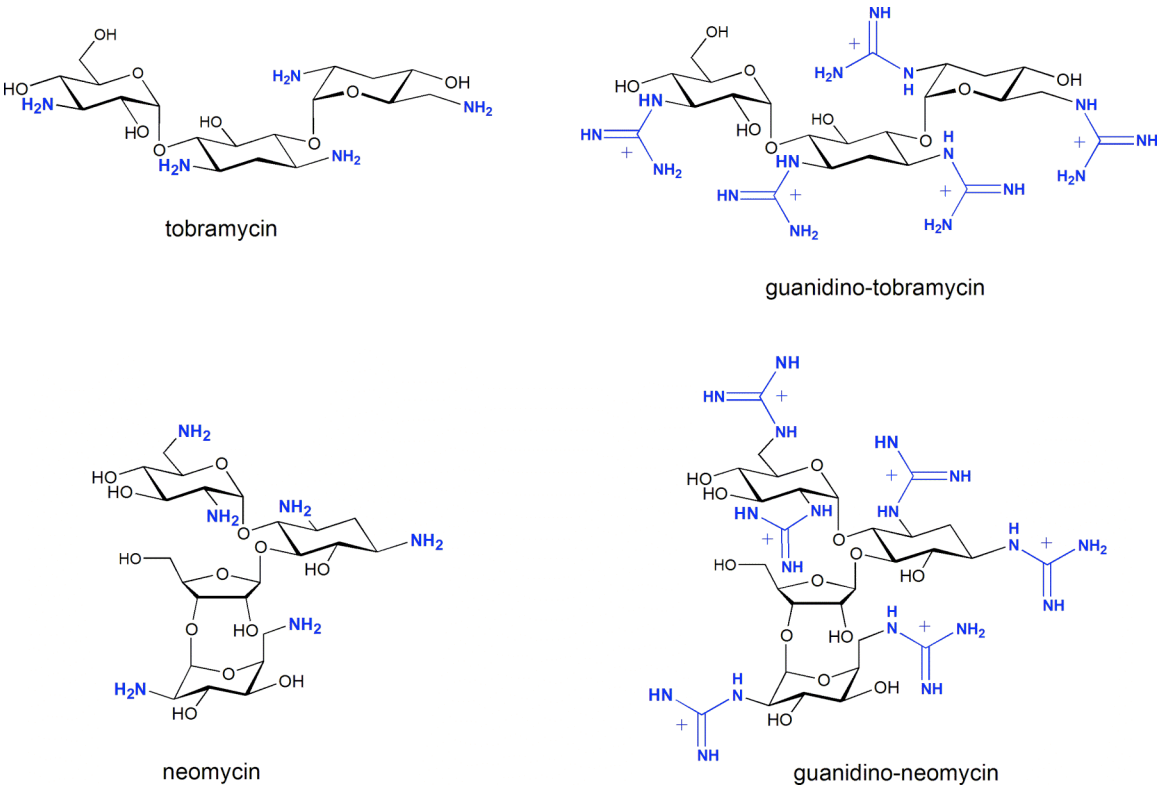


Figure 1.12. Tobramycin and neomycin show little to no cellular uptake, but when guanidinylated, they are able to efficiently cross cell membranes and localize within the cell. Guanidino-neomycin which contains six guanidinium groups shows a higher degree on internalization than guanidino-tobramycin which only contains five guanidinium groups.

Guanidinoglycosides (Figure 1.12) may prove to be ideal for drug delivery as they are small, relatively simple to produce and easily functionalized. A large variety of cleavable linkers and therapeutics can be covalently attached to guanidino-neomycin via the highly accessible 5′ position of ring III through known chemistry,[160] allowing a large number of possible pro-drug candidates to be quickly produced and tested.

Additionally, the high affinity of guanidinoglycosides for RNA, and specifically HIV RNA may afford further advantages. Subsequent to uptake and release from vesicles into the cytoplasm, guanidinoglycosides can preferentially bind to viral

RNA.[161,162] Attached therapeutics, designed to target viral RNA or enzymes, will have the advantage of being linked to a molecule with high affinity for viral RNA. This may allow them to be in close proximity to the molecules which they were designed to target. The extremely efficient uptake and high RNA affinity of guanidinoglycosides highlights their potential as highly active drug delivery agents.

While this potential is beginning to be explored, it is critical to understand the mechanism of uptake. Evidence suggests that uptake is a receptor mediated process where arginine rich proteins and guanidinoglycosides bind to cell surface proteoglycans including those containing heparan sulfate and are taken up through endocytosis.[155,157] Understanding of this receptor-guanidinoglycoside interaction is vital for the design of future pharmaceutically active compounds.

A detailed understanding the biology and biosynthesis of heparan sulfate proteoglycans is a critical step in the comprehension of possible interactions which may take place between small molecules such as guanidinoglycosides and these cell surface glycosaminoglycans.

1.10 Heparan sulfate Proteoglycans

Heparan sulfate appeared early in metazoan evolution, and because of this, many of the structural motifs which characterize heparan sulfate have been preserved in modern organisms. Many of the biological functions which are associated with heparan sulfate also occurred early in evolution, or have depended on the evolution of heparan sulfate

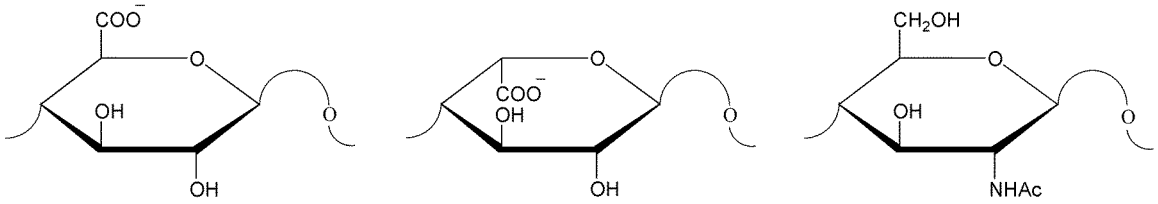


Figure 1.13. Heparan sulfate is made up of alternating uronic acid or L-iduronic acid and D-glucosamine residues which are modified in a variety of ways, giving heparan sulfate a rich diversity of possible structures.

binding ligands. There are hundreds of known heparin-binding proteins, many whose interactions have profound physiological consequences, and almost every known cell has the capacity to produce heparan sulfate, making it a universal and extremely significant receptor.[163]

Heparan sulfate polysaccharides represent a glycosaminoglycan which is characterized by alternating uronic acid (GlcA) or L-iduronic acid (IdoA) and D-Glucosamine (GlcN) subunits (Figure 1.13).[164] These alternating subunits undergo a highly orchestrated series of modifications, allowing for a rich variety of specific structures.

The heparan sulfate proteoglycan superfamily is formed by heparan sulfate chains and their covalent complexes with specific core proteins. Within this composite glycoprotein superfamily are two major subfamilies, the syndecans and glypicans.[165]

The four syndecans (syndecan-1, -2, -3, and -4) which make up the syndecan subfamily have protein cores with characteristic structural domains (Figure 1.14). Syndecans all have a variable ectodomain which is exposed to the extracellular

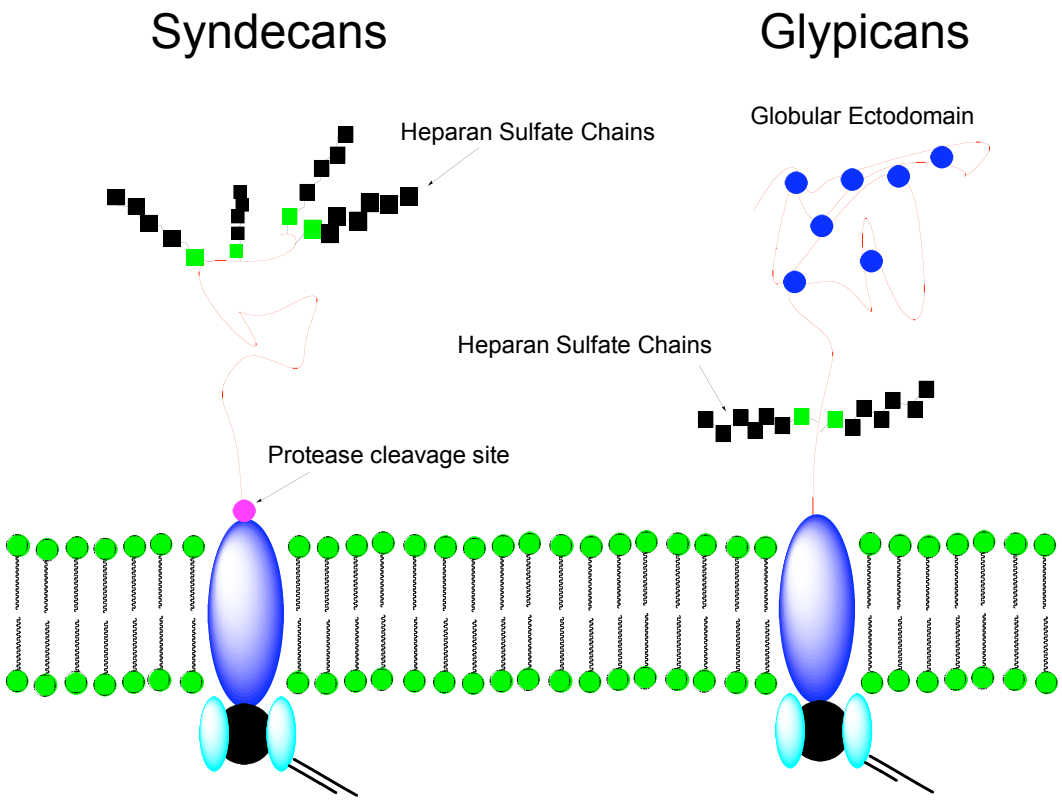


Figure 1.14. The syndecans and glypicans are the two major subfamilies of the heparan sulfate proteoglycan superfamily. Syndecans and glypicans both have a hydrophobic transmembrane segment which tethers them to the cell and a cytoplasmic domain with substrates for kinase activity. Syndecans, however have a protease cleavage site which allows for release of the ectodomain and attach heparan sulfate chains. Glypicans are characterized by a globular ectodomain with multiple cysteine residues.

environment, and contains between three and five heparan sulfate chains in conserved motifs. Some tissues express hybrid syndecans with both heparin and chondroitin sulfate bound to the ectodomain. This ectodomain is tethered to plasma membrane via a short, hydrophobic transmembrane segment.[166]

Syndecans contain a protease cleavage site located next to the junction where the ectodomain meets the transmembrane segment. Proteolysis results in the release of the ectodomain and attached heparan sulfate chains from the cell surface.[167]

Consequently, anything attached to the heparan sulfate chains would be shed from the cell surface as well. Membrane bound syndecans can also be internalized via endocytosis and subsequently degraded by lysosomes, suggesting that the membrane or cytoplasmic domain can interact with proteins involved in endocytosis.[167] When the cytoplasmic domain is endocytoticly taken up by the cell, any molecules bound to the heparan sulfate chains will be taken up into vesicles as well. The cytoplasmic domain also binds to cytoskeletal proteins and contains sequences which serve as substrates for cellular kinases.[168,169] This points to a role for syndecans as signaling molecules, and suggests that they may play a large role in regulation of cellular functions.

The six members of the glypican subfamily can be distinguished by their cysteine-rich globular ectodomains and the presence of two or three heparan sulfate chains (Figure 1.14). These chains are attached between the globular ectodomain and the proteins which anchor it to the outer membrane.[170] While there is no known release mechanism for the glypicans, they can be taken up via endocytosis and degraded by lysosomes in a very similar manner to the syndecans.[171]

The individual syndecans, glypicans and other subfamilies of the heparan sulfate proteoglycan superfamily are expressed in a tissue-specific manner. Because ligand binding depends on the specific fine structure and composition of heparan sulfate chains on the proteoglycans, this allows for tissue (and sometimes cell) specific binding of extracellular molecules.[172]

Heparan sulfate proteoglycans can be analyzed in terms of disaccharide composition. The same set of disaccharides exists in most tissues, but the content can

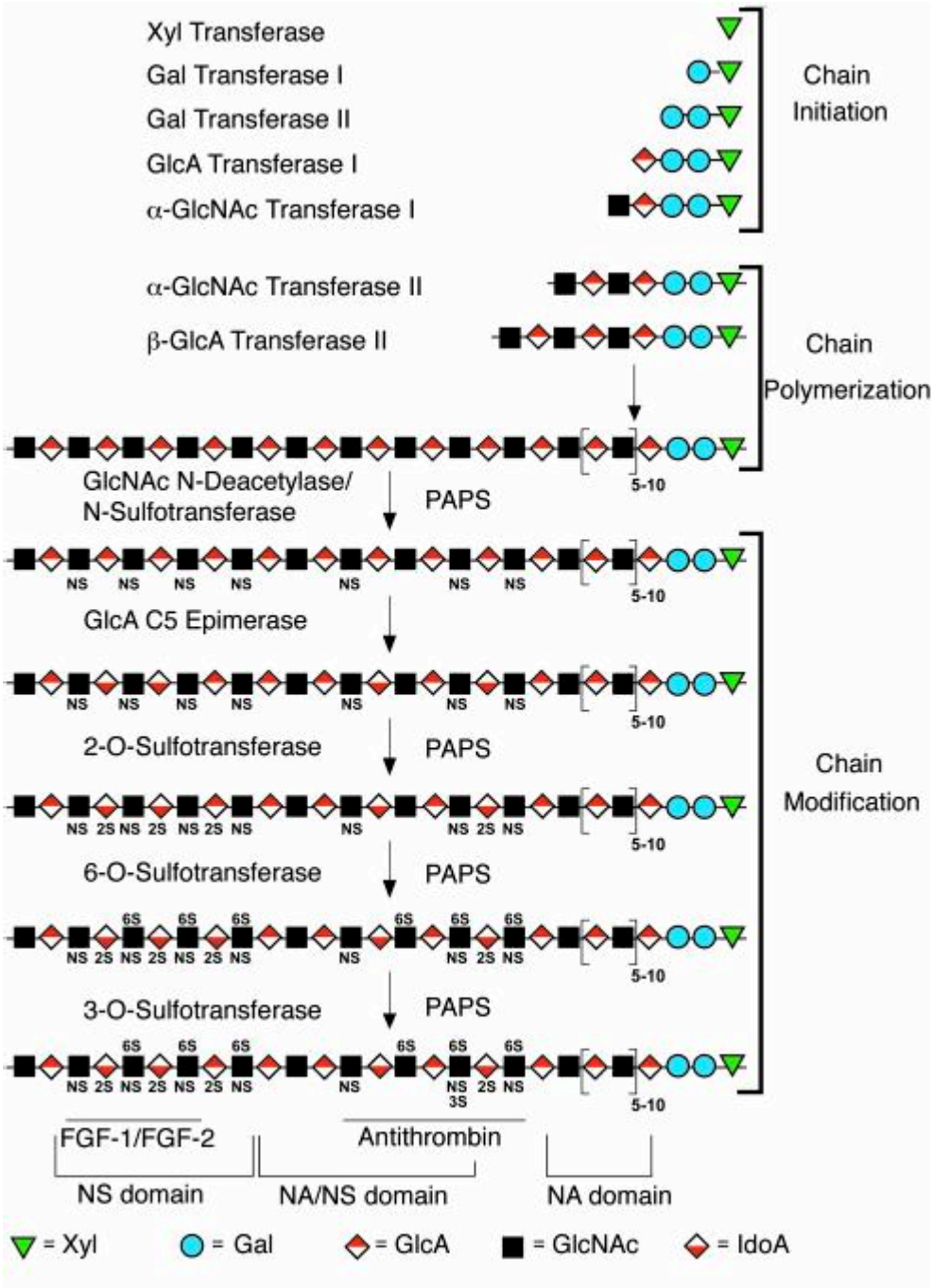


Figure 1.15. Scheme of heparan sulfate chain biosynthesis. A highly orchestrated process involving multiple enzymes allows for varied domains, each of which can have vastly different affinities for different targets. While multiple enzymes are responsible for chain modification, and the sequences of enzyme activities are not fully elucidated, specific domains with individual sulfation patterns are created. (figure modified from reference[163])

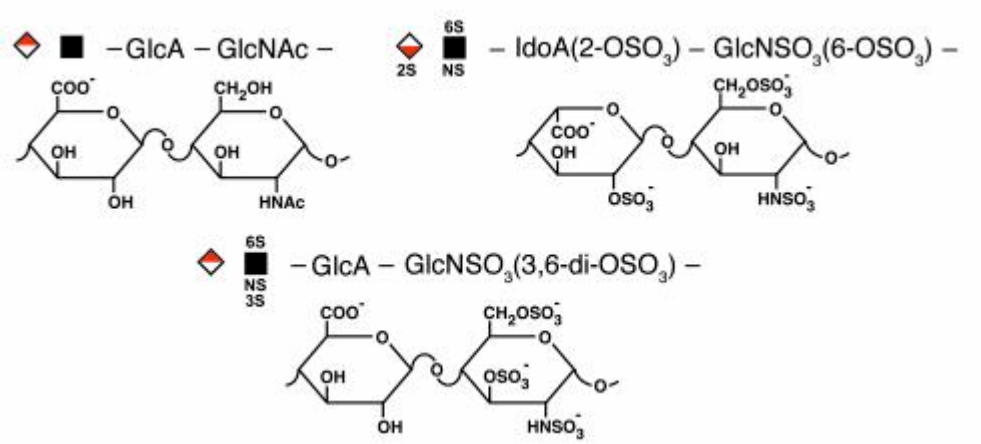


Figure 1.16 Disaccharides with varying degrees of sulfation are incorporated into the heparan sulfate chain. Deacetylation, epimerization and sulfation at each position is carried out by a specific enzyme. This allows for a wide range of sequences, corresponding to specific ligand binding domains. (figure modified from reference 163)

vary by a large amount. Disaccharide composition and their arrangement within the different domains are not enough to define ligand binding sites. Binding occurs only to specific sets of variably modified disaccharides.[173] Many enzymes in both the golgi apparatus and cytoplasm are required for heparan sulfate biosynthesis as well as for the fine structure of the proteoglycans.[174] Other enzymes exist which are essential for transport of both intermediates and fully formed heparan sulfate chains.[175]

Biosynthesis of the heparan sulfate chain is initiated through assembly of a linkage tetrasaccharide (Figure 1.15). The first sugar of the tetrasaccharide, xylose, is *O*-linked to a serine on specific sites of the core polypeptide defined by Ser-Gly residues flanked by one or more acidic residues.[176] Once the xylose is bound to a specific serine residue, GlcAβ1, 3Galβ1, 3Galβ1 and 4Xyl β1 are added sequentially until the tetrasaccharide is complete.[177]

At this point, a single α–1,4-linked GlcNAc unit is then added to the chain (Figure 1.16). This commits the intermediate to the assembly of heparan sulfate. If a β-1,4-linked GalNAc is added instead, this commits the growing chain to chondroitin sulfate.[176] After this initiation, the growing chain begins to polymerize with the addition of alternating GlcAβ1,4 and GlcNAc α1,4 residues.[178] During polymerization the growing chain undergoes a series of modifications, including *N*-deacetylation, *N*-sulfation, epimerization and *O*-sulfation. The concerted actions of the modifying enzymes results in the formation and organization of specific domains on the heparan sulfate chains. The genes of these enzymes all show tissue specific and developmentally regulated expression.[179]

The great variability and high cell type specificity of heparan sulfate proteoglycans enhances its appeal as a target for potential therapeutics. These proteoglycans appear on the surface of virtually all known cells and interact specifically with literally hundreds of proteins and other ligands. This raises the question of whether it is possible to design ligands to interact with heparan sulfate proteoglycans on specific cell types. The specificity of this interaction and subsequent uptake could allow for highly selective drug delivery systems.

1.11 Guanidinoglycosides as Heparan Sulfate and Viral RNA binders

Cell surface heparan sulfate and RNA are two examples of polyanionic oligomers which exist within the context of the cell. The repeating saccharides which make up heparan sulfate have tracts of highly sulfated regions where each sulfate group

contributes a negative charge to the molecule. The backbone of RNA is made up of negatively charged phosphodiester groups.

In this dissertation, we will show that guanidinoglycosides bind to both cell surface heparan sulfate proteoglycans and to RNA. However, we will show that this binding is reasonably selective in manner. Guanidinoglycosides bind to heparan sulfate but not to chondroitin sulfate, a similar sulfated proteoglycan. When conjugated to an intercalating moiety, they will bind to RNA with a much higher affinity than DNA.

In this dissertation, we will show that guanidinoglycosides possess significant potential as future delivery agents. Guanidino-neomycin is able to effectively internalize within living cells in a heparan sulfate dependent manner. Once taken up into the cell via heparan sulfate proteoglycan mediated endocytosis, guanidinoglycosides are released into the cytoplasm. We will show that guanidinoglycosides can facilitate the delivery of large, bioactive cargos into the cell and that this delivery is also dependent on heparan sulfate. Once it has been internalized, these compounds are then released into the cytoplasm.

Additionally, we will show that aminoglycosides and guanidinoglycosides which have been covalently linked to an ethidium based moiety bind with high affinity to the RRE RNA. These molecules show a strong preference for the RNA over DNA and are selective for the RRE over other RNA.

This highlights the potential of guanidinoglycosides as both potential delivery agents and building blocks for anti-HIV therapeutics. Through relatively simple chemistry, guanidinoglycosides can be bound to therapeutic agents for delivery or small molecules to enhance their RNA selectivity.

1.12 References

(1) http://www.unaids.org/epi/2005/doc/docs/PR_EpiUpdate_Nov05_en.pdf.

(2) http://www.unaids.org/bangkok2004/GAR2004_html/ExecSummary_en/Execsumm_en.pdf.

(3) Mitsuya, H.; Yachoan, R.; Broder, S. *Science* **1990**, *249*, 1533-1544.

(4) Emini, E.A. (Ed) *"The Humman Immunodeficiency Virus: Biology, Immunology and Therapy"* **2002**, Princeton University Press, Princeton, N.J.

(5) Moore, R.D.; Chaisson, R.E. *AIDS* **1999**, *13*, 1933-1942.

(6) Baliant, G.A. *Pharm. Therap.* **2001**, *89*, 17-27.

(7) Macchi, B.; Mastino, A. *Pharmacol. Res.* **2002**, *46*, 473-482.

(8) Claessens, Y.-E.; Chiche, J.-D.; Mira, J.-P.; Carious, A. *Critical Care* **2003**, *7*, 226-232.

(9) Garica-Lerma, J.C.; Heneine, W. *J. Clin. Virol.* **2001**, *21*, 197-212.

(10) Chen, R.; Quinones-Mateu, M.E.; Masnky, L.M. *Curr. Pharm. Design* **2004**, *10*, 4065-4070.

(11) Check, E. *Nature* **2003**, *424*, 361.

(12) Gulick, R.M. *Clin. Microbiol. Infect.* **2003**, *9*, 186-196.

(13) Hermann, T. *Angew. Chem. Int. Ed.* **2000**, *39*, 1891-1905.

(14) Tor, Y. *ChemBioChem* **2003**, *4*, 998-1007.

(15) Vaishnav, Y.N.; Wong-Staal, F. *Science* **1998**, *280*, 1880-1884.

(16) Frankel, A. D.; Young, J.A. *Annu. Rev. Biochem.* **1998**, *67*, 1-25.

(17) Rosen, C.A.; Pavlakis, G.N. *AIDS* **1990**, *4*, 499-509.

(18) Cullen, B.R. *Microbiol. Rev.* **1992**, *56*, 375-394.

(19) Karn, J. *J. Mol. Biol.* **1999**, *293*, 235-254.

(20) Rana, T.M.; Jeang, K.-T. *Archiv. Biochem. Biophys.* **1999**, *365*, 175-185.

(21) Pollard, V.W.; Malim, M.H. *Ann. Rev. Microbiol.* **1998**, *52*, 491-532.

(22) Hope, T.J. *Archiv. Briochem. Biophys.* **1999**, *365*, 186-191.

(23) Berkhout, B.; van Wamel, J.L. *J. Virol.* **1996**, *70*, 6723-6732.

(24) Hooper, I.R. *Aminoglycoside Antibiotics. Handbook of Experimental Pharmacology* **1982**, *Umezawa, S., Hooper, I.R. (eds), Springer-Verlag, New York. vol 62*, 1-35.

(25) Busscher, G.F.; Rutjes, F.P.J.T.; van Delft, F.L. *Chem. Rev.* **2005**, *105*, 775-792.

(26) Fourmy, D.; Recht, M.I.; Blanchard, S.C.; Puglisi, J.D. *Science* **1996**, *274*, 1367-1371.

(27) Moazed, D.; Noller, H.F. *Nature* **1987**, *327*, 389-394.

(28) von Ahsen, U.; Davies, J.; Schroeder, R. *Nature* **1991**, *353*, 368-370.

(29) von Ahsen, U.; Davies, J.; Schroeder, R. *J. Mol. Biol.* **1992**, *226*, 935.

(30) Stage, T.K.; Hertel, K.J.; Uhlenbeck, O.C. *RNA* **1995**, *1*, 95.

(31) Wang, H.; Tor, Y. *J. Am. Chem. Soc.* **1997**, *119*, 8734.

(32) Rogers, J.; Chang, A.H.; von Ahsen, U.; Schroeder, R.; Davies, J. *J. Mol. Biol.* **1996**, *259*, 916.

(33) Chia, J.-S.; Wu, H.-L.; Wang, H.-W.; Chen, D.-S.; Chen, P.-J. *Biomed. Sci.* **1997**, *4*, 208.

(34) Mikkelsen, N.E.; Brannvall, M.; Virtanen, A.; Kirsebom, L.A. *Proc. Natl. Acad. Sci. USA* **1999**, *259*, 916-925.

(35) Kirk, S.R.; Tor, Y. *Bioorg. Med. Chem.* **1999**, *7*, 1979-1991.

(36) Michael, K.; Wang, H.; Tor, Y. *Bioorg. Med. Chem.* **1999**, *7*, 1361-1371.

(37) Hermann, T.; Westhof, E. *J. Mol. Biol.* **1998**, *276*, 903-912.

(38) Tor, Y.; Hermann, T.; Westhof, E. *Chem. Biol.* **1998**, *5*, R277-R283.

(39) Lynch, S.R.; Puglisi, J.D. *J. Mol. Biol.* **2001**, *306*, 1037-1058.

(40) Recht, M.I.; Douthwaite, S.; Puglisi, J.D. *EMBO J.* **1999**, *18*, 3133-3138.

(41) Verhelst, S.H.; Michiels, P.J.A.; van der Marel, G.A.; van Boeckel, C.A.A.; van Boom, J.H. *ChemBioChem* **2004**, *5*, 937-942.

(42) Mei, H.-Y.; Galan, A.A.; Halim, N.S.; Mack, D.P.; Moreland, D.W.; Sanders, K.B.; Troung, H.N.; Czarnik, A.W. *Bioorg. Med. Chem. Lett.* **1997**, *5*, 2755.

(43) Zapp, M.L.; Stern, S.; Green, M.R. *Cell* **1993**, *74*, 969-978.

(44) Pierson, T. C.; Doms, R. W. *Curr Top Microbiol Immunol* **2003**, *281*, 1-27.

(45) Cocchi, F.; DeVico, A.L.; Garzino-Derno, A.; Arya, S.K.; Gallo, R.C.; Lusso, P. *Science* **1995**, *270*, 1811-1815.

(46) van Kooyk, Y.; Geijtenbeek, T. B. *Nat Rev Immunol* **2003**, *3*, 697-709.

(47) Bobardt, M. D.; Saphire, A. C.; Hung, H. C.; Yu, X.; Van der Schueren, B.; Zhang, Z.; David, G.; Gallay, P. A. *Immunity* **2003**, *18*, 27-39.

(48) Dvorin, J.D.; Malim, M.H. *Curr. Top. Microbiol. Immunol.* **2003**, *281*, 179-208.

(49) Arts, E.J.; LeGrice, S.F.J. *Prog. Nucl. Acids. Res. Mol. Biol.* **1998**, *58*, 339-393.

(50) Isel, C.; Ehresmann, C.; Keith, G.; Ehresmann, B.; Marquet, R. *J. Mol. Biol.* **1995**, *247*, 236-250.

(51) Isel, C.; Keith, G.; Ehresmann, B.; Ehresmann, C.; Marquet, R. *Nucl. Acids. Res.* **1998**, *26*, 1198-1204.

(52) Wakefield, J.K.; Wolf, A.G.; Morrow, C.D. *J. Virol.* **1995**, *69*, 6021-6029.

(53) Kim, S.H.; Sussman, J.L. *Nature* **1976**, *260*, 645-646.

(54) Puglisi, E.V.; Puglisi, J.D. *Nature Struc. Biol.* **1998**, *5*, 1033-1036.

(55) Robinson, H.; Wang, A.H.-J. *Nucl. Acids. Res.* **1996**, *24*, 676-682.

(56) Arya, D.P.; Xue, L.; Willis, B. *J. Am. Chem. Soc.* **2003**, 10148-10149.

(57) Campbell, E.M.; Hope, T.J. *Adv. Drug Deliv. Rev.* **2003**, *16*, 761-771.

(58) Greene, W.C.; Peterlin, B.M. *Nat. Med.* **2002**, *8*, 673-680.

(59) McDonald, D.; Vodicka, M.A.; Lucero, G.; Svitkina, T.M.; Borisy, G.G. *J. Cell Biol.* **2002**, *159*, 441-452.

(60) Fassati, A.; Gorlich, D.; Harrison, I.; Zaytseva, L.; Mingot, J.M. *EMBO J.* **2003**, *22*, 3675-3685.

(61) Schroder, A.R.; Shinn, P.; Chem, H.; Berry, C.; Ecker, J.R.; Bushman, F *Cell* **2002**, *110*, 521-529.

(62) Pereira, L.A.; Bently, K.; Peeters, A.; Churchill, M.J.; Deacon, N.J. *Nucl. Acids. Res.* **2000**, *28*, 663-668.

(63) Jones, K.A.; Peterlin, B.M. *Annu. Rev. Biochem.* **1994**, *63*, 717-743.

(64) Strebel, K. *AIDS* **2003**, *17*, 525-534.

(65) Daelemans, D.; Afonina, E.; Nilsson, J.; Werner, G.; Kiems, J.; de Clercq, E.; Pavlakis, G.N.; VanDamme, A.M. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14440-14445.

(66) Frankel, A. D.; Pabo, C.O. *Cell* **1988**, *55*, 1189-1193.

(67) Green, M.; Loewenstein, P.M. *Cell* **1988**, *55*, 1179-1188.

(68) Frankel, A. D. *Curr. Opin. Gen. Devo.* **1992**, *2*, 293-298.

(69) Aboul-ela, F.; Karn, J.; Varani, G. *J. Mol. Biol.* **1995**, *253*, 313-332.

(70) Hwang, S.; Tamilarasu, N.; Ryan, K.; Huq, I.; Richter, S.; Still, W.C.; Rana, T.M. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 12997-13002.

(71) Mei, H.-Y.; Mack, D.P.; Galan, A.A.; Halim, N.S.; Heldsinger, A.; Loo, J.A.; Moreland, D.W.; Sannes-Lowery, K.A.; Sharmeen, L.; Troung, H.N.; Czarnik, A.W. *Bioorg. Med. Chem.* **1997**, *5*, 1173-1184.

(72) Froeyen, M.; Herdewijn, P. *Curr. Top. Med. Chem.* **2002**, *2*, 1123-1145.

(73) Krebs, A.; Ludwig, V.; Boden, O.; Gobel, M.W. *ChemBioChem* **2003**, *4*, 972-978.

(74) Baba, M. *Curr. Top. Med. Chem.* **2004**, *4*, 871-882.

(75) Hamy, F.; Brondani, V.; Florsheimer, A.; Stark, W.; Blommers, M.J.; Klimkait, T. *Biochemistry* **1998**, *37*, 5086-5095.

(76) Dassonville, L.; Hamy, F.; Colson, P.; Houssier, C.; Bailly, C. *Nucl. Acids. Res.* **1997**, *25*, 4487-4492.

(77) Hamy, F.; Felder, E.R.; Heizmann, G.; Lazdins, J.; Aboul-ela, F.; Varani, G.; Karn, J.; Klimkait, T. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 3548-3553.

(78) Klimkait, T.; Felder, E.R.; Albrecht, G.; Hamy, F. *Biotechnol. Bioengin.* **1998-99**, *61*, 155-168.

(79) Mei, H.-Y.; Cui, M.; Heldsinger, A.; Lemrow, S.M.; Loo, J.A.; Sannes-Lowery, K.A.; Sharmeen, L.; Czarnik, A.W. *Biochemistry* **1998**, *37*, 14204-14212.

(80) Du, Z.; Lind, K.E.; James, T.L. *Chem. Biol.* **2002**, *9*, 707-712.

(81) Litovchick, A.; Evdokimov, A.G.; Lapidot, A. *FEBS Lett.* **1999**, *445*, 73-79.

(82) Litovchick, A.; Evdokimov, A.G.; Lapidot, A. *Biochemistry* **2000**, *39*, 2838-2852.

(83) Litovchick, A.; Lapidot, A.; Eisenstein, M.; Kalinkovich, A.; Borkow, G. *Biochemistry* **2001**, *40*, 15612-16523.

(84) Cabrera, C.; Gutierrez, A.; Barretina, J.; Blanco, J.; Litovchick, A.; Lapidot, A.; Clotet, B.; Este, J.A. *Antiviral Res.* **2002**, *53*, 1-8.

(85) Lee, J.; Kwon, M.; Lee, K.H.; Jeong, S.; Hyun, S.; Shin, K.J.; Yu, J. *J. Am. Chem. Soc.* **2004**, *126*, 1956-1957.

(86) Gelus, N.; Bailly, C.; Hamy, F.; Klimkait, T.; Wilson, W.D.; Boykin, D.W. *Bioorg. Med. Chem.* **1999**, *7*, 1089-1096.

(87) Gelus, N.; Hamy, F.; Bailly, C. *Bioorg. Med. Chem.* **1999**, *7*, 1075-1079.

(88) Yu, X.; Lin, W.; Li, J.; Yang, M. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3127-3130.

(89) Blount, K.F.; Tor, Y. *Nucl. Acids. Res.* **2003**, *31*, 5490-5500.

(90) Mikkelsen, N.E.; Brannvall, M.; Virtanen, A.; Kirsebom, L.A. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 6155-6160.

(91) Wong, C.H.; Hendrix, M.; Priestley, E.S.; Greenberg, W.A. *Chem. Biol.* **1998**, *5*, 397-406.

(92) Hendrix, M.; Priestley, E.S.; Joyce, G.F.; Wong, C.H. *J. Am. Chem. Soc.* **1997**, *119*, 3641-3648.

(93) Vicens, Q.; Westhof, E. *Structure* **2001**, *9*, 647-658.

(94) Fourmy, D.; Recht, M.I.; Blanchard, S.C.; Dalquist, K.; Puglisi, J.D. *Science* **1996**, *274*, 1367-1371.

(95) Carter, A.P.; Clemons, W.M.; Brodersen, D.E.; Morgan-Warren, R.J.; Wimberly, B.T.; Ramakrishnan, V. *Nature* **2000**, *407*, 897-902.

(96) Faber, C.; Sticht, H.; Schweimer, K.; Rosch, P. *J. Biol. Chem.* **2000**, *275*, 20660-20666.

(97) Blount, K.F.; Zhao, F.; Hermann, T.; Tor, Y. *J. Am. Chem. Soc.* **2005**, *127*, 9818-9829.

(98) Zhao, F.; Zhao, Q.; Blount, K.F.; Han, Q.; Tor, Y.; Hermann, T. *Angew. Chem. Int. Ed.* **2005**, *44*, 5329-5334.

(99) Hammarskjold, M.-L. *Cell Dev. Biol.* **1997**, *8*, 83-90.

(100) Wang, X.; Huq, I.; Rana, T.M. *J. Am. Chem. Soc.* **1997**, *119*, 6444-6445.

(101) Tamilarasu, N.; Huq, I.; Rana, T.M. *J. Am. Chem. Soc.* **1999**, *121*, 1597-1598.

(102) Hope, T.J. *Archiv. Biochem. Biophys.* **1999**, *365*, 186-191.

(103) Holland, S.M.; Chavez, M.; Gerstberger, S.; Venkatesan, S. *J. Virol.* **1992**, *66*, 3699-3706.

(104) Tilley, L.S.; Malim, M.H.; Tewary, H.K.; Stockley, P.G.; Cullen, B.R. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 758-762.

(105) Battiste, J.L.; Tan, R.; Frankel, A. D.; Williamson, J.R. *Biochemistry* **1994**, *33*, 2741-2747.

(106) Bartel, D.P.; Zapp, M.L.; Green, M.R.; Szostak, J.W. *Cell* **1991**, *67*, 529-536.

(107) Schaal, H.; Klein, M.; Gehrmann, P.; Adams, O.; Scheid, A. *J. Virol.* **1995**, *69*, 3308-3314.

(108) Phuphuakrat, A.; Auewarakul, P. *AIDS Res. Human Retrovir.* **2003**, *19*, 569-574.

(109) Werstuck, G.; Zapp, M.L.; Green, M.R. *Chem. Biol.* **1996**, *3*, 129-137.

(110) Lacourciere, K.A.; Stiver, J.T.; Marino, J.P. *Biochemistry* **2000**, *39*, 5630-5641.

(111) Luedtke, N.W.; Tor, Y. *Angew. Chem. Int. Ed.* **2000**, *39*, 1788-1790.

(112) Luedtke, N.W.; Tor, Y. *Biopolymers/Nucleic Acid Sciences* **2003**, *703*, 103-119.

(113) Luedtke, N.W.; Liu, Q.; Tor, Y. *Biochemistry* **2003**, *42*, 11391-11403.

(114) Friedler, A.; Friedler, D.; Luedtke, N.W.; Tor, Y.; Loyter, A.; Gilon, C. *J. Biol. Chem.* **2000**, *275*, 23783-23789.

(115) Luedtke, N.W.; Hwang, J.S.; Glazer, E.C.; Gut, D.; Kol, M.; Tor, Y. *ChemBioChem* **2002**, *3*, 766-771.

(116) Park, W.K.C.; Auer, M.; Jaksche, H.; Wong, C.H. *J. Am. Chem. Soc.* **1996**, *118*, 10150-10155.

(117) Hamasaki, K.; Woo, M.-C.; Ueno, A. *Tet. Lett.* **2000**, *41*, 8327-8332.

(118) Wang, Y.; Hamasaki, K.; Rando, R.R. *Biochemistry* **1997**, *36*, 768-779.

(119) Ahn, D.-R.; Yu, J. *Bioorg. Med. Chem.* **2005**, *13*, 1177-1183.

(120) Battiste, J.L.; Mao, H.; Rao, N.S.; Tan, R.; Muhandiram, D.R.; Kay, L.E.; Frankel, A. D.; Williamson, J.R. *Science* **1996**, *273*, 1547-1551.

(121) Wilson, W.B.; Ratmeyer, L.; Cegla, M.T.; Spychala, J.; Boykin, D.W.; Demeunynyck, M.; Lhomme, J.; Krishnan, G.; Kennedy, D.; Vinayak, R.; Zon, G. *New. J. Chem.* **1994**, *18*, 419-423.

(122) Kirk, S.R.; Luedtke, N.W.; Tor, Y. *J. Am. Chem. Soc.* **2000**, *122*, 980-981.

(123) Weiss, M.A.; Narayana, N. *Biopolymers/Nucleic Acid Sciences* **1999**, *48*, 167-180.

(124) Bayer, T.S.; Booth, L.N.; Knudsen, S.M.; Ellington, A.D. *RNA* **2005**, *11*, 1848-1857.

(125) Luedtke, N.W.; Baker, T.J.; Goodman, M.; Tor, Y. *J. Am. Chem. Soc.* **2000**, *122*, 12035-12036.

(126) Baker, T.J.; Luedtke, N.W.; Tor, Y.; Goodman, M. *J. Org. Chem.* **2000**, *65*, 9054-9058.

(127) Paillart, J.-C.; Shehu-Xhilaga, M.; Marquet, R.; Mak, J. *Nature Rev. Microbiol.* **2004**, *2*, 461-472.

(128) Huthoff, H.; Berkhout, B. *Biochemistry* **2002**, *41*, 10439-10445.

(129) Muriaux, D.; De Rocquigny, H.; Roques, B.-P.; Paoletti, J. *J. Biol. Chem.* **1996**, *271*, 33686-33692.

(130) Laughrea, M.; Jette, L.; Mak, J.; Kleiman, L.; Liang, C.; Wainberg, M.A. *J. Virol.* **1997**, *71*, 3397-3406.

(131) Shen, N.; Jette, L.; Liang, C.; Wainberg, M.A.; Laughrea, M. *J. Virol.* **2000**, *74*, 5729-5735.

(132) Mujeeb, A.; Clever, J.L.; Billeci, T.M.; James, T.L.; Parslow, T.G. *Nat. Struc. Biol.* **1998**, *5*, 432-436.

(133) Jossinet, F.; Paillart, J.-C.; Westhof, E.; Hermann, T.; Skripkin, E.; Lodmell, J.S.; Ehresmann, C.; Ehresmann, B.; Marquet, R. *RNA* **1999**, *5*, 1222-1234.

(134) Lodmell, J.S.; Ehresmann, C.; Ehresmann, B.; Marquet, R. *J. Mol. Biol.* **2001**, *311*, 475-490.

(135) McPike, M.P.; Goodisman, J.; Dabrowaik, J.C. *Bioorg. Med. Chem.* **2002**, *10*, 3663-3672.

(136) McPike, M.P.; Sullivan, J.M.; Goodisman, J.; Dabrowaik, J.C. *Nucl. Acids. Res.* **2002**, *30*, 2825-2831.

(137) Sullivan, J.M.; Goodisman, J.; Dabrowaik, J.C. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 615-618.

(138) Ennifar, E.; Walter, P.; Ehresmann, B.; Ehresmann, C.; Dumas, P. *Nature Struc. Biol.* **2001**, *8*, 1064-1068.

(139) Ennifar, E.; Paillart, J.-C.; Marquet, R.; Ehresmann, B.; Ehresmann, C.; Dumas, P.; Walter, P. *J. Biol. Chem.* **2003**, *278*, 2723-2730.

(140) Futaki, S.; Suzuki, T.; Ohashi, W.; Yagami, T.; Tanaka, S.; Ueda, K.; Sugiura, Y. *J Biol Chem* **2001**, *276*, 5836-5840.

(141) Futaki, S.; Suzuki, T.; Ohashi, W.; Yagami, T.; Tanaka, S.; Ueda, K.; Sugiura, Y. *J. Biol. Chem.* **2001**, *276*, 5836-5840.

(142) Astriab-Fisher, A.; Sergueev, D.S.; Fisher, M.; Shaw, B.R.; Juliano, R.L. *Biochem. Pharmacol.* **2000**, *60*, 83-90.

(143) Josephson, L.; Tung, C.H.; Moore, A.; Weissleder, R. *Bioconj. Chem.* **1999**, *10*, 186-191.

(144) Lewin, M.; Carlesso, N.; Tung, C.H.; Tang, X.W.; Cory, D.; Scadden, D.T.; Weissleder, R. *Nat. Biotechnol.* **2000**, *18*, 410-414.

(145) Wunderbaldinger, P.; Joesphson, L.; Weissleder, R. *Bioconj. Chem.* **2002**, *13*, 264-268.

(146) Polyakov, V.; Sharma, V.; Dahlheimer, J.L.; Pica, C.M.; Luker, G.D.; Piwnica-Worms, D. *Bioconj. Chem.* **2000**, *11*.

(147) Torchilin, V.P.; Rammohan, R.; Weissig, V.; Levchenko, T.S. *Proc. Nat. Acad. Sci. USA* **2001**, *98*, 8786-8791.

(148) Futaki, S.; Nakase, I.; Suzuki, T.; Youjun, Z.; Sugiura, Y. *Biochemistry* **2002**, *41*, 7925-7930.

(149) Wender, P.A.; Rothbard, J.B.; Jessop, T.C.; Kreider, E.L.; Wylie, B.L. *J. Am. Chem. Soc.* **2002**, *124*, 13382-13383.

(150) Wender, P.A.; Mitchell, D.J.; Pattabiraman, K.; Pelkey, E.T.; Steinman, L.; Rothbard, J.B. *Proc. Nat. Acad. Sci. USA* **2000**, *97*, 13003.

(151) Derossi, D.; Calvet, S.; Trembleau, A.; Brunissen, A.; Chassaing, G.; Prochiantz, A. *J. Biol. Chem.* **1996**, *271*, 18188-18193.

(152) Rueping, M.; Mahajan, Y.; Sauer, M.; Seebach, D. *ChemBioChem* **2002**, *03*, 257-259.

(153) Umezawa, N.; Gelman, M.A.; Haigis, M.C.; Raines, R.T.; Gellman, S.H. *J. Am. Chem. Soc.* **2002**, *124*, 368-369.

(154) Rothbar, J.B.; Kreider, E.; Van Deusen, C.L.; Wright, L.; Wylie, B.L.; Wender, P. *J. Med. Chem.* **2002**, *45*, 3612-3618.

(155) Luedtke, N.W.; Charmichael, P.; Tor, Y. *J. Am. Chem. Soc.* **2003**, *125*, 12374-12375.

(156) Rothbard, J.B.; Jessop, T.C.; Lewis, R.S.; Murray, B.A.; Wender, P.A. *J. Am. Chem. Soc.* **2004**, *126*, 9506-9507.

(157) Fuchs, S.M.; Raines, R.T. *Biochemistry* **2004**, *43*, 2438-2444.

(158) Richard, J.P.; Melikov, K.; Vives, E.; Ramos, C.; Vebeure, B.; Gait, M.J.; Chernomordik, L.V.; Lebleu, B. *J. Biol. Chem.* **2003**, *278*, 585-590.

(159) Roden, L.; Ananth, S.; Campbell, P.; Curenton, T.; Ekborg, G.; Manzella, S.; Pillion, D.; Meezan, E. *Adv Exp Med Biol* **1992**, *313*, 1-20.

(160) Wang, H.; Tor, Y. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 3665-3670.

(161) Luedtke, N.W.; Baker, T.J.; Goodman, M.; Tor, Y. *J. Am. Chem. Soc.* **2000**, *122*, 12035-12036.

(162) Baker, T.J.; Luedtke, N.W.; Tor, Y.; Goodman, M. *J. Org. Chem.* **2000**, *65*, 9054-9058.

(163) Esko, J.D.; Lindahl, U. *J. Clin. Invest.* **2001**, *108*, 169-173.

(164) Kjellen, L.; Lindahl, U. *Annu. Rev. Biochem.* **1991**, *60*, 443-475.

(165) Bernfield, M., et. al. *Annu. Rev. Biochem.* **1999**, *68*, 729-777.

(166) Bernfield, M., et. al. *Annu. Rev. Cell. Biol.* **1992**, *8*, 365-393.

(167) Williams, K.J.; Fuki, I.V. *Curr. Opin. Lipidol.* **1997**, *8*, 253-262.

(168) Tumova, S.; Woods, A.; Couchman, J.R. *Int. J. Biochem. Cell Biol.* **2000**, *32*, 269-288.

(169) Rapraeger, A.C. *J. Cell. Biol.* **2000**, *149*, 995-998.

(170) Filmus, J. *Glycobiology* **2001**, *11*, 19R-23R.

(171) Yanagishita, M.; Hascall, V. *J. Biol. Chem.* **1992**, *267*, 9451-9454.

(172) Jackson, R.L.; Busch, S.J.; Cardin, A.D. *Physiol. Rev.* **1991**, *71*, 481-539.

(173) Spillmann, D.; Lindahl, U. *Curr. Opin. Struct. Biol.* **1994**, *4*, 677-682.

(174) Hirschberg, C.B.; Robbins, P.W.; Abeijon, C. *Annu. Rev. Biochem.* **1998**, *67*, 49-69.

(175) Berninsone, P.M.; Hirschberg, C.B. *Curr. Opin. Struct. Biol.* **2000**, *10*, 542-547.

(176) Esko, J.D.; Zhang, L. *Curr. Opin. Struct. Biol.* **1996**, *6*, 663-670.

(177) Sugahara, K.; Kitagawa, H. *Curr. Opin. Struct. Biol.* **2000**, *10*, 518-527.

(178) Duncan, G.; McCormick, C.; Tufaro, F. *J. Clin. Invest.* **2001**, *108*, 511-516.

(179) Lindahl, U.; Kusche-Gullberg, M.; Kiellen, L. *J. Biol. Chem.* **1998**, *273*, 24979-24982.

Chapter 2

2.1 Introduction: Entering the Cell

Facilitating the delivery of molecules into living cells is of substantial interest in potential applications to drug delivery and medicinal chemistry, as well as generating insights into cellular functions.[1-6] Most contemporary therapeutics are either natural products or low molecular weight synthetic organic molecules. In recent years, high throughput screening and advances in the understanding of human diseases have yielded leads for new therapeutic approaches and helped to identify high molecular weight biomolecules as potential therapeutic agents. A major concern with many new drug candidates and high molecular weight therapeutics is delivery, the ability to be taken up into cells and released into the cytoplasm where they are able to access their cellular target.

While many approaches using "physical" control, such as polymeric microspheres, micelles and hydrogels,[7-10] help improve the pharmacokinetic features of bioactive molecules, they do not address a major inherent challenge of medicinal chemistry: the ability of a therapeutic agent to cross the cell membrane. Hence, there is a strong need for systems which facilitate the cellular uptake of bioactive molecules and do not rely on passive diffusion.

While it has been known that some polybasic proteins can enhance the cellular uptake of biomolecules,[11] recent advances in the science and applications of these tools[12] has opened a new field for natural and synthetic cellular uptake vehicles. The potential of Protein Transduction Domains (PTDs) was first realized with the

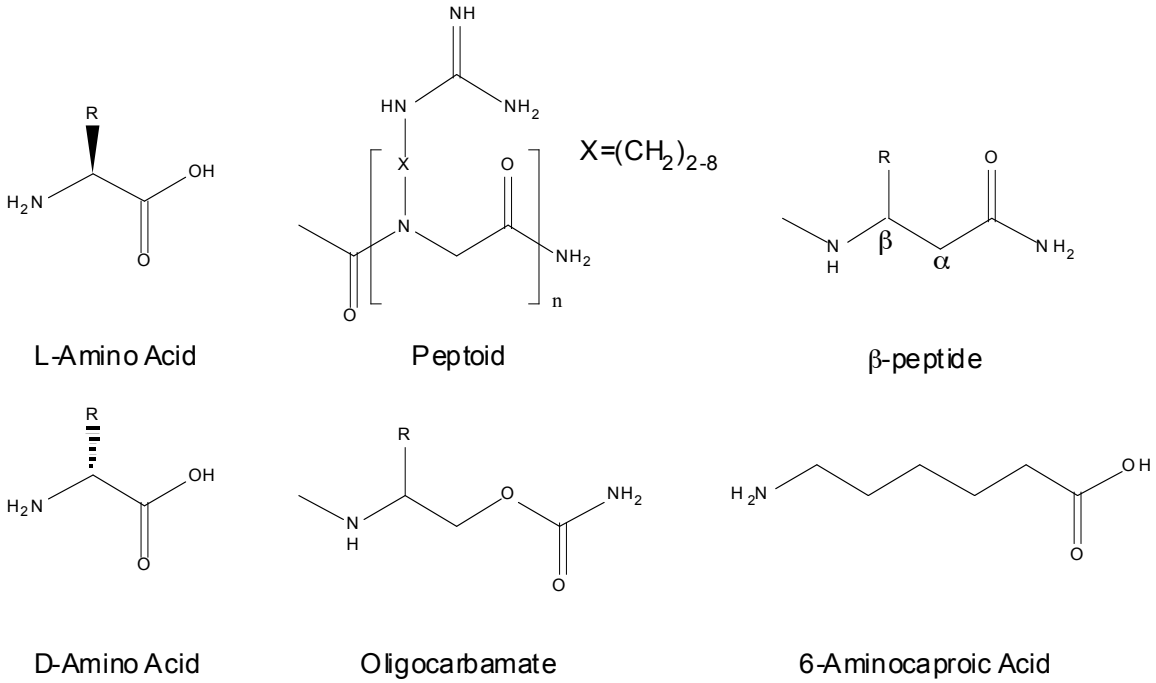


Figure 2.1 Guanidinium group-containing oligomers have been made with multiple different linear and branched scaffolds. These compounds can efficiently enter and carry large cargos across membranes and into cells. The common structural feature among these molecules is the presence of multiple guanidinium groups.

homeodomain of Antennapedia[13], whose DNA binding domain was shown to mediate cellular uptake in numerous cell types. Since then, many proteins and peptides have been shown to facilitate cellular uptake. A large number of these membrane-permeable peptides contain clustered arginine residues.[14] This suggests that the guanidinium functional group is a major contributor to cellular uptake. Branched chain arginine-rich oligomers,[15] as well as many oligomeric guanidinium-rich analogs (Figure 2.1) have been found to exhibit useful cellular uptake properties, further implicating the guanidinium group in cellular uptake.

Recently, "guanidinoglycosides" (Figure 2.2), a novel family of synthetic RNA ligands based on natural aminoglycosides scaffolds, have been shown to exhibit high

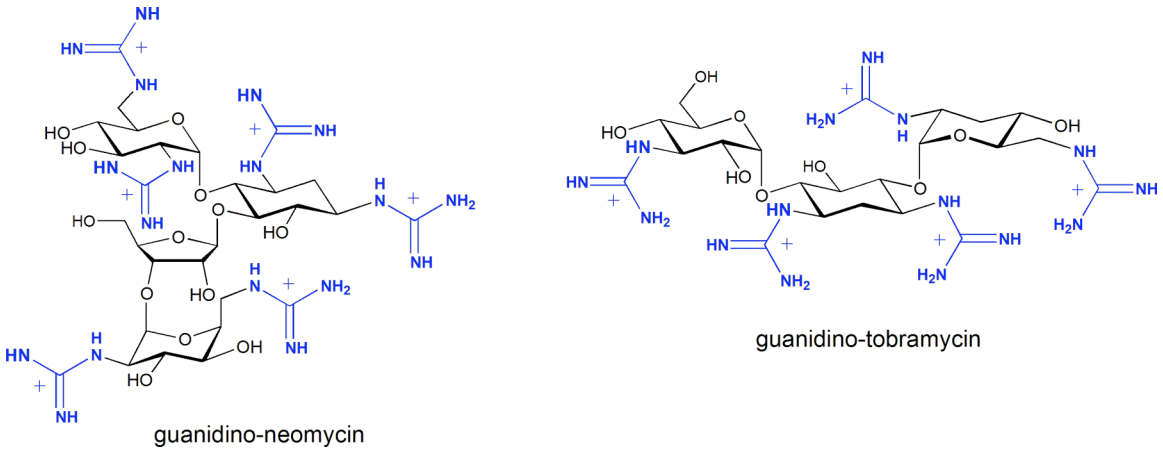


Figure 2.2. Guanidinoglycosides are a novel family of small molecules where all of the amino groups from aminoglycosides have been converted to guanidinium groups. Guanidino-neomycin and guanidino-tobramycin have high affinity and selectivity to RNA targets which are recognized by arginine-rich domains, and also show efficient cellular uptake properties.

affinity and selectivity to RNA targets that are naturally recognized by arginine-rich domains.[16,17] Additionally, guanidinoglycosides also display effective cellular uptake properties.[18]

The mechanism of cellular uptake by guanidine-rich molecules such as arginine-rich peptides and guanidinoglycosides remains unclear. Multiple models have been proposed, and multiple mechanisms may be acting simultaneously.[12,19] Receptor-mediated transport appears unlikely due to the structural diversity of guanidinium-rich molecules which show uptake activity Recent studies proposing an endocytotic mechanism have suffered because of differing results with fixed and live cells.

A new model has been proposed where positively charged guanidinium groups interact with negatively charged cell surface proteoglycans.[20,21] In this model:

- Guanidinium-rich oligomers bind to cell surface heparan sulfate proteoglycans.
- The proteoglycan–ligand complex is taken into the cell via receptor mediated endocytosis.
- Heparan sulfate is degraded by heparanases and the guanidinium-rich oligomers are released within the endocytotic vesicle.
- The effective increase in guanidinium concentration promotes vesicular leakage and allows these compounds to escape into the cytoplasm.

In this chapter, we will show that the guanidinoglycosides bind to cell surface heparan sulfate proteoglycans, and can effectively inhibit the binding of heparan sulfate specific proteins. Additionally, guanidino-neomycin is efficiently taken up into live cells in a heparan sulfate dependent manner. Aminoglycosides, neomycin and tobramycin are not internalized; suggesting that the interactions between the positively charged guanidinium groups of guanidinoglycosides and the negatively charged sulfates on heparan sulfate proteoglycans plays a critical role in uptake.

We will also show that guanidinoglycosides can facilitate uptake of very large cargo into living cells, and that the cargo molecules are released into the cytoplasm where they retain their biological activity. The delivery activity of guanidinoglycosides is also dependent on cell surface heparan sulfate proteoglycans. This highlights the potential of guanidinoglycosides as delivery agents.

2.2 Guanidinoglycosides Bind to Cell Surface Heparan Sulfate

To probe whether guanidinoglycosides are able to bind cell surface heparan sulfate proteoglycans, cells were incubated with various compounds, including guanidinoglycosides and exposed to a toxin-protein conjugate which specifically binds heparan sulfate. Cells with compound bound to cell surface heparan sulfate proteoglycans were protected from the toxin and were not killed while cells where no compound bound to the proteoglycans were not protected from toxin and did not survive (Figure 2.3).

Chinese hamster ovarian (CHO) cells are grown on tissue culture plates in HAM's F-12 media with 5% fetal bovine serum and incubated with compounds to be assayed for two days. This allows potential inhibitors to act, and pre-existing heparan sulfate to fully turn over.

Cells are then exposed to toxin. These toxins act by binding cell surface heparan sulfate proteoglycans, and internalizing into the cytoplasm, killing the cell. After three days of exposure to toxins, Alamar blue is added to the media. Living cells metabolize Alamar blue and display an increase in fluorescence. Dead cells can not metabolize the dye and thus do not display an increased fluorescent signal.

Cells expressing heparan sulfate die from the toxins and show low fluorescent signals. When inhibitors bind to heparan sulfate proteoglycans, they protect cells from the toxins and these cells are not killed. Cells can be protected from the toxins through

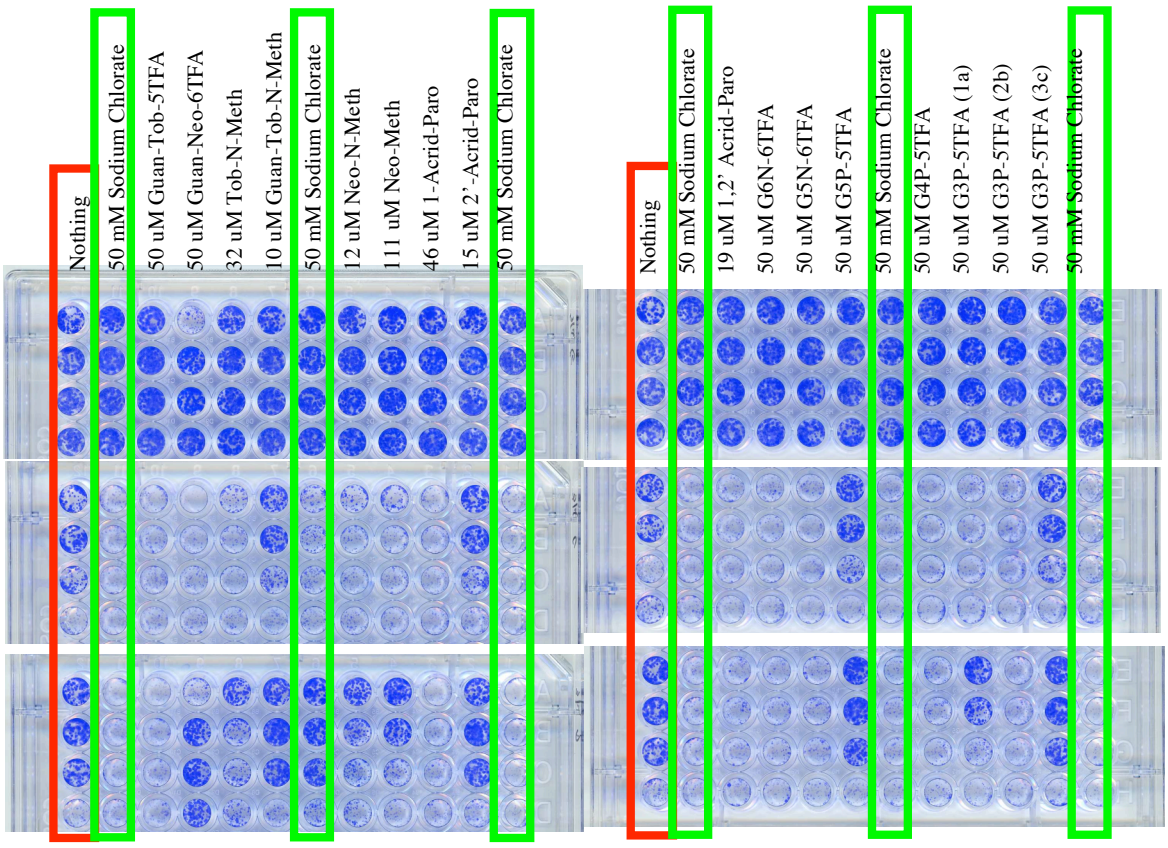


Figure 2.3. Viability assay where compounds are tested for their ability to block the binding of toxins bound to Heparan Sulfate specific binding proteins. Living cells are stained dark blue. The top section of each plate represents cells treated with compounds but not heparan sulfate binding toxins to assay for inherent compound toxicity. The middle and lower sections represent assays with two different heparan-sulfate binding protein-toxin conjugates. Compounds are diluted by a factor of two starting from the top well, down to the fourth well of each section.

lack of heparan sulfate or through compounds binding to the heparan sulfate and blocking the interactions which lead to toxin activity on the cell.

This assay allows for a preliminary structure activity relationship to be developed for heparan sulfate binding molecules. Multiple similar compounds can be assayed simultaneously and their activity can be directly compared.

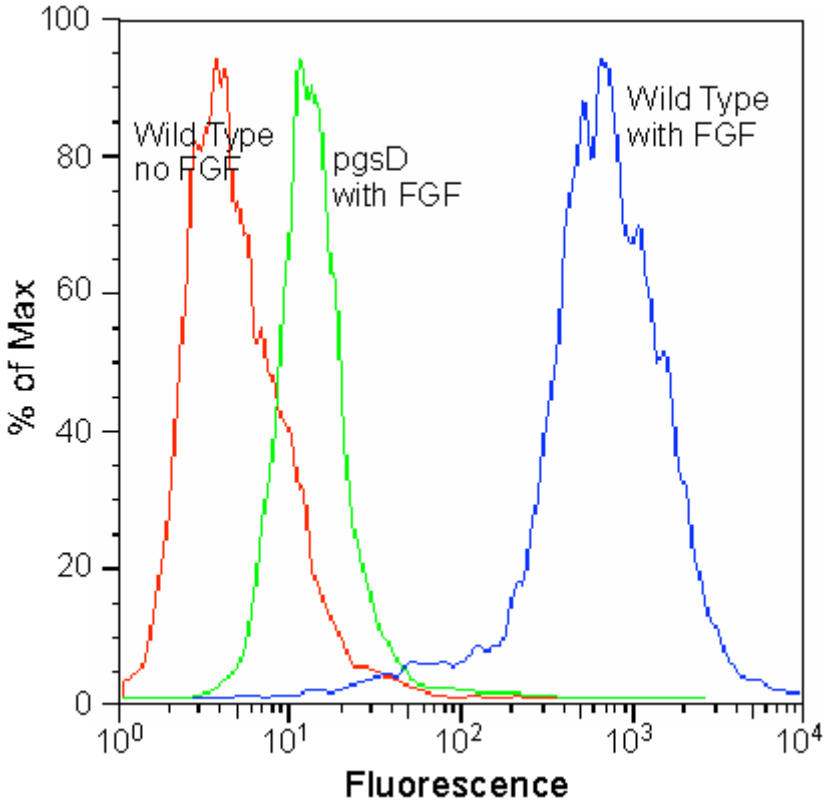


Figure 2.4. Controls for FGF inhibition experiments. Histograms show that wild type cells which do have cell surface heparan sulfate, strongly bind FGF to give a large increase in fluorescent signal whereas pgsD-677 (heparan sulfate deficient) cells do not. Flow cytometry profiles show that FGF binding depends on heparan sulfate, and that secondary avidinylated fluorophores do not have any cell surface binding properties.

While partially guanidinylated guanidinoglycosides (Figure 2.2) show little to no protection, fully guanidinylated guanidinoglycosides show protection at relatively low concentrations. For example, guanidino-neomycin begins to show protection at 12.5 μM (G6N-6TFA and Guan-Neo-6TFA) while the partially guanidinylated (G5N-6TFA), which has guanidinium groups substituted for five of its six amines, does not show protection at 50 μM. Interestingly, guanidino-tobramycin (Guan-Tob-5TFA), which has five guanidine groups, one less than guanidino-neomycin shows no protection in this assay. Additionally, guanidino-tobramycin does show heparan sulfate

binding activity, but not as much as guanidino-neomycin (discussed below). This is in agreement with previous results that show that guanidino-tobramycin is taken up into living cells, although not as well as guanidino-neomycin.[18] While providing the ability to screen large numbers of compounds for FGF-inhibition, this high throughput screen is not very sensitive and is limited to compounds with high activity.

Because of the insensitivity of the high throughput assay, a more quantitative heparan sulfate binding assay has been developed. In this assay, cells are treated with biotinylated fibroblast growth factor (FGF), a protein which selectively binds to heparan sulfate on the cell surface. Chinese hamster ovarian (CHO-K1) cells are grown to confluency, lifted from tissue culture plates with EDTA, washed with phosphate buffered saline and incubated in suspension with biotinylated FGF for 10 minutes at 37^{o}C.[22] Cells are then treated with a streptavidin bound Phycoerythrin-cychrome fluorophore (streptavidin-PE-Cy5), and analyzed with fluorescent activated cell sorting (FACS). FGF binds to wild type CHOK1, giving an average fluorescent signal of roughly 10^3 with PE-Cy5. Mutant cell lines lacking *N*-acetylglucosaminyl and glucuronosyltransferase activity (CHO-pgsD-677) do not have the ability to catalyze the polymerization of heparan sulfate. They do, however express roughly 3-4 fold increased amounts of Chondroitin Sulfate. These heparan sulfate lacking CHO-pgsD-677 do not bind FGF, and therefore show a greatly reduced fluorescent signal (Figure 2.4).

When wild type cells are treated with the secondary fluorophore, without being pre-treated with FGF-biotin, they do not show a fluorescent signal, indicating that the

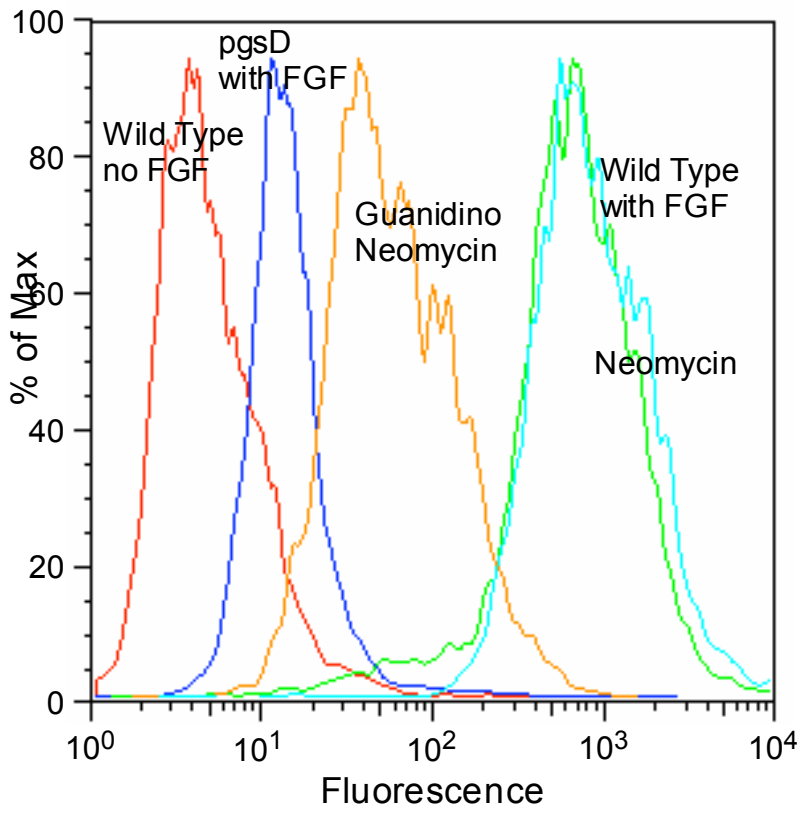


Figure 2.5. Neomycin and Guanidino-Neomycin competition experiments where FGF inhibition is monitored by flow cytometry. 25 μM Guanidino-Neoymcin shows a significant inhibition of FGF binding while Neomycin shows no effect at the same concentration

cells possess no significant inherent autofluorescence and that the secondary does not have any inherent affinity to the cell surface (Figure 2.4).

Molecules which bind to cell surface heparan sulfate proteoglycans may compete with FGF and inhibit binding, causing a decreased fluorescent signal while molecules which do not bind to heparan sulfate will not compete off FGF and will not decrease the fluorescent signal from FGF binding. While this assay does not have the higher throughput capabilities of other assays, it does allow for quantitative determination of heparan sulfate binding through FGF inhibition.

When cells are incubated with a compound which binds to heparan sulfate, and competes for FGF binding, fewer FGF molecules can then bind to the cell surface proteoglycans. This causes a decrease in the mean fluorescence per cell which can be quantitatively measured, allowing for determination of IC_{50} values of FGF blocking.

To examine their ability to bind to cell surface heparan sulfate and block FGF binding, neomycin, guanidino-neomycin, tobramycin and guanidino tobramycin were each dissolved in phosphate buffered saline (PBS) with 1-1000 FGF-biotin, and incubated with wild type cells in suspension for 1 hour at 4^{o}C. After 1 hour, cells were pelleted at 1500 rpm and washed 3 times with PBS. Cells were then suspended in FACS buffer (PBS with 0.1% Bovine Serum Albumen) with 0.1 % streptavidin-PE-Cy5 and incubated at room temperature for 20 minutes. After 3 washes with PBS, cells were suspended in FACS buffer and analyzed via flow cytometry (Figure 2.5).

Cells incubated with 25 μM guanidino-neomycin show roughly a 15 fold signal decrease, while neomycin, the parent aminoglycoside, has no effect on the fluorescent signal at this concentration. This indicates that guanidino-neomycin can effectively inhibit FGF binding to wild type cells, most likely through interactions with cell surface heparan sulfate proteoglycans. Neomycin, however does not bind to cell surface heparan sulfate proteoglycans to inhibit FGF binding.

A similar trend is observed when tobramycin and guanidino-tobramycin are assayed for their ability to block FGF binding to wild type cells (Figure 2.6). While tobramycin shows no inhibitory effect at 25 μM, at this concentration, guanidino-tobramycin is able to reduce FGF binding by 2-3 fold. These data suggests that the guanidinium group is directly responsible for the inhibitory potency of the

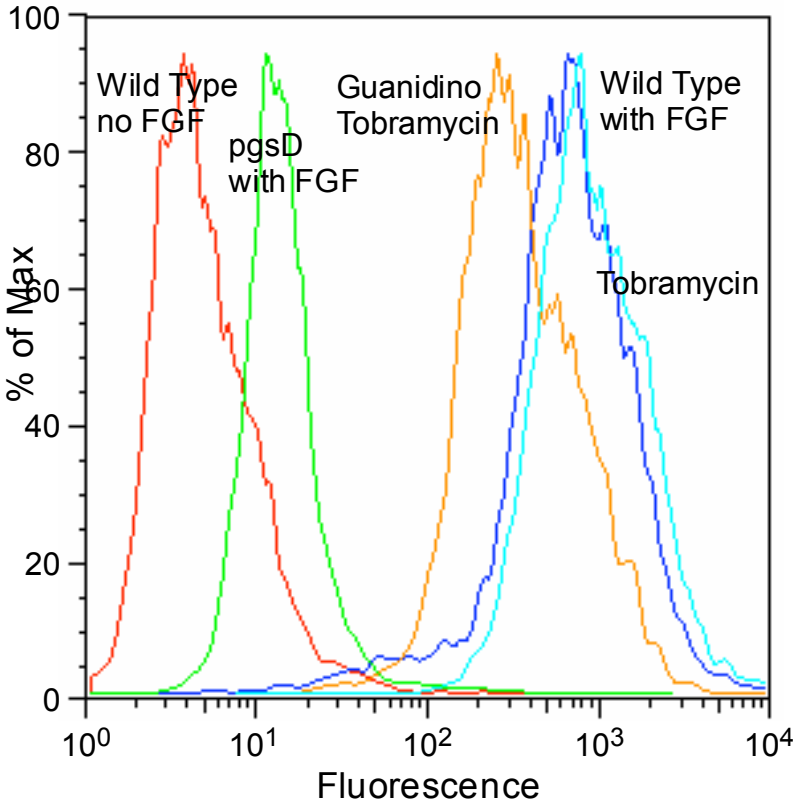


Figure 2.6. FGF competition experiments with Tobramycin and Guanidino-Tobramycin show similar effects as those with the Neomycin derivatives. Guanidino-Tobramycin is able to inhibit FGF binding at 25 μM but to a lesser degree than Guanidino-Neomycin. Tobramycin shows no ability to block FGF biding to wild type cells.

guanidinoglycosides, but also that this activity may depend on the number of guanidinium groups on guanidinoglycosides.

Guanidino-neomycin has six guanidinium groups while guanidino-tobramycin has only five. When compared for FGF inhibitory activity, the neomycin derivative is roughly 5-fold more active than the tobramycin derivative, even though it contains only one additional charged group. Previous studies have show that guanidino-neomycin has much better cellular uptake than guanidino-tobramycin.[18] When coupled, these

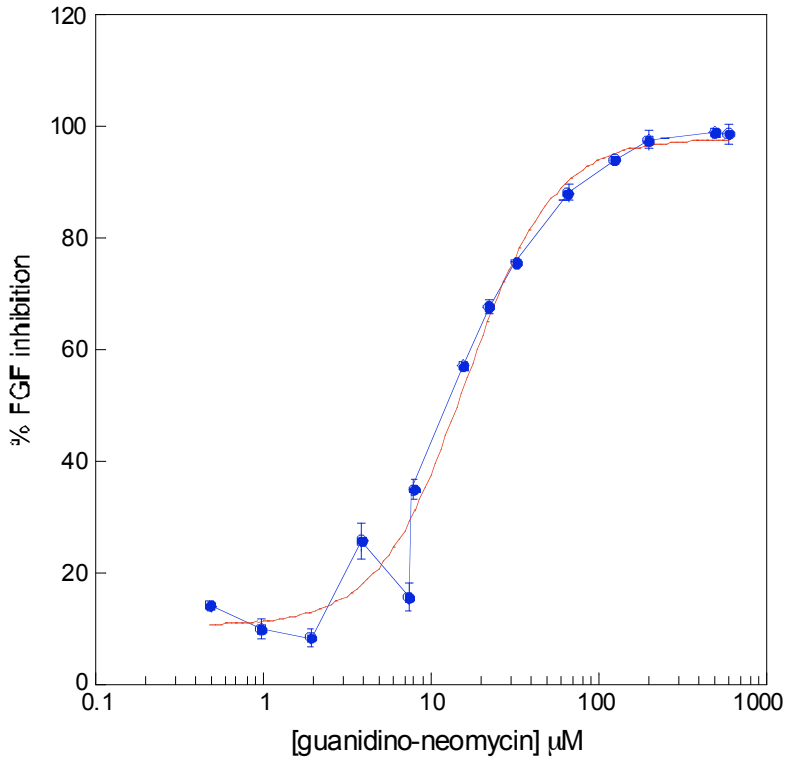


Figure 2.7. Basic FGF inhibition experiments performed with varying concentrations of guanidino-neomycin reveal an IC_{50} value of 20 μM for FGF blocking. Because the Fibroblast Growth Factor binds to cells in a heparan sulfate dependent manner, this indicates that guanidino-neomycin is binding to cell surface heparan sulfate proteoglycans. The blue dots are data points with error bars and the red curve represents a curve fit for the data.

data provide evidence to support a theory that guanidinoglycoside uptake is dependent on heparan sulfate proteoglycans.[20]

Guanidino-neomycin shows an IC_{50} value for FGF blocking of roughly 20 μM (Figure 2.7). Because FGF shows heparan sulfate specific binding, this indicates that guanidino-neomycin has either very high affinity for all cell surface proteoglycans or that it is selective for heparan sulfate proteoglycans.

2.3 Direct Measurements of Fluorescent Tagged Amino and Guanidinoglycosides

To further characterize the mechanism of cell surface binding and uptake of guanidinoglycosides, three compounds which were previously prepared by Nathan Luedtke[18] (Figure 2.8) were used. Neomycin, guanidino-neomycin and a short peptide with nine arginine residues, $Cys(Arg)_9$, were each conjugated to BODIPY.[18] BODIPY is a fluorescent probe with high quantum efficiency at physiological pH. By linking small molecules to fluorescent tags, we are able to directly observe the behavior of BODIPY-glycosides and BODIPY-$Cys(Arg)_9$. Neomycin-BODIPY, guanidino-neomycin-BODIPY and $Cys(Arg)_9$-BODIPY (as well as tobramycin-BODIPY and guanidino-tobramycin-BODIPY) were previously synthesized and purified via HPLC.[18] All compounds were stored as dry solid at -78^{o}C and freshly dissolved in water prior to each experiment. All BODIPY-glycoside conjugates are slightly to moderately hygroscopic, therefore concentrations were determined in methanol using ultraviolet absorbance with ε_{502} = 76,000 $cm^{-1}M^{-1}$.

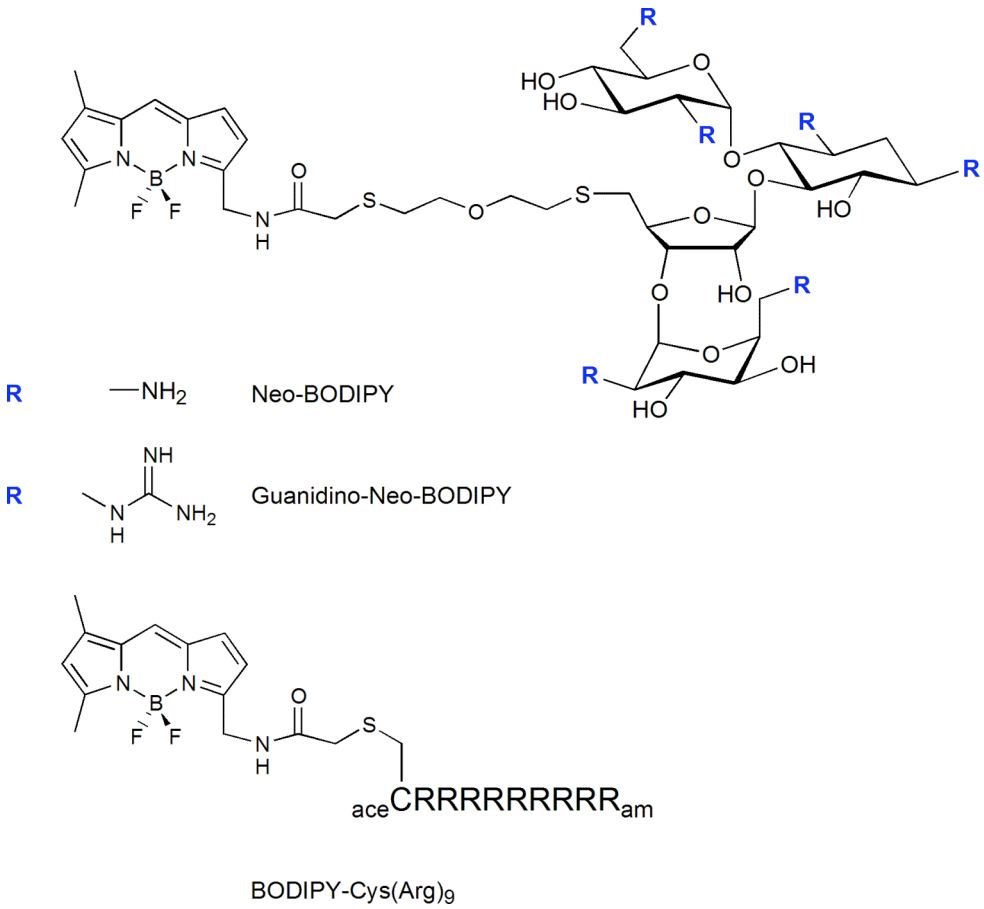


Figure 2.8. BODIPY tagged compounds used for direct uptake experiments. Neomycin-BODIPY and guanidino-neomycin-BODIPY were used to examine the importance of the guanidinium group on cellular uptake with BODIPY-Cys(Arg)$_9$ being used as a cellular uptake standard, allowing for the comparison of peptide and aminoglycoside based backbones.

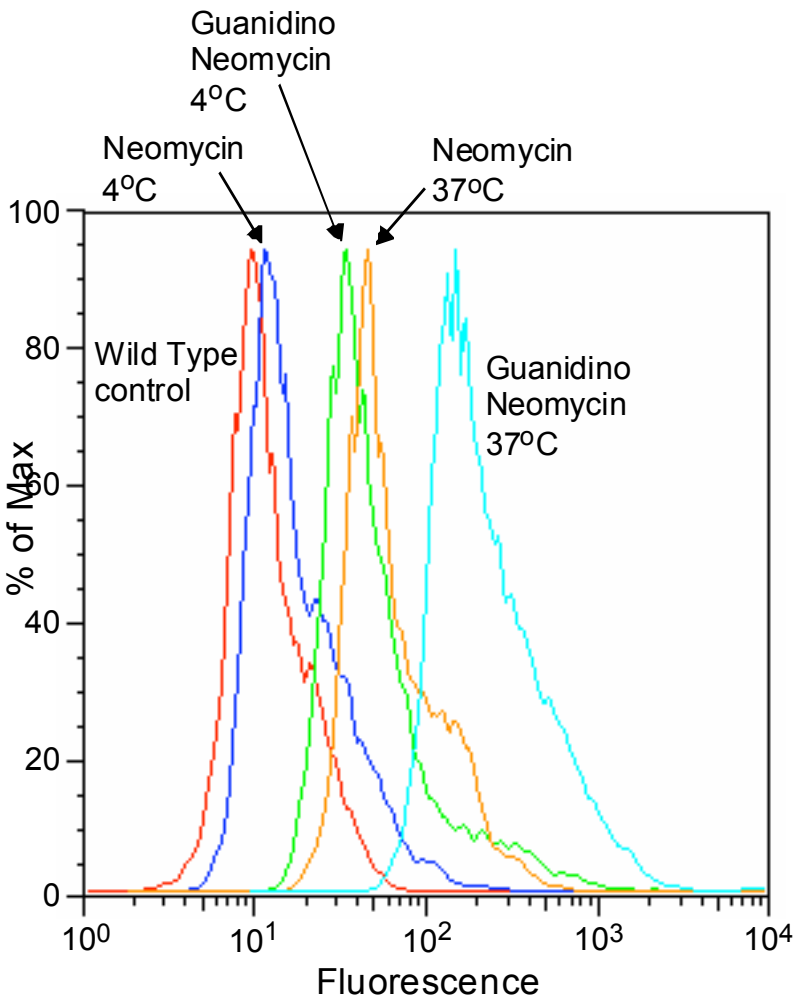


Figure 2.9. Neomycin-BODIPY and Guanidino-Neomycin-BODIPY both show stronger signals when incubated with wild type cells (1 μM Neomycin/Guanidino-Neomycin) at 37°C than at 4°C. The same trend is observed when Tobramycin is compared to Guanidino-Tobramycin.

While it is known that guanidinoglycosides interact with the surface of cells, it is important to define the interactions as surface binding, uptake or a combination of the two. To differentiate cell surface binding from endocytosis, cells were incubated with guanidino-neomycin-BODIPY and neomycin-BODIPY (1 μM) at both 4°C and 37°C, and analyzed via flow cytometry (Figure 2.9). Endocytosis is greatly inhibited in cells

at 4°C but not at 37°C. Cells were grown to confluency on six well plates, washed with phosphate buffered saline (PBS) and incubated with the fluorescent neomycin and guanidino-neomycin derivates in HAMS F12 media for one hour on ice or at 37°C under an atmosphere with 5% CO_2. Cells were then washed with PBS, released with trypsin or EDTA (10 mM) and FACS analyzed.

Cells incubated with neomycin-BODIPY at 4°C show virtually no increase in fluorescent signal. However, when endocytosis is not inhibited, at 37°C, cells incubated with BODIPY-neomycin show roughly a 10 fold increase in fluorescence (figure 2.9). The same trend of a large increase in signal when going from 4°C to 37°C is also seen with guanidino-neomycin-BODIPY. Cells incubated with guanidino-neomycin-BODIPY on ice show a significant increase in fluorescence when compared to untreated or neomycin-BODIPY treated cells.

A log fold increase in mean cellular fluorescence when going from 4°C to 37°C is indicative that roughly 90% of the fluorescent signal seen with the compounds at 37°C is due to uptake rather than surface binding. It is interesting to note that at the same concentrations neomycin-BODIPY the same log shift is seen, however the guanidino-neomycin-BODIPY binds to and is taken into cells roughly ten times better than neomycin-BODIPY.

While these data establish that guanidino-neomycin is being effectively internalized into cells, the mechanism by which it enters cells has not been fully elucidated. Guanidinoglycosides effectively inhibit binding of FGF to cells expressing heparan sulfate, providing circumstantial evidence of interactions between guanidinium

groups and cell surface heparan sulfate. However, more data is necessary to elucidate the mechanism of cellular uptake of guanidinoglycosides.

To examine whether cell surface binding and subsequent uptake of guanidinoglycosides is dependent on cell surface proteoglycans, uptake studies where cells lacking glycosaminoglycans are compared to wild type cells were performed.

In the assembly of glycosaminoglycans, the first committed sugar transfer reaction is the attachment of xylose to a serine residue of the core protein. This reaction initiates the assembly of glycosaminoglycans *in vivo* and the enzyme responsible for this addition is xylosyltransferase. CHO-pgsA-745 cells are a mutant line with highly decreased xylosyltransferase activity. This decrease in enzyme activity leads to a phenotype where glycosaminoglycans heparan sulfate and chondroitin sulfate are not expressed at any comparable level when compared to wild type cells.[23] This cell line provides a powerful tool for studying the requirements of proteoglycans in multiple systems, including uptake of fluorescent derivatives.

Wild type CHO-K1 cells and CHO-pgsA-745 cells were incubated for 1 hour at 37^{o}C under an atmosphere of 5% CO_2 in a solution of HAMS-F12 media with 1μM neomycin-BODIPY. After washing with phosphate buffered saline (PBS) five times, cells were released from tissue culture plates with 10mM EDTA in PBS for 10 minutes at 37^{o}C/5% CO_2. Released cells were pelleted at 1500 rpm for 4 minutes, resuspended in FACS buffer (1% BSA in PBS) and analyzed via fluorescent activated cell sorting (FACS).

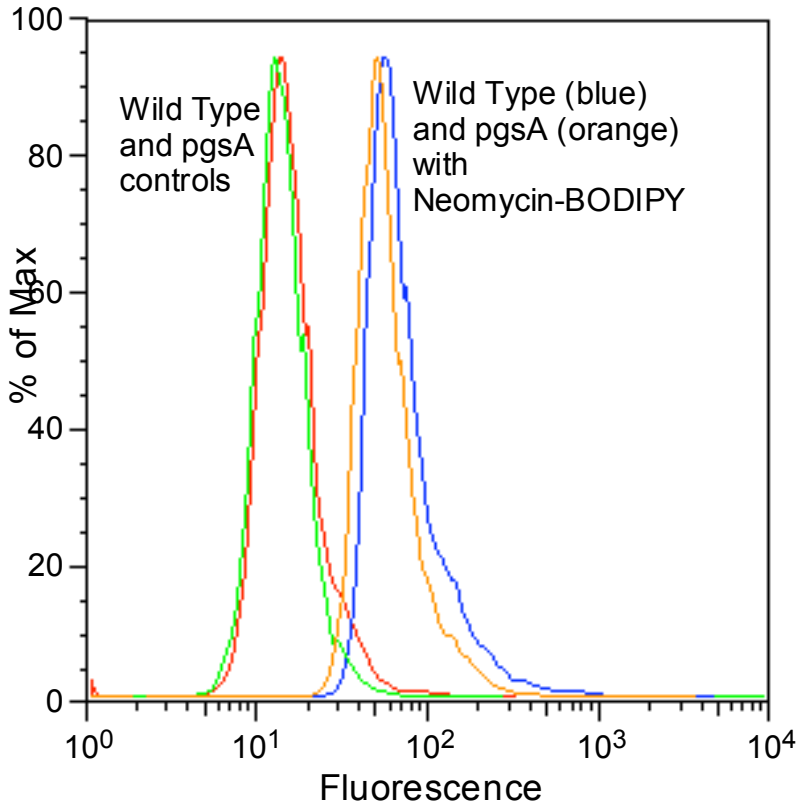


Figure 2.10. Neomycin-BODIPY shows no discrimination between wild type cells and cells lacking heparan and chondroitin sulfate proteoglycans when incubated at 37°C for 1 hour at 1 μM. Interactions between neomycin and the surface of cells is likely not dependent on cell surface proteoglycans.

Wild type and pgsA-745 cells which were not treated with any compound both display a weak fluorescent signal, but wild type cells treated with neomycin-BODIPY show roughly a log fold increase in fluorescence (Figure 2.10). While pgsA-745 cells treated with neomycin-BODIPY show an increased fluorescent signal when compared to untreated cells, there is virtually no difference in signal between treated wild type and pgsA-745 cells (Figure 2.10).

The fact that there is an increased fluorescent signal with treated cells does indicate that there are some interactions between neomycin-BODIPY and the cell

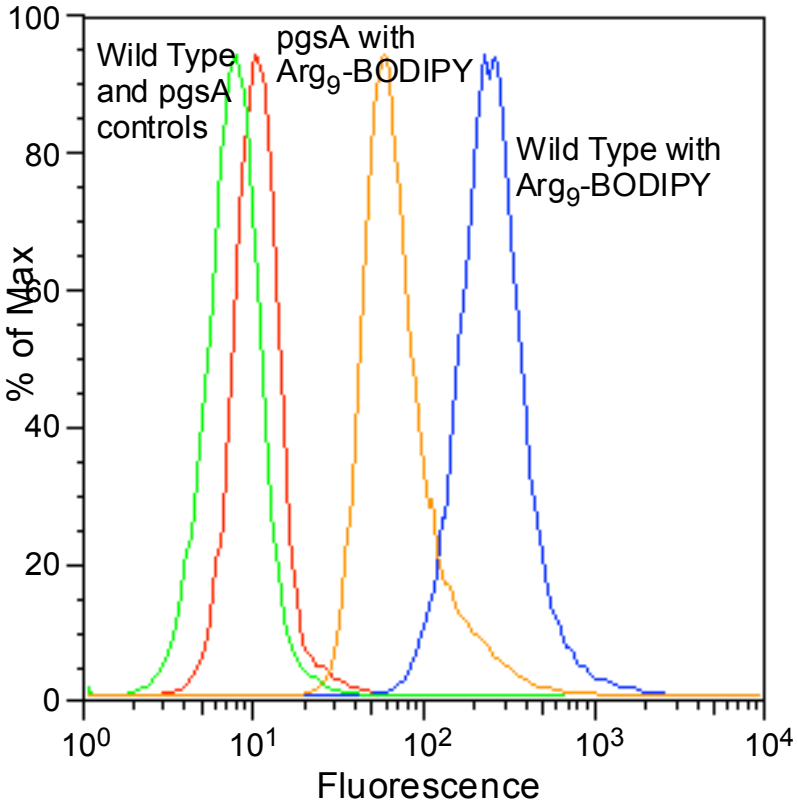


Figure 2.11. Arg_9-BODIPY shows a much stronger signal with wild type cells than with proteoglycan deficient cell lines, indicating that binding and uptake of Arg_9 have a strong dependence on cell surface proteoglycans.

surface. However, the lack of a difference in signal between the wild type and the proteoglycan deficient cells lines shows that these interactions are not dependent on cell surface proteoglycans. This indicates that the interactions between neomycin and both wild type and pgsA-745 cells are independent of cell surface heparan and chondroitin sulfate proteoglycans.

However, when the same experiment is performed with Arg_9-BODIPY there is almost a log shift between the signals of wild type cells and pgsA-745 cell lines (Figure 2.11). The signal from pgsA-745 cells treated with Arg_9-BODIPY is roughly

comparable to the signals from the same cell line treated with neomycin-BODIPY. Mean values from these pgsA-745 cells treated with each of these compounds are comparable. However, while wild type cells treated with neomycin-BODIPY do not show any change in signal, there is roughly an 8-fold signal increase for the same cell lines when treated with Arg_9-BODIPY (Figure 2.11). The magnitude of the increased signal shows that there are approximately 8 times more molecules of Arg_9-BODIPY bound to wild type cells than glycosaminoglycan deficient cells. These data demonstrates that the Arg_9 peptide interacts with living cells in a highly proteoglycan dependent manner.

While this data supports a previously reported mechanism of interaction between Arg_9 and cell surface heparan sulfate proteoglycans,[20] it does not address additional, heparan sulfate independent interactions which are possibly taking place. Interactions with phospholipids[24] or other non-specific/non-proteoglycan dependent interactions may be taking place.

Evidence for these other interactions is shown in the binding of Arg_9-BODIPY to proteoglycan-deficient cells. The increase in signal seen in pgsA-745 cells when incubated with Arg_9-BODIPY is indicative of interactions between the peptide and the cells. Because these cells lack cell surface proteoglycans, there must be a second, proteoglycan-independent mode of binding.

To examine the proteoglycan dependence of guanidino-neomycin interactions with the cell surface, guanidino-neomycin-BODIPY was incubated with wild type and proteoglycan deficient cell lines. Initial results show wild type cells have roughly a 2 fold greater signal than pgsA-745 cells when treated with guanidino-neomycin-

BODIPY (Figure 2.12) which would indicate that while the interactions between guanidino-neomycin and cells is proteoglycan dependent, it shows a weaker dependence on cell surface proteoglycans than the interactions between Arg_9 and the cell surface.

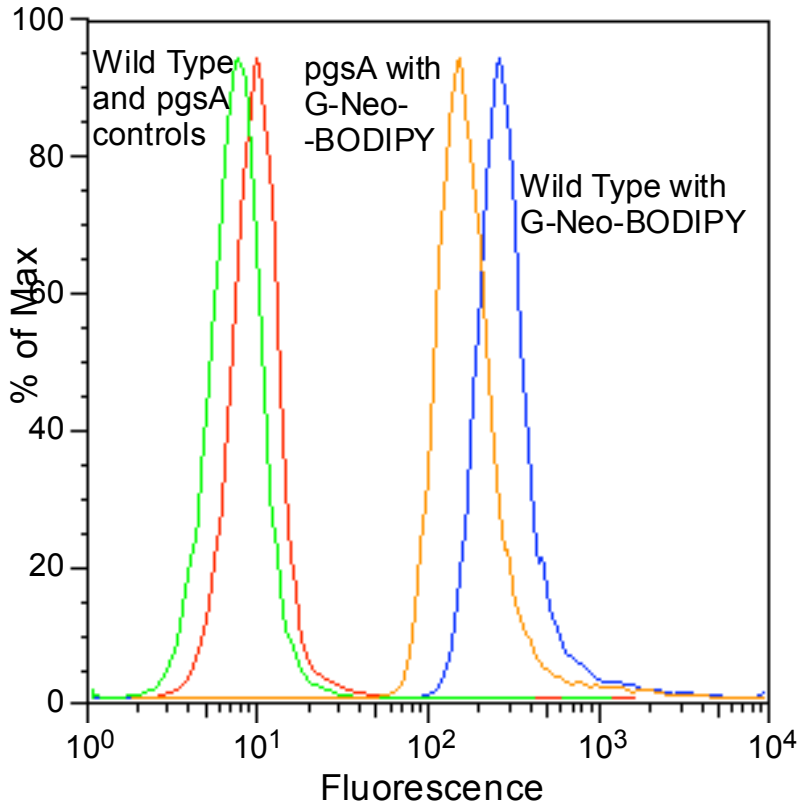


Figure 2.12. When incubated with wild type and pgsA-745 cells and analyzed with flow cytometry, guanidino-neomycin-BODIPY shows selectivity for wild type cells over proteoglycan deficient cells, however, the proteoglycan dependence seen with guanidino-neomycin-BODIPY is not as strong as that seen with Arg_9-BODIPY.

However, previous data has shown that guanidinoglycosides to bind to heparan sulfate, inhibiting FGF binding to the cell surface while neomycin does not (figure 3). Additionally, guanidino-neomycin-BODIPY appears to be taken up into cells much better than neomycin-BODIPY (Figure 2.9). Because the mechanism of binding Arg_9-

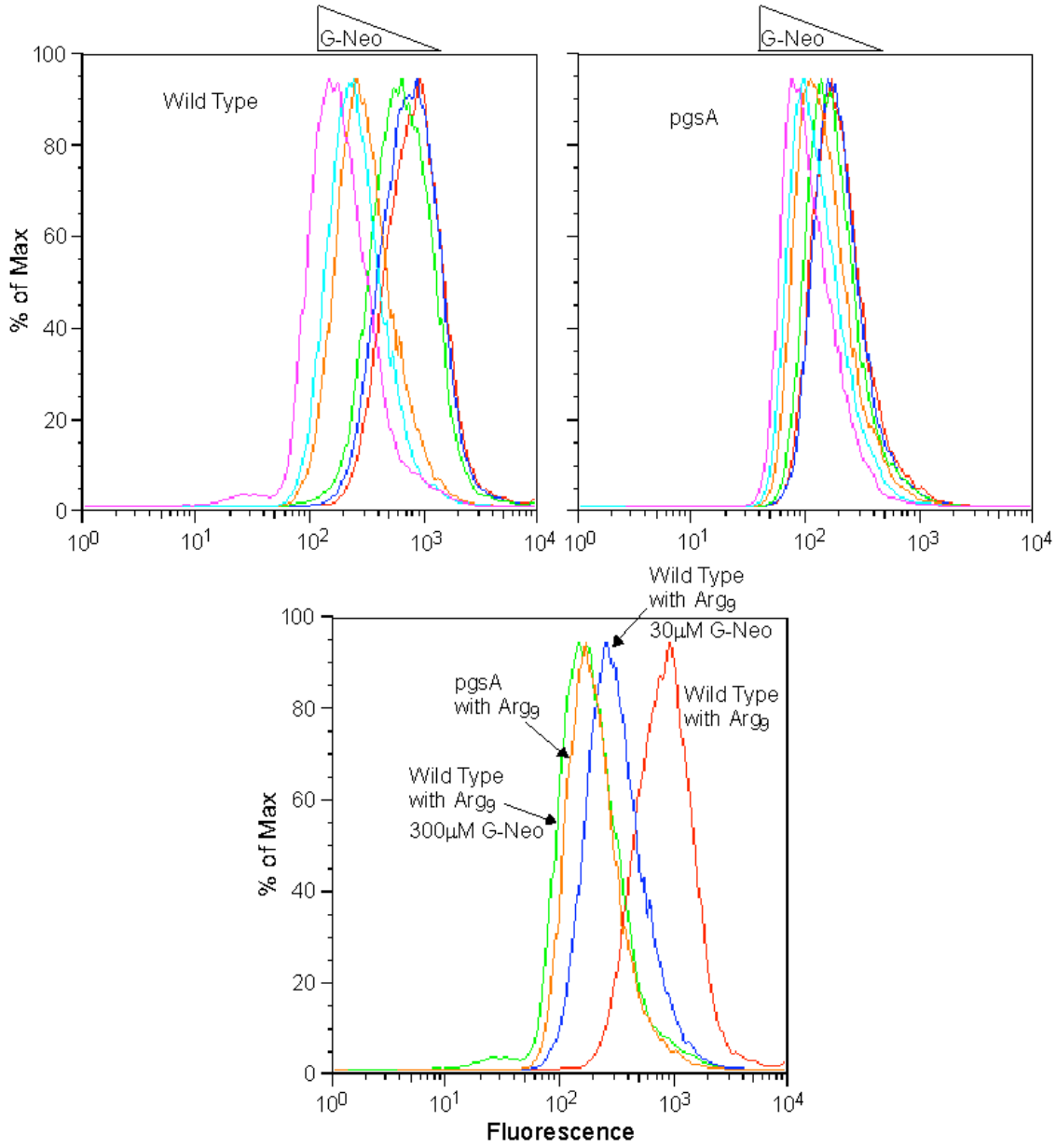


Figure 2.13. Wild type and pgsA-745 cells were incubated in HAMS F-12 media with 0.5 μM Arg_9-BODIPY and increasing concentrations of guanidino-neomycin (1,10,30,100 and 300 μM). After 1 hour at 37°C in an atmosphere of 5% CO_2, cells were washed five times with phosphate buffered saline, lifted with EDTA and analyzed via flow cytometry. While guanidino-neomycin does not effectively compete for Arg_9-BODIPY binding on cells lacking proteoglycans, it shows a strong competition effect with wild type cells. This shows that guanidino-neomycin and Arg_9-BODIPY share a common, proteoglycan dependent mode of cell surface binding which is independent of any other non-proteoglycan dependent binding activity.

BODIPY to the cell surface appears to be at least partially proteoglycans dependent (Figure 2.12), it is important to examine whether guanidino-neomycin can effectively compete for the proteoglycan dependent portion of Arg_9 binding.

To assay the ability of guanidino-neomycin to effectively compete for Arg_9-BODIPY binding to cell surface proteoglycans, wild type CHO-K1 and proteoglycan deficient pgsA-745 cells were incubated with 0.5μM Arg_9-BODIPY and increasing concentrations of guanidino-neomycin (Figure 2.13).

Cells lacking heparan and chondroitin sulfate proteoglycans (pgsA-745) show little to no change in Arg_9-BODIPY signal, even in the presence of 300 μM guanidino-neomycin (Figure 2.13). However, guanidino-neomycin does effectively reduce the fluorescent signal when used as a competitor for Arg_9-BODIPY in wild type cells which do express heparan and chondroitin sulfate. It is interesting to note that the signal reduction effectively saturates well above the signal of untreated cells. In fact, the signal at saturation, which occurs at concentrations of 30μM guanidino-neomycin and higher (100 and 300μM) is almost identical to the signal of pgsA-745 cells treated with Arg_9-BODIPY (Figure 2.11).

This indicates that guanidino-neomycin is able to completely compete for all Arg_9-BODIPY proteoglycan dependent binding sites. However, any secondary binding sites which are independent of cell surface proteoglycans are not affected by guanidino-neomycin, indicating that guanidino-neomycin and Arg_9-BODIPY share proteoglycans dependent binding sites on the cell surface.

2.4 Heparan Sulfate Selectivity of Biotinylated Derivatives

This raises the question of why guanidino-neomycin-BODIPY does not show significant proteoglycan dependence in direct cellular uptake experiments (Figure 2.12) even though guanidino-neomycin will bind to heparan sulfate to inhibit FGF binding (Figure 2.5) and will bind to heparan or chondroitin sulfate to interfere with Arg_9-BODIPY binding and uptake (Figure 2.13). One possible explanation for this

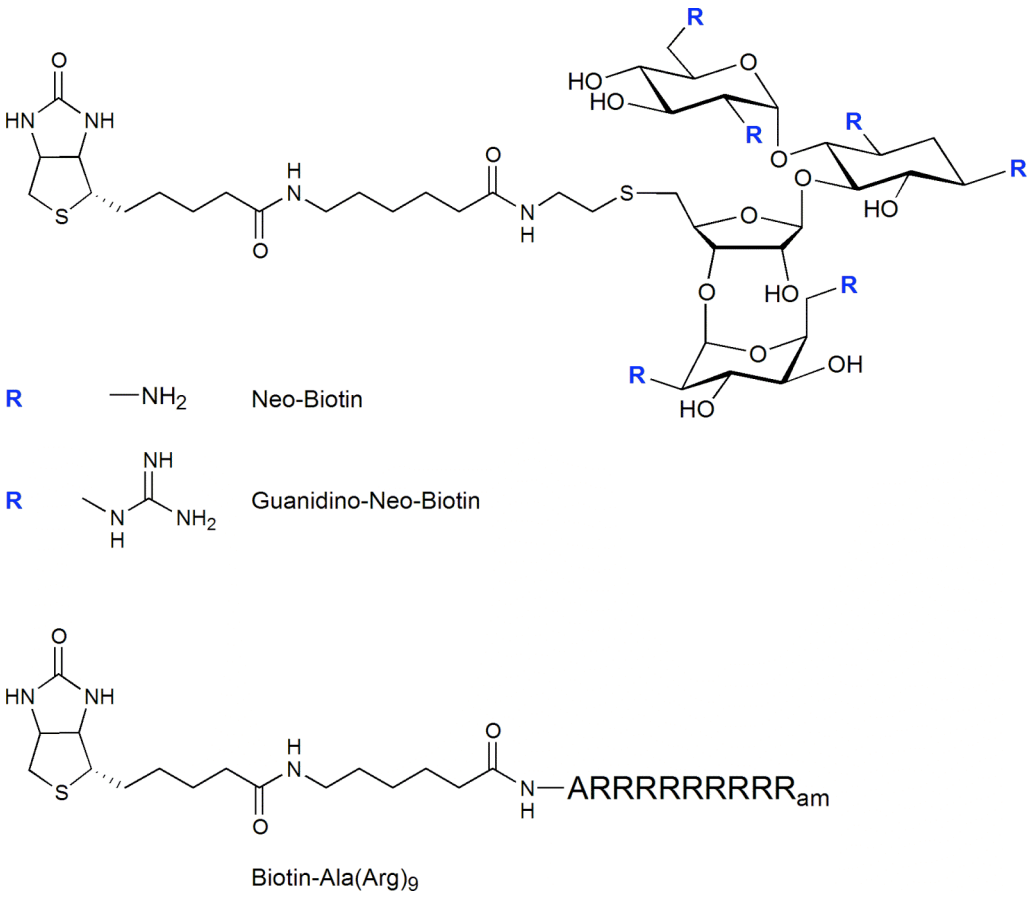


Figure 2.14. Biotinylated neomycin, guanidino-neomycin and arginine-9 peptide provide molecules which can be bound to a variety of streptavidin conjugated molecules and assayed for cell surface binding and uptake using multiple different assays.

Figure 2.15. Synthesis of neomycin-biotin. a) 6-aminohexanoic acid, thionyl chloride, methanol, -10°C (99%) b) biotin, HOBT, HATU, **48**, **49**, triisopropylethylamine, DMF (87%) c) **50**, sodium hydroxide, methanol (44%) d) **51**, **4**, HOBT, HATU, triisopropylethylamine, DMF (63%) e) **52**, trifluoroacetic acid, dichloromethane, triisopropylsilane.

contradictory data is that there are multiple mechanisms of cell surface binding and subsequent uptake for guanidine-rich molecules such as guanidinoglycosides and arginine-rich peptides. Another possibility is that the BODIPY fluorophore is somehow

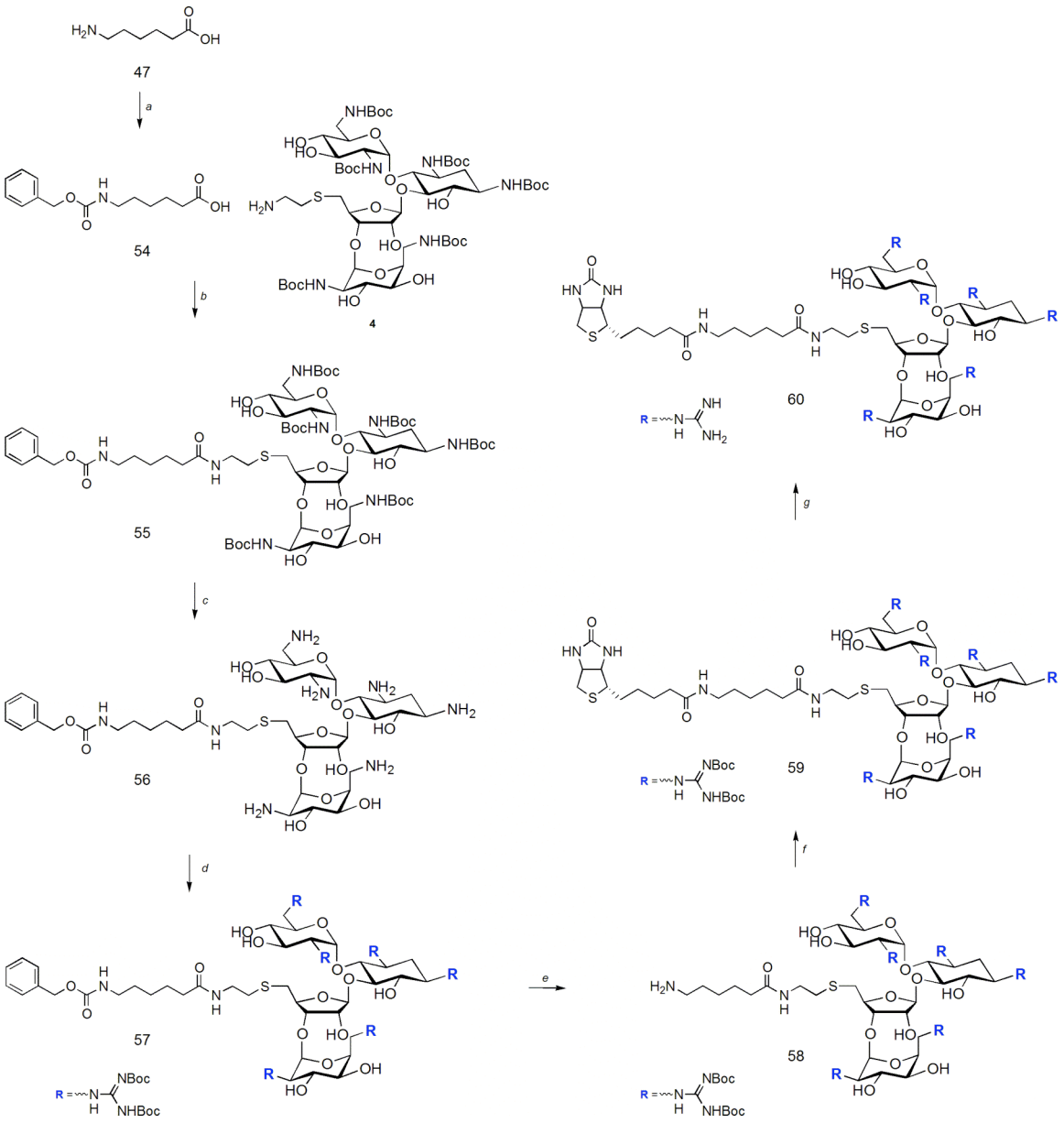


Figure 2.16. Synthesis of guanidino-neomycin-biotin. a) **47**, sodium carbonate, Cbz-Cl, methanol, 4°C (58%) b) **54**, **4**, HOBT, HATU, triisopropylethylamine, DMF (96%) c) **55**, trifluoroacetic acid, dichloromethane, triisopropylsilane (90%), d) **56**, **24**, triethylamine, methanol (55%) e) **57**, palladium on carbon, H_2, methanol (85%), f) **58**, **49**, HOBT, HATU, triisopropylethylamine, DMF (84%) g) **59**, trifluoroacetic acid, dichloromethane, triisopropylsilane (70%).

interfering with the assay through non-specific interactions with membrane phospholipids or membrane bound molecules.

In either case, this highlights the need for molecules which are versatile enough to be used in a variety of different assays to achieve convergent, not divergent data. To that end, we have designed and synthesized a set of biotinylated molecules: neomycin-biotin, guanidino-neomycin-biotin and Arg_9-biotin (Figure 2.14).

Biotin is a water-soluble B-complex vitamin which has an extraordinarily high affinity for the glycoprotein avidin. It also has a comparable affinity for streptavidin, a tetrameric protein which is a derivative of avidin. The biotin-streptavidin complex is extremely stable over a wide range of temperatures and pH, making it an ideal system for generalized assay development. By conjugating biotin to molecules such as Arg_9, neomycin and guanidino-neomycin we have developed a set of molecules which can be used in a variety of assays with different streptavidin bound proteins and fluorophores.

There is considerable evidence that guanidinylated compounds such as arginine-rich peptides have the ability to mediate the cellular uptake of large molecules such as carbonic anhydrase (29 kDa), β-galactosidase (120 kDa), DNA oligonucleotides, magnet beads (>50 nm diameter), chelators with radioisotopes, and liposomes (>200 nm diameter),[14,25-30] as well as other proteins and fluorophores. Based on competition data with Arg_9, (Figure 2.13) it is likely that guanidinoglycosides will exhibit similar delivery activity. However, when switching from a small conjugated fluorophore system of analysis to a large streptavidinylated fluorophore, differentiation of cell surface binding and uptake must be re-validated.

To examine if biotinylated guanidino-neomycin can bind to the surface of cells and carry large molecules across the membrane, a conjugate of guanidino-neomycin-biotin and streptavidin-Phycoerythrin-cychrome (Streptavidin-PE-Cy5) was prepared and wild type cells were incubated under 2 mL of the complex incubated at 4^{o}C and 37^{o}C for 1 hour. At 37^{o}C endocytosis is not inhibited, but at 4^{o}C, cells are unable to absorb compounds through endocytosis. A solution of 1 μM guanidino-neomycin-biotin in 10mL HAMS F-12 media was prepared. To this solution, 10 μL of Streptavidin-PE-Cy5 (0.2 mg/mL) was added and this mixture was allowed to incubate at room temperature for 40 minutes to ensure that the biotin-streptavidin interaction was complete, and then diluted to working concentrations. All experiments performed with streptavidin-PE-Cy5 were incubated in this manner to ensure identical concentrations and guanidino-neo/PE-Cy5 ratios.

CHO-K1 wild type cells were washed once with phosphate buffered saline, and incubated with guanidino-neomycin-biotin streptavidin-PE-Cy5 in HAMS F12 (60 nM). Media for 4^{o}C incubation was pre-chilled on ice. Cells were then incubated at 4^{o}C or 37^{o}C for 1 hour before being washed five times with phosphate buffered saline, lifted with 10 mM EDTA, and analyzed via flow cytometry.

As with the BODIPY compounds, there is roughly a log fold increase in signal from cells incubated at 4^{o}C to those incubated at 37^{o}C (Figure 2.17). The molecular weight of guanidino-neomycin-BODIPY compounds is roughly 2 kDa, with the majority of the mass coming from guanidinoglycoside. Guanidino-neomycin-biotin has roughly the same mass as the BODIPY compound; 2 kDa, but when complexed to

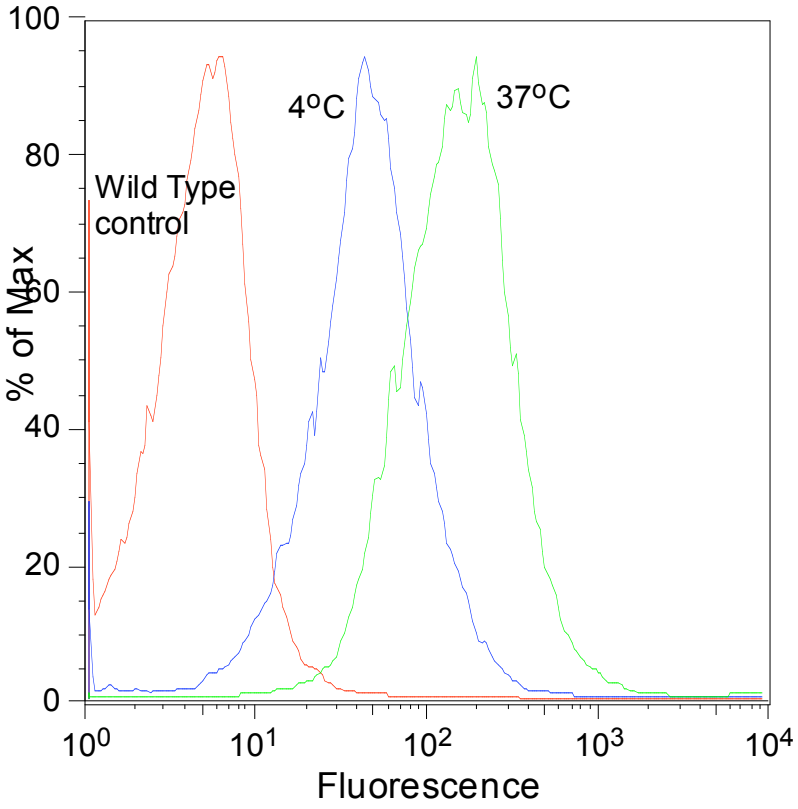


Figure 2.17. Wild type cells were incubated with guanidino-neomycin-biotin bound to streptavidine-PE-Cy5 at 4°C and 37°C for one hour, washed with phosphate buffered saline, released with EDTA and analyzed with flow cytometry. Endocytosis is inhibited at 4°C but not at 37°C. When guanidino-neomycin-biotin is bound to a large fluorophore (streptavidin-PE-Cy5), it retains both cell binding and uptake properties. The increase in signal between incubations at 4°C and 37°C indicates that the guanidino-neomycin-biotin Streptavidin-PE-Cy5 complex is being taken into CHO-K1 cells via endocytosis.

streptavidin-PE-Cy5 the complex has a molecular weight of a little over 300 kDa. This signal increase indicates that guanidino-neomycin can, when conjugated to large cargo not only bind to the surface of cells, but also cross cell membranes and deliver large cargos.

To address the mechanism of uptake of guanidino-neomycin, wild type and proteoglycans deficient cell lines were incubated with increasing concentrations of guanidino-neomycin-biotin streptavidin-PE-Cy5 (10, 30, 100, 300 nM and 1 μM).

After 1 hour at 37°C in a 5% CO_2 atmosphere, cells were washed five times with phosphate buffered saline, released with 10 mM EDTA for 10 minutes at 37°C and analyzed via flow cytometry.

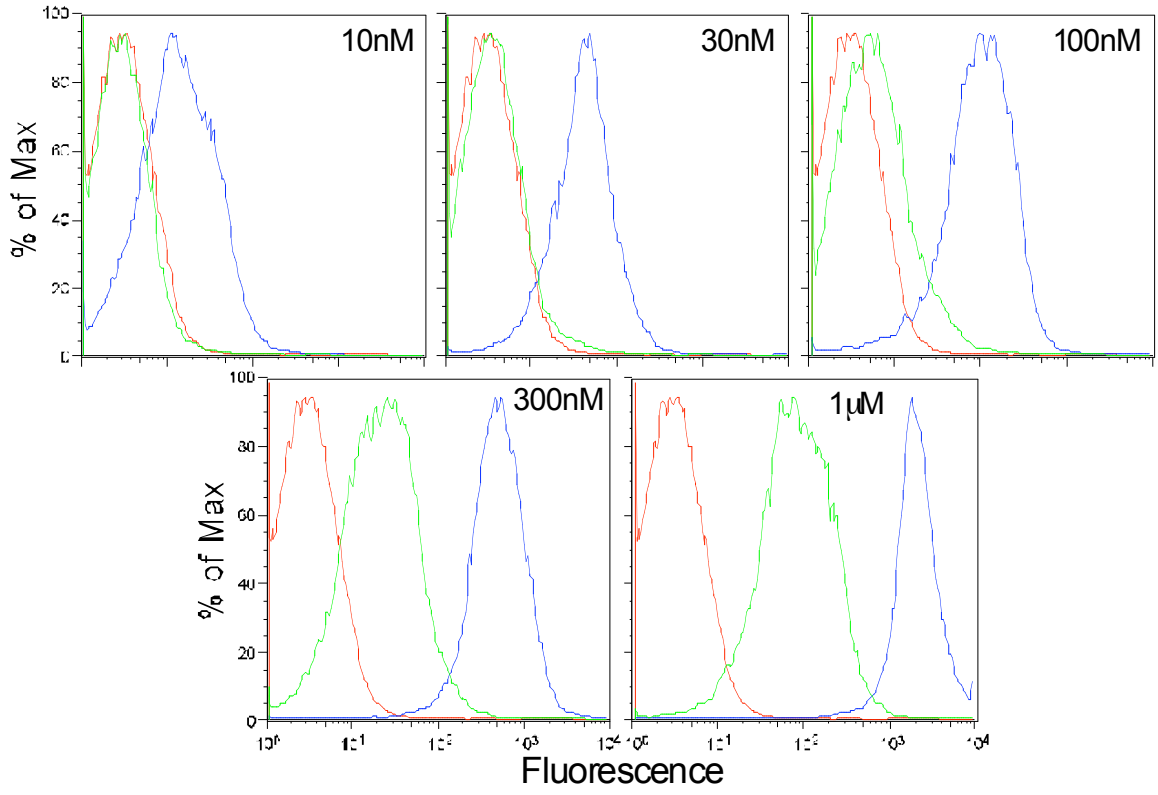


Figure 2.18. Wild type (blue) and pgsA-745 (green) cells were incubated with the guanidino-neomycin-biotin-streptavidin-PE-Cy5 conjugate at different concentrations for one hour at 37°C. After washing with phosphate buffered saline and release with EDTA, cells were analyzed with flow cytometry. Guanidino-Neomycin shows a strong preference for wild type over proteoglycan deficient cell types, at lower concentrations (30-100 nM). However, at higher concentrations (300 nM – 1 μM) a second, non-proteoglycan dependent mechanism of uptake is seen.

Using this system, a much clearer picture of the mechanism for uptake of guanidinoglycosides begins to emerge. When incubated with the guanidino-neomycin PE-Cy5 system, wild type cells (CHO-K1) show a response at concentrations as low as

10 nM (figure 16). However, cells lacking heparan and chondroitin sulfate proteoglycans (CHO-pgsA-745) do not start to show an increase in fluorescent signal until the guanidino-neomycin concentration reaches between 100-300 nM (Figure 2.18).

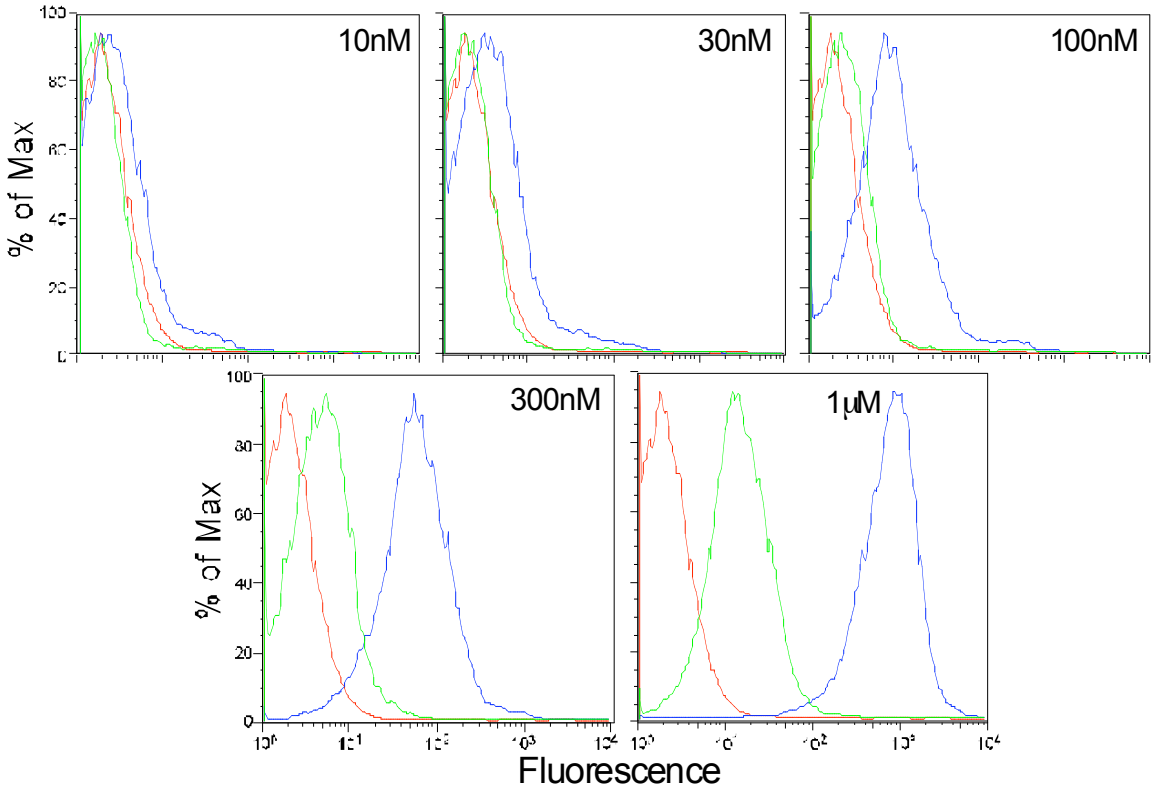


Figure 2.19. Neomycin-biotin streptavidin-PE-Cy5 shows a proteoglycan dependence similar to that seen with the guanidino-neomycin system. While, the neomycin systems appears to have roughly a 10-fold lower affinity for wild type cells at all but the highest (1 μM) concentration, the fluorescent signal for pgsA-745 cells lines is roughly equivalent to that of guanidino-neomycin.

At higher concentrations (300 nM and 1 μM) there is a significant increase in the guanidino-neomycin fluorescent signal in the pgsA-745 cell lines. It is interesting to note, that the magnitude of difference in fluorescent signals between the

wild type and proteoglycans deficient cells lines appears to be roughly fixed at 1.5 orders of magnitude. However, as concentrations of guanidino-neomycin were increased it is likely that secondary receptors may become saturated and the signals may converge. These data shows that the mechanism of uptake of guanidino-neomycin is dependent on cell surface proteoglycans at low concentrations. However, at higher concentrations there appears to be a second, non-proteoglycan dependent binding event.

When the same experiment is performed with neomycin-biotin streptavidin-PE-Cy5, little to no signal is seen from wild type cells to almost 100 nM. At higher concentrations, this neomycin conjugate does appear to bind to wild type cells (figure 17). However, only at 1 μM concentrations does the signal approach that of the guanidino-neomycin system with a mean value of roughly 50% of the value measured with guanidino-neomycin at the same concentration. At lower concentrations, the signal from neomycin is roughly equal to 10% of the guanidino-neomycin signal at equivalent concentration points. It is possible that while guanidino-neomycin has a higher affinity for cells expressing surface proteoglycans than neomycin, at high enough ligand concentrations saturating conditions are reached, lowering the apparent cellular discrimination between amino and guanidino-glycosides.

When comparing the fluorescent signals from the neomycin and guanidino-neomycin systems with pgsA-745 cells, it was found that they are roughly equivalent, suggesting that neomycin and guanidino-neomycin have approximately the same affinity for cells lacking heparan and chondroitin sulfate proteoglycans. Based on this data, we propose a model of binding where guanidinoglycosides interact with cell surface proteoglycans. These interactions facilitate cellular uptake and enable

guanidinoglycosides to be much more efficiently incorporated into the cell than their aminoglycoside parent molecules.

There appears to be a second, non-proteoglycan dependent mechanism of cell surface binding which is not discriminatory between amino and guanidino-glycosides.

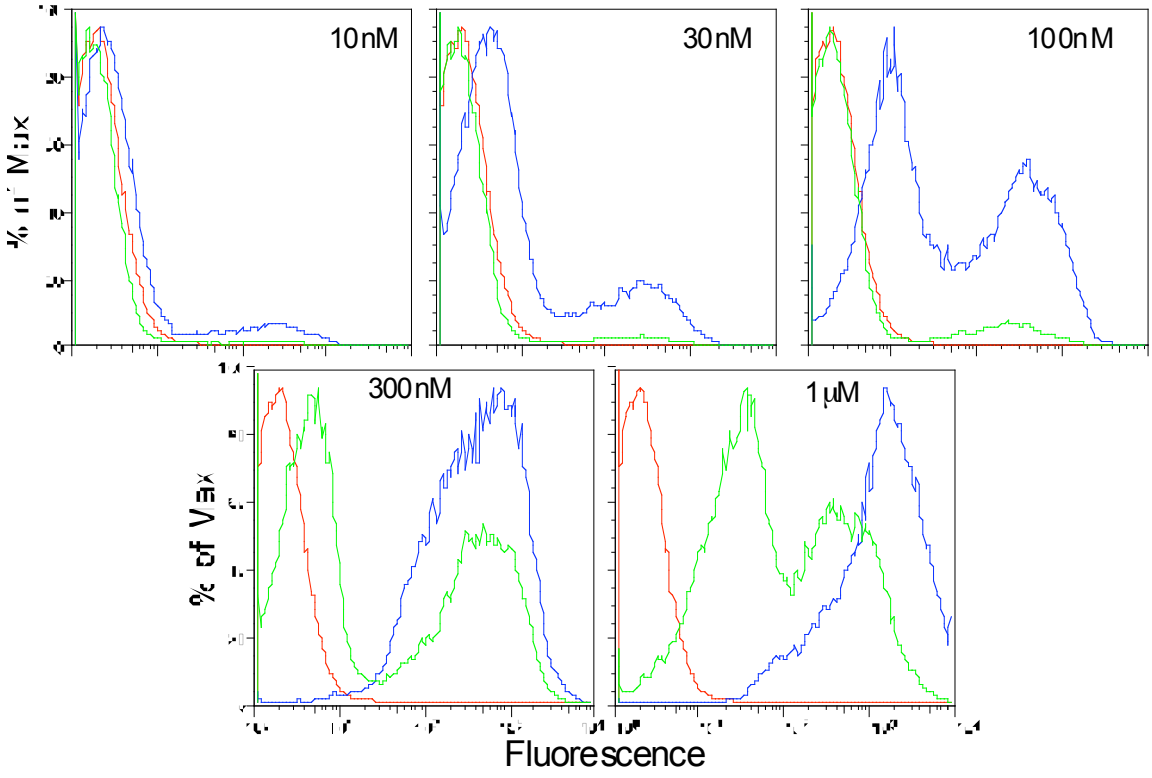


Figure 2.20. The arginine-9-biotin streptavidin-PE-Cy5 system shows proteoglycan dependence on cell surface binding and internalization. However, multiple populations are observed with wild type and proteoglycan deficient cell lines indicating multiple Arg_9 receptors on the cell surface.

This second binding mode, however, occurs at higher concentrations than the proteoglycan dependent mode. It is possible that there is another class of receptors which interact with guanidinoglycosides and other charged molecules. This secondary

mode of binding and uptake has a lower affinity for guanidinoglycosides than heparan sulfate. Additionally, it appears to be less selective as well, binding to aminoglycosides as well as guanidinoglycosides. There is evidence that these interactions take place between charged phospholipids in the cell membrane and guanidinium groups.[31]

2.5 The Arginine-9 Peptide Has Multiple Mechanisms of Uptake

It is important to compare the proteoglycans dependence of guanidinoglycoside uptake to that of arginine-rich peptides. We chose to use the Arg_9 peptide because it has been utilized as a standard for polyarginine peptide uptake into cells. A mechanism of arginine-rich peptide uptake which has been gaining considerable support involves Arg_9 interactions between the peptide and cell surface heparan sulfate proteoglycans facilitating receptor mediated endocytosis.[20] A key data set which supports this theory is based on comparison of individual cells rather than large populations. Wild type CHO cells and cells lacking cell surface proteoglycans were incubated with fluorescent tagged Arg_9 peptide and examined via fluorescence microscopy.[20] These data provide evidence of the proteoglycans dependence of Arg_9 uptake. However, when examined via flow cytometry, where it is possible to examine large populations of cells, slightly different results were achieved.

Wild type and proteoglycan deficient cells were incubated with Arg_9-biotin streptavidin-PE-Cy5 for 1 hr at 37^{o}C in an atmosphere of 5% CO_2. After 5 washes with phosphate buffered saline, cells were lifted with EDTA and examined via flow cytometry. As with neomycin and guanidino-neomycin, the Arg_9 peptide binds to wild

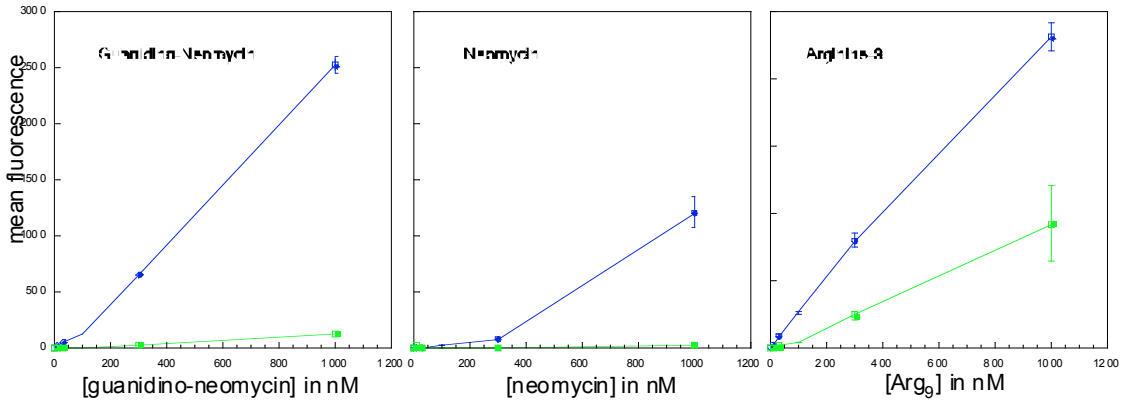


Figure 2.21. Mean fluorescent values for guanidino-neomycin-, neomycin- and Arg_9- PE-Cy5 conjugates with wild type and proteoglycan deficient cell lines demonstrate that guanidino-neomycin has a much higher proteoglycan selectivity than Arg_9 with a comparable affinity to wild type cells.

type cells at low to mid nanomolar concentrations (Figure 2.20). However, unlike neomycin and guanidino-neomycin, Arg_9 appears to have multiple binding modes within the same population of cells. Multiple populations are observed when Arg_9 is incubated with wild type cells (Figure 2.20). When this experiment was repeated (data not shown) two main populations of Arg_9 were consistently seen. This is indicative of multiple mechanisms of binding, something not seen with either neomycin or guanidino-neomycin.

As with the wild type cells, pgsA-745 cells also display multiple populations when incubated with Arg_9 (Figure 2.20). While little to no Arg_9 binding is observed with proteoglycans deficient cells at concentrations up to 100 nM, at higher concentrations a second population begins to emerge. The first population appears to be highly proteoglycans dependent, with a signal much weaker than that of wild type cells. The second population, however does not appear to be very dependent on cell

surface proteoglycans, presenting a signal which is roughly equal to the signal seen with wild type cells.

While Arg_9 does show proteoglycans dependence in cell surface binding and uptake, these experiments provide evidence of multiple mechanisms of cell surface binding and/or internalization. At higher peptide concentrations (above 300 nM) where one class of Arg_9 receptors may become saturated, populations begin to converge, but at lower concentrations, multiple populations are observed, indicating multiple mechanisms of Arg_9 binding to both wild type and proteoglycans deficient cells.

This highlights the advantage of a fluorescent activated cell sorting analysis. By using flow cytometry, it is possible to get data on thousands of cells in one experiment rather than examining the single cells or smaller populations used with microscopic techniques. This approach minimizes errors due to aberrations in binding or multiple populations.

Analysis of the mean fluorescence values for titrations of guanidino-neomycin-, neomycin-, and Arg_9-biotin stretpavidin-PE-Cy5 into wild type and proteoglycan deficient pgsA-745 cells provides insight into the differences in binding and uptake mechanisms. Guanidino-neomycin and neomycin constructs both display proteoglycan dependence (Figure 2.21). However, guanidino-neomycin has a considerably higher efficiency of uptake than neomycin. At micromolar concentrations, neomycin begins to show a large increase in signal, and while this appears to be proteoglycan dependent, it appears at relatively high concentrations were cell surface proteoglycans may start to become saturated.

2.6 Analysis of the Heparan Sulfate Dependent Uptake of Guanidino Neomycin

Arg_9 displays efficient uptake in wild type cells (Figure 2.21), but it is clear, even at lower concentrations that it also has a non-proteoglycan dependent interaction with cells. The signal seen with Arg_9 when incubated with pgsA-745 cells is much higher than the signals from neomycin and guanidino-neomycin at the same concentrations. While Arg_9 may be incorporated into wild type cells at slightly higher levels than guanidino-neomycin, it is likely that this is due to both proteoglycan dependent and independent interactions with cells.

A potential source for error could be in the assumption that, by genetically removing the activity of an enzyme from a cell, there are no other effects. There is potential that a secondary effect of genetic modification could affect small molecule binding and/or uptake. To that end, a plasmid containing the gene for the xylotransferase enzyme was transfected into pgsA-745 cells. These cells regain the ability to attach xylose to serine residues on the core proteoglycan protein, thus allowing for fully expressed heparan sulfate and chondroitin sulfate proteoglycans.

These cells were then assayed for guanidinoglycoside binding and uptake. PgsA-745 and "restored" pgsA-745 cells were incubated with 60 nM guanidino-neomycin-biotin streptavidin-PE-Cy5 in HAMS F-12 for 1 hour at 37^{o}C in an atmosphere with 5% CO_2. After 5 washes with phosphate buffered saline and release with 10 mM EDTA, cells were analyzed via flow cytometry (Figure 2.22).

When incubated with guanidino-neomycin at 60 nM, pgsA-745 cells show almost complete loss of fluorescent signal. However, when xylotransferase activity is

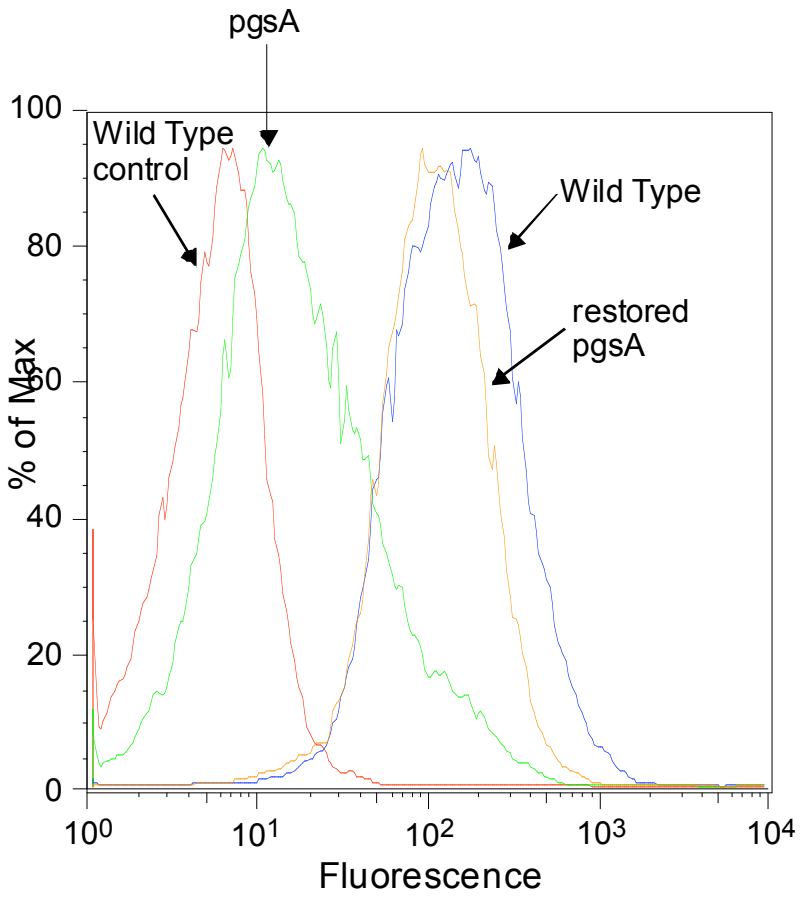


Figure 2.22. Wild type, pgsA-745 and "restored" pgsA-745 cells were incubated with guanidino-neomycin-biotin-streptavidin-PE-Cy5 for one hour at 37°C before being washed with phosphate buffered saline, released with EDTA and analyzed with flow cytometry. When the gene for xylotransferase is re-introduced into pgsA-745 cells, they are able to recover guanidino-neomycin binding activity. The fluorescent signal of "restored" pgsA-745 cells is almost equivalent to that of wild type cells, indicating that cell surface proteoglycans are responsible for the cellular binding activity of guanidino-neomycin.

restored, re-establishing the expression of cell surface heparan sulfate and chondroitin sulfate proteoglycans, the signal is increased to roughly the equivalent of the signal seen with wild type cells. These data support the proposed mechanism of uptake where guanidinoglycosides bind to cell surface proteoglycans, and are internalized into cells

via receptor mediated endocytosis, and not through a secondary, non-proteoglycan dependent mechanism.

There are, however, multiple types of cell surface proteoglycans. Heparan sulfate serves as a co-receptor for many factors, and has been implicated in multiple signaling pathways, while chondroitin sulfate appears to serve mainly a biophysical role, with sulfates and other negatively charged groups interacting with cations and absorbing water to aid in the plasticity of connective tissues.

While they have different biological roles, heparan and chondroitin sulfate have similar structural features. Chondroitin sulfate and heparan sulfate polymerization start with the same linkage tetrasaccharide being *O*-linked to specific serines on the core proteins. A single α(1,4)-linked GlcNAc unit is then added to the tetrasaccharide which commits this intermediate to the formation of heparan sulfate. If a β(1,4)-linked GalNAc is added to the tetrasaccharide, chondroitin sulfate formation is initiated instead.

The rest of the growing glycosaminoglycan chain consists mostly of a repeating disaccharide unit which is added one sugar at a time. In chondroitin sulfate, the disaccharide is composed of D-glucuronic acid and *N*-acetyl-D-galactosamine, and in heparan sulfate it is D-glucosamine (or L-iduronic acid) and *N*-acetyl-D-gluocsamine (Figure 1.13). Heparan sulfate and chondroitin sulfate both contain similar sulfated saccharides with slightly different linkages, however, their activities and known binders are markedly different.

There are currently hundreds of known heparan sulfate binding proteins, many of which serve critical physiological roles. Molecules such as fibroblast growth factor

(FGF) bind specifically to heparan sulfate proteoglycans (Figure 2.4), and will not bind to chondroitin sulfate even in the absence of heparan sulfate chains. Ligand binding depends on the composition of heparan sulfate chains on the proteoglycans, giving these chains potentially very high ligand selectivity.

To further examine the mechanism and potential proteoglycan selectivity of guanidino-neomycin binding and internalizing into living cells, a mutant Chinese hamster ovarian cell line that does not produce heparan sulfate was examined for guanidino-neomycin binding activity. The CHO-pgsD-677 mutant lacks both *N*-acetylglucosaminyl- and glucuronosyltransferase, two enzymes required for the polymerization of heparan sulfate chains. While they lack the ability to produce heparan sulfate, pgsD-677 cells accumulate roughly 3- to 4- fold more chondroitin sulfate than wild type CHO-K1 cells.

When incubated with wild type, pgsA-745 and pgsD-677 cells at 60 nM, neomycin-biotin streptavidin-PE-Cy5 shows little to no difference in signal, indicative of a lack of binding to any of the cell lines (Figure 2.23). This is not unexpected as neomycin does not begin to show significant binding to wild type cells or proteoglycan deficient cells until roughly 100 and 300 nM respectively. The lack of an increased signal when incubated with pgsD-677 cells indicates that neomycin has no large inherent affinity for chondroitin sulfate.

Guanidino-neomycin, however does show a strong proteoglycan dependence, binding to and inducing almost 100-fold increased signal in wild type cells when compared to cells lacking surface proteoglycans. When pgsD-677 cells are incubated with 60 nM guanidino-neomycin, they display no signal increase over pgsA-745 cells

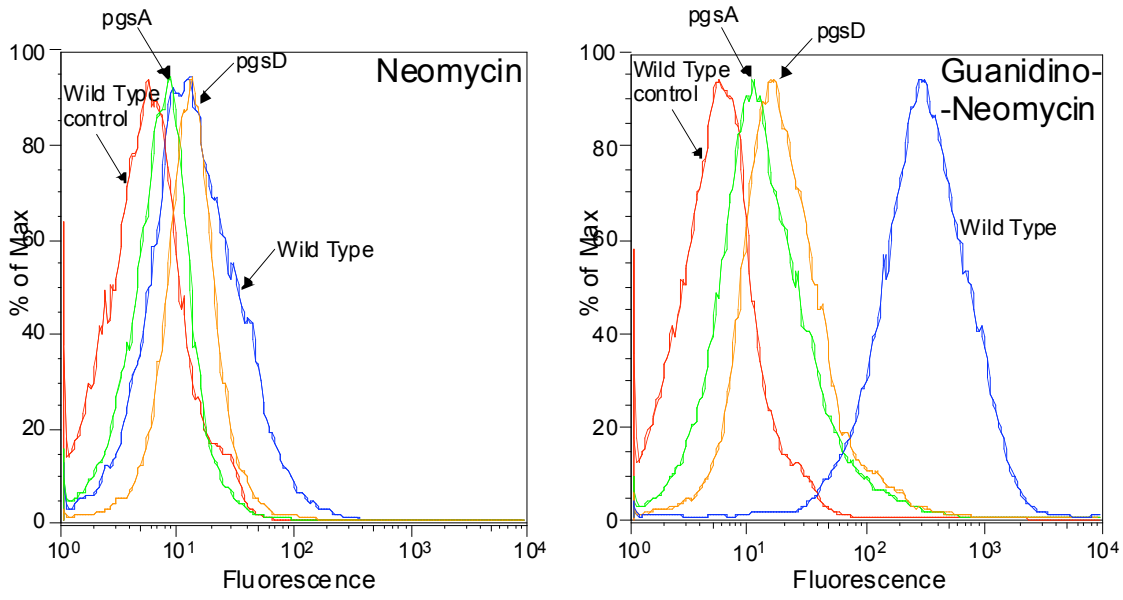


Figure 2.23. While neomycin-biotin streptavidin-PE-Cy5 does not bind to cells at 60 nM regardless of the levels of cell surface proteoglycans, guanidino-neomycin-biotin streptavidin-PE-Cy5 will bind to an internalize into cells in a heparan sulfate dependent manner.

(Figure 2.23). This shows that the binding of guanidino-neomycin to the cell surface is not only proteoglycan dependent, but specific to heparan sulfate proteoglycans.

Heparan sulfate proteoglycans are expressed in a tissue specific manner with specific chain structure varying based on the type of cell on which it is expressed. Ligand binding by proteoglycans is a selective process depending on the fine structure of the heparan sulfate chain. The disaccharides of heparan sulfate are distinguished by the presence of GlcA/IdoA and GlcN residues. These residues can be sulfated but the degree and positioning on the sulfate groups varies greatly and helps to define binding sites for specific ligands.

It is unlikely that there is a specific binding domain on heparan sulfate for guanidinoglycosides. By examining cell lines with mutations within the heparan sulfate

biosynthesis pathway, however, we are able to better define the specific binding requirements for guanidinoglycosides. To that end, guanidinoglycoside binding and uptake were examined with a mutant cell line with overall reduced sulfation on heparan sulfate proteoglycans.

Heparan sulfate *N*-sulfotransferase catalyzes the transfer of sulfate groups from adenosine 3′-phosphate, 5′-phosphosulfate to the free amino groups of glucosamine residues in heparan sulfate. This enzyme activity is a key determinant in the overall extent of sulfation of heparan sulfate. CHO-pgsE-606 is a Chinese hamster ovarian cell line which is 3-5 fold defective in *N*-sulfotransferase activity.[32] This decreased activity gives rise to heparan sulfate proteoglycans in which roughly 20% of the glucosamine residues are *N*-sulfated, as opposed to 40% in the wild type. Additionally, while spacing between *N*-sulfated glucosamine residues is generally between 1-4 disaccharides in wild type cells, it is increased to over 7 in pgsE-606 cells.

The extent of *O*-sulfation is also reduced in pgsE-606. The decrease in *O*-sulfation parallels the decrease in *N*-sulfation, providing evidence that while there are multiple *O*-sulfotransferases, it appears that they share a preference for *N*-sulfated substrates. This is supported by the fact that *O*-sulfation generally takes place on residues near *N*-sulfated glucosamine residues.[33] This leads to a phenotype of pgsE-606 cells where there is an overall decrease of sulfation on heparan sulfate chains without the loss of sulfate groups on any specific residues within the chain.[32]

Heparan sulfate chains generally contain tracts of sulfated saccharides whose specific sequence and sulfation allow for highly specific binding domains which are recognized by binding proteins such as FGF (Figures 1.15 and 1.16). In the pgsE-606

mutant, the numbers of these sulfated binding domains are reduced. However, the fine structure and specific sulfation on the remaining tracts does not appear to be affected.[32]

PgsE-606 cells were incubated with Guanidino-neomycin-biotin streptavidin-PE-Cy5 for 1 hour at 37°C in a 5% CO_2 atmosphere, released with 10 mM EDTA and analyzed via flow cytometry. Very little signal difference is seen between pgsE-606 and wild type cell lines at guanidino-neomycin concentrations ranging from 10 nM to 1 μM.

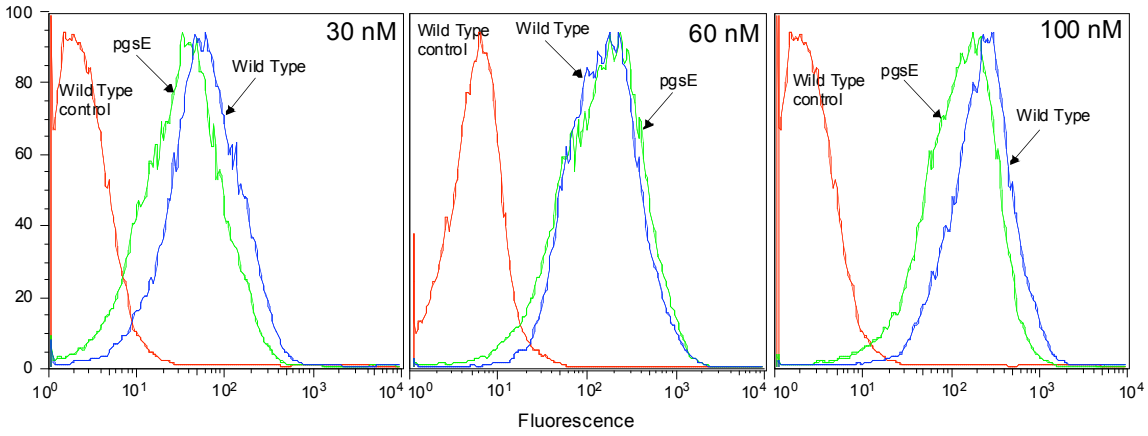


Figure 2.24. pgsE-606 cells express roughly half the sulfation levels of wild type cells. This decreased sulfation is due to fewer sulfated domains as opposed to sulfation differences within these domains. This causes guanidino-neomycin to show almost no selectivity in binding between the wild type and lowered sulfation cell lines, with the difference in binding being almost exactly proportional to the difference in overall sulfation.

Repetitions of this experiment reveal that guanidino-neomycin binds to pgsE-606 cells with roughly the same affinity as it binds to wild type cells (Figure 2.24).

These results are indicative of a model of guanidinoglycoside binding to heparan sulfate proteoglycans on the surface of cells where lowered sulfation levels have little effect on ligand binding. Guanidinoglycosides will recognize sulfated tracts

or domains within the heparan sulfate chain. As fewer of these tracts are expressed on the surface of cells, guanidinoglycoside binding may be proportionally affected; however, the inherent affinity of guanidinoglycosides for sulfated domains within the heparan sulfate chain remains unchanged.

It is not possible to examine the heparan sulfate dependent effects of saturating amounts of guanidinoglycosides on wild type and pgsE-606 cells because as the concentrations of guanidino-neomycin are increased, a second, non-proteoglycan dependent mode of binding begins to emerge (figure 2.18). Signals from binding to this other class of receptors will overshadow proteoglycan dependent signals at concentrations above 1 μM. This raises the question of how guanidinoglycoside binding to cells expressing heparan sulfate would be affected by loss of sulfation at specific positions on the heparan sulfate chain. To examine how loss of specific sulfate groups influences binding, guanidinoglycosides were incubated with CHO-pgsF-17 cells.

This mutant Chinese hamster ovarian cell line has no 2-*O*-sulfotransferase activity, thus making it defective in 2-*O*-sulfation of iduronic acid. Heparan sulfate chains are composed of disaccharides of D-glucuronic acid (GlcA) or L-iduronic acid (IdoA) and D-glucosamine (GlcN) units (Figure 1.13). There is no GlcA sulfation, and only one possible IdoA sulfation, at the 2-position. Therefore, a loss of 2-*O*-sulfotransferase activity would lead to heparan sulfate chains where there is no sulfation on at least one sugar out of every disaccharide, greatly decreasing overall sulfate density.

Heparan sulfate chains in the pgsF-17 cell lines will retain roughly the same overall numbers of sulfated domains. However, because of the lack of 2-*O*-sulfation, each of these sulfated tracts will have an overall decrease in sulfate density. There is a possibility that guanidinoglycosides, which are much smaller than heparan sulfate binding proteins (FGF ≈ 20 kDa) may recognize a shorter segment on the heparan sulfate chain, and thus may be sensitive to loss of sulfation on alternating sugar residues.

To examine the possibility of this sensitivity, guanidino-neomycin-biotin streptavidin-PE-Cy5 was incubated with pgsF-17 cells for 1 hour at 37^{o}C in HAMS-F12 media under an atmosphere of 5% CO_2. After washing five times with PBS and releasing with 10 mM EDTA, cells were analyzed via flow cytometry.

The fluorescent signal from guanidino-neomycin binding is roughly half as strong in pgsF-17 cells than the fluorescent signal seen in wild type cell lines, indicating that guanidino-neomycin has a lower affinity for pgsF-17 cells than for wild type cells. This highlights the importance of sulfate group density on heparan sulfate chains for guanidino-neomycin binding (Figure 2.25).

When compared to the pgsE-606 cell lines, heparan sulfate proteoglycans on pgsF-17 cells express more overall sulfate groups. However, because there is no decrease in the number of sulfated tracts, the overall sulfate density on the pgsF-17 heparan sulfate proteoglycans is lower than that on pgsE-606 cells. The data showing that there is a greater decrease in guanidino-neomycin binding to 2-*O*-sulfotransferase deficient cell lines than to cells with decreased *N*-sulfotransferase

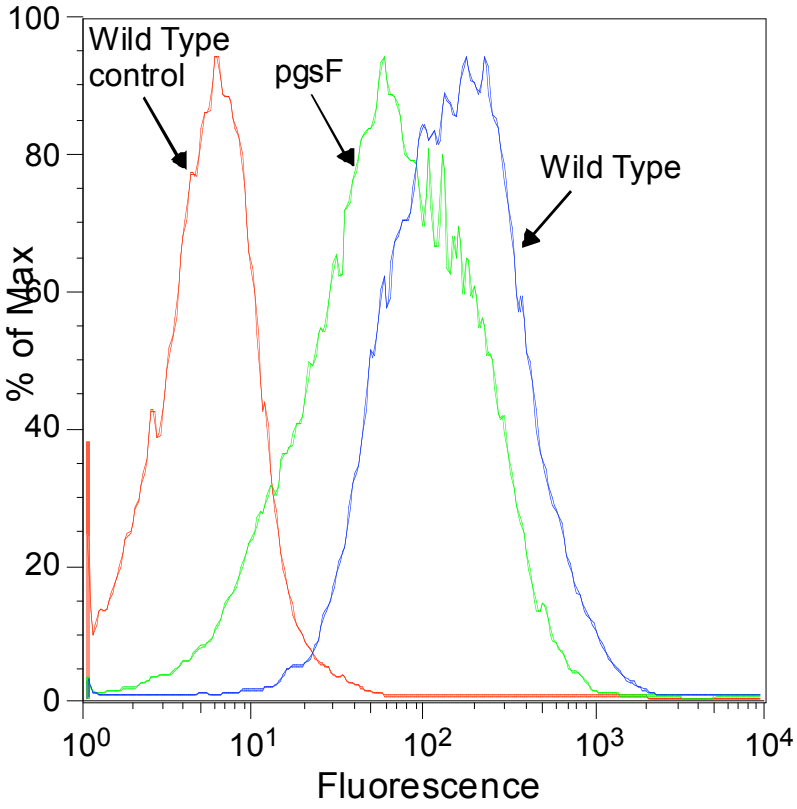


Figure 2.25. Sulfate density on heparan sulfate proteoglycans plays a key role in the binding of guanidinoglycosides. Cell lines lacking 2-*O*-sulfotransferase activity express similar heparan sulfate proteoglycan chains as wild type cells, with the difference being an overall decrease in sulfate density. Guanidino-neomycin binds to this mutant cell line similar to wild type cells.

activity highlights the necessity for sulfate density in guanidino-neomycin binding to cell surface proteoglycans.

These data help to refine the model of guanidinoglycoside binding to cells. We have shown that the binding of guanidino-neomycin to the surface of cells is dependent on heparan sulfate proteoglycans. There is evidence that at higher concentrations, a second class of non-heparan-sulfate dependent binding receptors does exist. However, at guanidinoglycoside concentrations below the high nanomolar, binding is exclusively dependent on heparan sulfate.

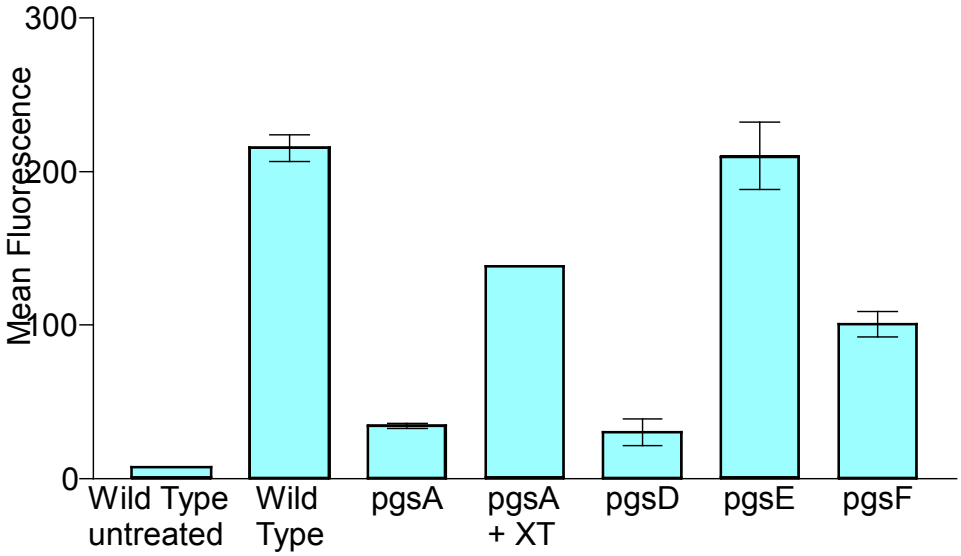


Figure 2.26. Chinese hamster ovarian mutant cell lines incubated with guanidino-neomycin reveal vastly different binding properties. Untreated wild type cells show no inherent fluorescence, however when incubated with 60 nM guanidino-neomycin-biotin streptavidin-PE-Cy5 there is a marked increase in fluorescent signal. Cells lacking heparan sulfate (pgsD) or both heparan and chondroitin sulfate proteoglycans (pgsA) show a greatly reduced signal, highlighting the dependence of guanidinoglycoside binding on heparan sulfate. When the gene for xylosyltransferase is transfected back into pgsA cells, they are able to recover their guanidinoglycoside binding capabilities. No significant change is seen in guanidinoglycoside binding to cells where overall sulfation is reduced (pgsE), however when 2-*O*-sulfation of heparan sulfate is specifically inhibited, (pgsF) there is a large signal decrease.

Additionally, binding is at least partially dependent on the fine structure of the heparan sulfate saccharide chain. When sulfated domains are modified so that there is no 2-*O*-sulfation on Iduronic acid, the overall sulfate density of these domains is lowered without reducing the overall numbers of these domains. This loss of sulfate density causes a significant decrease in guanidinoglycoside binding and subsequent uptake.

Flow cytometry experiments performed with guanidino-neomycin at 4^{o}C and 37^{o}C provide evidence that the majority of the fluorescent signal is due to binding and uptake of the guanidinoglycoside as opposed to exclusively cell surface binding.

However, this data fails to address whether guanidinoglycosides can deliver cargo into the cell so that this cargo retains its function and can perform a given function such as enzymatic activity within the cell. For guanidinoglycosides to serve as efficient delivery vehicles, linked cargo must retain its inherent bioactivity.

2.7 Guanidinoglycosides Can Delivery Large Bioactive Cargo into the Cell

It is important to establish that guanidinoglycosides can not only mediate uptake of different molecules, but allow these molecules to enter the cell and be released into the cytoplasm. Molecules trapped within vesicles, shuttled to lysosomes, or quickly transported out of the cell will be unable to retain their inherent function. There is evidence that the mechanism of uptake of arginine-rich peptides involves binding to heparan sulfate proteoglycans followed by endocytosis. Once inside the cell, heparan sulfate is degraded by heparanases, promoting vesicular leakage which allows these peptides to escape into the cytoplasm.[20]

To determine whether guanidinoglycosides have the ability to carry large, bioactive cargo into the cell, neomycin-, guanidino-neomycin- and Arg_9-biotin were conjugated to saporin and assayed for their ability to kill cells. Saporin is a 30 kDa ribosome-inactivating protein from the seeds of the plant *Saponaria officinalis*.[34] Saporin will inactivate ribosomes and lead to cell death only if it is able to enter the cytoplasm of cells and come in proximity to active ribosomes within the cell. This makes a conjugated saporin system ideal to examine the delivery capabilities of guanidinoglycosides. If guanidino-neomycin is able to carry saporin into the cell

through vesicles and localize in the lysosome, saporin would likely become inactivated. However, if guanidino-neomycin has the ability to carry cargo into the cell for release into the cytoplasm, it could retain its activity without being degraded by the lysosome.

To examine the ability of guanidino-neomycin to facilitate the delivery of bioactive molecules into the cell, it was conjugated to saporin and this conjugate was assayed for its ability to kill cells. Saporin-streptavidin or Saporin (without streptavidin) (Advanced Targeting Systems, San Diego, California) was allowed to complex with neomycin-biotin, guanidino-neomycin biotin or Arg_9-biotin. Saporin-streptavidin is a covalent complex of Saporin and streptavidin which allows the toxin to be tethered to biotinylated molecules. As a control, Saporin, which has no inherent affinity for biotin was also used. Cells were exposed to media with differing concentrations of these conjugates for four days.

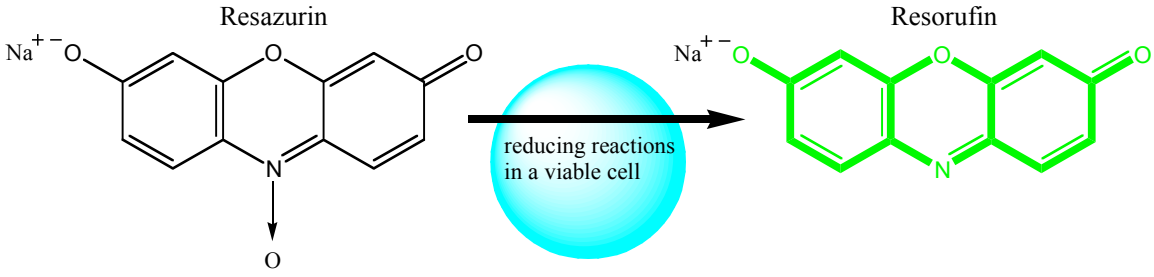


Figure 2.27. CellTiter-Blue (Promega) contains Resazurin, a redox indicator dye that will be reduced by viable cells into Resorufin. Resofurin excitation at 488 nm results in emission at 575 nm. This allows for quantitative determination of ratios of living cells between different populations.

After four days, CellTiter-Blue (Promega) was added, and allowed to incubate for 4 hours. CellTiter-Blue contains a buffered solution of resazurin, a redox indicator

dye that can be added directly to cells in culture. Living cells can convert the oxidized form of the dye (resazurin) into a reduced form (resorufin) which has a distinct excitation/emission profile. Excitation at 488 nm results in emission at 575 nm (Figure 2.27). Because only living cells will metabolize resazurin into resorufin, this fluorescence is directly proportional to viable cell numbers. Analysis of fluorescence measurements from each well allow for determination of the relative proportions of living cells.

Wild type and pgsA cells were examined with guanidino-neomycin-biotin streptavidin-saporin at concentrations ranging from 82 pM to 84 nM. Control experiments were performed with guanidino-neomycin-biotin in the absence of saporin and also with guanidino-neomycin-biotin being incubated with saporin which was not streptavidinylated. While guanidino-neomycin shows no inherent toxicity to cells within the concentration range studied, unconjugated saporin does retain the ability to kill cells (Figure 2.28). However, cell populations were only reduced by roughly 50% at the highest saporin concentrations (84 nM).

The conjugated guanidino-neomycin-biotin-streptavidin-saporin system, in contrast to the unconjugated guanidino-neomycin saporin, induces cell death at much lower concentrations. IC_{50} values for cell death with wild type cells are in the low (1-5) nanomolar range (Figure 2.28). While the guanidino-neomycin-saporin system does show some cell lysis activity at higher concentrations with pgsA cell, it is roughly equivalent to the activity of unconjugated saporin, indicating that cell death here is a result of the inherent activity of saporin as opposed to delivery from guanidinoglycosides.

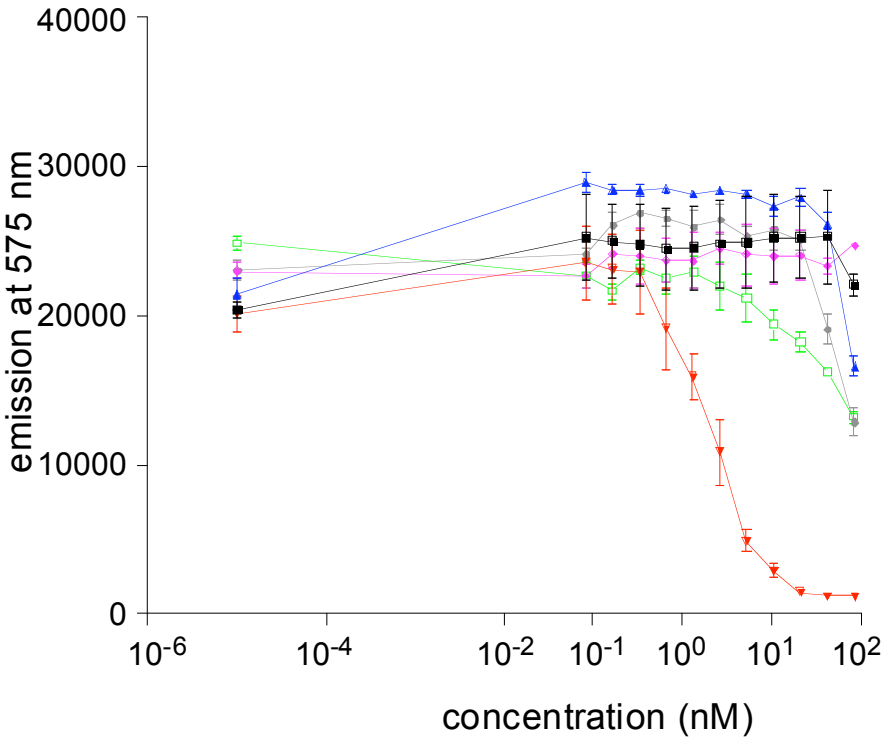


Figure 2.28. Titrations of guanidino-neomycin-biotin into wild type (black) and proteoglycan deficient (pink) cells show a lack of toxicity. When guanidino-neomycin-biotin is incubated with saporin and titrated into wild type (blue) and pgsA cells, the unconjugated saporin appears to have some ability to kill cells as concentrations approach 100 nM. A complex of guanidino-neomycin-biotin with streptavidin-saporin shows roughly the same ability to kill pgsA cells (green) as the uncomplexed saporin, indicating little to no guanidinoglycoside mediated delivery in pgsA cells. However, when the same guanidino-neomycin-biotin streptavidin-saporin system is titrated into wild type cells, this complex kills cells with an IC_{50} in the low nM. Wild type and pgsA-745 cells were grown in tissue culture plates in HAMS F-12 media with 10% Fetal Calf Serum, exposed to compounds and controls in media and incubated for four days at 37°C in a 5% CO_2 atmosphere. After 4 days, CellTiter-Blue (Promega) was added to media and the cells were incubated at 37°C under an atmosphere of 5% CO_2 for an additional four hours before taking fluorescence measurements.

These data highlights the ability of guanidinoglycosides to deliver large cargo (streptavidin-saporin is roughly 130 kDa) into cells in a proteoglycan dependent manner. Additionally, delivered molecules are not shuttled to lysosomes and degraded, but instead released into the cytoplasm, retaining their inherent activity.

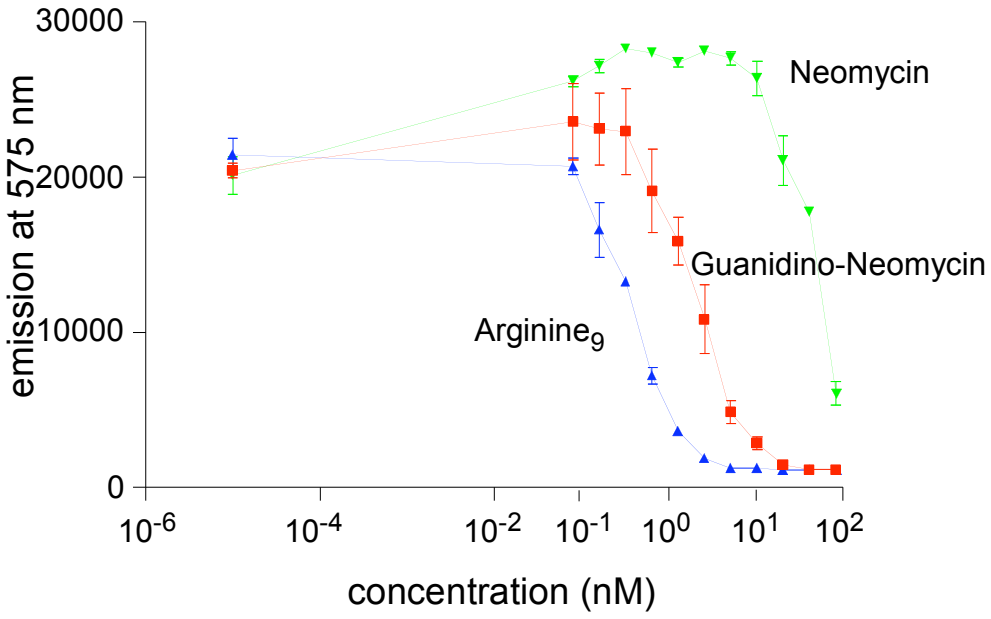


Figure 2.29. The Arg_9-saporin complex is able to efficiently kill cells with IC_{50} values in the high picomolar. IC_{50} values for guanidino-neomycin-saporin are roughly 10 fold higher, and the activity of neomycin-saporin is only slightly more active than the unconjugated saporin toxin by itself.

When wild type cells are titrated with a neomycin-biotin streptavidin-saporin complex, cell death occurs at 30-40 fold higher concentrations than with the guanidino-neomycin-saporin complex (Figure 2.29). While this neomycin-saporin system does have the ability to kill cells, it is only slightly more active than unconjugated saporin, indicating that neomycin is a poor cellular delivery agent compared to its guanidinylated derivate.

Interestingly, Arg_9-biotin appears to have a slightly higher (roughly 10-fold) activity than guanidino-neomycin-biotin when conjugated to streptavidin-saporin (Figure 2.29). However, earlier studies have shown that guanidino-neomycin and Arg_9 bind to and are taken up by wild type cells at roughly equivalent concentrations (Figure

2.21), so it may be possible that efficiency of uptake will vary depending on the specific systems studied.

It is also possible that a second class of receptors is facilitating uptake of the Arg_9 saporin system into wild type cells. While this would enhance the efficacy of delivery with the poly-arginine system, the presence of multiple receptors makes the Arg_9 a less selective delivery molecule when compared to guanidino-neomycin.

2.8 Summary and Discussion

Facilitating the delivery of bioactive compounds into living cells is a key issue in the design of novel therapeutic agents. Many drug candidates which show potential *in vitro* have little to no *in vivo* activity because they are not able to cross the cellular membranes. By linking these molecules to compounds which can efficiently cross cell membranes and internalize inside cells, it may be possible to significantly enhance the efficacy of these potential therapeutics, affording them clinical relevance.

Proteins and peptides with multiple arginine groups (Tat, Rev, Arg_9, and others) have been shown to efficiently cross cell membranes and deliver cargos upwards of hundreds of kilo Daltons into the cytoplasm. Recently, we have found that guanidinoglycosides also display effective cellular uptake properties, with uptake efficiency comparable to Arg_9, a common oligo-arginine transduction peptide.

Understanding the mechanism of guanidinoglycoside and arginine-rich peptide internalization into living cells is important from both basic biological and medicinal chemistry perspectives. However, while arginine-rich peptide internalization has been

examined, the mechanistic understanding of this uptake is still incomplete with multiple models being proposed, and multiple mechanisms potentially acting simultaneously. Here, we show evidence that there are at least two modes of uptake for the Arg_9 peptide, one which is dependent on cell surface proteoglycans, and one which appears to be proteoglycan independent.

While we have shown that guanidinoglycosides share a common, proteoglycan dependent mechanism of uptake with Arg_9, they do not share any secondary non-proteoglycan dependent interactions. Cellular binding and uptake of guanidino-neomycin appears to be exclusively dependent on cell surface heparan sulfate proteoglycans. Heparan sulfate and chondroitin sulfate proteoglycans are both composed of similar sulfated saccharides, yet guanidino-neomycin binds exclusively to heparan sulfate proteoglycans. In fact, guanidino-neomycin displays no difference in signal between cells lacking both proteoglycans and cells lacking only heparan sulfate but expressing 3-4 fold excess chondroitin sulfate.

The fine structure of the heparan sulfate chains plays a key role in the uptake activity of guanidino-neomycin. Heparan sulfate oligosaccharide chains contain tracts of highly sulfated regions. When the overall number of these tracts is decreased, there is little to no effect on guanidinoglycoside binding. However, when sulfation at specific residues within these domains is inhibited, there is roughly a 2 fold decrease in guanidinoglycoside signal. This highlights the importance of sulfate density in the binding and uptake of guanidinoglycosides.

While these data shed light on the cellular requirements for internalization of guanidinoglycosides, it is important to note that guanidinoglycosides can carry large,

bioactive cargo into the cell. We have shown that guanidino-neomycin can facilitate the transport of large molecular weight complexes, hundreds of times as large as the guanidinoglycoside across the cell membrane. Streptavidin-phycoerythrin-cychrome-5, a ~300 kDa fluorophore is carried across the cell membrane by guanidino-neomycin via heparan sulfate mediated endocytosis

While the signal of a fluorophore such as streptavidin-PE-Cy5 does not depend on its cellular localization, the ribosome inactivating toxin saporin must be delivered into the cytoplasm of a cell to be active. If it were sequestered into vesicles or transported to the lysosome, saporin would have no activity. When complexed to guanidino-neomycin, saporin can efficiently internalize and kill cells.

We have shown that guanidino-neomycin is internalized into living cells in a heparan sulfate proteoglycan dependent manner. Additionally, guanidino-neomycin facilitates the uptake of large, bioactive molecules via heparan sulfate mediated endocytosis.

Novel drug delivery agents are needed, and low molecular-weight guanidinoglycosides provide an ideal candidate for development of future drug delivery tools. The heparan sulfate exclusivity of guanidinoglycosides allows for design of specifically targeted compounds. Because Arg_9 displays multiple modes of binding, it is possible that arginine-rich peptides may have inherently lower selectivity than guanidinoglycosides.

More research needs to be done into understanding the mechanism of cellular internalization, tissue specificity and therapeutic possibilities of guanidinoglycosides

Their proteoglycan dependent delivery activity highlights the enormous potential of low molecular weight guanidinium-containing glycosides as future drug-delivery agents.

2.9 Addendum: Other Molecules Bind Heparan Sulfate

The advantage of a high throughput assay is the ability to screen a large number of compounds for a specific activity reasonably quickly. It is for this reason that the assay described earlier (figure 2.3) was developed. Large numbers of compounds were screened for their ability to bind to cell surface heparan sulfate and block toxins which are specific for these cell surface proteoglycans.

The National Cancer Institute (NCI) maintains a set of 1990 compounds which were narrowed down from a set of over 140,000 compounds. Chem-X software (Oxford Molecular Group) uses defined centers (hydrogen bond acceptor, hydrogen bond donor, positive charge, aromatic, hydrophobic, acid, and base) and defined distance intervals to create a particular finite set of pharmacophores. Based on this, compounds with five new pharmacophores and five or fewer rotatable bonds were included in the NCI diversity set. When tested for heparan sulfate binding and toxin inhibition, the NCI diversity set revealed a unique active compound.

bis-4-amino-2-methyl-6-quinolyl-urea (Surfen / NSC 12155) was originally synthesized in 1933, and was late discovered to be an oncogenic agent with heparin neutralizing properties as well.[35] More recently, due to it's inclusion in the NCI diversity set, Surfen has been shown to inhibit anaphylatoxin C5a receptor binding,[36]

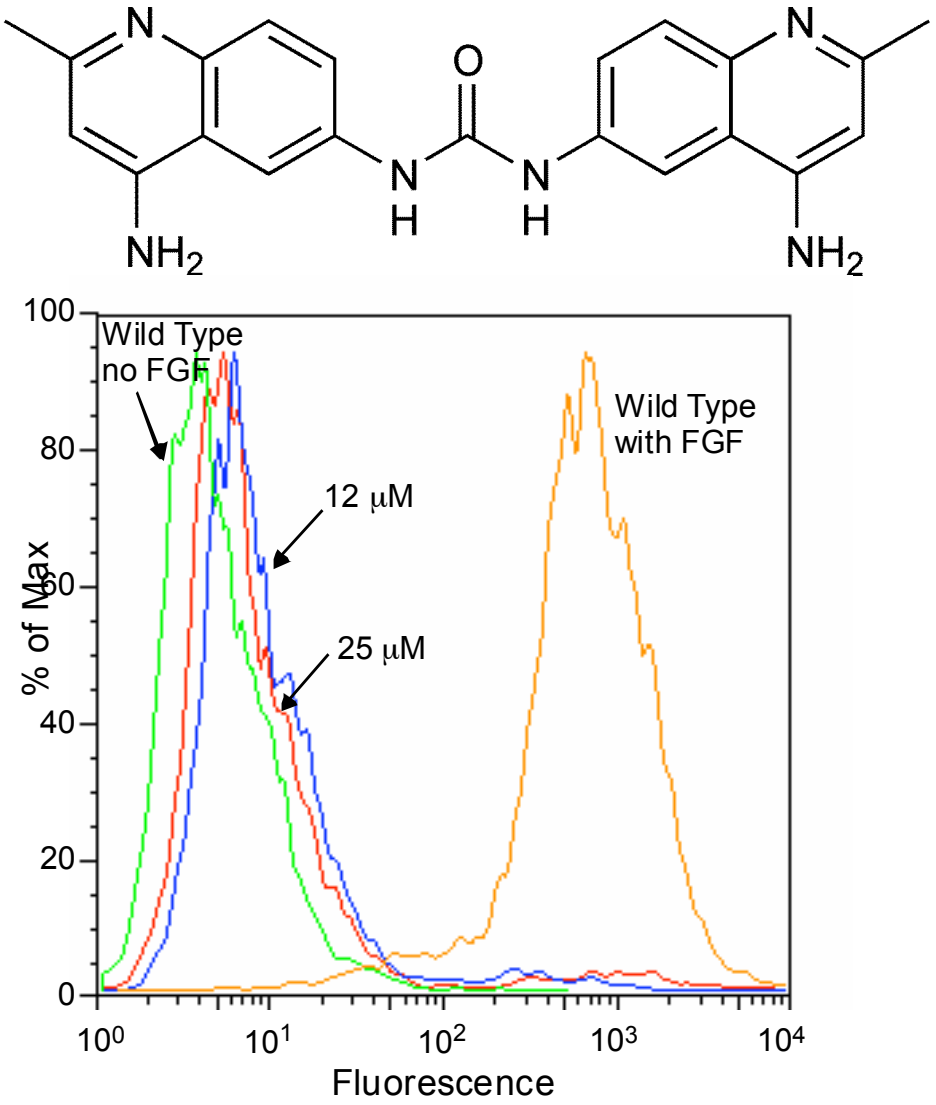


Figure 2.30. Surfen (NSC 12155) is a molecule included in the National Cancer Institute diversity set which displays efficient blocking of heparan sulfate binding proteins. When assayed for its ability to block FGF binding to the cell, there is complete inhibition at concentrations as low as 12 μM.

and may show potential as an anti-anthrax agent as well.[37-39] When the set was assayed with the high throughput assay described above (Figure 2.3), Surfen emerged as a potent toxin inhibitor (data not shown).

To quantify its ability to bind to cell surface heparan sulfate proteoglycans, Surfen was examined with the flow cytometry based FGF assay described above

(Figure 2.4). Surfen was found to fully block FGF binding concentrations as low at 12 μM (Figure 2.30). To discover the properties of Surfen which contribute to its ability to block heparan sulfate binding proteins, a library of derivatives was designed, synthesized and assayed for their ability to bind cell surface heparan sulfate proteoglycans.

The scheme for synthesis of Surfen allows for relatively simple variation of multiple structural features of the molecule. Modifications of the amino group at the 4 position included 4-H, 4-OH and 4-OMe. The length of the linker between the individual monomers was modified from the urea functionality seen with the parent Surfen to an oxalyl moiety which, by including an additional carbonyl both increases the length and changes the relative conformation of the quinolyn based monomers. Molecules based on the quinolyn monomers were also studied with amino, urea and *N*-acetyl groups at the 6 position.

The library of Surfen derivatives was screened for heparan sulfate and toxin protection activities. Replacement of the 4-amino group with an -H, -OH or -OMe resulted in compounds which displayed no protection against toxin at concentrations up to 100 μM. Additionally, molecules with modifications at the 6 position, based on a quinolyn monomer likewise showed no activity through the same concentrations (data not shown).

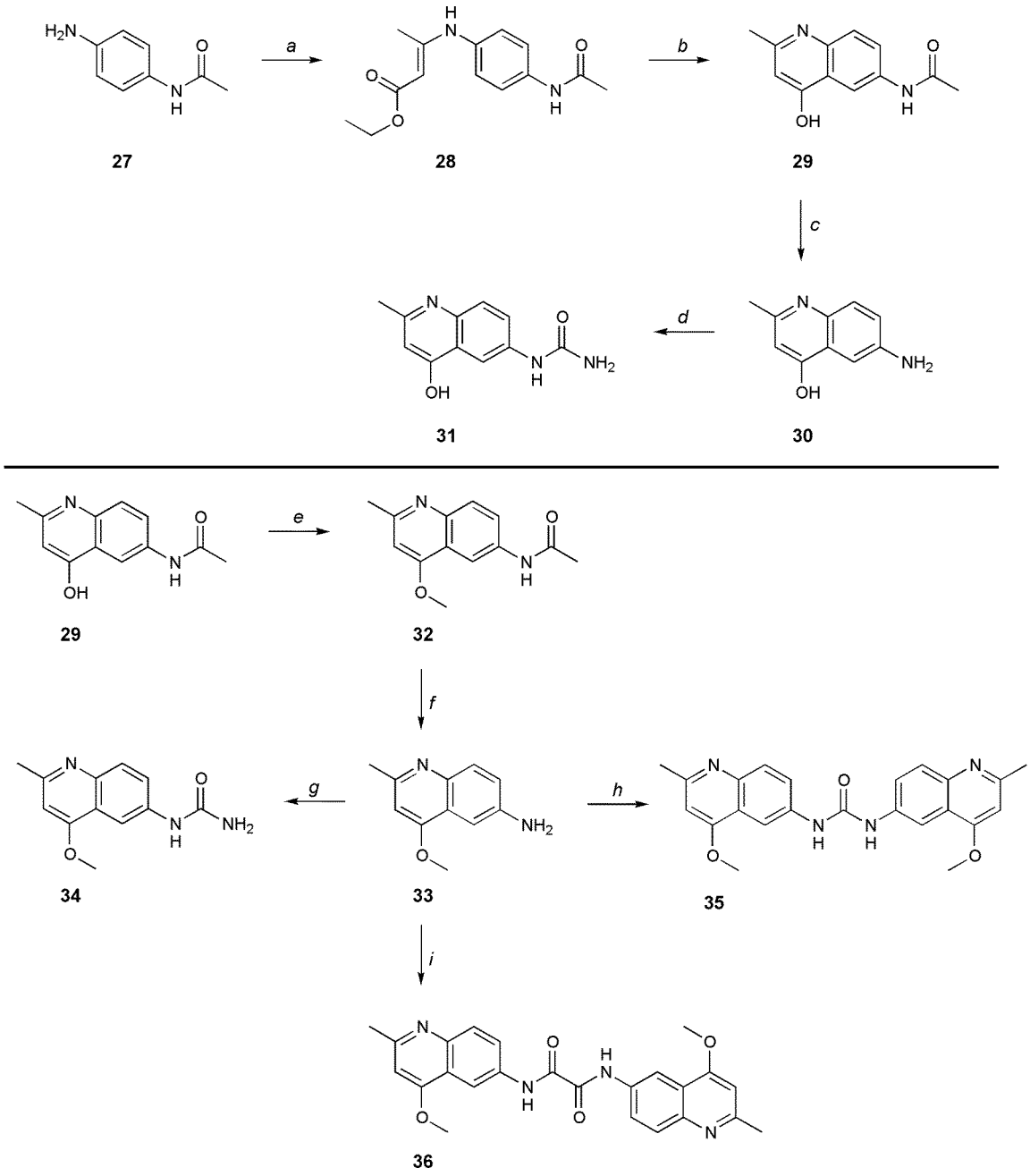


Figure 2.31. Synthesis of the building blocks for the Surfen derivates allows for a wide variety of similar molecules to be synthesized. a) **27** and ethyl acetoacetate, reflux (91%) b) **28** and Dowtherm A, reflux 90% c) **29** and HCl, 90°C (81%) d) **30** and KCN in acetic acid (49%) e) **29** and dimethyl sulfate in toluene, reflux (76%) f) **32** and HCl, 90°C (99%) g) **33** and KCN in acetic acid (46%) h) **33** and triphosgene in acetic acid (95%) i) **33** and oxalyl chloride in acetic acid (46%)

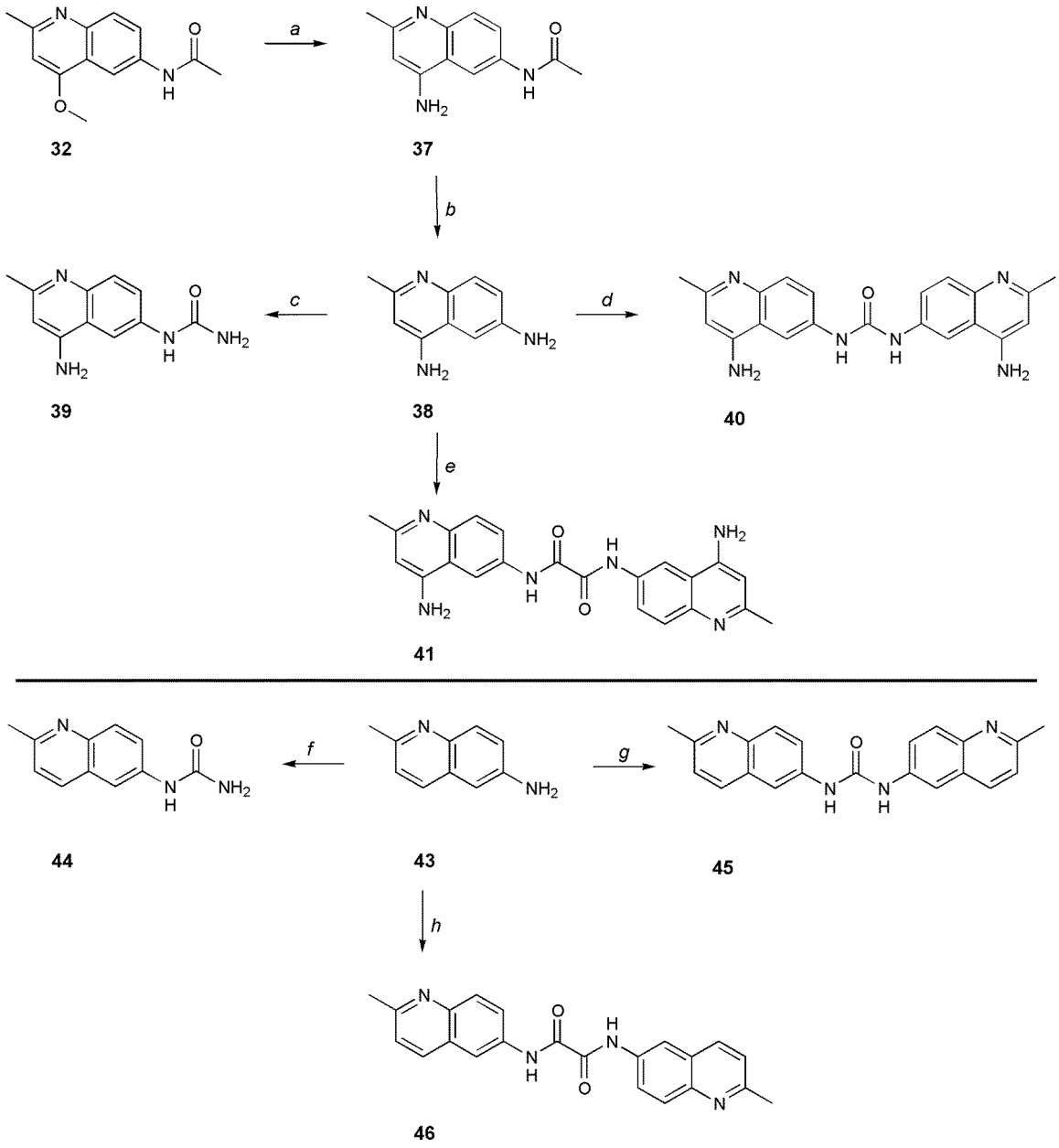


Figure 2.32. Synthesis of Surfen, oxalyl-Surfen and de-aminated derivatives. a) **32** and NH_4OAc, 135°C (91%) b) **37** and HCl, 90°C (80%) c) **38** and KCN in acetic acid (43%) d) **38** and triphosgene in acetic acid (90%) e) **38** and oxalyl chloride in acetic acid (45%) f) **43** and KCN in acetic acid (69%) g) **43** and triphosgene in acetic acid (48%) h) **43** and oxalyl chloride in acetic acid (49%)

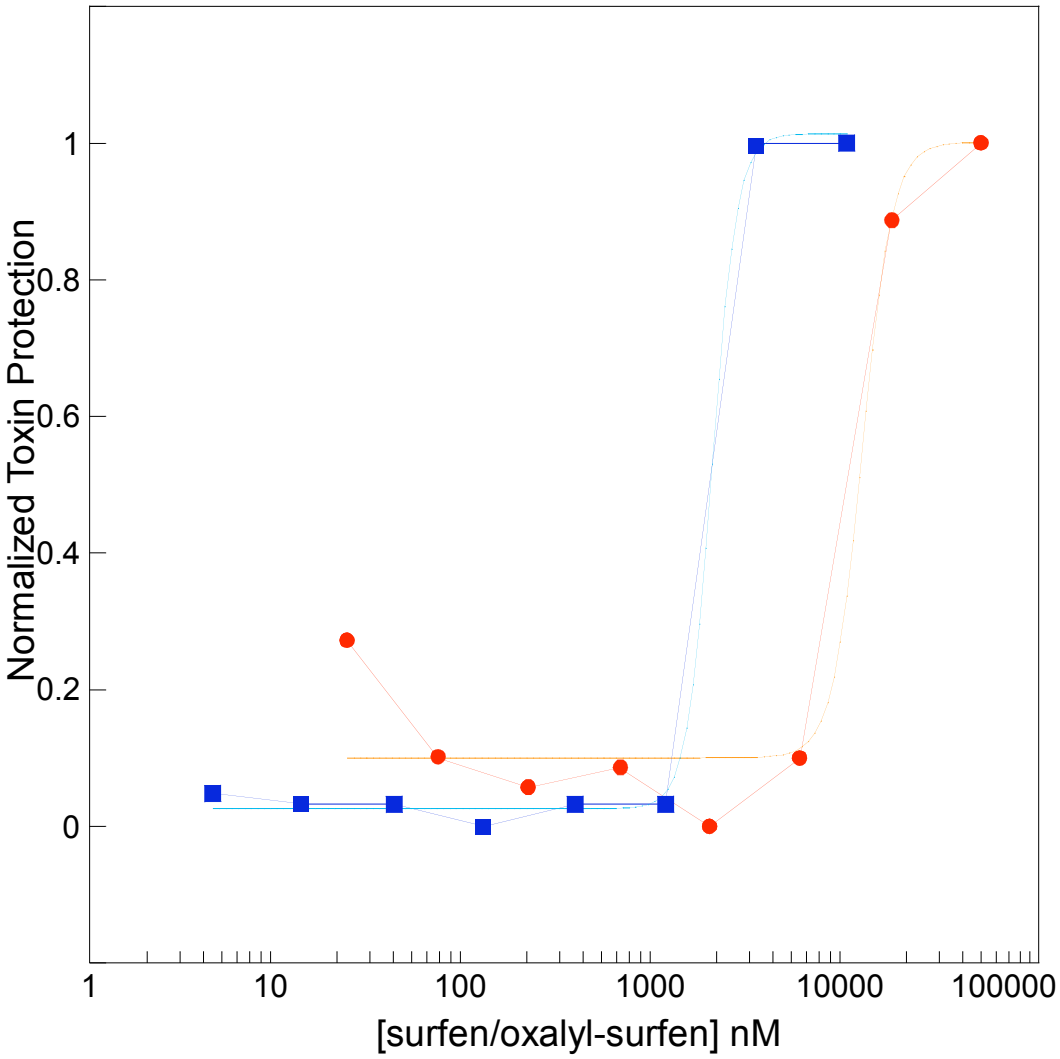


Figure 2.33. A graph of the activity of oxalyl surfen (blue squares) and surfen (red circles) in a toxin protection assay described above (Figure 2.3). Curve fits show that Oxalyl Surfen (light blue curve) displays roughly 10 fold better protection against heparan sulfate binding than Surfen (orange curve) in a preliminary high throughput screen. None of the other Surfen-based derivatives showed any protection through 100 μM.

The only Surfen derivative which displayed activity comparable to the parent molecule was the oxalyl Surfen where the urea moiety was extended to an oxalyl-urea. Oxalyl Surfen shows protection at roughly 10-fold lower concentrations than the parent Surfen molecule (Figure 2.33). This preliminary screen indicates that while the amino

functional group and symmetrical dimerization of Surfen are critical for its heparan sulfate binding and toxin inhibitory activities, the distance between the quinolyn moieties may not be critical and in fact, a longer linker may enhance activity.

These data is preliminary, at best, and a more quantitative approach to studying the Surfen based derivatives must be performed. To that end, a selected group of molecules will be assayed for their ability to block FGF-biotin binding to cell surface heparan sulfate proteoglycans (Figure 2.4). Using this method, we hope to derive quantitative IC_{50} values for heparan sulfate binding and FGF blocking activities for a variety of Surfen derivatives. This should better enable us to generate a defined structure-activity relationship and understand specifically which pharmacophores impart the heparan sulfate binding characteristics to Surfen.

Additionally, to follow up on the possible enhancement of activity shown with the oxalyl derivative, we will synthesize Surfen molecules with extended linker lengths. Malonyl, succinyl and glutaryl chlorides can be used similarly to oxalyl chloride to generate molecules with 1, 2 or 3 methylenes between the carbonyl groups on the Surfen based dimers (Figure 2.34).

Another avenue of potential study would be a “trivalent” Surfen. Currently, the compounds which show activity all contain two quinolyn based rings. Monomers of Surfen show no activity to high concentrations, regardless of the functional group at the 6-position, indicating that the “divalency” is required for the heparan sulfate binding of Surfen. Potentially, a Surfen derivative with three quinolyn monomers covalently attached will show enhanced activity. These compounds will allow us to

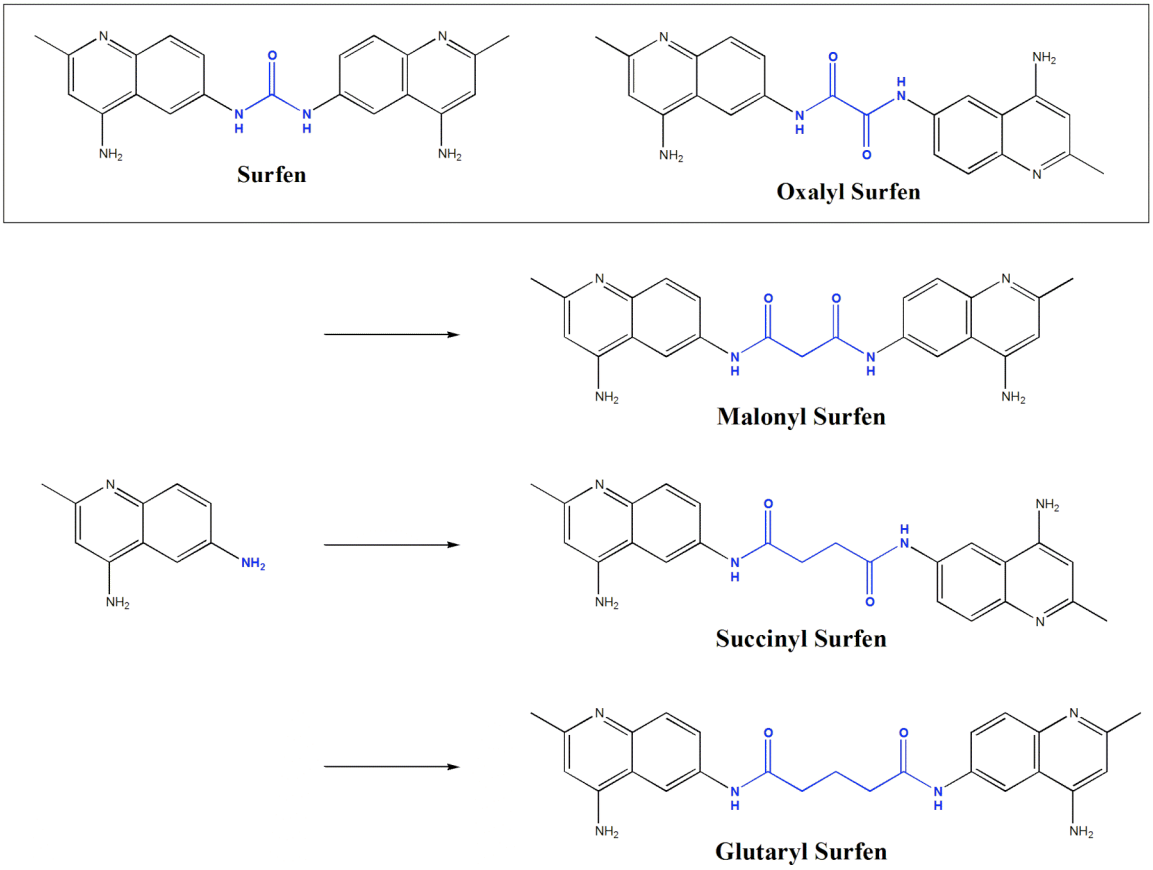


Figure 2.34. Surfen and oxalyl Surfen both bind to heparan sulfate and protect cells from toxins which require cell surface proteoglycans. To study the effects of linker length and relative quinolyn conformations, three other Surfen derivatives, malonyl, succinyl and glutaryl Surfen, have been designed.

study the roles of linker length, relative quinolyn conformations and numbers on heparan sulfate binding. There is also a possibility that we may generate molecules with enhanced heparan sulfate binding characteristics, higher stability or lower *in vivo* toxicity.

2.10 Comparing Families of Newly Discovered Heparan Sulfate Binders

Another set of molecules which was screened for heparan sulfate binding and toxin inhibition included aminoglycoside and guanidinoglycoside based compounds covalently linked to an intercalator based on ethidium bromide (For details regarding structure and synthesis of these compounds, see Chapter 3). While these molecules were designed to specifically target and bind to viral RNA sequences, when they were assayed for their heparan sulfate binding activities, surprising results were achieved.

The high throughput assay revealed that all compounds containing the methidium functional group showed some heparan sulfate and toxin inhibitory activity (Figure 2.3). This activity was independent of whether the conjugated molecule was an aminoglycoside or a guanidinoglycoside, indicating that the methidium moiety may have some intrinsic heparan sulfate binding characteristics. However, at high concentrations (over 100 μM) the methidium compounds begin to show toxicity toward cells.

It is interesting to note that this is not a general feature seen amongst all intercalating systems. Paromamine derivatives with acridine moieties at the 1, 2′ or both 1 and 2′ positions show no activity in this assay through concentrations of almost 50 μM.

Due to the indications that heparan sulfate binding activity is intrinsic to the methidium containing molecules, more quantitative measurements were taken with two of the methidium-containing molecules. Tobramycin-N-methidium and guanidino-

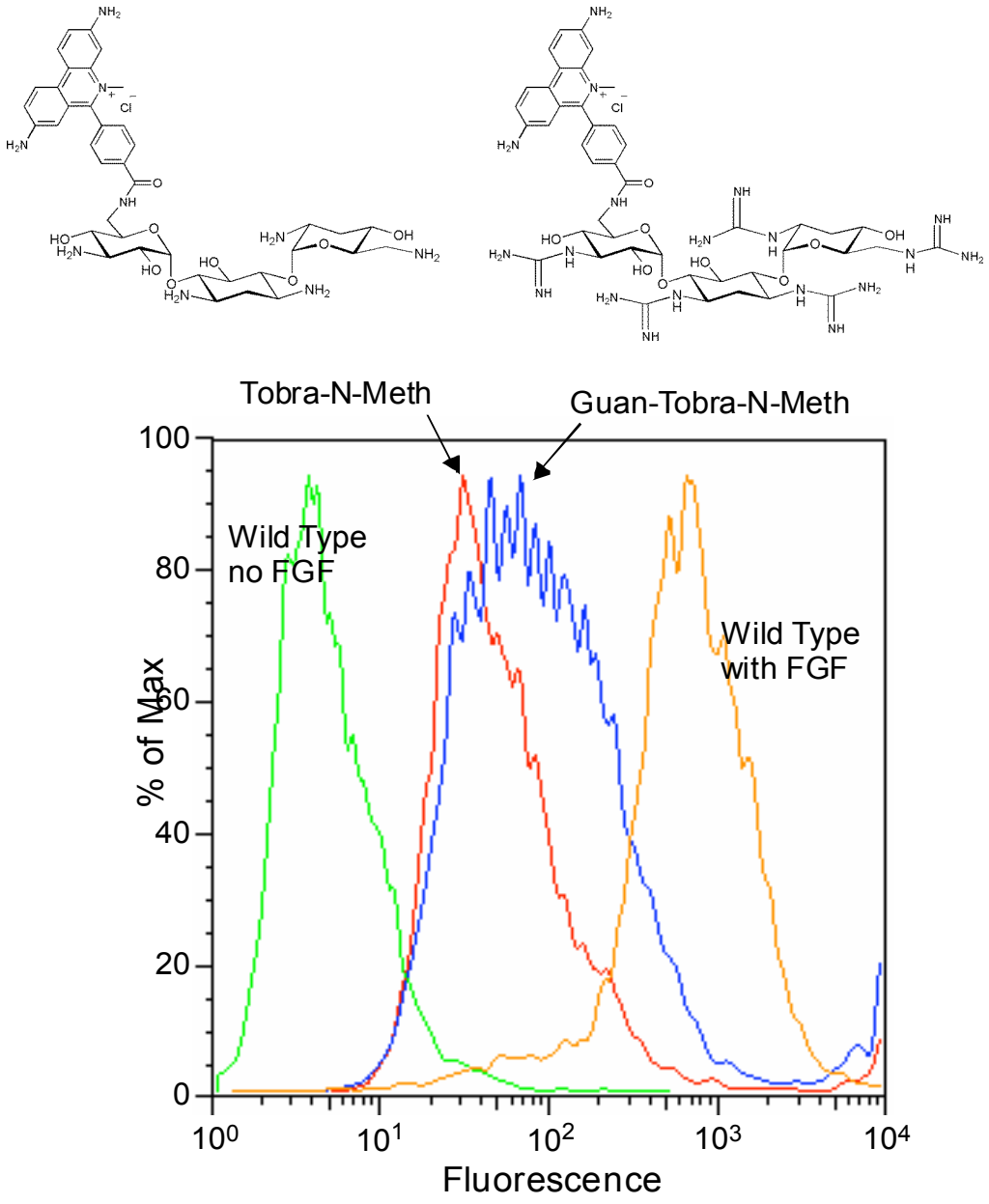


Figure 2.35. Aminoglycosides and guanidinoglycoside covalently linked to methidium show roughly equivalent cell surface heparan sulfate binding and FGF blocking activity. At 50 μM, both compounds are able to reduce FGF binding to the surface of cells by over 90%. This highlights the methidium moiety as an efficient heparan sulfate binder.

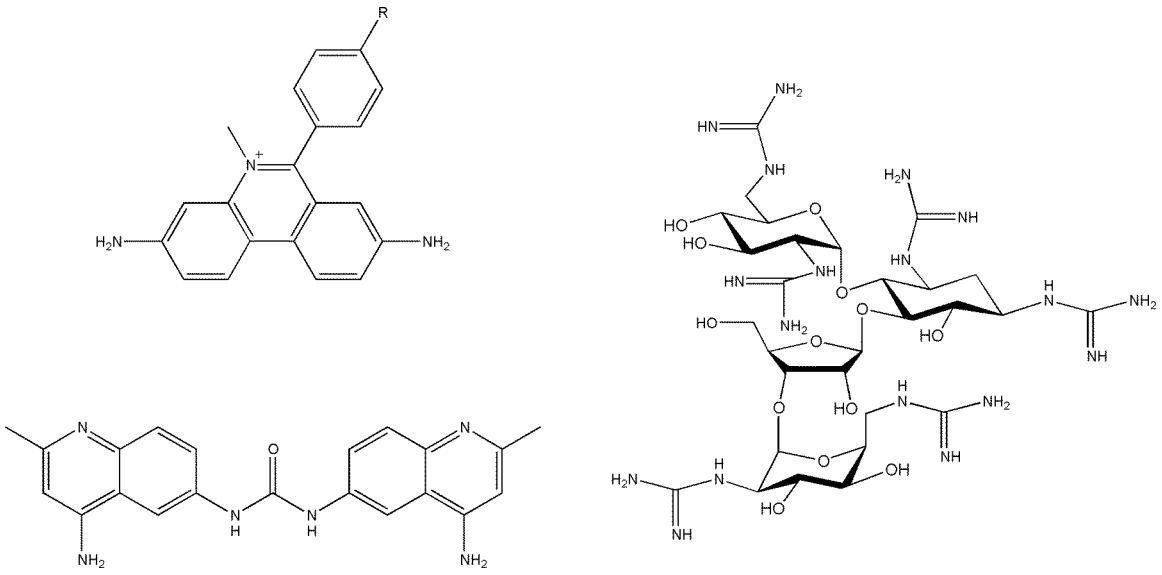


Figure 2.36. The two main classes of heparan sulfate binding molecules which we have studied. Surfen and methidum based molecules all contain planar aromatic structures of two or three fused rings, along with exocyclic amino groups which appear to be critical at least for Surfen activity. The other class of molecules includes the guanidinoglycosides, oligomeric guanidino sugars, where multiple guanidinium groups are critical for both binding to cell surface heparan sulfate proteoglycans and subsequent internalization.tobramycin-N-methidium were assayed for their ability to bind to cell surface proteoglycans and inhibit FGF binding to cells.

Both compounds reduce the signal from cell-associated FGF by over 90% at 50 μM. However, it is the non-guanidinylated derivative, tobramycin-N-methidium which shows 2-3 fold higher activity.

While this data is preliminary, it highlights the need for further study of ethidium bromide like molecules as potential cell surface heparan sulfate binders. Modifications of these molecules may provide selective heparan sulfate binders with higher affinity and lower toxicity than the parent molecules.

Based on a preliminary screen for heparan sulfate binding molecules, certain trends begin to emerge. There appears to be two main classes of molecules which bind

to heparan sulfate and inhibit the interactions between cell surface heparan sulfate proteoglycans and proteins which are specific for these glycosaminoglycans.

Guanidinium rich oligomers including arginine-rich peptides (and multimers with non-peptide backbones) and guanidinoglycosides form one class of heparan sulfate binding small molecules. Guanidinoglycosides are composed of oligomeric guanidino sugars where the amino groups from aminoglycosides are replaced with guanidinium groups. At physiological pH, these groups will be protonated, giving guanidinoglycosides an overall highly positive charge.

Aminoglycosides show no heparan sulfate binding and FGF inhibitory activity. However, when guanidinylated, these molecules show significant heparan sulfate binding and inhibition of FGF interactions with cell surface proteoglycans. The dependence on the guanidinium group for activity highlights it as the main component responsible for the interactions between guanidinoglycosides and heparan sulfate.

The other class includes the aromatic, amine containing Surfen and methidium containing molecules. Uptake studies have not been performed with these molecules, however they both show the ability to bind to cell surface heparan sulfate proteoglycans, and inhibit complex formation with heparan sulfate specific binding proteins such as FGF.

Similarities between Surfen and the methidium molecules include a planar structure of at least two fused aromatic rings. Surfen has a pair of quinolyn based rings and methidium has three rings fused together. It is possible that stacking events are taking place with the fused aromatic moieties being “sandwiched” between saccharides on one or more heparan sulfate chains. However, detailed structural analysis would

need to be performed to elucidate the specific interactions taking place between the ligands and heparan sulfate proteoglycans.

Another striking similarity between the Surfen and methidium molecules is the presence of exocyclic amines. We have seen that the presence of these amines is critical for the heparan sulfate binding properties of Surfen. Molecules with modifications at this position show no activity. De-aminated methidium molecules or molecules where the amines are replaced with other substituents have not been synthesized, limiting our ability to study the necessity of the amino group on the proteoglycan binding properties of the methidium molecules.

FGF inhibition experiments reveal that Surfen is roughly 10-20 fold more active than the different methidium derivatives. Although their structures are considerably different, one possible cause for this difference may be found in the charge state of the two molecules. The exocyclic amines of ethidium bromide are poor nucleophiles and only weakly basic with p*K*a's of 0.8 and 2.[40–42] At physiological pH (7.4), neither of these amines will be protonated, and will carry a neutral charge.

Surfen carries a phamacophore which is very similar to 4-aminopyridine and 4-aminoquinolyn, which have p*K*a's of 9.17[43] and 7.99[44] respectively. It is likely, therefore that Surfen will be protonated in physiological conditions, and will carry a positive charge. The interactions between the positively charged amines on Surfen and the negatively charged sulfates on heparan sulfate proteoglycan chains may provide a significant contribution to the FGF inhibition activity seen in Surfen.

While significant progress has been made in understanding some of the interactions which play a role in the binding of small molecules to heparan sulfate

proteoglycans on the surface of the cells, much has yet to be elucidated. A comprehensive structure activity relationship where derivatives of Surfen are tested for their ability to bind heparan sulfate and block FGF will provide a more detailed understanding of the pharmacophores which can be implicated in these interactions. Additionally, this study may provide direction for the design of novel compounds with improved heparan sulfate binding properties.

2.11 References

(1) Hallbrink, M.; Floren, A.; Elmquist, A.; Pooga, M.; Bartfai, T.; Langel, U. *Biochim Biophys Acta* **2001**, 1515, 101–109.

(2) Denicourt, C.; Dowdy, S. F. *Trends Pharmacol Sci* **2003**, 24, 216–218.

(3) Green, I.; Christison, R.; Voyce, C. J.; Bundell, K. R.; Lindsay, M. A. *Trends Pharmacol Sci* **2003**, 24, 213-215.

(4) Leifert, J. A.; Whitton, J. L. *Mol Ther* **2003**, 8, 13–20.

(5) Thierry, A. R.; Vives, E.; Richard, J. P.; Prevot, P.; Martinand-Mari, C.; Robbins, I.; Lebleu, B. *Curr Opin Mol Ther* **2003**, 5, 133–138.

(6) Vives, E.; Richard, J. P.; Rispal, C.; Lebleu, B. *Curr Protein Pept Sci* **2003**, 4, 125–132.

(7) Hoffman, A. S. *Adv Drug Deliv Rev* **2002**, 54, 3–12.

(8) Panyam, J.; Labhasetwar, V. *Adv Drug Deliv Rev* **2003**, 55, 329–347.

(9) Lu, Y.; Chen, S. C. *Adv Drug Deliv Rev* **2004**, 56, 1621–1633.

(10) Peppas, N. A. *Adv Drug Deliv Rev* **2004**, 56, 1529–1531.

(11) Ryser, H. J. *Science* **1968**, 159, 390–396.

(12) Dietz, G. P.; Bahr, M. *Mol Cell Neurosci* **2004**, 27, 85–131.

(13) Perez, F.; Joliot, A.; Bloch-Gallego, E.; Zahraoui, A.; Triller, A.; Prochiantz, A. *J Cell Sci* **1992**, 102 (Pt 4), 717–722.

(14) Futaki, S.; Suzuki, T.; Ohashi, W.; Yagami, T.; Tanaka, S.; Ueda, K.; Sugiura, Y. *J Biol Chem* **2001**, 276, 5836–5840.

(15) Futaki, S.; Nakase, I.; Suzuki, T.; Youjun, Z.; Sugiura, Y. *Biochemistry* **2002**, 41, 7925–7930.

(16) Baker, T. J.; Luedtke, N. W.; Tor, Y.; Goodman, M. *J Org Chem* **2000**, 65, 9054–9058.

(17) Luedtke, N. W.; Baker, T. J.; Goodman, M.; Tor, Y. *J. Am. Chem. Soc.* **2000**, 122, 12035-12036.

(18) Luedtke, N. W.; Carmichael, P.; Tor, Y. *J Am Chem Soc* **2003**, 125, 12374–12375.

(19) Futaki, S. *Adv Drug Deliv Rev* **2005**, 57, 547–558.

(20) Fuchs, S.M.; Raines, R.T. *Biochemistry* **2004**, 43, 2438–2444.

(21) Tyagi, M.; Rusnati, M.; Presta, M.; Giacca, M. *J Biol Chem* **2001**, 276, 3254–3261.

(22) Bai, X.; Wei, G.; Sinha, A.; Esko, J. D. *J Biol Chem* **1999**, 274, 13017–13024.

(23) Esko, J. D.; Stewart, T. E.; Taylor, W. H. *Proc Natl Acad Sci U S A* **1985**, 82, 3197–3201.

(24) Jonathan B. Rothbard, Theodore C. Jessop, Richard S. Lewis, Bryce A. Murray and Paul A. Wender *J Am Chem Soc* **2004**, 126, 9506–9507.

(25) Astriab-Fisher, A., Sergueev, D.S., Fisher, M., Shaw, B.R., and Juliano, R.L. 60 **2000**.

(26) Josephson, L., Tung, C.H., Moore, A., and Weissleder, R. *Bioconjugate Chemistry* **1999**, 10, 186–191.

(27) Lewin, M., Carlesso, N., Tung, C.H., Tang, X.W., Cory, D., Scadden, D.T., and Weissleder, R. *Nature Biotechnology* **2000**, 18, 410–414.

(28) Wunderbaldinger, P., Joesphson, L., and Weissleder, R. *Bioconjugate Chemistry* **2002**, 13, 264–268.

(29) Polyakov, V., Sharma, V., Dahlheimer, J.L., Pica, C.M., Luker, G.D., and Piwnica-Worms, D. *Bioconjugate Chemistry* **2000**, 11.

(30) Torchilin, V.P., Rammohan, R., Weissig, V., and Levchenko, T.S. *Proceedings of the National Academy of Sciences* **2001**, 98, 8786–8791.

(31) Rothbard, J. B.; Jessop, T. C.; Lewis, R. S.; Murray, B. A.; Wender, P. A. *J Am Chem Soc* **2004**, 126, 9506–9507.

(32) Bame, K. J.; Esko, J. D. *J Biol Chem* **1989**, 264, 8059–8065.

(33) Jacobsson, I; Lindahl, U *J. Biol. Chem*. **1985**, 255, 5094–5100.

(34) Stirpe, F.; Gasperi-Campani, A.; Barbieri, L.; Falasca, A.; Abbondanza, A.; Stevens, W. A. *Biochem J* **1983**, 216, 617–625.

(35) Hunter, D. T., Jr.; Hill, J. M. *Nature* **1961**, 191, 1378–1379.

(36) Lanza, T. J.; Durette, P. L.; Rollins, T.; Siciliano, S.; Cianciarulo, D. N.; Kobayashi, S. V.; Caldwell, C. G.; Springer, M. S.; Hagmann, W. K. *J Med Chem* **1992**, 35, 252–258.

(37) Soelaiman, S.; Wei, B. Q.; Bergson, P.; Lee, Y. S.; Shen, Y.; Mrksich, M.; Shoichet, B. K.; Tang, W. J. *J Biol Chem* **2003**, 278, 25990–25997.

(38) Montecucco, C.; Tonello, F.; Zanotti, G. *Trends Biochem Sci* **2004**, 29, 282–285.

(39) Panchal, R. G.; Hermone, A. R.; Nguyen, T. L.; Wong, T. Y.; Schwarzenbacher, R.; Schmidt, J.; Lane, D.; McGrath, C.; Turk, B. E.; Burnett, J.; Aman, M. J.; Little, S.; Sausville, E. A.; Zaharevitz, D. W.; Cantley, L. C.; Liddington, R. C.; Gussio, R.; Bavari, S. *Nat Struct Mol Biol* **2004**, 11, 67–72.

(40) Luedtke, N. W.; Liu, Q.; Tor, Y. *Chemistry* **2005**, 11, 495–508.

(41) Luedtke, N. W.; Liu, Q.; Tor, Y. *Bioorg Med Chem* **2003**, 11, 5235–5247.

(42) Zimmerman, I; Zimmerman, H W Z. *Naturforsch* **1976**, 31, 656.

(43) Angyal, S J; Angyal, C L *J. Chem. Soc*. **1952**, 1461.

(44) Albert, A; Goldacre, R; J, Phillips *J. Chem. Soc*. **1948**, 2240.

Chapter 3

3.1 Introduction

Aminoglycosides antibiotics are a diverse family of RNA binding pseudo amino-sugars which are known to target specific sequences of ribosomal RNA and interfere with prokaryotic protein biosynthesis via translational inhibition and misreading of the genetic code.[1,2] When bound to the ribosomal decoding site, neomycin class aminoglycosides induce a conformational change in the RNA which causes non-cognate tRNA to be recognized by the ribosome, resulting in translational miscoding.[3-6]

The interactions between aminoglycosides and RNA can be very specific. This conformational change in three adenosine residues of an internal loop within the prokaryotic A-site is not seen when paromomycin is bound to the eukaryotic A-site.[7] Change of a single nucleotide, A1408 to G distinguishes between prokaryotic and eukaryotic ribosomes (Figure 3.1). This one nucleotide change induces a 20-50 fold difference in binding affinity for paromomycin between the two ribosomal RNA sequences, making the interactions between paromomycin and the prokaryotic A-site RNA among the most selective RNA-aminoglycoside interactions.

For the most part, however, aminoglycosides are highly promiscuous RNA binders. While there is specific binding to the decoding region on prokaryotic 16S rRNA,[8] aminoglycosides bind to a variety of other RNAs including group I introns,[9,10] the hammerhead[11,12] and hepatitis delta ribozymes,[13,14] RNase P,[15] transfer RNA[16] and

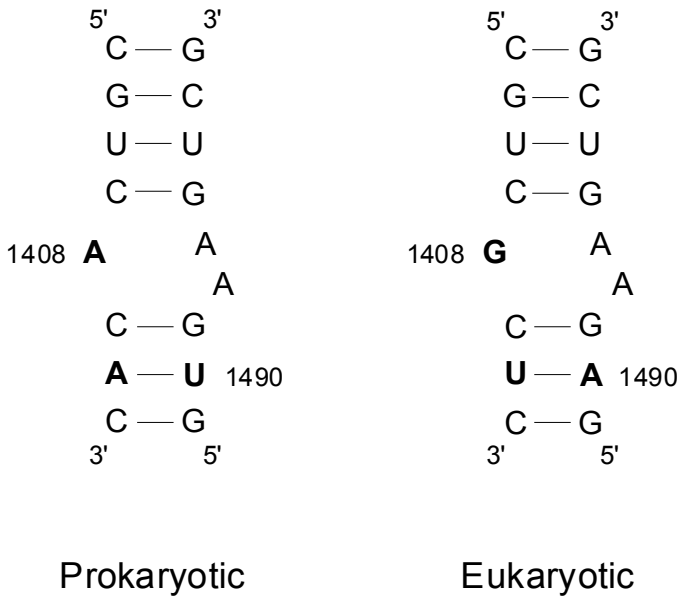


Figure 3.1 Sequence of the decoding region A site RNA from prokaryotic (*E.coli*) and eukaryotic (*Tetrahymena thermophila*) ribosomal RNAs (numbered according to *E.coli* sequence. Changing nucleotide 1408 from an A to a G confers a marked resistance to aminoglycoside antibiotics.

other large and small RNA targets.[16] This promiscuity shows that while aminoglycosides have the potential to bind to different RNA targets in a reasonably selective manner, they show an overall lack of specificity.

High resolution structural analysis of RNA-aminoglycoside interactions has mostly focused on 16S ribosomal RNA[3,4] and RNA aptamers. Analysis of these structures has yielded important insight into the elucidation of the specific requirements for selective binding of aminoglycosides to their RNA targets.

Aminoglycosides with common chemical groups of rings I and II (Figure 3.2) (paromomycin and gentamycin) make similar contacts with the A-site RNA with both base stacking and hydrogen bonding interactions.[4] The contacts made between rings I and II and specific bases on the A-site RNA likely comprise a common binding mode

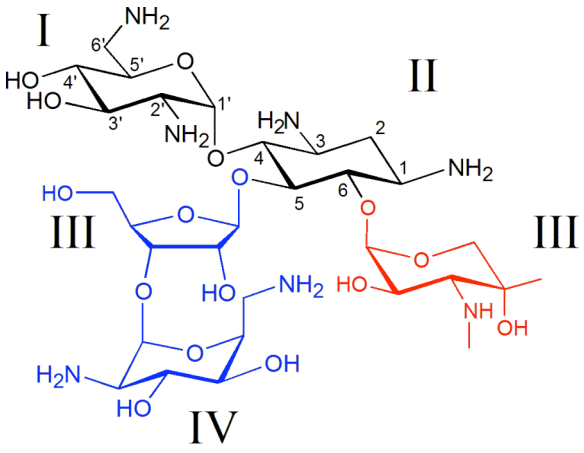


Figure 3.2 Structure of the rings and linkages of aminoglycoside antibiotics that bind in the A site of 16S rRNA. Rings I and II (black) are conserved among all aminoglycosides. Gentamicin class aminoglycosides (rings I and II and red) have a 4-6 linkage on the deoxystreptamine ring while the neomycin class (rings I and II and blue) have a 4-5 linkage to ring II.

for aminoglycosides which target the decoding site, as these rings are highly conserved among most members of the aminoglycoside family.

Ring III (and IV) of neomycin class aminoglycosides are linked at the 5 position on ring II, while ring III of gentamicin is connected to position 6 of ring II. Because the RNA interactions with rings I and II are conserved among aminoglycosides, and the linkages between rings II and III vary, this results in different RNA binding sites for ring III of gentamicin and rings III and IV of paromomycin. Ring III of gentamycin and rings III and IV of paromomycin each direct specific aminoglycoside-RNA interactions with conserved base pairs on the A-site.

It appears that while the interactions may be slightly different, each ring is making a contribution to the overall affinity and selectivity of the binding interaction. This shows that a single RNA target may be able to specifically accommodate multiple

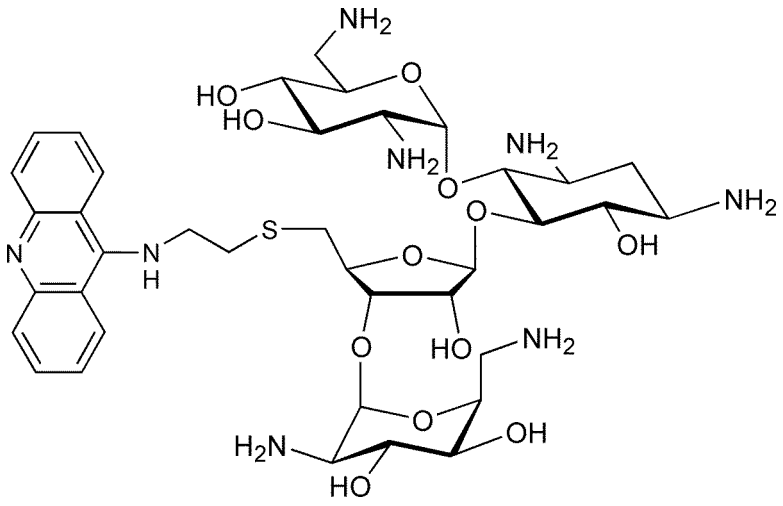


Figure 3.3. Neo-acridine is a potent inhibitor of the Rev-RRE interaction and binds to the RRE with high affinity. However, it shows very little RNA selectivity, prompting the search for novel intercalator-aminoglycoside conjugates with increased specificity for the RRE RNA.

different binding ligands as long as those ligands contain pharmacophores which have specific interactions with the nucleic acid.

When analyzing possible aminoglycoside-RNA interactions, it is important to understand the structural motifs of the RNA as well as the pharmacophores on the ligand. Small molecules which are not selective for a specific RNA may still recognize a short sequence or structural motif within the RNA. Ring II of aminoglycosides (2-deoxystreptamine) is the most conserved element among aminoglycoside antibiotics. Studies have shown that the 2-deoxystreptamine ring interacts specifically with G-U steps within a widened major groove of the RNA.[17] The major groove must be opened up by the presence of bulged residues to allow the ring to bind.

While it may be possible to design pseudo oligosaccharides, similar to aminoglycosides, to bind to specific RNA targets based on the sequence and structural selectivity of the building blocks, most synthetic molecules show a lack of selectivity

when compared to aminoglycosides. To that end, it may be possible to enhance ligand selectivity by conjugating aminoglycosides to other molecules which may have a preference for certain RNA sequences or, more likely, structural motifs.

Because aminoglycosides prefer bulged RNA, with widened major grooves, early studies in our lab focused on designing ligands targeted to sequences with the HIV viral genome. Aminoglycosides have been shown to bind to viral RNA, blocking the formation of RNA-peptide and RNA-protein complexes in both the HIV Tat-TAR[18] and Rev-RRE[19-22] systems, thus enhancing their attractiveness as targets for modified aminoglycoside based ligands.

Intercalating agents, like aminoglycosides, target bulged duplex regions of RNA.[23,24] Potentially, the combination of the ionic interactions and hydrogen bonding seen with aminoglycoside and an intercalating moiety may create more RRE selective RNA binding molecules.

A conjugate of neomycin and acridine was designed and synthesized by Sarah Kirk (Figure 3.3).[26] Neomycin contains six basic amines, none of which were used to conjugate to the acridine intercalator. By not linking at one of these positions, all of the amino groups are able to retain their charge and maximize the potential electrostatic interactions between the aminoglycoside and RNA. The primary hydroxyl on ring III was used as the site for aminoglycoside modification because it is an accessible handle which does not appear to be necessary for RNA binding.[12] Acridine was chosen as the initial intercalator because it is a small, planar fused ring system which is known to intercalate between the base pairs of DNA and RNA.[25]

Neo-acridine displaces the Rev peptide from the RRE-RNA roughly 11 times more effectively than neomycin B,[26] and is one of the most effective known inhibitors of the Rev-RRE binding interaction, with an K_i of 1.5 nM for Rev displacement. The affinity of neo-acridine to the RRE is roughly two fold lower than the affinity of the Rev peptide.

However, footprinting assays show that it has two binding sites on the 66 mer RRE construct used. One site is shared with the Rev peptide and is seen at ligand concentrations of 0.5-2 μM. The other is found towards another bulged nucleotide and is roughly 5-10 μM. A lack of a single, high affinity binding site is suggestive that neo-acridine may lack a high degree of selectivity between different RNA molecules. This lack of selectivity is experimentally observed. When the RRE affinity of neo-acridine was examined in the presence of both RNA and DNA competitors[27] binding to the RRE was significantly decreased. Neo-acridine binds to multiple RNAs and also to DNA, not exhibiting a very high RRE specificity.

When the linker length between the neomycin and the acridine is shortened, affinity for the RRE is not affected, however selectivity is greatly improved when DNA or tRNA are used as competitors. Conversely, when the linker length is extended, there is a marked decrease in RRE specificity.[28] Decreasing linker lengths provides fewer degrees of structural freedom, limiting the variety of contacts that an aminoglycoside-intercalator conjugate can make. This loss of flexibility limits the number of RNA molecules to which the conjugate can bind to with high affinity, increasing its selectivity. However, even when conjugated to via a short linker, acridine does not show very high inherent selectivity for large bulged RNA areas like that seen in the

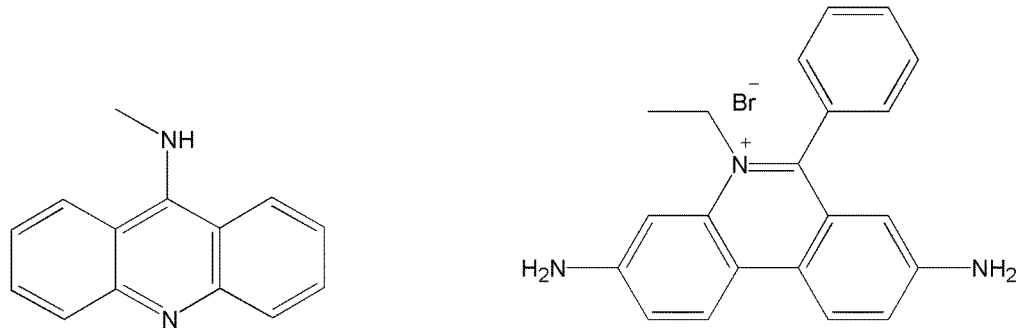


Figure 3.4. Acridine intercalators are small fused ring systems which bind to a large variety of different RNA and DNA molecules with little to no selectivity. Ethidium bromide, in contrast, has a preference for bulged duplex nucleic acids and has sub micro molar affinity for the RRE RNA. When bound to an aminoglycoside ethidium may impart a higher RRE selectivity.

RRE, prompting the search for new intercalators (Figure 3.4) which can be bound to aminoglycosides and will provide molecules with higher specificity.

In this chapter we will show that aminoglycosides and guanidinoglycosides conjugated to the larger, methidium moiety have a very high affinity for the RRE RNA, efficiently displacing the Rev peptide with low IC_{50} values in the low to mid nanomolar. All of the conjugates display a lower affinity to tRNA, poly A-U RNA and calf thymus DNA, indicating at least a moderate RRE selectivity. Gel shift experiments with the RRE show that a selected compound, guanidino-tobramycin-N-methidium, does not show any selectivity between RRE constructs where two bulged bases have been removed or three guanines were replaced with insoines and the wild type sequence. A wild type DNA RRE sequence, however, shows markedly reduced binding of the guanidinoglycoside-methidium conjugate when compared to RNA.

This data is indicative of a model of binding where these conjugates may show a moderate selectivity for the RRE over other RNA molecules, but show a very high RNA over DNA selectivity.

3.2 Design and Synthesis of Aminoglycoside-Methidium Conjugates

A larger, extended planar ring system may preferentially recognize and bind to bulky RNA bulges like that found in the RRE, where there are two purine-purine base pairs with a bulged U between them. Fortuitously, such a molecule, with known nucleic acid binding properties has previously been discovered.

3,8-diamino-5-ethly-6-phenylphenanthridinium bromide, more commonly known as ethidium bromide, is a common stain for double stranded DNA and RNA.[29] The nucleic acid binding properties of ethidium bromide have been studied by monitoring the changes in emission upon addition of nucleic acids.[30] Studies have shown that ethidium binds to DNA and RNA with affinities ranging from 0.2-200 μM (figure 3.4).

Ethidium has been shown to not only bind to the surface of nucleic acids, but intercalate between the base pairs as well,[31-33] showing a preference for duplexes containing bulged nucleotides over similar duplexes without bulges.[24] Ethidium bromide displays a high affinity for the RRE RNA (K_d = 0.2 μM) and can efficiently displace the Rev peptide.

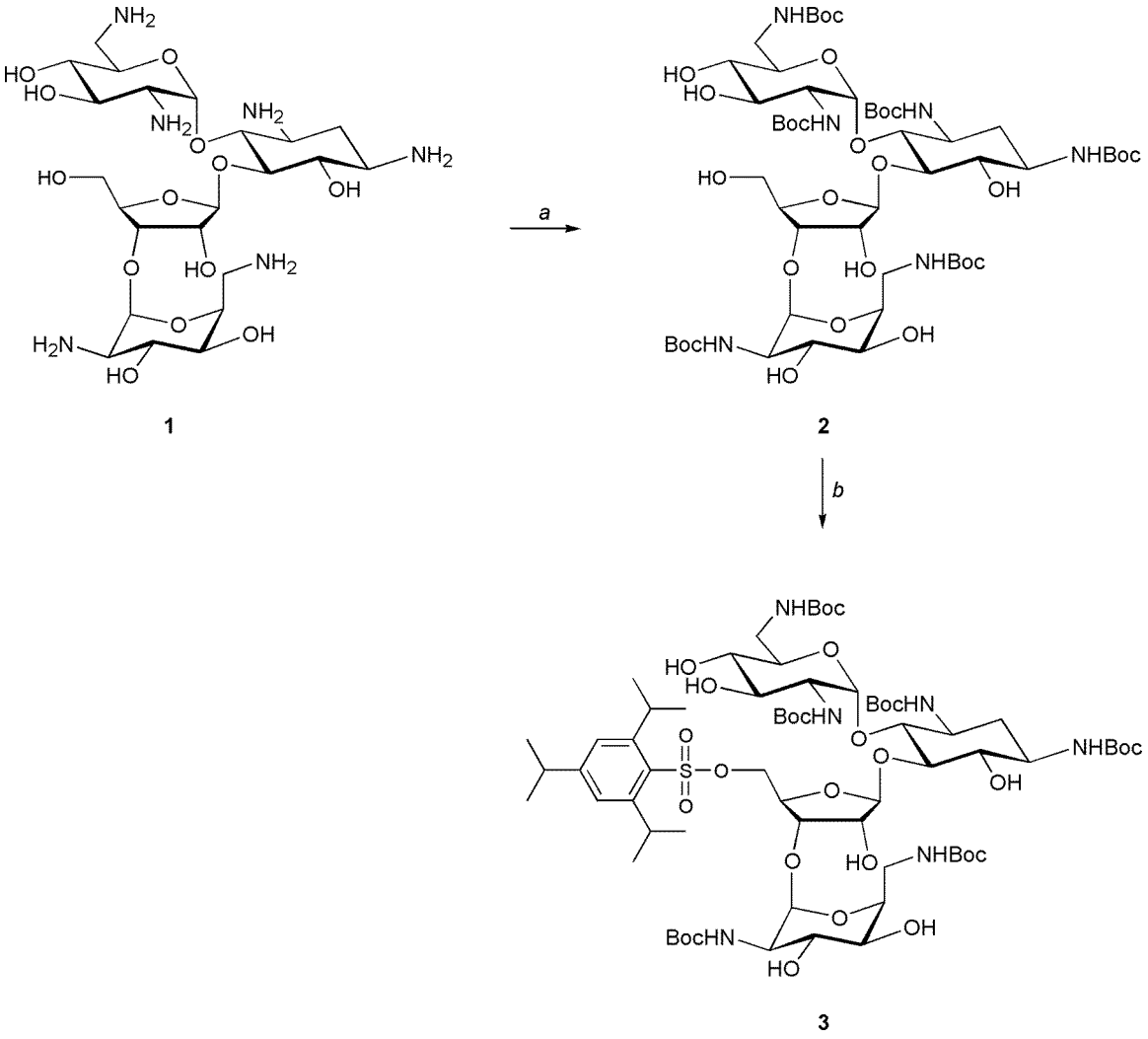
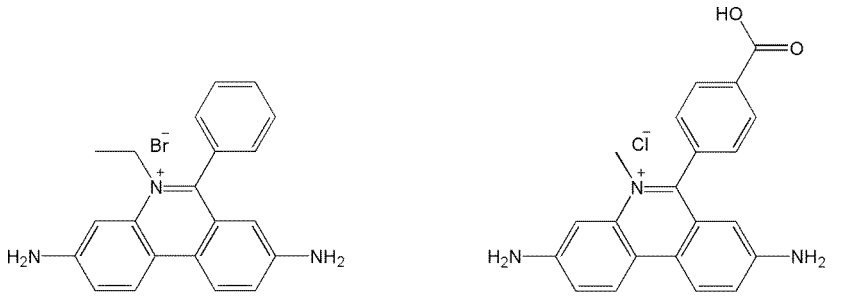


Figure 3.5. Synthesis of different neomycin-intercalator conjugates is facilitated by the existence of an intermediate which can be conjugated to a variety of linkers. a) **1** and Boc_2O in DMF/H_2O (5/1) (89%) b) **2** and 2,4,6-triisopropylbenzenesulfonyl chloride in pyridine (51%)

Ethidium bromide also has a phenyl ring attached at the 6 position (Figure 3.4). The presence of this bulky group may limit the angles at which ethidium can enter the nucleic acid duplex when intercalating. This constraint may impart an increased RNA selectivity by discouraging the binding of ethidium to some RNA sequences more than



Ethidium

6-phenyl-3,8-diamino-5-
-ethylphenanthridinium Br

Methidium

6-(4-carboxyphenyl)-3,8-diamino-5-
-methylphenanthridinium Cl

Figure 3.6. A commercially available ethidium bromide based molecule was conjugated to neomycin using the same 2-aminoethanethiol linker used with the original neo-acridine compound. The main difference between the methidium intercalator used an ethidium bromide is the presence of a methyl group as opposed to an ethyl group at the 5-position of the molecule. A carboxylic acid on the phenyl ring allows for conjugation via amide bond formation.

others. Because of its size, RRE affinity, and potential for increased selectivity, ethidium bromide was chosen as a model for intercalators to attach to aminoglycosides.

The six amines of neomycin were boc protected and the primary hydroxyl on ring III was activated with a triisopropylbenzene sulfonyl leaving group to afford a molecule which could be conjugated with multiple different linker and intercalator systems, allowing for a wide variety of compounds to be relatively easily synthesized (figure 3.5).

The 2,4,6-triisopropylbenzenesulfonyl group on the primary hydroxyl of ring III can be displaced with a variety of linkers. Here, we use an azide (sodium azide) which we later reduce to an amine to afford the minimal linker length, or a thiol (2-aminoethane thiol) which gives a slightly longer linker which retains the primary amino functional group. However, other linkers can be used to provide a variety of molecules.

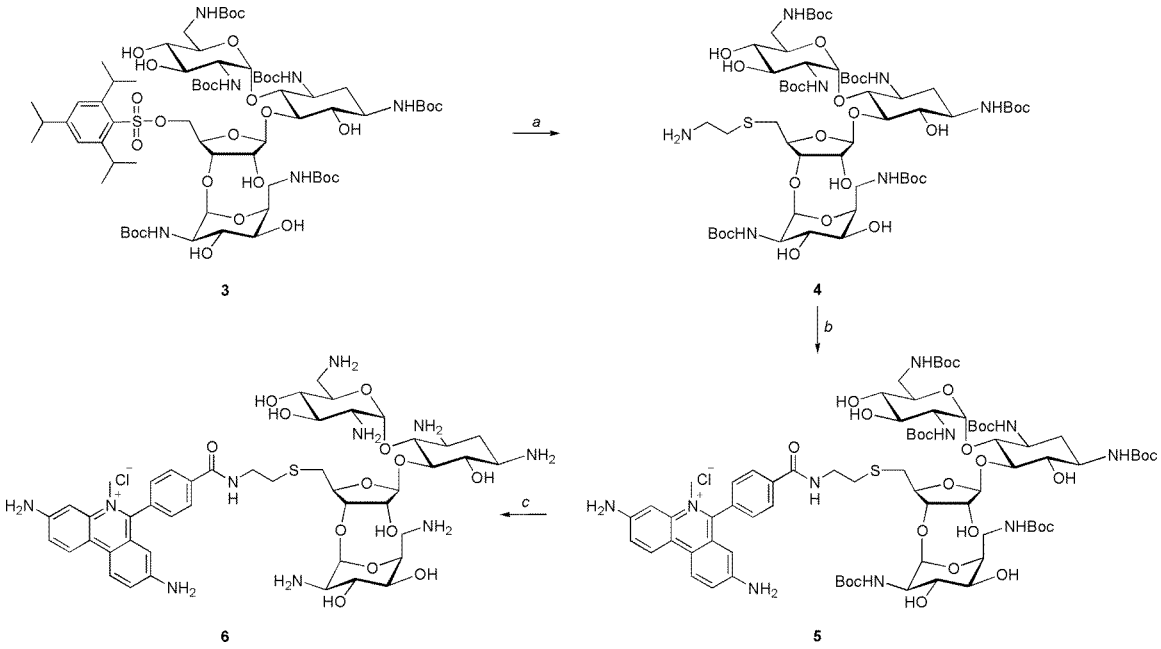


Figure 3.7. Synthesis of a neomycin-methidium conjugate. a) **3**, 2-aminoethanethiol HCl, NaOEt (81%) b) **4**, HOBT, TBTU, 6-(4-carboxyphenyl)-3,8-diamino-5-methylphenanthridinium Cl (**22**), DMF (82%) c) **5**, trifluoroacetic acid/dichloromethane (1/1).

Tobramycin is an aminoglycoside with a similar core structure to the neomycin class antibiotics. However, tobramycin has only three rings, with ring III being connected to the 2-deoxystreptamine core with a 4-6 linkage as opposed to the 4-5 linkage seen in the neomycin class aminoglycosides. Additionally, there are five amines on tobramycin, while neomycin has six. These differences make tobramycin an attractive scaffold for the design of a second aminoglycoside-intercalator conjugate.

Neomycin and tobramycin each contain one primary hydroxyl. On neomycin, this functional group is on ring III whereas it is on ring I in tobramycin (Figure 3.11). Structural analysis of neomycin class aminoglycosides with the A-site RNA reveals that the 6′ position on ring I plays a critical role in binding, making specific contacts with the RNA. However, the primary hydroxyl on ring III does not appear to make specific

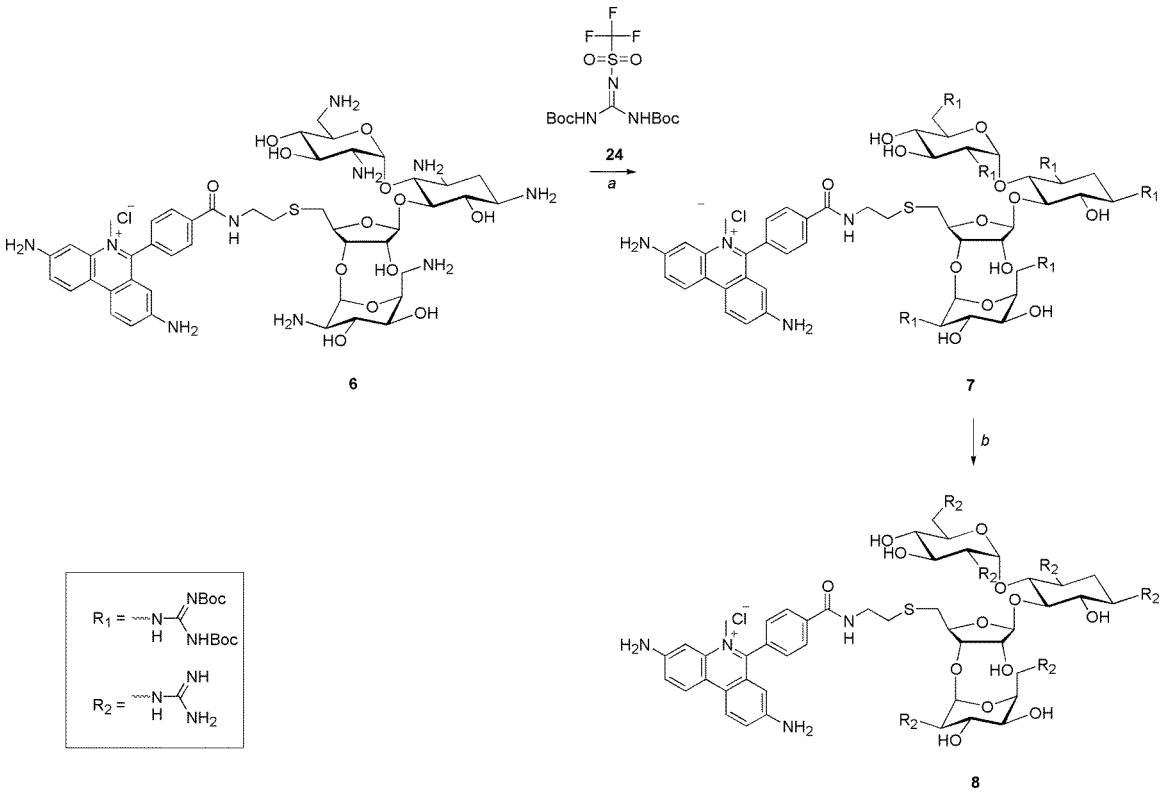


Figure 3.8. Guanidinylation of neomycin-methidium afforded a compound where the amines on the aminoglycoside portion of the molecule were converted to guanidinium groups while the aromatic amines on the methidium moiety remain unconverted. a) **6**, **24**, triethylamine, methanol/chloroform (3/1) b) **7**, trifluoroacetic acid/chloroform (1/1).

contacts with the RNA and is actually directed away from the RNA helix. If this is a general trend for aminoglycoside binding to RNA, modifications at these positions may have markedly different effects on both selectivity and affinity, and could partially explain some of the differences between the neomycin-acridine and tobramycin-acridine conjugates. This is something which should be taken into consideration when analyzing binding results from modified neomycin and tobramycin based molecules.

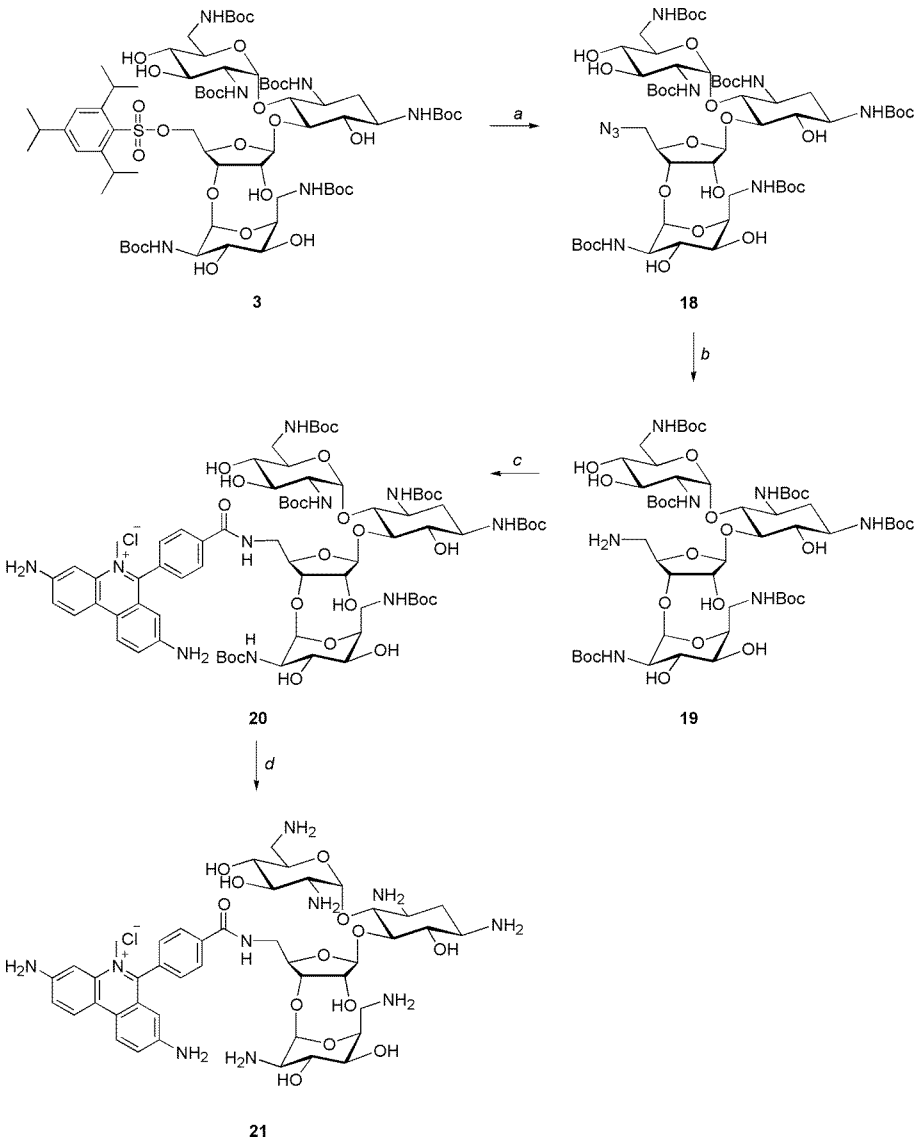


Figure 3.9. Synthesis of neomycin-N-methidium. A shorter linker may provide a molecule with increased selectivity. a) **3**, sodium azide, DMF/H_2O (5/1) (80%) b) **18**, palladium on carbon, methanol (85%) c) **19** HOBT, TBTU, 6-(4-carboxyphenyl)-3,8-diamino-5-methylphenanthridinium Cl (**22**), DMF (13%) d) **20**, trifluoroacetic acid/dichloromethane (1/1).

Acridine has been conjugated to tobramycin via the 6″ hydroxyl, and tobramycin-N-acridine has a 300-fold higher IC_{50} for Rev displacement than

tobramycin.[28] This result is slightly surprising as conjugation of acridine to Neomycin increases its affinity by roughly 65-fold. However, tobramycin-N-acridine has shows an increase in selectivity over the neomycin-acridine conjugates. With the acridine compounds, as the affinity for RRE increases, there appears to be a decrease in RRE selectivity.

To enable as direct as possible comparison between the acridine and methidium based intercalating moieties, neomycin-methidium was synthesized (Figure 3.7). Additionally, all of the amino groups on the neomycin portion of the aminoglycoside--methidium conjugate were guanidinylated (figure 3.8). Guanidinylation of aminoglycosides has been shown to increase their RNA affinity and anti-HIV properties.[34] A guanidinylated neomycin-methidium may show increased RRE affinity and/or selectivity.

It is interesting to note that while full conversion of the aminoglycoside amines to guanidinium groups is observed, none is seen with the aromatic amines at the 3 and 8 positions of the methidium moiety. The exocyclic amines of ethidium are poor nucleophiles and are weakly basic (pKa_3 = 0.8, pKa_8 = 2), necessitating the use of highly reactive electrophiles when to modify them.[35] While these amines have been converted to guanidium groups, the chemistry requires using a protected S-methyl-isothiourea with mercury dichloride. Attempts at guanidinylation using the reagent used with aminoglycosides[36] were unsuccessful.

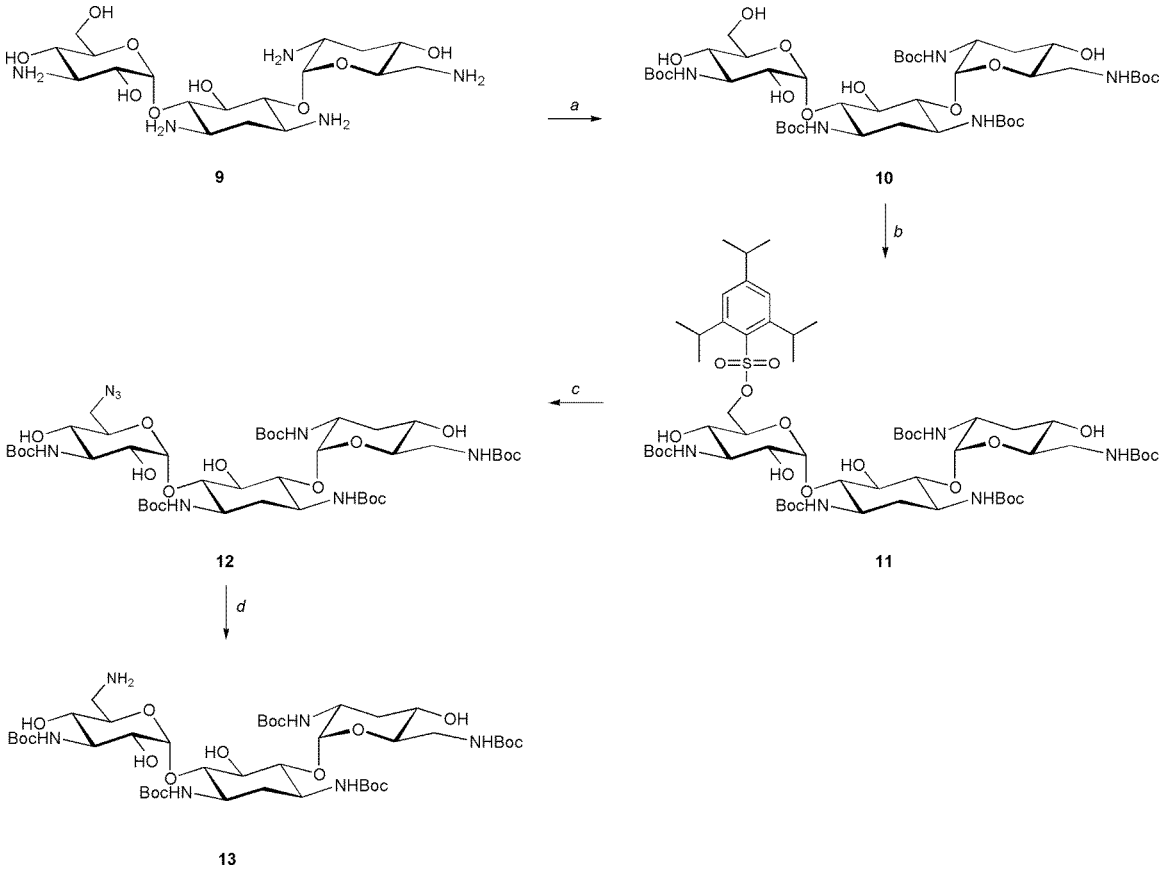


Figure 3.10. Protection and activation of tobramycin leads to the formation of an aminoglycoside to which multiple linkers and conjugates can be added. a) **9**, Boc_2O, DMSO/H_2O (5/1) (91%) b) **10**, 2,4,6-triisopropylbenzenesulfonyl chloride, pyridine (51%) c) **11**, sodium azide, DMF/H_2O (4/1) 96% d) **12**, palladium on carbon, methanol (49%)

To probe the effect of linker length, a neomycin-N-methidium was also designed and synthesized. Instead of displacing the tirisopropylbeneze sulfonyl group at the primary (5) position of ring III with an aminoethanethiol, sodium azide was used and subsequently reduced to an amine, resulting in a shorter linker. Reduced linker length gave increased selectivity in the neo-acridine conjugates.[26] A similar trend may be seen with the methidium conjugates as well.

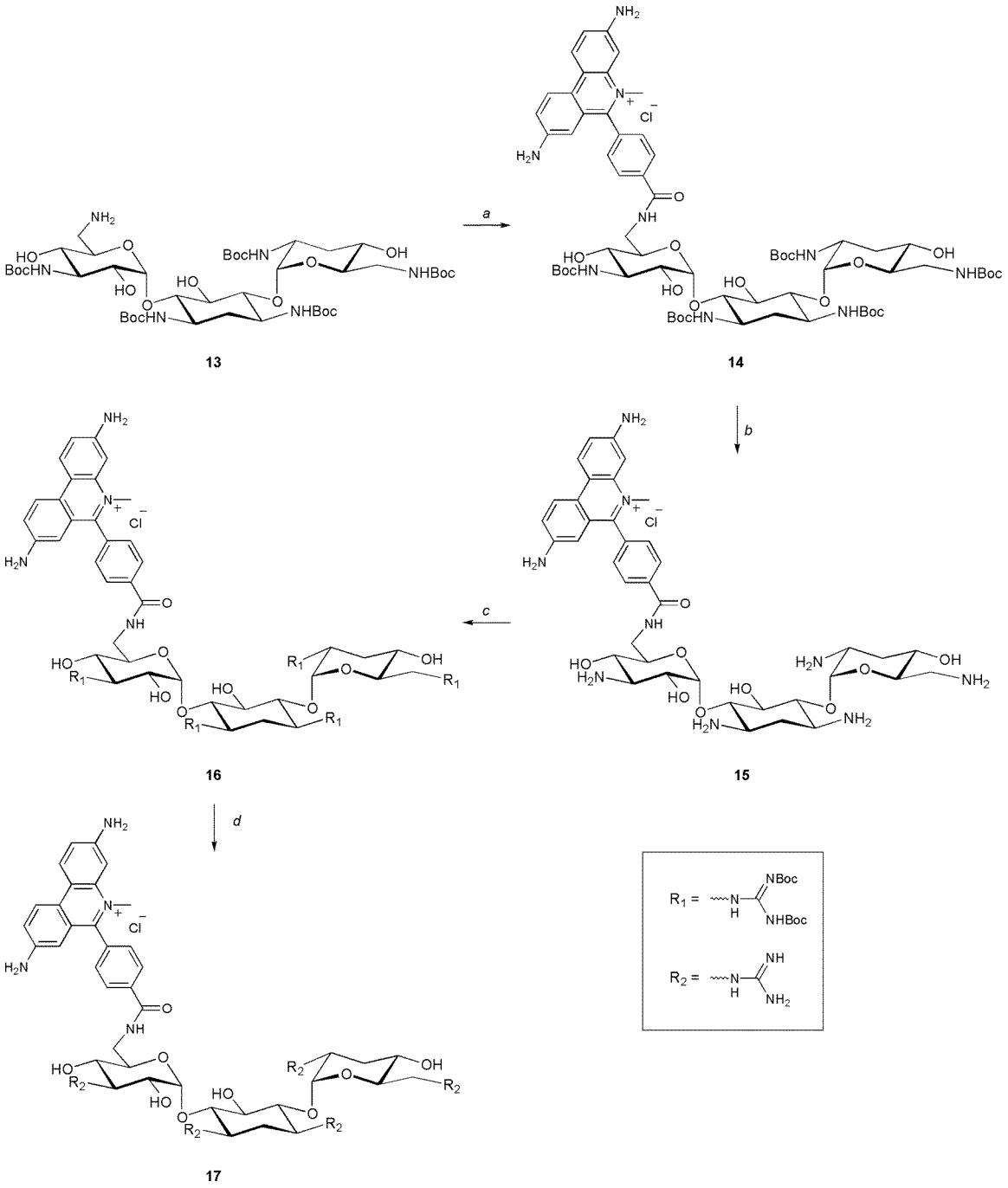
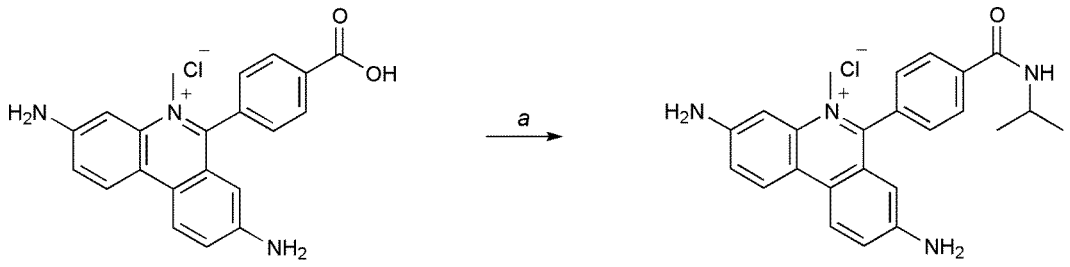


Figure 3.11. Tobramycin is conjugated to methidium at the primary (6′) position of ring I, providing tobramycin-N-methidium which is then guanidinylated. a) **13**, HOBT, TBTU, 6-(4-carboxyphenyl)-3,8-diamino-5-methylphenanthridinium Cl (**22**), DMF (91%) b) **14**, trifluoroacetic acid/dichloromethane (1/1) (83%) c) **15**, **24**, methanol/chloroform (10/3) (25%) d) **16**, trifluoroacetic acid/dichloromethane (1/1).



22 **23**

Figure 3.12. Synthesis of an isopropylamine-methidium conjugate. This compound allows for precise determination of extinction coefficients for all aminoglycoside-methidium conjugates. a) isopropylamine, HOBT, TBTU, 6-(4-carboxyphenyl)-3,8-diamino-5-methylphenanthridinium Cl (**22**), DMF.

To examine potential changes in affinity and selectivity between different aminoglycosides, a tobramycin-methidium conjugate was designed and synthesized (Figure 3.11). Due to the fact that the primary hydroxyl was used as the site for conjugation of both the neomycin and tobramycin derivatives, there is a major structural difference between the two classes of molecules.

When examining the biophysical properties of a group of related compounds which may potentially display similar activities, it is important to quantify the concentrations of compound in each experiment. The presence of a chromophore on a molecule allows for precise concentration measurements (Figure 3.12)

3.3 Nucleic Acid Binding Properties of Aminoglycoside-Methidium Conjugates

To examine the ability of the methidium conjugated amino and guanidinoglycosides to bind to the RRE and displace the Rev peptide, compounds were tested with a fluorescence anisotropy assay. A fluorescein label was attached to

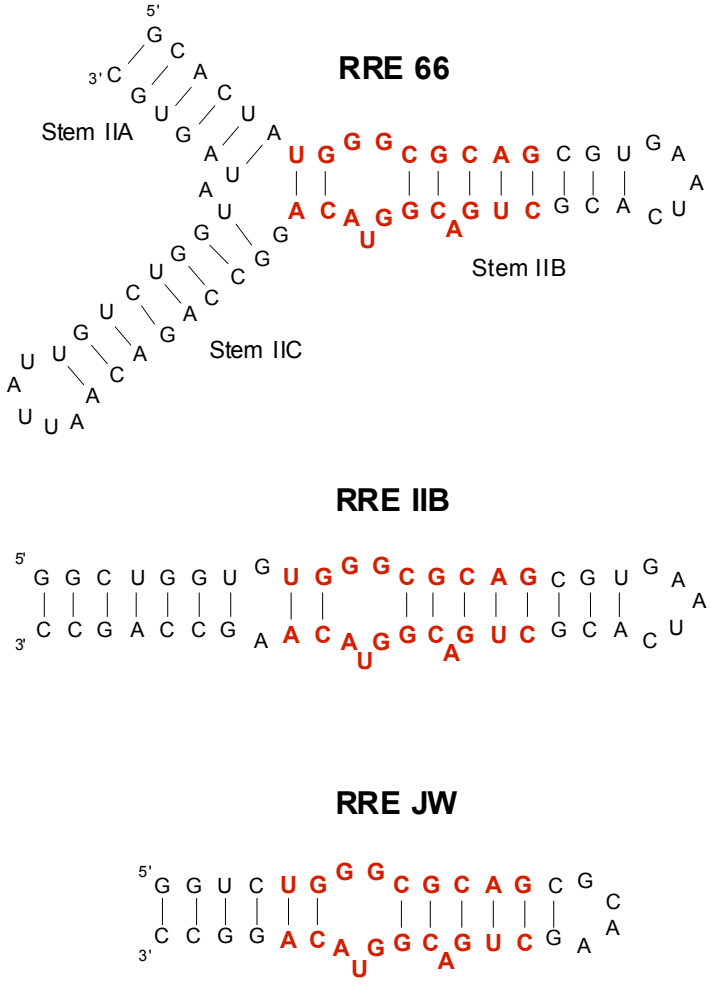


Figure 3.13. The three main wild type RRE constructs used for evaluating small molecule binding include a 66 nucleotide RRE which closely mimics the wild type sequence. The core sequence (Red) serves as a high affinity Rev binding site. Minimization of this sequence gives a 47 nucleotide fragment RRE IIB which can be further minimized to a construct which retains Rev binding, but is small enough to be used for NMR structural studies.

the Rev peptide at the C-terminus. Observations of the anisotropy changes of the fluorescent tag allow for quantitative measurements of free versus RNA bound peptide.

Following the formation of a complex between the fluorescent tagged Rev peptide (Rev-Fl) and the RRE RNA, inhibitors can be titrated into the solution and we are able to observe a change in anisotropy which is directly proportional to the ratios of

free to bound peptide. Once the peptide has been fully competed off of the RNA, the anisotropy value measured will saturate at the same level as that seen with the free peptide.

Multiple different RRE constructs exist, but there are three which have been used predominantly in our lab (Figure 3.13). The first is a 66-mer which contains the three stems (stem IIA, stem IIB and stem IIC) which make up the high affinity Rev binding site and include the three helix junction found in the viral genome. The second is a minimized construct based on stem IIB.[37] This is a smaller RNA molecule which, while retaining the purine-purine base pairing and large bulge seen in the high affinity Rev binding site, it does not have the three helix junction. The third construct is a further minimized construct based on the RRE IIB where the stem is shortened on both sides of the high affinity Rev binding site.[38]

Conserved in all constructs is a core element (bold sequence in Figure 3.13) which includes a unique bulged sequence of two purine-purine (G-A and G-G) base pairs with a bulged U in between them. This element serves as the foundation of the high affinity Rev binding site, and contributes the same Rev binding energy in all three of the RRE constructs.

When unlabeled Rev peptide is used as an inhibitor to compete off the Rev-Fl binding to the RRE 66 construct, it displays an IC_{50} value of 12±5 nM. However, when the aminoglycoside and guanidinoglycoside-methidium conjugates are used as Rev-Fl competitors, they are shown to cause dissociation of the Rev-Fl RRE complex as well as or better than the Rev peptide (Figure 3.14).

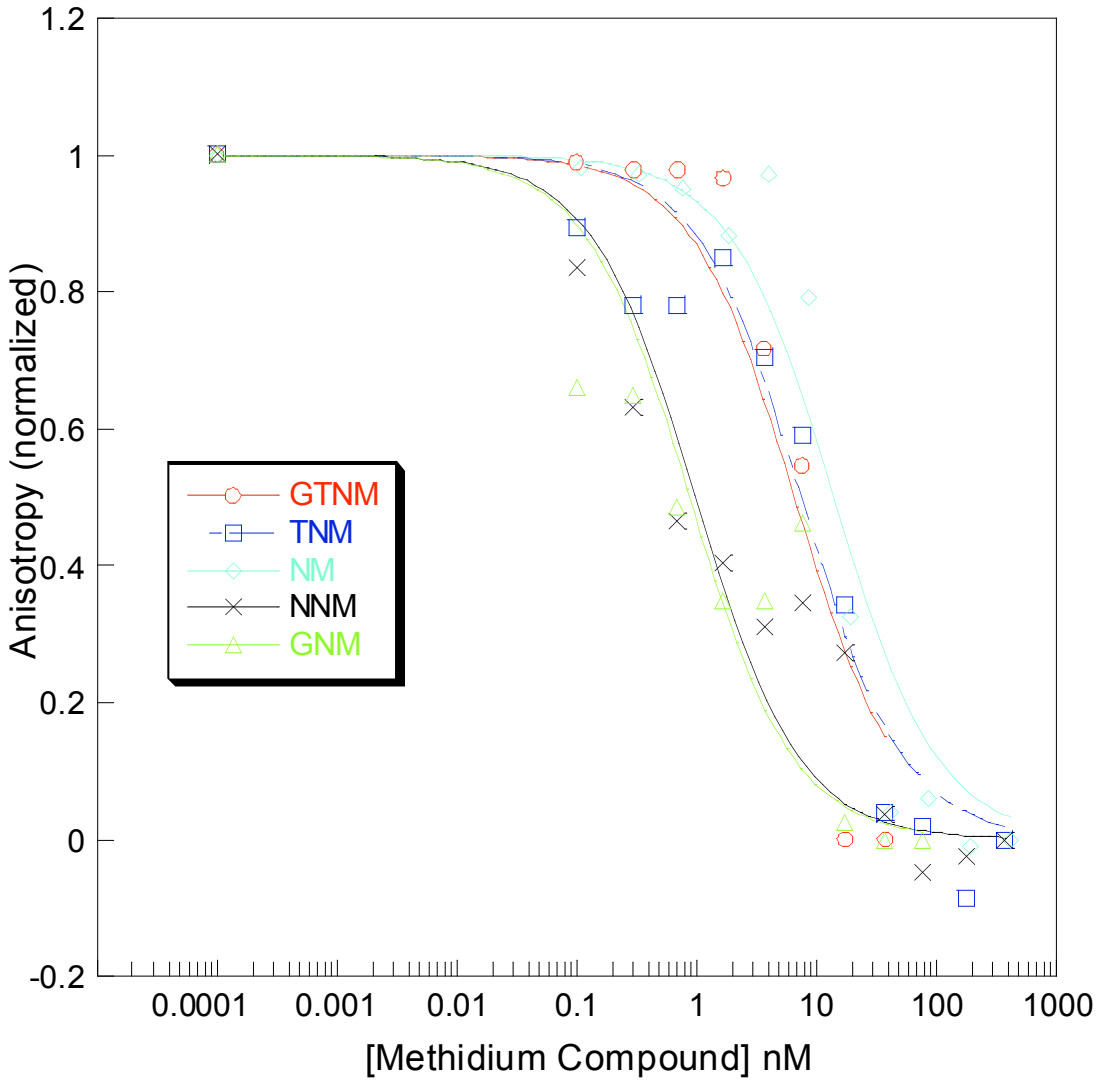


Figure 3.14. Rev-Fl displacement anisotropy experiments with aminoglycoside-methidium and guanidinoglycoside-methidium conjugates. Guanidinylation and reduction of the linker length of neomycin-methidium both increase the Rev-Fl inhibition activity of the compound to roughly 10 fold more active than the unlabeled Rev peptide. Guanidinylation of tobramycin-N-methidium does not appear to affect its Rev-Fl competition activity.

Unexpectedly, linker length appears to play a large role in RRE binding and Rev-Fl displacement. Neomycin-Methidium has the aminoethanethiol linker between the aminoglycoside and methidium and it has the highest IC_{50} value of all of the compounds tested. In contrast, Neomycin-N-methidium is composed of the same two

	Rev Displacement Anisotropy	Direct Titrations tRNA	Direct Titrations CT DNA	Direct Titrations Poly AU RNA
Neo-Meth	19 ± 3 nM	2.1 ± 1 μM	1.5 ± 0.5 μM	0.9 ± 0.3 μM
Neo-*N*-Meth	1 ± 0.2 nM	2.2 ± 1 μM	2 ± 0.7 μM	1.2 ± 0.3 μM
Guan-Neo-Meth	0.9 ± 0.3 nM	2.0 ± 1 μM	3 ± 0.9 μM	2.0 ± 0.4 μM
Tob-*N*-Meth	10 ± 4 nM	1.9 ± 1 μM	3 ± 0.7 μM	1.1 ± 0.4 μM
Guan-Tob-*N*-Meth	9 ± 3.5 nM	2.0 ± 1 μM	4 ± 0.8 μM	3.3 ± 0.8 μM

Figure 3.15 Nucleic Acid binding properties of methidium conjugates reveals that the methidium moiety contributes to a high affinity for RRE and selectivity over other nucleic acids. When titrated into a complex of RRE and a fluorescent tagged Rev peptide, the compounds show the ability to displace the peptide with activity in the low nanomolar. Direct titrations of tRNA, CT DNA and poly AU RNA into the different compounds reveals little difference in binding from one molecule to the other, and little selectivity between the different nucleic acids.

building blocks with the linker minimized, and it shows an IC_{50} value roughly 10-fold lower.

The effects of guanidinlyation are not as clear. When neomycin-methidium is guanidinylated, there is the same increase in Rev-Fl displacement activity as seen when the linker length is shortened, signifying that the guanidinium group plays a large role in binding. However, when tobramycin-N-methidium is compared with guanidino-tobramycin-N-methidium, virtually no difference in activity is observed (Figure 3.15).

To determine the specificity of the methidium compounds for RRE over other nucleic acid sequences, attempts were made to use a novel solid-phase assay which was developed in our lab.[27] In this assay, a biotinylated RRE is bound to streptavidin-immobilized agarose beads which are insoluble in water/buffer. The fluorescently tagged Rev-Fl is then bound to the complex via the RRE to make a ternary complex. Rev-RRE inhibitors are then incubated with this complex. Comparison of the

fluorescence in the supernatant versus the fluorescence of the agarose beads results in quantitative ratios of free Rev-Fl and Rev-Fl bound to RRE and associated with the beads.

Compounds can be titrated in the absence of competitor nucleic acids, and data generated from this should be roughly equivalent to that gained from the peptide displacement anisotropy assay. However, the advantage of this assay is that it can be done in the presence of other nucleic acids. If a Rev-RRE inhibitor also to other nucleic acids, it will show decreased Rev-RRE inhibition activity in their presence.

Aminoglycoside-methidium and guanidinoglycoside-methidium conjugates were assayed for their ability to bind to the RRE RNA and displace the Rev-Fl peptide into solution using the solid phase assay. Two RRE constructs, the RRE 66 and RRE IIB (Figure 3.16), were used with slightly (2-3 fold) higher activity being seen for the inhibitors with the RRE IIB.

However, when this experiment was performed in the presence of competitor nucleic acids (poly A-U RNA, tRNA, and calf thymus DNA), varying results were observed. Aggregation of the methidium compounds resulted is inconsistent IC_{50} values. Because of this, we were not able to determination of relative RRE specificity of the different methidium conjugates.

While selectivity experiments based on displacement of Rev-Fl from different RRE constructs in the presence of competitor nucleic acids did not prove possible, we were able to take advantage of the fluorescent properties of the methidium moiety to directly measure the affinity of the methidium conjugates to different nucleic acids.

As it binds to nucleic acids, ethidium bromide displays an increase in fluorescence emission. Measuring these spectral changes as a function of nucleic acid concentration allows us to generate binding isotherms and calculate binding affinities. Small molecules such as ethidium bromide and the methidium-amino/guanidinoglycoside show consistency between direct binding and Rev displacement. This allows us to make relatively direct comparisons between results generated from peptide displacement anisotropy and direct nucleic acid binding.

Direct nucleic acid titrations resulted in almost equivalent affinity values for each of the methidium conjugates. This trend was observed when compounds were titrated with poly A-U RNA, tRNA and calf thymus DNA. K_d values for nucleic acid titrations ranged from roughly 1 μM to over 25 μM. These affinities are much weaker than those seen with the different RRE constructs. This indicates that aminoglycoside and guanidinoglycoside-methidium conjugates possess selectivity for the RRE over other nucleic acids (Figure 3.15).

3.4 Measuring the Affinity of Methidium Conjugates to Modified RRE RNA

To understand the features of the RRE RNA which confer selectivity of the methidium conjugates, several small constructs based on the RRE were designed and assayed for binding of methidium molecules. Mutation of the minimized RRE JW construct may reveal which elements of the high affinity Rev binding site are essential for ligand binding. Because similar results were seen with all of the methidium

molecules, gel shift experiments with one compound (guanidino-tobramycin-N-methidium) are shown.

The different RRE constructs were all based on the short RRE JW molecule (Figure 3.16.). The RRE JW RNA is a wild type oligonucleotide which contains the high affinity Rev binding site. In the ΔAU construct, the bulged U found between the two purine-purine base pairs, and the bulged A two base pairs away were both removed. Three adjoining G residues within the high affinity Rev site, two of which are involved in purine-purine base paring were replaced with inosines in the 3I RRE. The dRRE JW construct is a DNA version of the wild type RRE JW sequence.

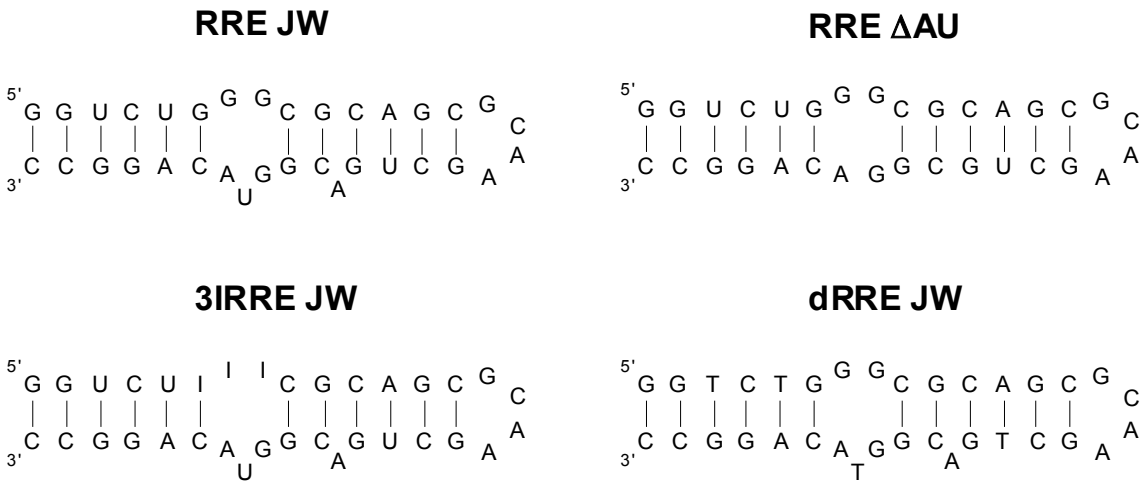


Figure 3.16. Minimized RRE constructs based on the RRE JW all contain the high affinity Rev binding site with the bulged U and A removed (RRE δAU), three G residues within the high affinity site replaced with inosines (3I RRE JW), or the wild type RNA sequence replaced with DNA.

Non-denaturing gel shift experiments were performed with both the RRE JW and dRRE JW constructs. Guanidino-tobramycin-N-methidium was titrated into solutions containing each RNA (with trace 5′-^{32}P RNA) in buffer. The RNA is resolved on a native polyacrylamide gel. As the compound binds to the RNA, the band for free RNA becomes less intense and a new band representing RNA with compound bound begins to form (Figure 3.17). This second band has a slower mobility than the free RNA.

Guanidino-tobramycin-N-methidium binds to the RRE JW at concentrations between 10 nM and 100 nM (Figure 3.17). This compound also binds to the ΔAUconstruct at the same concentrations. This indicates that removal of the bulged bases does not affect RRE affinity of the methidium conjugates. These results are slightly surprising as ethidium bromide, the parent molecule to the linker, shows a preference for RNA sequence with bulged nucleotides. It is possible that the presence of two purine-purine base pairs next to each other opens the RNA enough to accommodate the methidium moiety.

A third band begins to appear at concentrations above 100 nM. This band runs with a much slower mobility than the free RNA band or the RNA-small molecule band. This is indicative of a second binding site for the aminoglycoside-methidium conjugate on both RRE constructs. However, it is impossible to determine affinity values for this site as at concentrations of 1 μM and above guanidino-tobramycin-N-methidium aggregates with the RNA and gets stuck in the well.

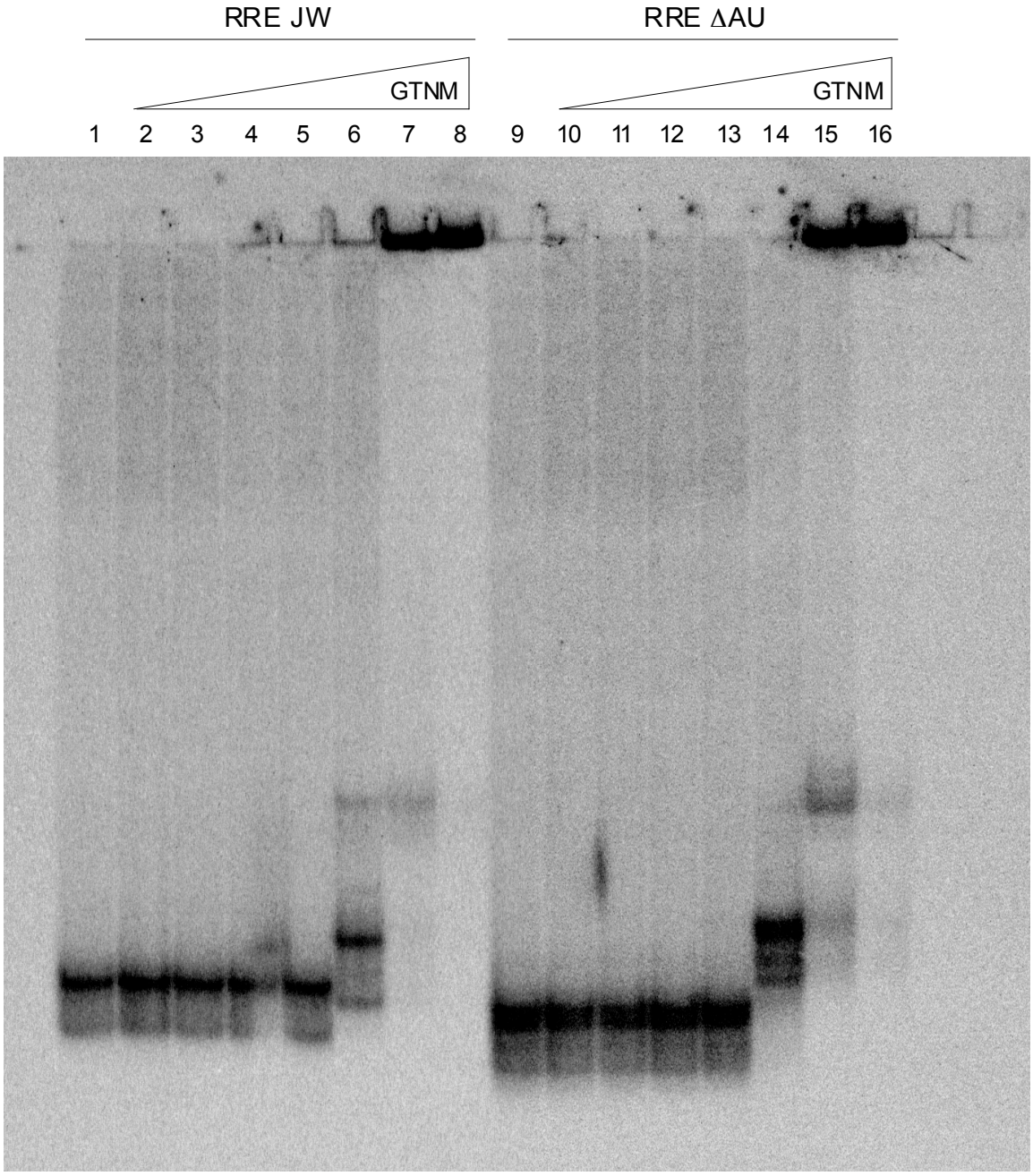


Figure 3.17. Titrations of guanidino-tobramycin-N-methidium into wild type RRE JW and RRE ΔAU RNA reveals that deletions of the two bulged bases does not affect the RNA affinity of the small molecule. Lanes 1-8 contain RRE JW RNA and lanes 9-16 are RRE ΔAU RNA with trace 5′-^{32}P labeled RNA. Lanes 1 and 9 being RNA only controls. Lanes 2-8 and 10-16 have 10 pM, 100 pM, 1 nM, 10 nM, 100 nM, 1 μM and 10 μM respectively. IC_{50} values of between 10 and 100 nM agree with fluorescence data showing that the methidium conjugates have a high affinity for RRE RNA. There also appears to be a second binding site which begins to appear at concentrations above 100 nM.

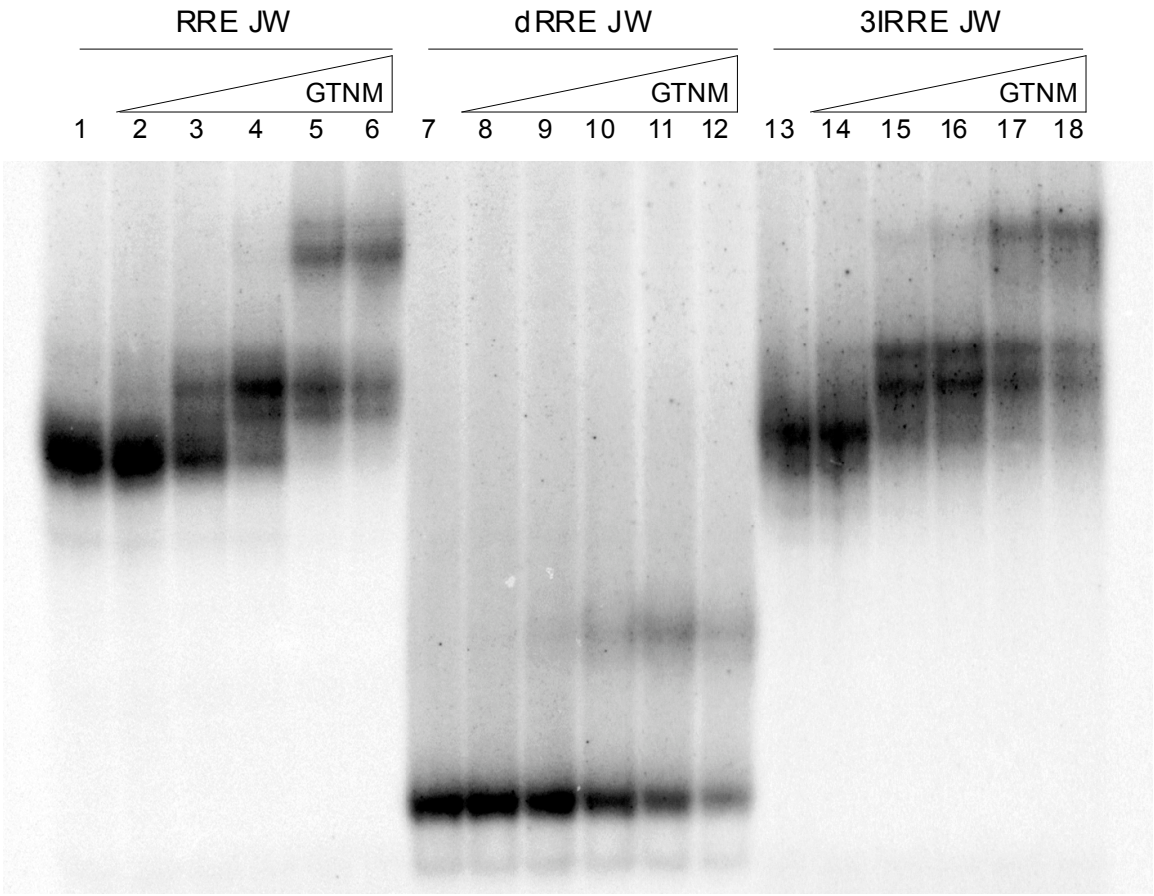


Figure 3.18. Non-denaturing polyacrylamide gel experiments with guanidino-tobramycin-N-methidium being titrated into RRE JW, dRRE JW and 3IRRE JW. Ligand affinity is decreased by over 1,000-fold when the sequence is changed to DNA. However, it is slightly increased in the 3IRRE JW construct. The second binding site seen with both the wild type and 3I RNA molecules appears to be roughly equivalent indicating that only the first site was affected by the change. Lanes 1-6 contain RRE JW RNA, lanes 7-12 contain dRRE JW DNA and lanes 13-18 contain 3I RRE JW RNA, all with trace 5′-^{32}P label. Lanes 1, 7 and 13 represent RNA (and DNA) controls. Lanes 2-6, 8-12 and 14-18 have 1 nM, 10 nM, 100 nM, 1 μM and 10 μM respectively.

While removing the two bulges does not appear to have an effect of the affinity of the tobramycin-methidium conjugate, replacing three adjacent guanine residues with inosines actually increases affinity while changing the sequence from RNA into DNA greatly reduces it (Figure 3.18).

As seen above, guanidino-tobramycin-N-methidium binds to the wild type RRE JW once at between 10 nM and 100 nM and a second time at roughly 1 μM. At 10 nM roughly 70% of the signal is seen in the free RNA band and only 30% is from the RNA-ligand complex. However, at the same concentration with the 3I RRE JW, the free RNA band has almost completely lost all signal intensity and a new band corresponding to the RNA-ligand complex appears.

As with the wild type RRE JW, a third 3I RRE JW band begins to appear at ligand concentrations of 1 μM. While affinities for the first binding site were not equivalent between the two RNA molecules, they do appear to be roughly the same for the second binding site. This is not implausible as a small change in one part of the RNA should not greatly affect affinities at other binding sites.

The greatest change in binding is seen when the RNA sequence of the RRE JW is completely replaced with DNA. This change induces over a 1,000-fold difference in binding affinities for the tobramycin-methidium conjugate. While the ligand has a K_d with the RNA sequence of between 10 and 100 nM, it only begins to bind the DNA construct at 10 μM. These data shows that the guanidinoglycoside-methidium conjugates have a very strong preference for RNA over DNA.

3.5 Discussion

Previous work has shown that conjugation of aminoglycosides to intercalating moieties can yield molecules with vastly increased RNA affinities. However, an increase in affinity appears to be coupled with a decrease in selectivity for one RNA

target over another. It is this selectivity which is one of the major hurdles encountered when attempting to synthesize novel molecules with a specific RNA target in mind.

Variation of linker lengths with both aminoglycoside-acridine and aminoglycoside-methidium conjugates has been shown to have a large effect on the selectivity of aminoglycoside-intercalator conjugates for the RRE RNA.[28] As the length of the linker is increased, there is a marked decrease in selectivity. Conversely, shorter linkers appear to impart increased selectivity.

We have shown that the composition of the intercalator also plays a role in both the affinity and selectivity of aminoglycoside and guanidinoglycoside-intercalator conjugates. While acridine is a small, fairly promiscuous intercalator, ethidium bromide has a larger intercalating area and has been shown to preferentially bind to RNA molecules with bulged nucleotides. Additionally, the orthogonal phenyl ring on ethidium likely imparts selectivity by limiting the potential interactions between the conjugates and RNA. For these reasons, conjugates of methidium and aminoglycosides were designed, synthesized and assayed for their RRE affinity and selectivity.

All aminoglycoside and guanidinoglycoside-methidium conjugates were able to effectively bind to the RRE RNA and compete off a fluorescent tagged Rev peptide. IC_{50} values for this activity ranged from roughly 1-20 nM. These values are up to 1000-fold lower than neomycin ($IC_{50} \approx 1\ \mu M$).

Direct competition experiments, where the ligands were assayed for their Rev-Fl displacement IC_{50} values in the presence of competing nucleotides (poly A-U RNA, tRNA and calf thymus DNA), were not possible. However, direct titrations with these

nucleic acids revealed that the aminoglycoside and guanidinoglycoside-methidium conjugates bind them weaker than the RRE RNA, indicating some level of selectivity.

To determine the internal specificity of the RRE for these conjugates, the features of the RNA which impart selectivity, a representative compound; guanidino-tobramycin-N-methidium was tested for its affinity to a group of modified RRE constructs. K_d values for wild type RNA were found to be between 10 and 100 nM, an increase over both ethidium bromide and tobramycin.

Values with modified RRE constructs yielded some interesting results. Little to no affinity difference is observed for the guanidinoglycoside-methidium conjugate between the wild type RRE and a construct where the two bulged bases are removed. Contrary to binding seen with ethidium bromide, the presence of bulged nucleotides does not appear to greatly affect binding. It is possible that, while the methidium moiety prefers to bind to regions with bulged nucleotides, the two purine-purine base pairs within the high affinity Rev binding site are opening up the RNA enough to easily accommodate the intercalator.

Modification of three adjacent guanine residues within this high affinity site to inosine bases results in an increase in ligand affinity for the higher affinity RRE binding site. It is possible that the presence of modified nucleotides involved in purine-purine base pairing is further perturbing the local helical structure to better fit the guanidinoglycoside portion of the molecule or that the lack of amines at the 2-position of the nucleotide bases somehow allows for tighter binding of the methidium moiety.

Similar to ethidium bromide, the guanidinoglycoside-methidium conjugate appears to have two main binding sites on the RRE RNA. The first, higher affinity site

is located within the high affinity Rev binding site, and is affected by changes therein. Binding at the second site is much weaker, and has not been characterized.

The most striking change in the affinity of guanidino-tobramycin-N-methidium is seen when the RRE JW RNA is converted into the corresponding DNA construct. There is over a 1000-fold difference in affinity between the two constructs with the ligand binding to the RNA at between 10 and 100 nM and showing almost no DNA binding at 10 μM.

In conclusion, we have designed and synthesized a small family of aminoglycoside and guanidinoglycoside conjugates bound to an ethidium bromide based intercalator with varying linker lengths. These compounds are able to bind to the RRE RNA and displace the Rev peptide with IC_{50} values in the low nanomolar. Additionally, they show stronger binding to the RRE RNA than to other RNA (poly-AU and tRNA) and DNA (calf thymus) nucleotides, indicating a degree of selectivity for the RRE RNA. These guanidinoglycoside-methidium conjugates appear to be binding within the high affinity Rev binding site of the RRE, as changes within this region have an effect on binding affinities.

Another factor which affects binding affinity is the nucleic acid composition of the sequence. When the exact same sequence is assayed in both RNA and DNA for ligand binding, guanidinoglycoside-methidium conjugates show a very high specificity for the RNA over the DNA sequence.

We have shown that aminoglycoside-methidium conjugates can be potent Rev inhibitors, binding to the RRE RNA with high affinity and reasonable selectivity. However, small changes in nucleic acid composition and sequence can affect ligand

binding. More structural and biochemical analysis of multiple RNA-ligand complexes must be performed for us to begin to gain insight into the nature of RNA-small molecule selectivity.

3.6 References

(1) Davies, J.; Gorini, L.; Davis, B. D. *Mol Pharmacol* **1965**, 1, 93-106.

(2) Davies, J.; Davis, B. D. *J Biol Chem* **1968**, 243, 3312-3316.

(3) Fourmy, D.; Recht, M. I.; Blanchard, S. C.; Puglisi, J. D. *Science* **1996**, 274, 1367-1371.

(4) Yoshizawa, S.; Fourmy, D.; Puglisi, J. D. *Embo J* **1998**, 17, 6437-6448.

(5) Fourmy, D.; Yoshizawa, S.; Puglisi, J. D. *J Mol Biol* **1998**, 277, 333-345.

(6) Jiang, L.; Patel, D. J. *Nat Struct Biol* **1998**, 5, 769-774.

(7) Lynch, S. R.; Puglisi, J. D. *J Mol Biol* **2001**, 306, 1037-1058.

(8) Moazed, D.; Noller, H. F. *Nature* **1987**, 327, 389-394.

(9) von Ahsen, U.; Davies, J.; Schroeder, R. *Nature* **1991**, 353, 368-370.

(10) von Ahsen, U.; Davies, J.; Schroeder, R. *J Mol Biol* **1992**, 226, 935-941.

(11) Stage, T. K.; Hertel, K. J.; Uhlenbeck, O. C. *RNA* **1995**, 1, 95-101.

(12) Wang, H; Tor, Y *J. Am. Chem. Soc.* **1997**, 119, 8734.

(13) Rogers, J.; Chang, A. H.; von Ahsen, U.; Schroeder, R.; Davies, J. *J Mol Biol* **1996**, 259, 916-925.

(14) Chia, J. S.; Wu, H. L.; Wang, H. W.; Chen, D. S.; Chen, P. J. *J Biomed Sci* **1997**, 4, 208-216.

(15) Mikkelsen, N. E.; Brannvall, M.; Virtanen, A.; Kirsebom, L. A. *Proc Natl Acad Sci U S A* **1999**, 96, 6155-6160.

(16) Kirk, S. R.; Tor, Y. *Bioorg Med Chem* **1999**, 7, 1979-1991.

(17) Yoshizawa, S.; Fourmy, D.; Eason, R. G.; Puglisi, J. D. *Biochemistry* **2002**, 41, 6263-6270.

(18) Mei, H. Y.; Mack, D. P.; Galan, A. A.; Halim, N. S.; Heldsinger, A.; Loo, J. A.; Moreland, D. W.; Sannes-Lowery, K. A.; Sharmeen, L.; Truong, H. N.; Czarnik, A. W. *Bioorg Med Chem* **1997**, 5, 1173-1184.

(19) Zapp, M. L.; Stern, S.; Green, M. R. *Cell* **1993**, 74, 969-978.

(20) Zapp, M. L.; Young, D. W.; Kumar, A.; Singh, R.; Boykin, D. W.; Wilson, W. D.; Green, M. R. *Bioorg Med Chem* **1997**, 5, 1149-1155.

(21) Tok, J. B.; Cho, J.; Rando, R. R. *Biochemistry* **1999**, 38, 199-206.

(22) Park, W K C; Auer, M; Jaksche, H; Wong, C-H *J. Am. Chem. Soc*. **1996**, 118, 10150.

(23) Ratmeyer, L.; Vinayak, R.; Zhong, Y. Y.; Zon, G.; Wilson, W. D. *Biochemistry* **1994**, 33, 5298-5304.

(24) Wilson, W. B.; Ratmeyer, L.; Cegla, M. T.; Spychala, J.; Boykin, D. W.; Demeunynyck, M.; Lhomme, J.; Krishnan, G.; Kennedy, D.; Vinayak, R.; Zon, G. *New. J. Chem*. **1994**, 18, 419.

(25) Gale, E. F. (Ed.) *The Molecular Basis of Antibiotic Action* **1981**, John Wiley & Sons, London.

(26) Kirk, S. R.; Luedtke, N. W.; Tor, Y *J. Am. Chem. Soc*. **2000**, 122, 980-981.

(27) Luedtke, N. W.; Tor, Y. *Angew Chem Int Ed Engl* **2000**, 39, 1788-1790.

(28) Luedtke, N. W.; Liu, Q.; Tor, Y. *Biochemistry* **2003**, 42, 11391-11403.

(29) Watkins, T. I.; Woolfe, G. *Nature* **1952**, 169, 506-507.

(30) LePecq, J. B.; Paoletti, C. *J Mol Biol* **1967**, 27, 87-106.

(31) Chu, W. C.; Liu, J. C.; Horowitz, J. *Nucleic Acids Res* **1997**, 25, 3944-3949.

(32) Tsai, C. C.; Jain, S. C.; Sobell, H. M. *J Mol Biol* **1977**, 114, 301-315.

(33) Jain, S. C.; Tsai, C. C.; Sobell, H. M. *J Mol Biol* **1977**, 114, 317-331.

(34) Baker, T. J.; Luedtke, N. W.; Tor, Y.; Goodman, M. *J Org Chem* **2000**, 65, 9054-9058.

(35) Luedtke, N. W.; Liu, Q.; Tor, Y. *Chemistry* **2005**, 11, 495-508.

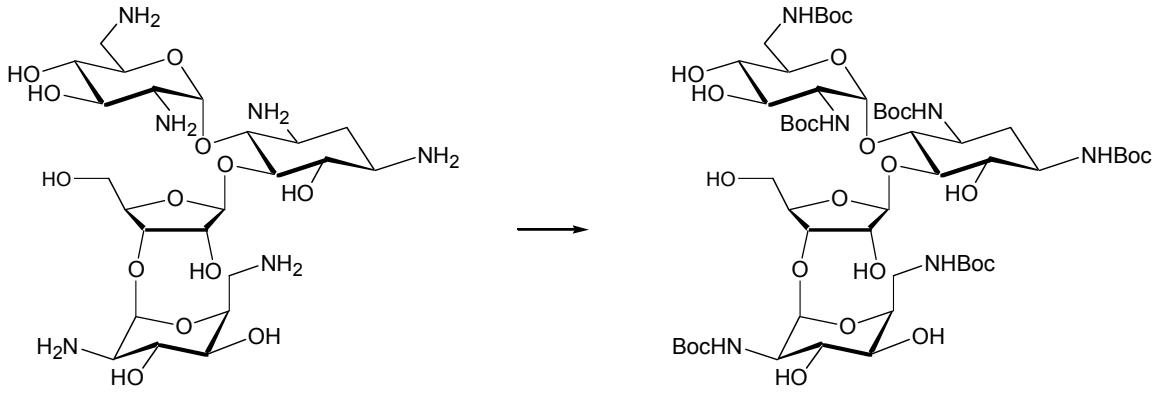
(36) Luedtke, N. W.; Baker, T. J.; Tor, Y; Goodman, M. *J. Am. Chem. Soc*. **2000**, 122, 12035-12036.

(37) Kjems, J.; Calnan, B. J.; Frankel, A. D.; Sharp, P. A. *Embo J* **1992**, 11, 1119-1129.

(38) Battiste, J. L.; Mao, H.; Rao, N. S.; Tan, R.; Muhandiram, D. R.; Kay, L. E.; Frankel, A. D.; Williamson, J. R. *Science* **1996**, 273, 1547-1551.

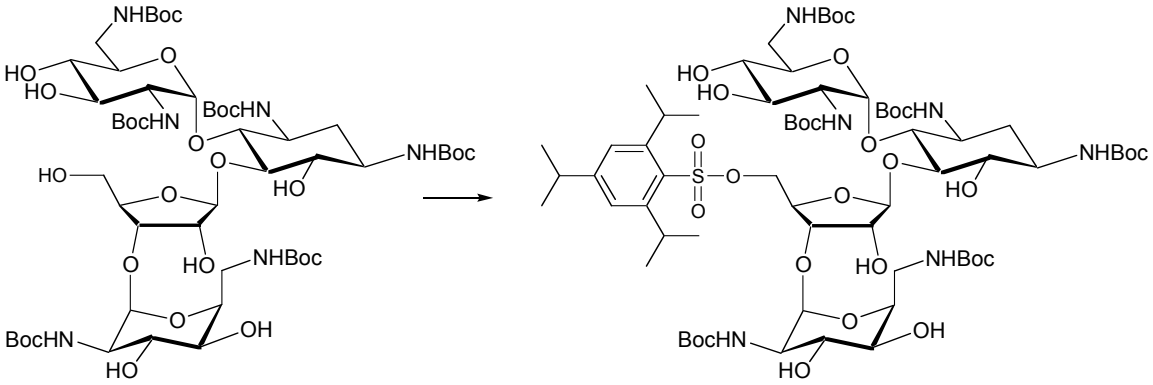
Chapter 4

4.1 Synthesis of Aminoglycoside-Methidium conjugates



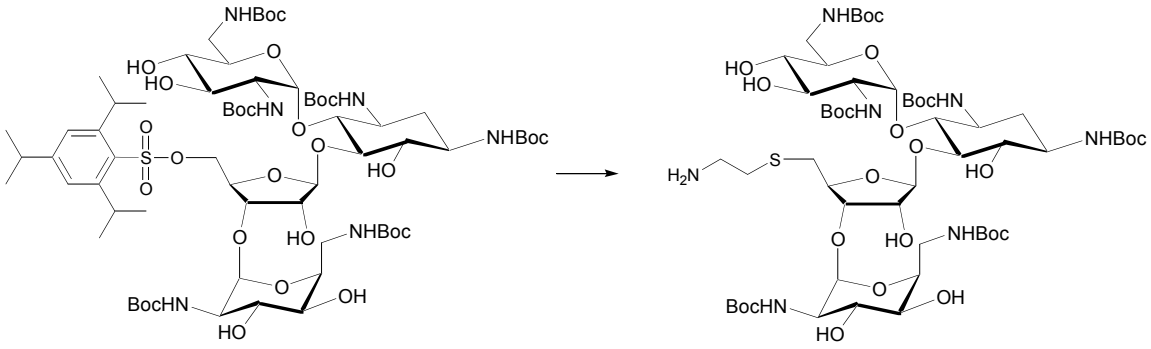
Compound **2**. A solution of neomycin sulfate (a mixture of neomycin B and neomycin C sulfate) (0.51 g, 0.83 mmol, 1 equiv.) in a mixture of DMF (10 mL), water (2 mL) and triethylamine (1 mL) was treated with di-*tert*-butyldicarbonate (2.27 g, 10.4 mmol, 12.5 equiv.). The solution was stirred for three hours until all starting material was consumed as monitored by TLC (7.5 % methanol in dichloromethane). All volatiles were removed in vacuo and the residue was partitioned between water (150 mL) and ethyl acetate (200 mL). The aqueous layer was separated and extracted with ethyl acetate (2 × 100 mL). The combined organic layer was washed with saturated sodium chloride, dried over Na_2SO_4 and concentrated in vacuo. Flash chromatography (1-4% methanol in dichloromethane) afforded the desired product as a while solid (0.90 g, 89%): R_f 0.4, 10% methanol in dichloromethane: ^{1}H NMR (400 MHz, methanol-d_4) δ 6.75 (d, 2H), 5.30 (s, 1H), 5.16 (s, 1H), 4.93 (s, 10H) 4.89 (s, 1H), 4.18 (s, 2H), 3.96 (s,

1H), 3.92-3.80 (m, 3H), 3.78 (s, 1H), 3.73-3.62 (m, 4H), 3.53-3.47 (m, 6H), 3.36-3.33 (m, 16H) 3.31-3.20 (m, 5H), 1.94 (q, 1H), 1.56 (m, 1H), 1.46-1.36 (s, 54 H).



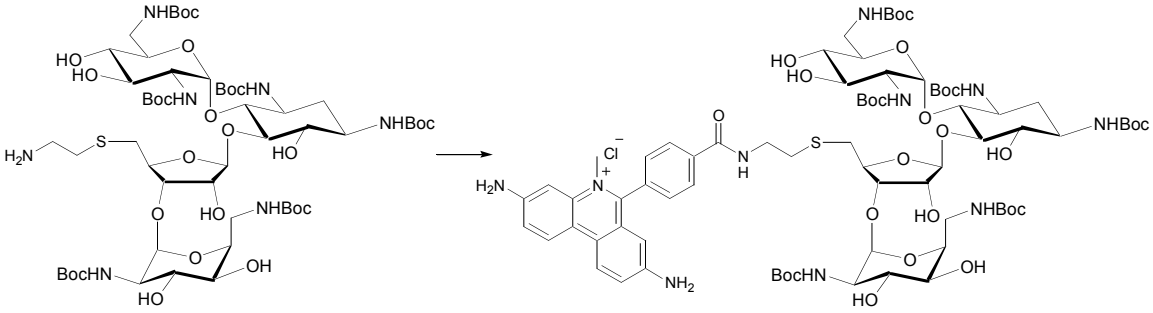
Compound **3**. A solution of (**2**) (0.86 g, 0.71 mmol, 1 equiv.) in pyridine (18 mL) was treated with 2,4,6-triisopropylbenzenesulfonyl chloride (11 g, 36.4 mmol, 51.5 equiv.). The reaction was stirred under argon for 36 hours until all starting material was consumed as monitored by TLC (R_f 0.5, 10% methanol in dichloromethane), followed by removal of pyridine in vacuo. The reaction residue was dissolved in 300 mL of ethyl acetate and washed with pH 3 water (5 × 100 mL) to remove all traces of pyridine. The organic layer was then washed with saturated sodium chloride, dried over Na_2SO_4, and removed in vacuo. Flash chromatography (0-4% methanol in dichloromethane) afforded the desired product as a white solid (0.532 g, 51%): R_f 0.5, 10 % methanol in dichloromethane: ^{1}H NMR (400 MHz, methanol-d_4) δ 8.54 (s, 1H), 7.86 (m, 1H), 7.45 (m, 1H), 7.28 (s, 2H), 6.66-6.36 (m, 4H), 6.10 (d, 1H), 5.45 (s, 1H), 5.18 (s, 1H), 4.86 (s, 44H), 4.60 (br, 1H), 4.36 (m, 1H), 4.26 (m, 2H), 4.15 (m, 4H),

3.88 (s, 1H), 3.72 (m, 2H), 3.58 (m, 1H), 3.5 (m, 4H), 3.42-3.23 (m, 35H), 2.95 (m, 1H), 1.95 (d, 1H), 1.61-1.07 (m, 99H).



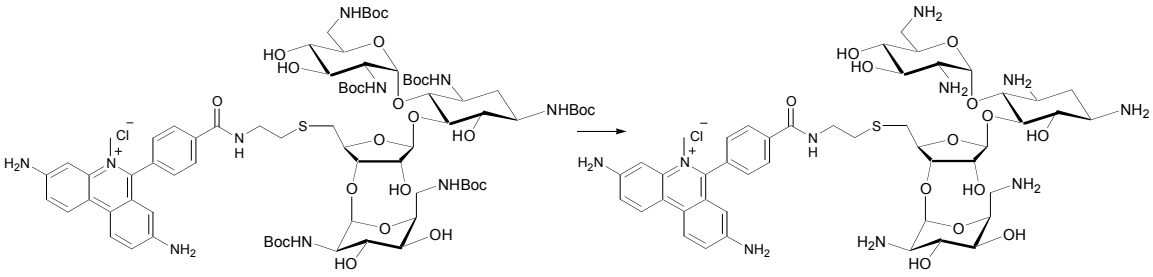
Compound **4**. Under a balloon filled with Argon, sodium metal (1 g) was dissolved in ethanol (26 mL). 2-aminoethanethiol hydrochloride (860 mg, 7.6 mmol) was added and the solution was stirred for 15 minutes. A white precipitate formed. Separately a solution of (**3**) (190 mg, 0.128 mmol, 1 equiv.) in ethanol (4 mL) was prepared and this solution was then added dropwise to the reaction flask. The reaction was stirred under argon for 16 hours and monitored by TLC (R_f 0.4, 10% methanol in dichloromethane). Upon completion chloroform (200 mL) was added and the reaction was washed with pH 5 water (4 × 100 mL). The aqueous layer was washed with chloroform (2 × 100 mL) and all organic layers were combined, washed with saturated sodium chloride (2 × 150 mL), dried over Na_2SO_4, and evaporated in vacuo. Flash chromatography (3% methanol in dichloromethane) afforded the desired product as a white solid (0.132 g, 81%): R_f 0.35, 10% methanol in dichloromethane: ^{1}H NMR (400 MHz, methanol-d_4) δ 5.41 (s, 1H), 5.18 (s, 1H), 4.95 (s, 1H), 4.27 (m, 1H), 4.03 (m, 2H), 3.90 (m, 2H), 3.74

(m, 2H), 3.59-3.44 (m, 6H), 3.36-3.22 (m, 8H), 3.00 (m, 2H), 2.89 (m, 2H), 2.80 (m, 2H), 1.99 (m, 1H), 1.60 (s, 1H), 1.46-1.33 (m, 58H). Exact mass calculated for $C_{55}H_{99}N_7O_{24}S$: 1273.65, found 1274.4 $[M + H]^+$, 1296.4 [M + Na]



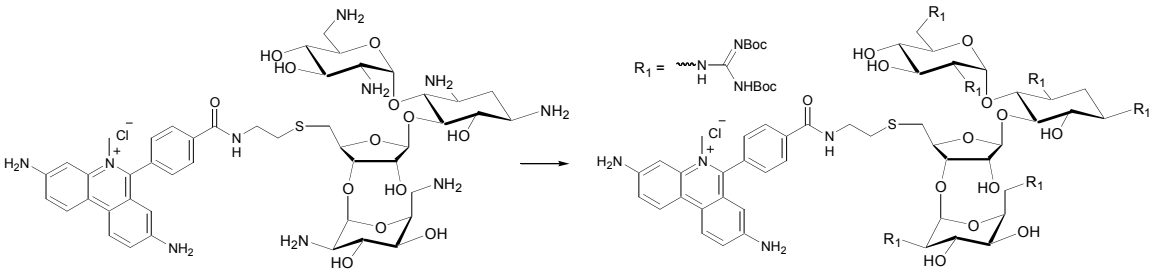
Compound **5**. To a solution of (**4**) (102 mg, 0.08 mmol, 1.2 equiv.) in dimethylformamide (1 mL), 1H-benzotriazolium (TBTU) (26 mg, 0.08 mmol, 1.2 equiv.) and 1-hydroxybenzotriazole (HOBT) (13 mg, 0.08 mmol, 1.2 equiv) were added. 6-(4-carboxyphenyl)-3,8-diamino-5-methylphenanthridinium chloride (**22**) (25 mg, 0.066 mmol, 1 equiv.) was then added. The reaction was covered with foil to protect it from light and stirred under argon. After 3 hours TLC (10% methanol in dichloromethane) showed all starting material was gone and a new, red spot appeared R_f 0.6. All volatiles were removed in vacuo and the residue was dissolved in chloroform (50 mL). The solution was washed with saturated ammonium chloride (50 mL), saturated sodium bicarbonate (50 mL) and saturated sodium chloride (50 mL). All aqueous layers were combined and washed with chloroform (50 mL). All organic layers were combined and evaporated in vacuo to yield 153 mg of purple/red solid. Flash chromatography (5-13% methanol in dichloromethane) on the solid afforded the

desired product as a red solid (86.5 mg, 82%): R_f 0.6 10% methanol in dichloromethane: ^{1}H NMR (400 MHz, acetonitrile-d_3) δ 8.58-8.50 (q, 2H), 8.31 (m, 1H), 7.64 (d, 1H), 7.56 (d, 1H), 7.36 (s, 1H), 6.58 (s, 1H), 6.03 (m, 1H), 5.28 (s, 1H), 4.95 (d, 1H), 4.79 (m, 1H), 4.36-4.17 (m, 2H), 4.06 (s, 1H), 3.95 (s, 1H), 3.82-3.62 (m, 2H), 3.55-3.40 (m, 2H), 3.35 (m, 1H), 3.30 (m, 1H), 3.18-3.02 (m, 1H), 2.95 (m, 2H), 2.14 (m, 3H), 1.94 (s, 3H), 1.42 (m, 6H). Exact mass calculated for $C_{76}H_{115}N_{10}O_{25}S^+$: 1599.78, found 1599.5 $[M]^+$, 811.3 $[M + Na]^{2+}/2$

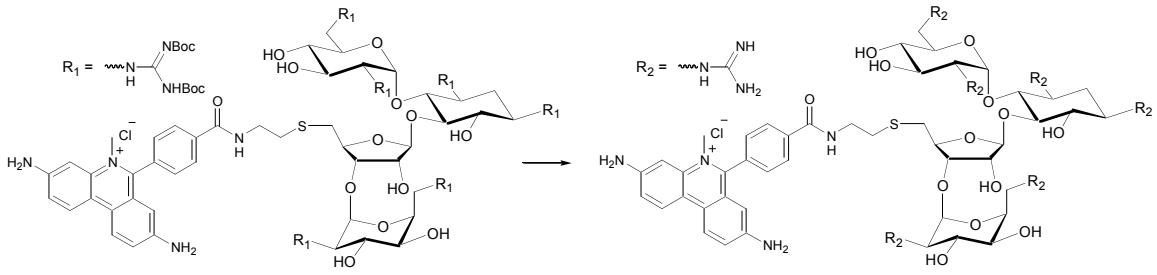


Compound **6**. Methanol (5 μL) was added to a solution of (**5**) (40 mg, 0.026 mmol) in 50/50 trifluoroacetic acid/dichloromethane (5 mL). The reaction was covered with foil and vigorously stirred for 1 hour. Toluene (5 mL) was then added to the reaction and all volatiles were removed in vacuo. The residue was dissolved in water (1 mL) and lyophilized to dryness. Analytical HPLC (C-18, 0-30% acetonitrile (0.1% TFA) in water (0.1 % TFA) over 30 minutes) showed one major peak at 16.23 minutes. HPLC purification (C-8, 5-13% acetonitrile (0.1% TFA) in water over 20 minutes then 13% isocratic acetonitrile for 10 minutes) gave one major peak at 20.53 minutes: ^{1}H NMR (400 MHz, methanol-d_4) δ 8.59 (dd, 1H), 8.23 (d, 1H), 7.71 (d, 1H), 7.58 (dd, 1H), 7.38

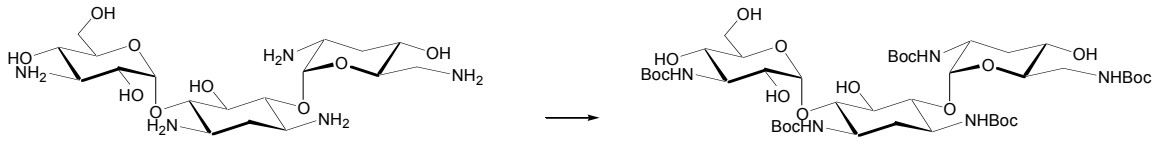
(d, 1H), 6.50 (d, 1H), 6.10 (d, 1H), 5.45 (d, 1H), 5.36 (d, 1H), 4.88 (s, 31H), 4.98 (m, 1H), 4.50 (t, 1H), 4.40 (m, 1H), 4.37 (m, 1H), 4.17 (m, 2H), 4.31 (s, 2H), 4.06 (d, 1H), 3.89 (t, 1H), 3.72 (m, 2H), 3.60 (t, 1H), 3.5-3.22 (m, 9H), 3.17 (m, 2H), 3.10 (s, 3H), 3.03-2.82 (m, 4H), 2.47 (m, 1H), 2.08 (q, 1H), 1.32-1.19 (m, 4H), 0.49 (m, 1H). Exact mass calculated for $C_{46}H_{67}N_{10}O_{13}S^{+}$: 99.46, found 999.5 $[M]^{+}$, 500.1 $[M + H]^{2+}/2$.



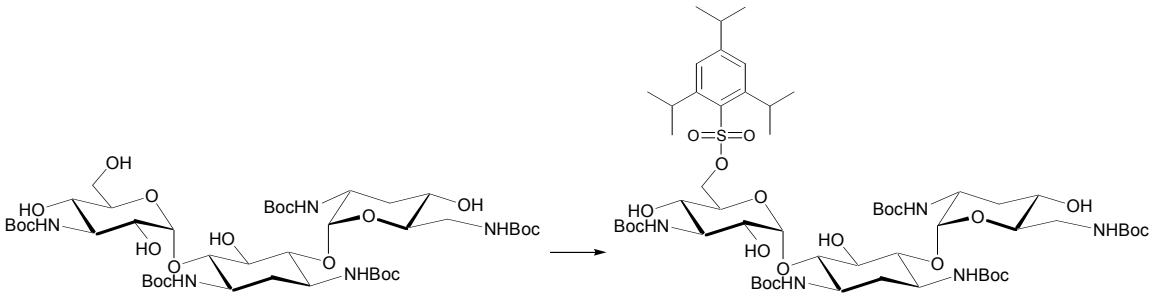
Compound **7**. Chloroform (0.2 mL), triethylamine (35 μL) and (**24**) (80 mg, 0.204 mmol, 13 equiv) were added to a solution of (**6**) (7 × TFA) (28.8 mg, 0.0158 mmol) in methanol (0.6 mL). The reaction was stirred under argon and covered in foil. Monitoring by TLC (4% methanol in dichloromethane) showed multiple new products forming, all with higher R_f than the starting material. Gradually these products all converged to the fastest moving spot. After 48 hours all solvent was removed in vacuo. Flash chromatography (2% methanol in dichloromethane) afforded the desired product as a red solid which was taken directly to boc-deprotection. Exact mass calculated for $C_{112}H_{175}N_{22}O_{37}S^{+}$: 2452.22, found 1226.7 $[M + H]^{2+}/2$, 1237.7 $[M + Na]^{2+}/2$.



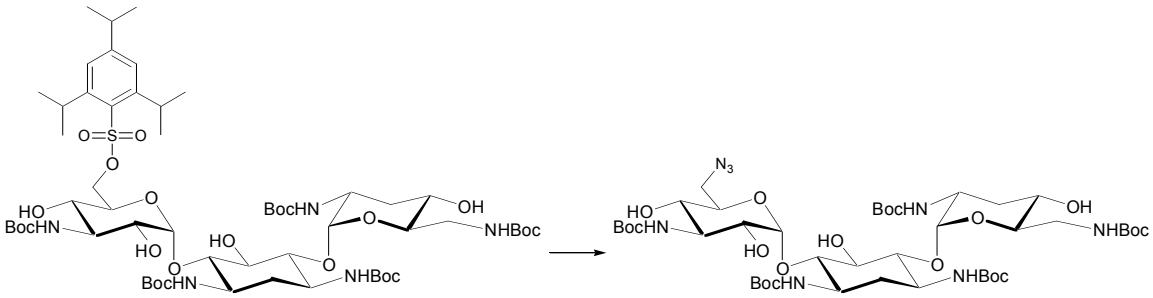
Compound **8**. Triisopropylsilane (0.6 mL) was added to a solution of crude (**7**) (from previous reaction) in 50/50 trifluoroacetic acid/chloroform (10 mL). The reaction was covered with foil and stirred for 4 hours. Water (50 mL) was added to the reaction which was then washed with chloroform (2 × 50 mL) and ethyl ether (2 × 50 mL). The aqueous layer was evaporated in vacuo and the remaining residue was resuspended in 0.1 % trifluoroacetic acid in water (1 mL) and lyophilized overnight. Analytical reverse phase HPLC (C-18, 5-30% acetonitrile (0.1% TFA) in water (0.1% TFA) over 30 minutes showed one peak at 21.38 minutes). HPLC purification (C-8, 5-13% acetonitrile (0.1% TFA) in water over 20 minutes then 13% isocratic acetonitrile for 10 minutes) gave one major peak at 24.42 minutes: ^{1}H NMR (400 MHz, methanol-d_4) δ 8.62 (q, 1H), 8.22 (d, 1H), 7.73 (d, 1H), 7.59 (dd, 1H), 7.19 (t, 1H), 6.48 (s, 1H), 6.04 (d, 1H), 5.09 (d, 1H), 5.04 (s, 1H), 4.92 (s, 35H), 4.41 (m, 1H), 4.16 (m, 2H), 4.05 (m, 2H), 3.98 (m, 1H), 3.82-3.41 (m, 9H), 3.33 (m, 17H), 3.12 (m, 5H), 2.99 (m, 2H), 2.92 (m, 2H), 2.63 (s, 1H), 2.07 (m, 1H), 1.76 (m, 1H), 1.29 (m, 3H). Exact mass calculated for $C_{52}H_{79}N_{22}O_{13}S^{+}$: 1251.59, found 1251.4 $[M]^{+}$, 626.3 $[M + H]^{2+}/2$, 417.9 $[M + 2H]^{3+}/3$, 313.7 $[M + 3H]^{4+}/4$.



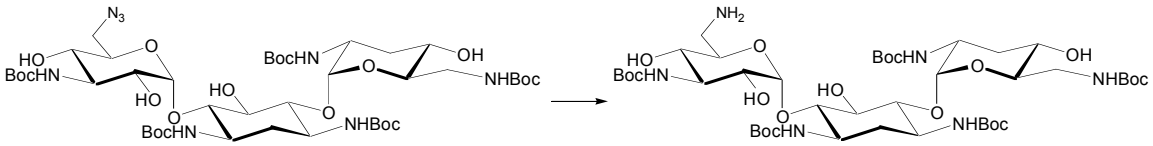
Compound **10**. A solution of di-*tert*-butyldicarbonate (5.60 g, 25.68 mmol, 6 equiv) was dissolved in DMSO (30 mL) and added dropwise to a solution of tobramycin (**9**) (2 g, 4.28 mmol, 1 equiv) in water (6 mL). The reaction was stirred for 18 hours until all starting material was consumed as monitored by TLC (10% methanol in dichloromethane). Concentrated ammonium hydroxide was added until a white precipitate formed. The precipitate was filtered and washed with water, yielding a white solid (3.76 g, 91%): R_f 0.6 in 10% methanol in chloroform, ^{1}H NMR (400 MHz, methanol-d_4) δ 5.10 (br, 2H), 4.87 (s, 9H), 3.93 (m, 1H), 3.79 (d, 1H), 3.70 (m, 2H), 3.6 (m, 3H), 3.51-3.35 (m, 11H), 3.31 (m, 2H), 2.62 (m, 12 H), 2.11 (d, 1H), 1.99 (m, 1H), 1.64 (q, 1H), 1.49-1.33 (m, 47H). Exact mass calculated for $C_{43}H_{77}N_5O_{18}$: 965.54, found 966.4 $[M + H]^+$.



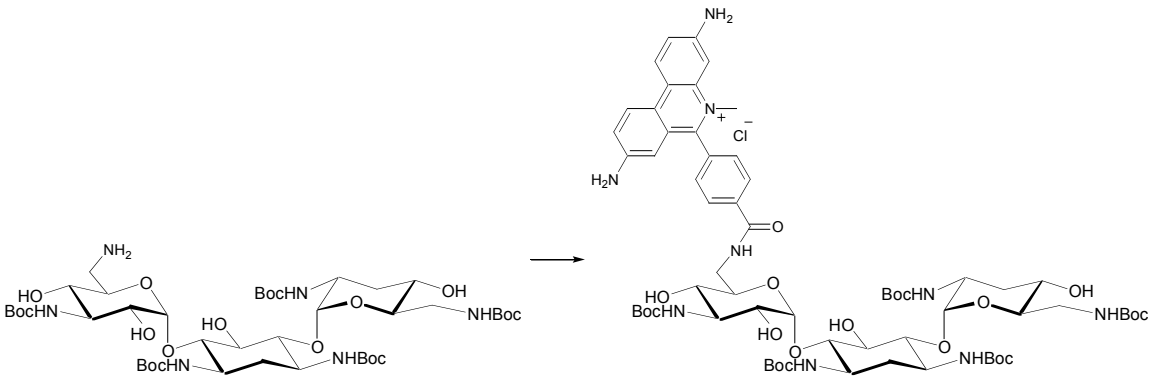
Compound **11**. A solution of (**10**) (1.50 g, 1.55 mmol, 1 equiv) and 2,4,6-triisopropylbenzenesulfonyl chloride (17 g, 56 mmol, 36 equiv) in anhydrous pyridine (30 mL) was stirred under argon for 48 hours. After all starting material was consumed, as monitored by TLC (10% methanol in dichloromethane), pyridine was removed in vacuo. The resulting dry reside was suspended in ethyl acetate (400 mL) and washed with pH 3 water (7 × 200 mL) and saturated sodium chloride (1 × 200 mL). The organic layer was evaporated in vacuo and the resulting solid was purified using flash chromatography (0-4 % methanol in dichloromethane). The product was purified as a white solid (0.97 g, 0.79 mmol , 51%): R_f 0.7 in 7% methanol in chloroform, ^{1}H NMR (400 MHz, DMSO-d_6) 7.23 (s, 3H), 6.90 (s, 1H), 6.58 (m, 3H), 6.42 (m, 1H), 5.73 (s, 1H), 4.96 (m, 3H), 4.92 (s, 2H), 4.27 (d, 1H), 4.07 (d, 1H), 4.01 (m, 4H), 3.58-3.33 (m, 21H), 2.84 (q, 1H), 1.85 (m, 1H), 1.75 (m, 1H), 1.36 (d, 41H), 1.28 (d, 28H), 1.19 (t, 29H). Exact mass calculated for $C_{58}H_{99}N_5O_{21}S$: 1231.68, found 1254.5 $[M + Na]^+$.



Compound **12**. A solution of (**11**) (0.33 g, 0.255 mmol, 1 equiv) in 80/20 DMF/water (10 mL) was heated on an oil bath to 75°C. Sodium azide (0.40 g, 6.16 mmol, 26.5 equiv) was added to the the solution and the reaction was stirred under argon for 24 hours. When all starting material was consumed, as monitored by TLC (70% ethyl acetate 30% hexanes), the reaction was removed from the bath and allowed to return to room temperature. Water (250 mL) was added and the reaction was washed with ethyl acetate (3 × 300 mL). The combined organic layer was washed with saturated sodium chloride (1 × 200 mL) and dried over Na_2SO_4. Flash chromatography (50-70% ethyl acetate in hexanes) yielded the product as a white solid (245 mg, 0.246 mmol, 96%): R_f 0.5 in 50/50 Ethyl Acetate/Hexanes, ^{1}H NMR (400 MHz, methanol-d_4) δ 5.05 (m, 3H), 4.86 (s, 13H), 4.17 (m, 2H), 4.12 (q, 1H), 3.75-3.23 (m, 29H), 2.99 (s, 1H), 2.83 (s, 1H), 2.11 (d, 2H), 2.03 (m, 2H), 1.82 (m, 2H), 1.43 (m, 59H), 1.23 (m, 2H), 0.93 (m, 1H). Exact mass calculated for $C_{43}H_{76}N_8O_{18}$: 992.54, found 993.5 $[M + H]^+$.

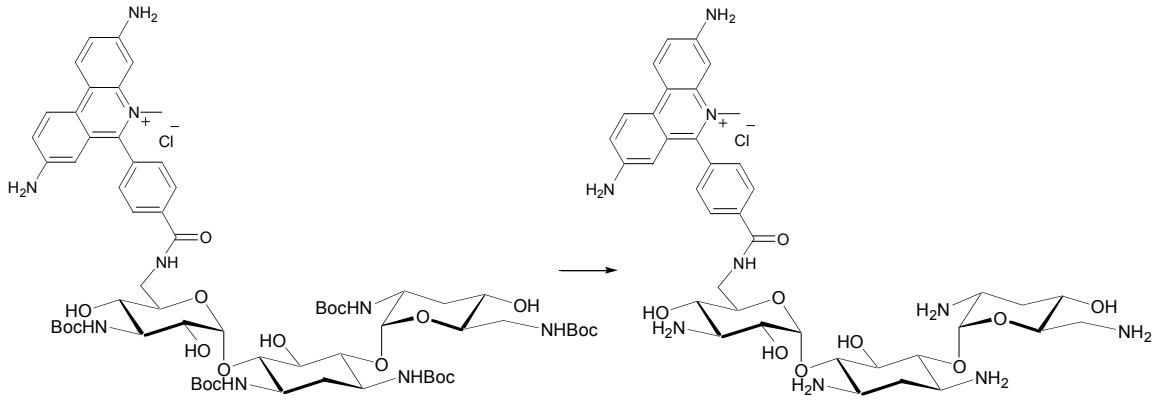


Compound **13**. A solution of (**12**) (0.230 g, 0.231 mmol) and palladium on carbon (0.115 g) in methanol (10 mL) was stirred under 1 atmosphere of hydrogen. After 18 hours palladium was filtered off and the reaction was concentrated in vacuo. Flash chromatography (4-12% methanol in dichloromethane) afforded the desired product as an off white solid (110 mg, 0.113 mmol, 49%): R_f 0.5 in 7% methanol in dichloromethane, ^{1}H NMR (400 MHz, methanol-d$_4$) δ 5.16 (m, 2H), 4.89 (s, 17H), 4.13 (m, 1H), 3.69 (t, 1H), 3.59 (m, 5H), 3.51-3.38 (m, 8H), 3.31 (m, 4H), 3.19 (t, 2H), 2.02 (m, 1H), 2.03 (br, 2H), 1.68 (q, 2H), 1.43 (s, 57H), 1.3 (d, 2H). Exact mass calculated for $C_{43}H_{78}N_6O_{18}$: 966.54, found 968 $[M + H]^+$, 990 $[M + Na^+]^+$, 1002 $[M + Cl]^-$.



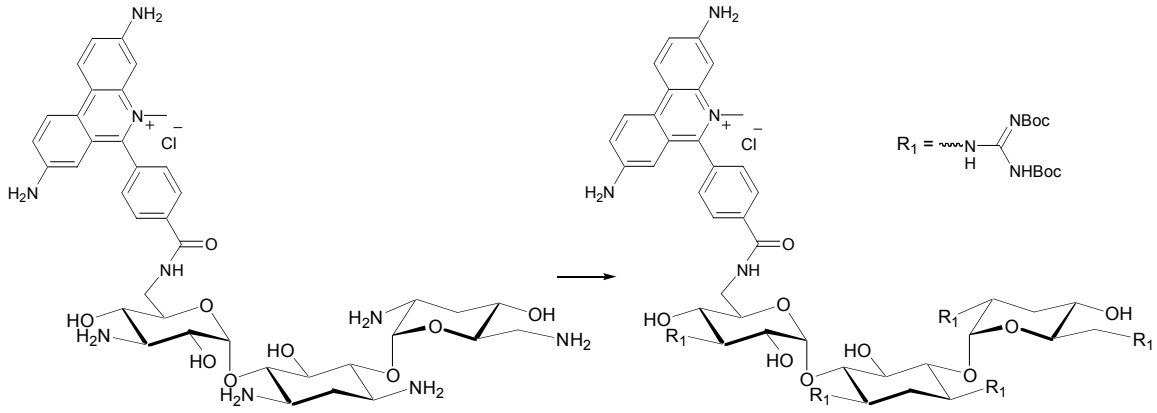
Compound **14**. To a solution of (**13**) (43 mg, 0.044 mmol, 1.35 equiv.) in dimethylformamide (3 mL), 1H-benzotriazolium (TBTU) (14 mg, 0.044 mmol, 1.35

equiv.) and 1-hydroxybenzotriazole (HOBT) (6.7 mg, 0.044 mmol, 1.35 equiv) were added. 6-(4-carboxyphenyl)-3,8-diamino-5-methylphenanthridinium chloride (**22**) (12.5 mg, 0.033 mmol, 1 equiv.) and 2,4,6-collidine (0.13 mL) were subsequently added and the reaction mixture was covered with foil and stirred. The reaction was followed by TLC (10% methanol in dichloromethane), and when all starting material (**22**) was consumed, chloroform (75 mL) was added and the solution was washed with saturated ammonium chloride (50 mL), saturated sodium bicarbonate (50 mL) and saturated sodium chloride (50 mL). The aqueous layers were combined and back extracted with chloroform (100 mL). The organic layers were combined dried over Na_2SO_4, and evaporated in vacuo. Flash chromatography (5-12% methanol in dichloromethane) yielded the desired product as a purple/red solid (38.8 mg, 0.030 mmol, 91%): R_f 0.5 in 10% methanol in chloroform, ^{1}H NMR (300 MHz, methanol-d_4) δ 8.45 (t, 2H), 8.16 (t, 2H), 7.63 (d, 1H), 7.6 (d, 1H), 7.49 (dd, 1H), 7.30 (s, 1H), 7.26 (s, 1H), 6.43 (s, 1H), 5.10 (br, 1H), 4.99 (d, 2H), 4.88 (br, 2H), 4.79 (s, 36H), 4.51 (s, 1H), 4.23 (t, 1H), 4.09 (m, 1H), 4.02 (s, 3H), 3.72-3.30 (m, 23H), 3.21 (s, 1H), 3.19 (m, 6H), 3.10-2.85 (m, 5H), 1.83 (br, 5H), 1.35 (m, 68H), 1.24-1.13 (m, 36H), 0.79 (m, 3H). Exact mass calculated for $C_{64}H_{94}N_9O_{19}$: 1292.67, found 1293 $[M]^+$.

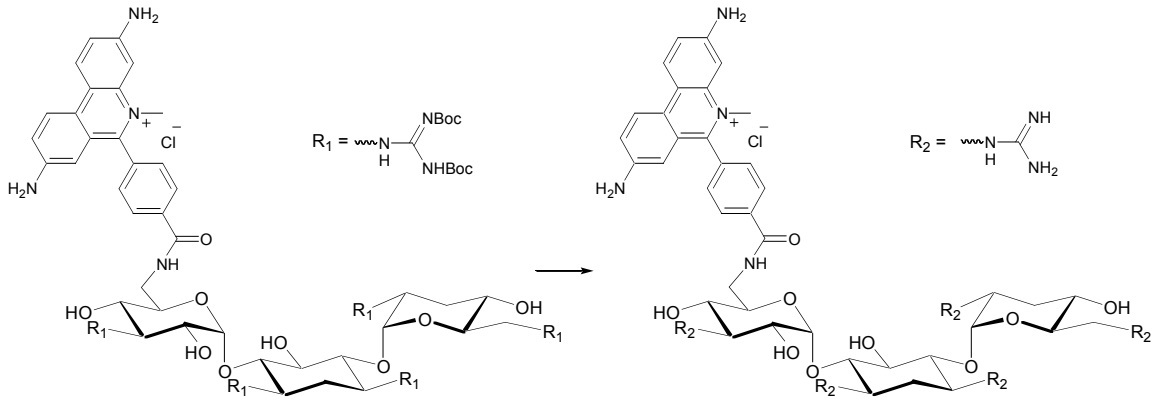


Compound **15**. Methanol (5 μL) was added to a solution of (**14**) (28.8 mg, 0.030 mmol) in trifluoroacetic acid (5 mL) and dichloromethane (5 mL). The reaction was covered with foil and stirred vigorously for 2 hours. Water (25 mL) was added and the reaction was washed with chloroform (1 × 25 mL) and ethyl ether (1 × 25 mL). The aqueous layer was evaporated in vacuo and the resulting solid was dissolved in 0.1% TFA in water (1 mL) and lyophilized to dryness (20 mg, 0.025 mmol, 83%). Analytical reverse phase HPLC (C-8, 13% isocratic acetonitrile (0.1% TFA) in water (0.1% TFA) over 35 minutes) showed one major peak at 8.73 minutes with a large shoulder/second peak immediately following. No conditions were found which gave good separation of the two peaks but when purified separately and re-injected, each peak behaved exactly the same as the crude sample. HPLC purification of the combined peaks (C-8, 5-13% acetonitrile (0.1% TFA) in water over 20 minutes then 13% isocratic acetonitrile for 10 minutes) gave one major peak at 18.64 minutes. ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.62 (dd, 2H), 8.19 (d, 2H), 7.74 (t, 2H), 7.52 (d, 1H), 7.33 (d, 1H), 7.26 (s, 1H), 6.30 (s, 1H), 5.81 (d, 1H), 4.97 (d, 1H), 3.93 (s, 2H), 3.79 (m, 1H), 3.77-3.62 (m, 4H), 3.60-

3.42 (m, 35H), 3.40-3.25 (m, 6H), 3.19 (m, 3H), 2.91 (m, 1H), 2.83 (m, 1H), 2.50 (2H), 2.41 (m, 1H), 2.09 (s, 3H), 2.0 (m, 2H), 1.78 (m, 2H). Exact mass calculated for $C_{39}H_{54}N_9O_9$: 792.40, found 792 $[M]^+$.

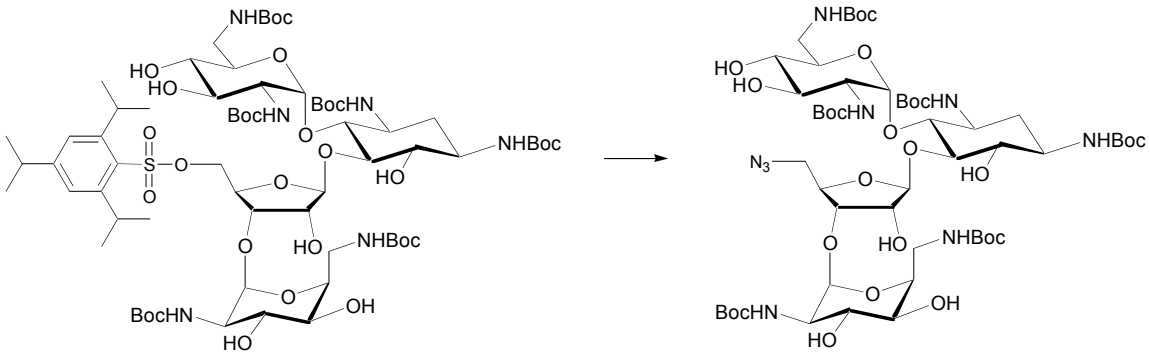


Compound **16**. Triethylamine (66 μL) was added to a solution of (**15**) (10 mg, 0.013 mmol, 1 equiv) in methanol (1 mL) and chloroform (0.3 mL). This solution was treated with (**24**) (74 mg, 0.189 mmol, 15 equiv) and stirred for 48 hours. The volatiles were then removed in vacuo and flash chromatography (4-12% methanol in dichloromethane) yielded the desired product as a red solid (6 mg, 3 μmol, 25%) which was taken directly to the next step.

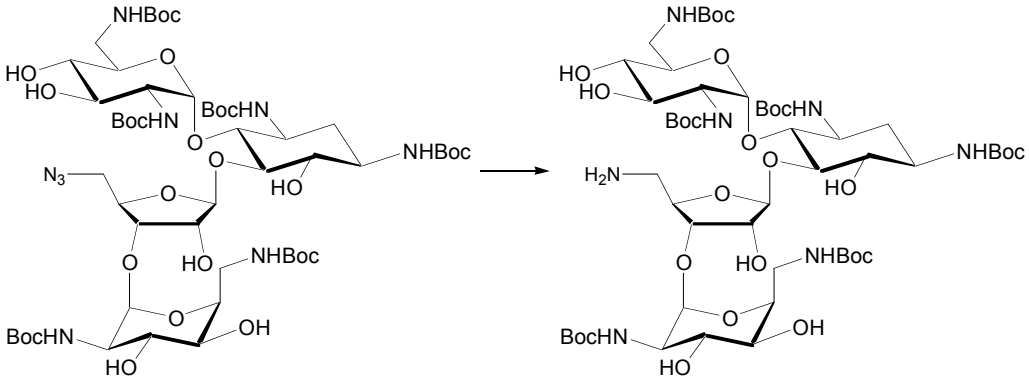


Compound **17**. Triisopropylsilane (40 μL) was added to a solution of crude (**16**) in trifluoroacetic acid (4 mL) and dichloromethane (4 mL). The reaction was covered with foil and stirred for 5 hours. Excess water (20 mL) was then added to the reaction which was then washed with chloroform (2 × 10 mL) and ethyl ether (2 × 10 mL). The aqueous layer was evaporated in vacuo and the resulting solid was suspended in 0.1% TFA in water (1 mL) and lyophilized to dryness. As with (**15**), analytical reverse phase HPLC (C-8, 12% isocratic acetonitrile (0.1% TFA) in water (0.1% TFA) over 15 minutes) showed a major peak (at 4 minutes) and a secondary "shoulder" peak (immediately following the main peak) which, when purified separately and re-injected, each split into the original peaks with identical ratios. HPLC purification of the combined peaks (C-8, 12% isocratic acetonitrile (0.1% TFA) in water over 25 minutes gave one major peak at 12.938 minutes. ^{1}H NMR (300 MHz, D_2O) δ 8.51 (d, 2H), 7.99 (d, 2H), 7.54 (m, 3H), 7.34 (m, 2H), 6.60 (s, 1H), 5.34 (d, 1H), 5.01 (s, 1H), 4.60 (m, 42H), 4.18 (m, 2H), 4.02 (s, 4H), 3.76-3.45 (m, 12H), 3.37 (m, 2H), 3.22 (m, 2), 3.09

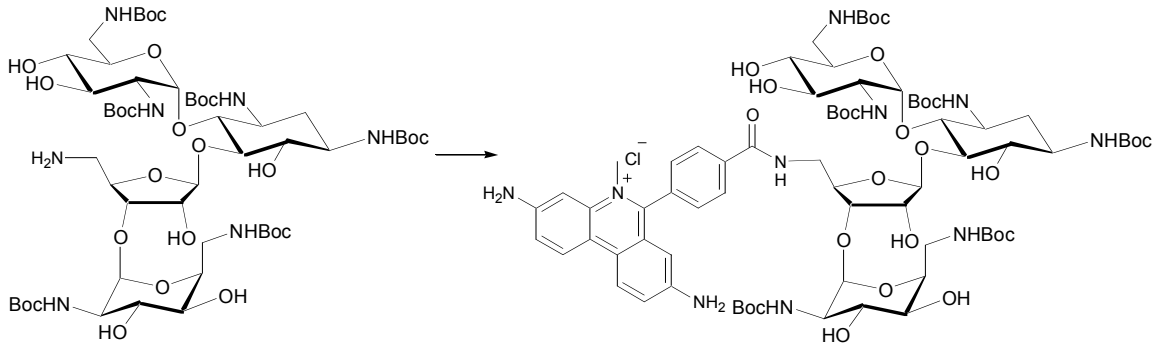
(s, 3H), 3.01 (q, 2H), 2.86 (d, 1H), 2.72 (m, 1H), 2.03 (m, 2H), 1.51 (m, 2H), 1.09 (t, 3H). Exact mass calculated for $C_{44}H_{64}N_{19}O_9$: 1002.51, found 1003 $[M]^+$.



Compound **18**. Sodium azide (304 mg, 4.68 mmol, 30 equiv) was added to a solution of (**3**) (231 mg, 0.156 mmol, 1 equiv) in 5/1 DMF/Water (6 mL) which was then heated to 70^{o}C on an oil bath. The reaction was stirred for 48 hours, removed from the oil bath and allowed to cool to room temperature. Water (250 mL) was added and the reaction was extracted with ethyl acetate (1 × 500 mL, 2 × 200 mL). All organic layers were combined and washed with saturated sodium chloride (1 × 200 mL), dried over Na_2SO_4 and evaporated in vacuo. Flash chromatography (3-6% methanol in dichloromethane) afforded the desired product as a white solid (151 mg, 122 μmol, 80%): R_f 0.4 in 7% methanol in dichloromethane, ^{1}H NMR (400 MHz, methanol-d_4) δ 6.69 (br, 1H), 6.43 (br, 1H), 5.43 (s, 1H), 5.12 (s, 1H), 4.90 (s, 1H), 4.79 (s, 11H), 4.32 (d, 1H), 4.26 (s, 1H), 3.99 (s, 1H), 3.83 (s, 2H), 3.81-3.64 (m, 3H), 3.58-3.39 (m, 8H), 3.38-3.21 (m, 9H), 3.18 (t, 1H), 1.92 (m, 1H), 1.42 (s, 55H), 1.30 (m, 1H). Exact mass calculated for $C_{53}H_{93}N_9O_{24}$: 1239.63, found 1240.5 $[M + H]^+$.

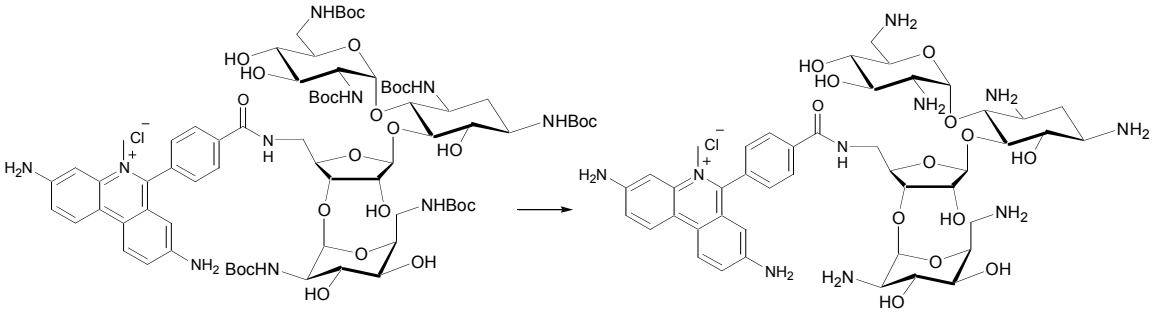


Compound **19**. A solution of (**18**) (100 mg, 0.081 mmol) and palladium on carbon (0.05 g) in methanol (10 mL) was stirred under 1 atmosphere of hydrogen for 48 hours. After all starting material was consumed the reaction was filtered to remove the palladium, and subsequently dried in vacuo to yield a white solid (84 mg, 69 μmol, 85%): R_f 0.2 in 10% methanol in dichloromethane. ^{1}H NMR (400 MHz, methanol-d_4) δ 6.70 (m, 1H), 6.49 (m, 1H), 6.17 (d, 1H), 5.48 (s, 1H), 5.33 (s, 1H), 5.14 (m, 1H), 4.91 (s, 1H), 4.83 (s, 10H), 4.73 (m, 1H), 4.65-4.55 (m, 2H), 4.51 (m, 1H), 4.44 (m, 1H), 4.34 (s, 3H), 4.30-4.07 (m, 4H), 3.89 (m, 2H), 3.80-3.68 (m, 2H), 3.55-3.40 (m, 6H), 3.35-3.11 (m, 10H), 3.07 (m, 1H), 1.91 (br, 1H), 1.43 (s, 54H), 1.30 (m, 2H), 1.18 (m, 1H). Exact mass calculated for $C_{53}H_{95}N_7O_{24}$: 1213.64, found 1214.5 $[M + H]^+$, 1236.4 $[M + Na]^+$.



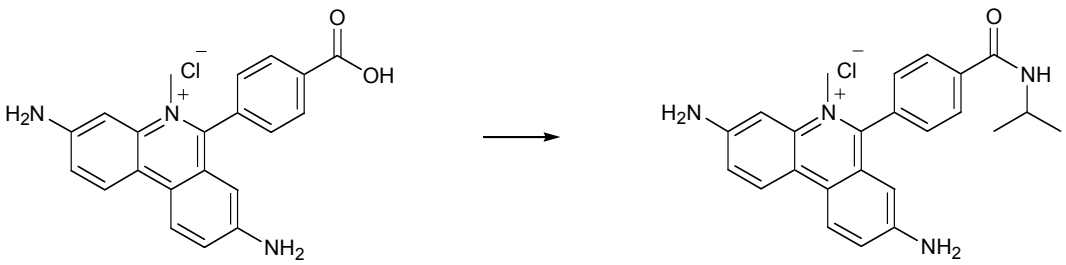
Compound **20**. To a solution of (**19**) (53.9 mg, 0.044 mmol, 1.35 equiv.) in dimethylformamide (3 mL), 1H-benzotriazolium (TBTU) (14 mg, 0.044 mmol, 1.35 equiv.) and 1-hydroxybenzotriazole (HOBT) (6.7 mg, 0.044 mmol, 1.35 equiv) were added. 6-(4-carboxyphenyl)-3,8-diamino-5-methylphenanthridinium chloride (**22**) (12.5 mg, 0.033 mmol, 1 equiv.) and 2,4,6-collidine (0.13 mL) were subsequently added and the reaction mixture was covered with foil and stirred. The reaction was followed by TLC (10% methanol in dichloromethane), and when all starting material (**22**) was consumed (5 hours), chloroform (75 mL) was added and the solution was washed with saturated ammonium chloride (50 mL), saturated sodium bicarbonate (50 mL) and saturated sodium chloride (50 mL). The aqueous layers were combined and back extracted with chloroform (100 mL). The organic layers were combined dried over Na_2SO_4, and evaporated in vacuo. Flash chromatography (5-13% methanol in dichloromethane) yielded the desired product as a red solid which was taken directly to removal of boc-protecting groups (6.2 mg, 4 μmol, 13%), R_f 0.3 in 7% methanol in dichloromethane, ^{1}H NMR (400 MHz, methanol-d_4) δ 8.61 (m, 2H), 8.35 (br, 2H), 8.03 (s, 1H), 7.71 (m, 3H), 7.57 (m, 1H), 7.37 (m, 2H), 7.26 (m, 2H), 6.56 (s, 1H), 5.43 (br,

1H), 5.34 (br, 1H), 5.16 (s, 1H), 4.80 (s, 50H), 4.59 (s, 3H), 4.37 (s, 1H), 4.25 (br, 1H), 4.14 (m, 2H), 4.10 (br, 1H), 3.92 (s, 1H), 3.77 (s, 1H), 3.72 (br, 1H), 3.64 (m, 1H), 3.59-3.48 (m, 3H), 3.37-3.34 (m, 21H), 2.19 (m, 1H), 2.03 (m, 1H), 1.95 (br, 1H), 1.59 (m, 1H), 1.43 (s, 22H), 1.28 (s, 7H), 1.03 (m, 1H), 0.89 (m, 1H).

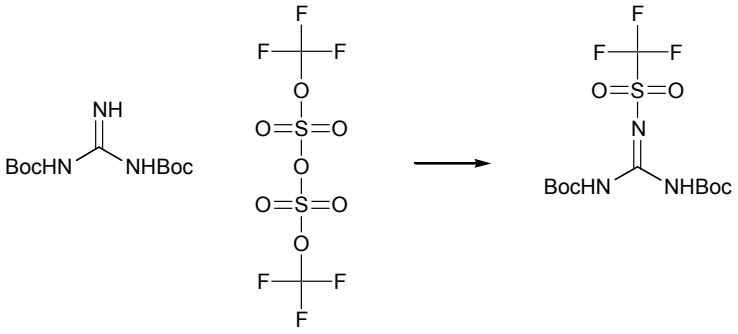


Compound **21**. Triisopropylsilane (30 μL) was added to a solution of crude (**20**) (6.2 mg, 4 μmol) in trifluoroacetic Acid (3 mL) and dichloromethane (3 mL). The reaction was covered in foil and stirred vigorously for 1 hour. Water (15 mL) was added and the reaction was washed with chloroform (1 × 15 mL) and ethyl ether (1 × 15 mL). The aqueous layer was evaporated in vacuo and the resulting solid was dissolved in 0.1% TFA in water (1 mL) and lyophilized to dryness. Analytical reverse phase HPLC (C-8, 5-15% acetonitrile (0.1% TFA) in water (0.1% TFA) over 30 minutes showed one peak at 21.03 minutes). HPLC purification (C-8, 5-15% acetonitrile (0.1% TFA) in water over 30 minutes) gave one major peak at 21.50 minutes. , ^{1}H NMR (400 MHz, DMSO-d_6 + 1 drop of D_2O) δ 8.61 (t, 2H), 8.17 (m, 1H), 8.09 (t, 2H), 7.78 (d, 2H), 7.57 (d, 1H), 7.37 (d, 1H), 7.12 (s, 1H), 6.31 (s, 1H), 5.77 (s, 1H), 5.27 (s, 1H), 5.06 (s, 1H), 4.43 (m, 1H), 4.22 (m, 1H), 4.16 (m,1H), 4.01 (s, 1H), 3.95 (s, 1H), 3.93 (br, 1H), 3.79

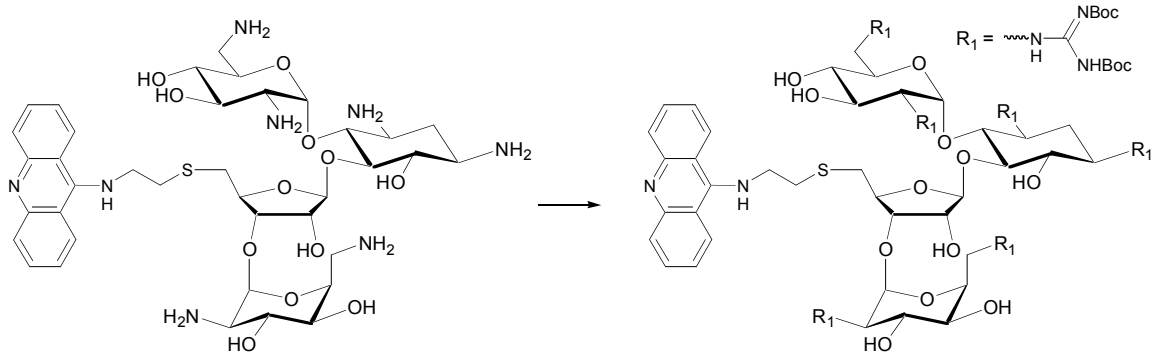
(m, 1H), 3.61 (m, 1H), 3.52 (s, 1H), 3.33 (s, 40H), 3.03 (m, 2H), 2.5 (s, 14H), 2.34 (m, 1H), 2.10 (s, 8H). Exact mass calculated for $C_{44}H_{63}N_{10}O_{13}$: 939.46, found 939 $[M]^+$.



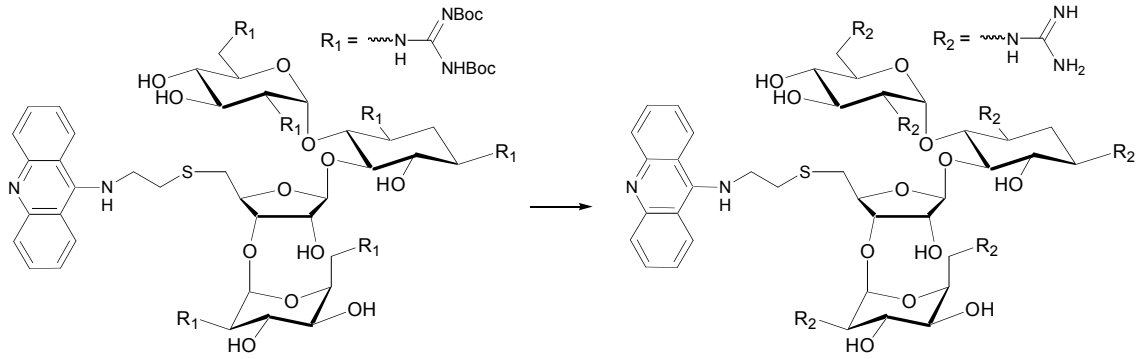
Compound **23**. To a solution of isopropylamine (5 μl, 85 μmol, 1.3 equiv) in dimethylformamide (5 mL), 1H-benzotriazolium (TBTU) (27 mg, 85 μmol, 1.3 equiv.) and 1-bydroxybenzotriazole (HOBT) (13 mg, 85 μmol, 1.3 equiv) were added. 6-(4-carboxyphenyl)-3,8-diamino-5-methylphenanthridinium chloride (**22**) (25 mg, 66 μmol, 1 equiv.) and 2,4,6-collidine (0.26 mL) were subsequently added and the reaction mixture was covered with foil and stirred for 18 hours. All volatiles were removed in vacuo and flash chromatography (3-15% methanol in dichloromethane) afforded the desired product as a red solid. Reverse phase HPLC purification (C-8, 5-20% acetonitrile (0.1% TFA) in water (0.1% TFA) over 30 minutes gave one major product at 27.3 minutes. ^{1}H NMR (400 MHz, methanol-d_4) δ 8.61 (q, 2H), 8.19 (d, 2H), 7.70 (d, 2H), 7.58 (dd, 1H), 7.39 (dd, 1H), 7.36 (s, 1H), 6.52 (d, 1H), 4.90 (s, 21H), 4.29 (m, 1H), 4.14 (s, 3H), 3.31 (m, 5H), 2.65 (s, 1H), 1.32 (s,3H), 1.31(s, 3H). Calculated mass for $C_{24}H_{25}N_4O$: 385.20 found at 385 $[M]^+$.



Compound **24**. Under argon, a solution of di-boc-guanidinium (5 g, 190.23 mmol, 1 equiv) and triethylamine (3 mL) in anhydrous dichloromethane (100 mL) was placed in a dry ice/acetone bath and allowed to equilibrate to -78^{o}C. While slowly stirring, trifluoromethane sulfonic anhydride (3.5 mL, 21 mmol, 1.1 equiv) was added dropwise with the temperature never exceeding -65^{o}C. The reaction was kept in the bath for 30 minutes and then removed from the bath and stirred for 4 hours. The reaction was then washed with water (2 x 50 mL), dried with Na_2SO_4, and all volatiles were removed in vacuo. Flash chromatography (100% dichloromethane) afforded the desired product as a hard white clay (6.21 g, 15.88 mmol, 83%): R_f 0.9 in Dichloromethane, ^{1}H NMR shows one peak at 1.53 in methanol-d_4.



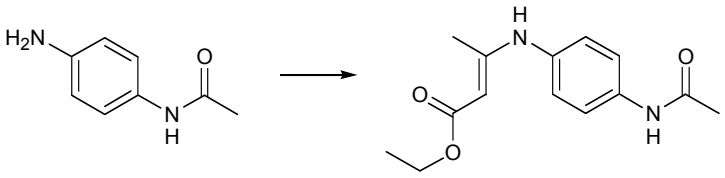
Compound **25** A solution of Neo-Acridine (from Sarah Kirk as described in her Ph.D. Thesis, University of California, San Diego, 2000) (15 mg, 0.018 mmol, 6 equiv) in methanol (8 mL) was treated with chloroform (3 mL), triethylamine (0.4 mL) and (**24**) (692 mg, 1.77 mmol, 100 equiv). Monitoring by TLC (7% methanol in dichloromethane) showed multiple new products forming, all with higher R_f than the starting material. Gradually these products all converged to the fastest moving spot. After 48 hours the reaction appeared complete and all solvent was removed in vacuo. Flash chromatography (2-6% methanol in chloroform) yielded the product as a yellow solid (40.8 mg) which was deprotected directly. ^{1}H NMR (400 MHz, methanol-d_4) δ 8.59 (d, 2H), 7.97 (m, 2H), 7.85 (q, 2H), 7.65 (m, 2H), 5.83 (d, 1H), 5.08 (2, 1H), 4.91 (s, 20H), 4.45-4.15 (m, 3H), 4.13-3.91 (m, 2H), 3.75-3.45 (m, 5H), 3.36 (s, 3H), 3.31 (m, 3H), 2.91 (m, 1H), 2.22 (m, 1H), 1.49 (m, 48H), 1.31 (m, 10H), 0.89 (m, 1H).



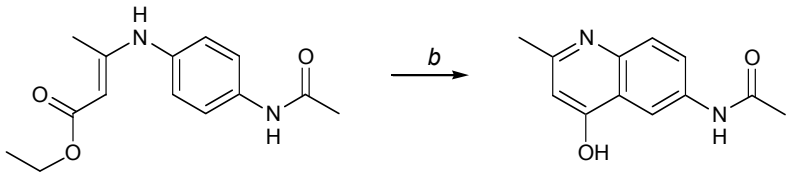
Compound **26**. Triisopropylsilane (10 μL) was added to a solution of (**25**) (40 mg) in dichloromethane (1 mL) and trifluoroacetic acid (1 mL). The reaction was covered with foil and vigorously stirred for 1 hour. When the reaction was complete, water (10 mL) was added and the solution was washed with chloroform (2 × 15 mL) and ethyl ether (2 × 15 mL). The aqueous layer was evaporated in vacuo and the remaining residue was resuspended in 0.1 % trifluoroacetic acid in water (1 mL) and lyophilized overnight. Analytical reverse phase HPLC (C-18, 5-14% acetonitrile (0.1% TFA) in water (0.1% TFA) over 15 minutes than 14% isocratic for 20 more minutes showed one peak at 24.85 minutes and a larger peak at 30.12 minutes). HPLC purification using the same conditions allowed both peaks to be purified with the peak at 24.85 minutes being a mix of 5/6 guanidinylated products and the peak at 30.12 minutes being the desired product as a yellow solid which was dried well, resuspended in 0.1% TFA in water (1 mL) and lyophilized overnight. After dissolution in water (0.8 mL), the concentration of (**26**) was quantified via ultraviolet spectroscopy (known ε's for acridine: $\varepsilon_{222} = 1.9 \times 10^4$, $\varepsilon_{266} = 5.6 \times 10^4$, $\varepsilon_{412} = 9.4 \times 10^3$, $\varepsilon_{438} = 7.9 \times 10^3$) (700 μM, 0.62 mg): ^{1}H NMR (400 MHz, methanol-d_4) δ 8.33 (d, 2H), 7.88 (t, 2H), 7.75 (d, 2H), 7.49 (t, 1H), 5.80 (d,

1H), 5.07 (d, 1H), 4.82 (s, 1H), 4.72 (s, 15H), 4.31 (t, 1H), 4.18 (m, 1H), 3.99 (m, 1H), 3.80 (t, 1H), 3.72 (t, 1H), 3.61 (m, 2H), 3.51-3.18 (m, 10H), 3.17-3.06 (m, 4H), 2.89 (dd, 1H), 2.63 (m, 3H), 2.13 (s, 13H), 1.93 (s, 3H), 1.88 (s, 1H), 1.52 (q, 1H), 1.12 (m, 1H). Exact mass for $C_{44}H_{70}N_{20}O_{12}S$: 1102.52, found 1103.36 $[M + H]^+$, 1125.32 $[M + Na]^+$.

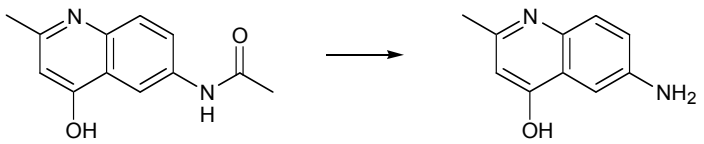
4.2 Synthesis of Surfen Derivatives



Compound **28**. Ethyl acetoacetate (75 mL, 593.25 mmol, 1.19 equiv) was added to a solution of *p*-aminoacetanilide (**27**) (75 g, 500.25 mmol, 1 equiv) in methanol (225 mL). The reaction was refluxed for 5 hours and then removed from heat. After cooling to room temperature a large amount of fluffy, white crystalline precipitate was formed. The product was filtered, washed with cold methanol and dried well. (119 g, 91%). ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.22 (s, 1H), 9.94 (s, 1H), 7.54 (d, 2H), 7.09 (d, 2H), 4.62 (s, 1H), 4.03 (q, 2H), 3.31 (s, 1H), 2.48 (m, 0.5H), 2.02 (s, 3H), 1.92 (s, 3H), 1.17 (t, 3H). Exact mass calculated for $C_{14}H_{18}N_2O_3$: 262.13, found 263.23 $[M + H]^+$, 285.01 $[M + Na]^+$, 261.03 $[M - H]^-$.

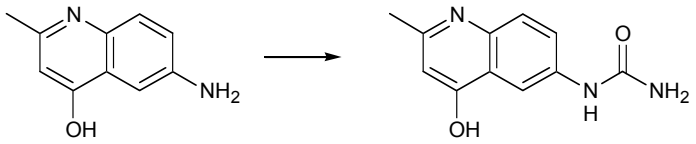


Compound **29**. Dowtherm A (250 mL) was put into a 2 neck round bottom flask sealed on one neck with a glass stopper, and slowly heated to reflux on a heating mantle. Once the solution achieved reflux (>257^{o}C) (**28**) (62.5 g, 239 mmol) was carefully added over a period of 30 minutes. The reaction was allowed to reflux for an additional 30 minutes and then removed from heat and allowed to cool to room temperature. An orange, granular precipitate was observed to form as the solution cooled. Once cool, the precipitate was filtered, washed with cold methanol and dried well. (46.4 g, 90%). ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.49 (s, 1H), 10.04 (s, 1H), 8.21 (d, 1H), 7.81 (dd, 1H), 7.41 (d, 1H), 5.82 (s, 1H), 3.31 (s, 1H), 2.49 (m, 1H), 2.30 (s, 3H), 2.04 (s, 3H). Exact mass calculated for $C_{12}H_{12}N_2O_2$: 216.09, found 217.23 $[M + H]^+$, 239.10, $[M + Na]^+$, 215.12 $[M – H]^-$.

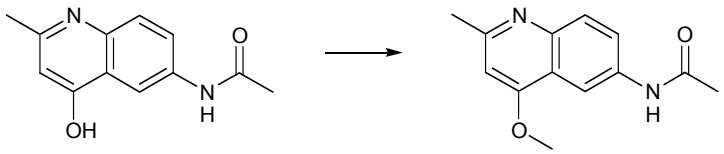


Compound **30**. A solution of (**29**) (500 mg, 2.31 mmol) in water (2.8 mL) and 37% HCl (4.5 mL) was heated to 90^{o}C and stirred for 5 hours. The solution was then removed

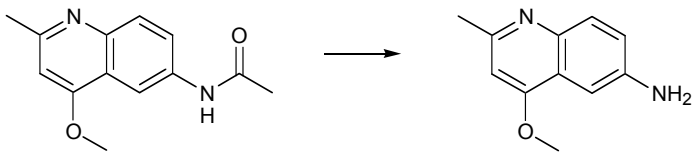
from heat and cooled to 0°C in an ice bath. A white precipitate formed, was filtered, washed with ice cold ethanol (2 × 50 mL) and dried well. (325 mg, 1.87 mmol, 81%). ^{1}H NMR (400 MHz, methanol-d_4) δ 7.49 (d, 1H), 7.44 (d, 1H), 7.04 (dd, 1H), 6.27 (s, 1H), 5.01 (s, 13H), 3.31 (m, 3H), 2.37 (s, 3H). Exact mass calculated for $C_{10}H_{10}N_2O$: 174.08, found 175.21 $[M + H]^+$.



Compound **31**. Potassium cyanate (116 mg, 1.44 mmol, 1 equiv) was added to a solution of (**30**) (250 mg, 1.44 mmol, 1 equiv) in glacial acetic acid (2 mL) and the reaction was stirred for 12 hours. A white precipitate formed, was filtered, washed with ethyl ether and dried. Flash chromatography (5-15% methanol in chloroform) afforded the product as a white solid (154 mg, 0.71 mmol, 49%): R_f 0.4, 10% methanol in chloroform: ^{1}H NMR (400 MHz, methanol-d_4) δ 8.06 (d, 1H), 7.87 (dd, 1H), 7.49 (d, 1H), 6.16 (s, 1H), 4.89 (s, 15H), 3.33 (m, 6H), 2.46 (s, 3H), 1.92 (s, 5H). Exact mass calculated for $C_{11}H_{11}N_3O_2$: 217.09, found 218.17 $[M + H]^+$, 240.12 $[M + Na]^+$, 216.10 $[M – H]^-$.

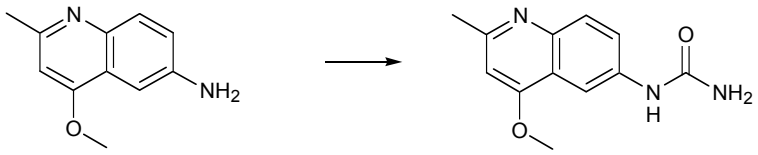


Compound **32**. Dimethyl sulfate (3.9 mL, 40.5 mmol, 1.75 equiv) was slowly added to a suspension of (**29**) (5 g, 23.15 mmol, 1 equiv) in dry toluene (25 mL). The reaction was refluxed for 5 hours until all starting material was consumed (as monitored by TLC 8:2:0.5 chloroform : methanol : ammonium hydroxide), and then removed from heat and allowed to cool to room temperature. The solid precipitate was filtered, washed with toluene and air dried. The dry solid was dissolved in warm water (65 mL), cooled to room temperature and alkalinized with 35% aqueous NaOH (~5 mL). A white solid precipitates from the reaction as it is alkalinized (4.03 g, 17.5 mmol, 76%): ^{1}H NMR (400 MHz, methanol-d_4) δ 8.46 (d, 1H), 7.79 (d, 1H), 7.74 (dd, 1H), 6.85 (s, 1H), 4.85 (s, 5H), 4.06 (s, 3H), 3.31 (m, 2H), 2.64 (s, 3H), 2.17 (s, 3H). Exact mass calculated for $C_{13}H_{14}N_2O_2$: 230.11, found 231.14 $[M + H]^+$.

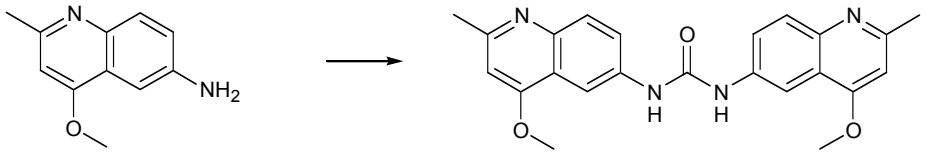


Compound **33**. A solution of (**32**) (500 mg, 2.17 mmol) in water (2.8 mL) and 37% HCl (4.5 mL) was heated to 90^{o}C and mixed for 3 hours until all starting material was consumed (as monitored by TLC 8:2:0.5 chloroform : methanol : ammonium hydroxide). Once complete, the reaction was removed from heat and cooled to room

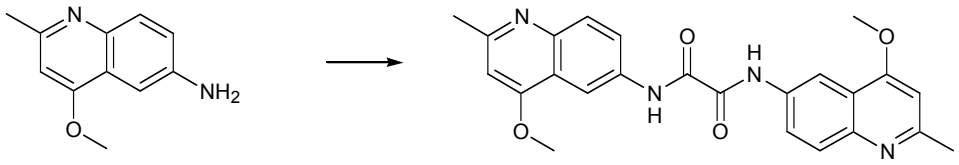
temperature. Once cool, 35% aqueous NaOH was added to alkalinize the solution and precipitate the product as a white solid which was pelleted, dissolved in methanol and dried well (404 mg, 2.15 mmol, 99%): ^{1}H NMR (400 MHz, methanol-d_4) δ 7.60 (d, 1H), 7.21 (d, 1H), 7.14 (dd, 1H), 6.60 (s, 1H), 4.89 (s, 2H), 3.93 (s, 3H), 3.31 (m, 0.5H), 2.52 (s, 3H). Exact mass calculated for $C_{11}H_{12}N_2O$: 188.09, found 189.17 $[M + H]^+$.



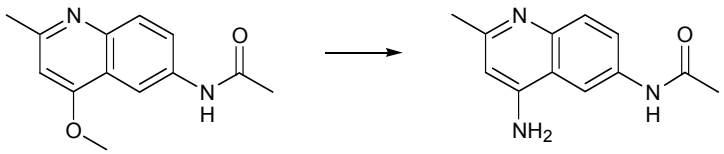
Compound **34**. Potassium cyanate (81 mg, 0.96 mmol, 1.5 equiv) was added to a solution of (**33**) (125 mg, 0.66 mmol, 1 equiv) in glacial acetic acid (2 mL). The reaction was stirred for 16 hours until all starting material was consumed (as monitored by TLC 10% methanol in chloroform). The reaction was dried down and purified via flash chromatography (2-12% methanol in chloroform) which yielded the desired product as a white solid (70 mg, 0.303 mmol, 46%): R_f 0.25, 10% methanol in dichloromethane: ^{1}H NMR (400 MHz, D_2O) δ 7.38 (d, 1H), 7.23 (m, 2H), 6.77 (s, 1H), 4.65 (s, 6H), 3.95 (s, 3H), 3.20 (m, 2H), 2.53 (s, 3H), 1.83 (s, 4H). Exact mass calculated for $C_{12}H_{13}N_3O_2$: 231.10, found 232.05 $[M + H]^+$, 229.99 $[M – H]^-$.



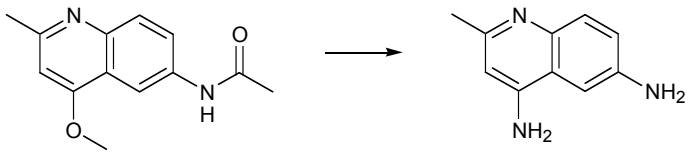
Compound **35**. Triphosgene (30.2 mg, 0.102 mmol, 1 equiv) was dissolved in 1,4-dioxane (0.2 mL) and slowly added to a solution of (**33**) (120 mg, 0.638 mmol, 6.25 equiv) in glacial acetic acid (0.7 mL). A white precipitate quickly forms, and the reaction is monitored by TLC (8:2:0.5 chloroform : methanol : ammonium hydroxide) and allowed to stir until all starting materials are consumed. When the reaction was complete, ethyl ether (5 mL) was added and the white solid was filtered, washed with ethyl ether and dried. This hydrochloride salt was dissolved in warm water (25 mL), cooled to room temperature and alkalinized with 2.5 N NaOH. As the solution became basic, the desired product precipitated out as a white solid (39 mg, 0.097 mmol, 95%). To purify and remove any salts from the compound, reverse phase C-18 flash chromatography (0-20% Acetonitrile (0.1% TFA) in water (0.1% TFA)) was performed and afforded the desired product as a white solid. ^{1}H NMR (400 MHz, methanol-d_4) δ 8.68 (d, 1H), 8.05 (dd, 1H), 7.92 (d, 1H), 7.34 (s, 1H), 4.89 (s, 7H), 4.34 (s, 3H), 3.31 (m, 5H), 2.86 (s, 3H). Exact mass calculated for $C_{23}H_{22}N_4O_3$: 402.17, found 403.13 $[M + H]^+$, 400.96 $[M - H]^-$.



Compound **36**. Oxalyl chloride (44 μL, 0.507 mmol, 1 equiv) was added to a solution of (**33**) (200 mg, 1.06 mmol, 2.1 equiv) in glacial acetic acid (1.2 mL) and a heavy white precipitate quickly formed. After 1 hour TLC (8:2:0.5 chloroform : methanol : ammonium hydroxide) showed no more starting material remained in the reaction. Ethyl ether (10 mL) was added, the solid was filtered, washed with ethyl ether and dried. The hydrochloride salt was dissolved in warm water (25 mL), allowed to cool to room temperature and precipitated as 2.5 N NaOH was added to alkalinize the solution. The precipitate was filtered, washed with water and dried. The white solid was then dissolved in warm DMF (6 mL) which was filtered. Ethyl ether was added and the product precipitated out of the solution as a white solid (101 mg, 0.235 mmol, 46%). To purify and remove any salts from the compound, reverse phase C-18 flash chromatography (0-20% Acetonitrile (0.1% TFA) in water (0.1% TFA)) was performed and afforded the desired product as a white solid. ^{1}H NMR (400 MHz, methanol-d_4) δ 9.01 (d, 1H), 8.43 (dd, 1H), 8.06 (d, 1H), 7.44 (s, 1H), 4.89 (s, 7H), 4.37 (s, 3H), 3.31 (m, 2H), 2.91 (s, 3H). Exact mass calculated for $C_{24}H_{22}N_4O_4$: 430.16, found 431.08 $[M + H]^+$.

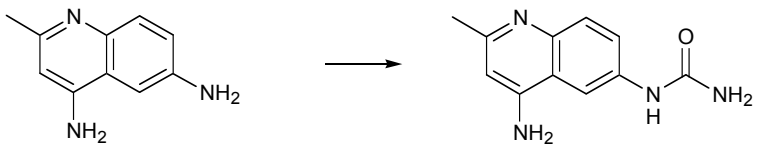


Compound **37**. Ammonium acetate (5 g, 64.94 mmol, 15 equiv) was heated to 135°C with (**32**) (1g, 4.34 mmol, 1 equiv) and stirred under argon for 3 hours until TLC (8:2:0.5 chloroform : methanol : ammonium hydroxide) showed no more starting material remaining. The reaction was then removed from heat and allowed to cool to room temperature and dissolved in water (70 mL). Charcoal (50 mg) was added and the solution was mixed for 20 minutes before the charcoal was filtered off. A white precipitates formed as the solution was alkalinized by the addition of 35% aqueous NaOH. (856 mg, 3.98 mmol, 91%). ^{1}H NMR (400 MHz, methanol-d_4) δ 8.24 (d, 1H), 7.69 (d, 1H), 7.57 (dd, 1H), 6.51 (s, 1H), 4.91 (s, 10H), 3.31 (m, 4H), 2.47 (s, 3H), 2.18 (s, 3H). Exact mass calculated for $C_{12}H_{13}N_3O$: 215.11, found 216.17 $[M + H]^+$.



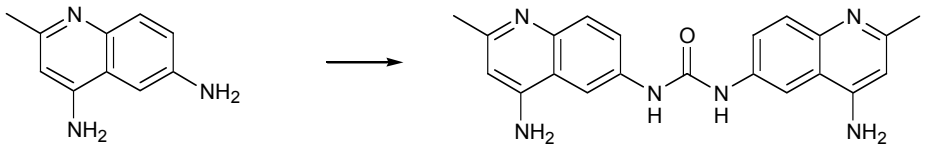
Compound **38**. Ammonium acetate (10.05 g, 130.44 mmol, 15 equiv) was heated to 135°C on an oil bath with (**32**) (2g, 8.70 mmol, 1 equiv) and stirred under argon for 5 hours until TLC (8:2:0.5 chloroform : methanol : ammonium hydroxide) showed no more starting material remaining. Once all starting material had been consumed, water

(10.4 mL) and 37% HCl (16.8 mL) were added and the temperature was reduced to 90°C. The reaction was allowed to stir for an additional 4 hours until TLC showed that no more intermediate (**37**) remained. The reaction was then removed from heat and cooled to 0°C in an ice bath. A precipitate formed, was isolated by filtration and dissolved in hot water (50 mL). Charcoal (35 mg) was added and to the solution and mixed for 25 minutes before being filtered off. The solution was alkalinized with 35% aqueous NaOH, and cooled to 0°C in an ice bath and a white precipitate formed which was filtered and dried (1.20 g, 6.96 mmol, 80%). ^{1}H NMR (400 MHz, D_2O) δ 6.97 (d, 1H), 6.71 (dd, 1H), 6.41 (d, 1H), 5.73 (s, 1H), 4.65 (s, 6H), 1.93 (s, 3H). Exact mass calculated for $C_{10}H_{11}N_3$: 173.10, found 174.23 $[M + H]^+$.



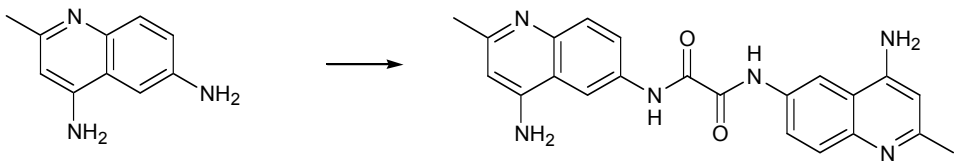
Compound **39**. Potassium cyanate (93 mg, 1.15 mmol, 1equiv) was added to a solution of (**38**) (200 mg, 1.15 mmol, 1 equiv) in glacial acetic acid (10 mL). The reaction was stirred for 16 hours until all starting material was consumed (as monitored by TLC 8:2:0.5 chloroform : methanol : ammonium hydroxide). The reaction was dried down and purified via flash chromatography (86% chloroform, 12% methanol, 2%

ammonium hydroxide) which yielded the desired product as a white solid (106 mg, 0.491 mmol, 43%): R_f 0.4, 8:2:1 chloroform : methanol : ammonium hydroxide: ^{1}H NMR (400 MHz, D_2O) δ 8.28 (d, 1H), 7.76 (dd, 1H), 7.66 (d, 1H), 6.58 (s, 1H), 4.89 (s, 10H), 3.31 (m, 4H), 2.58 (s, 3H). Exact mass calculated for $C_{11}H_{12}N_4O$: 216.10, found 217.05 $[M + H]^+$.



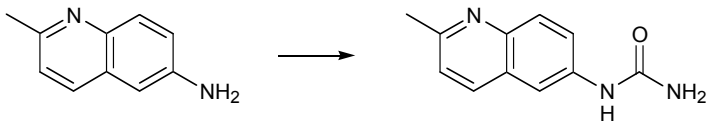
Compound **40**. Triphosgene (57 mg, 0.193 mmol, 1 equiv) was dissolved in 1,4-dioxane (0.2 mL) and slowly added to a solution of (**38**) (200 mg, 1.16 mmol, 6 equiv) in glacial acetic acid (6 mL). A white precipitate forms, and the reaction is monitored by TLC (8:2:0.5 chloroform : methanol : ammonium hydroxide) and allowed to stir until all starting materials are consumed. When the reaction was complete, ethyl ether (25 mL) was added and the white solid was filtered, washed with ethyl ether and dried. This hydrochloride salt was dissolved in warm water (125 mL), cooled to room temperature and alkalinized with 2.5 N NaOH. As the solution became basic, the desired product precipitated out as a white solid. The solid was added to warm DMF (25 mL) and the solution was hot filtered and allowed to cool slowly to room temperature. The desired product precipitated out of solution with ethyl ether, and was purified by reverse phase C-18 flash chromatography (0-20% Acetonitrile (0.1% TFA) in water (0.1% TFA)). Ultraviolet absorbance yielded an extinction coefficient ε_{272} =

30,000 in methanol. ^{1}H NMR (400 MHz, methanol-d_4) δ 8.31 (d, 1H), 8.09 (dd, 1H), 7.73 (d, 1H), 6.61 (s, 1H), 4.88 (s, 10H), 3.31 (m, 4H), 2.61 (s, 3H). Exact mass calculated for $C_{21}H_{20}N_6O$: 372.17, found 373.16 $[M + H]^+$, 371.04 $[M – H]^-$.

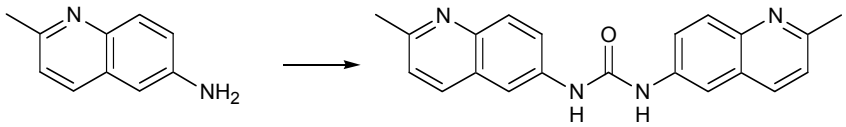


Compound **41** and **42**. Oxalyl chloride (52 μL, 0.592 mmol, 1 equiv) was added to a solution of (**38**) (215 mg, 1.24 mmol, 2.1 equiv) in glacial acetic acid (8 mL) and a heavy white precipitate quickly formed. After 1 hour TLC (8:2:0.5 chloroform : methanol : ammonium hydroxide) showed no more starting material remained in the reaction. Ethyl ether (25 mL) was added, the solid was filtered, washed with ethyl ether and dried. The hydrochloride salt was dissolved in warm water (75 mL), allowed to cool to room temperature and precipitated as 2.5 N NaOH was added to alkalinize the solution. The precipitate was filtered, washed with water and dried. The white solid was then dissolved in warm DMF (10 mL) which was filtered. Ethyl ether was added and the product precipitated out of the solution as a white solid. Reverse phase C-18 flash chromatography (0-20% Acetonitrile (0.1% TFA) in water (0.1% TFA)) did not separate two products which appeared as one spot using TLC analysis. HPLC analysis (C-18, 5% isocratic acetonitrile (0.1% TFA) in water for 5 minutes than 5-30% acetonitrile at 30 minutes) gave 2 peaks at 27.03 minutes and 28.35 minutes which were separated via HPLC using the same conditions. The first peak to elute (**41**) (20.4 mg,

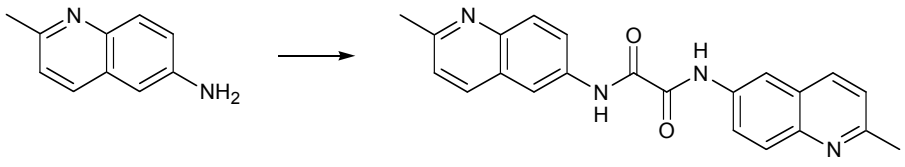
0.051 mmol) was dissolved in water (0.1% TFA) (6 mL). The second peak (**42**) (9.3 mg, 0.023 mmol) was also dissolved in water (0.1% TFA) (3 mL). ^{1}H NMR (400 MHz, methanol-d_4) of (**41**) showed: δ 8.70 (d, 1H), 8.21 (dd, 1H), 7.81 (d, 1H), 6.67 (s, 1H), 4.88 (s, 11H), 3.31 (m, 3H), 2.64 (s, 3H). NMR analysis of (**42**) showed two of each expected peak indicating an asymmetrical molecule: ^{1}H NMR (400 MHz, methanol-d_4) δ 8.70 (d, 1H), 8.64 (d, 1H), 8.21 (dd, 1H), 8.16 (dd, 1H), 7.79 (d, 1H), 7.64 (d, 1H), 6.66 (s, 1H), 6.33 (s, 1H), 4.87 (s, 13H), 3.31 (m, 5H), 2.62 (s, 3H), 2.52 (s, 3H). Exact mass calculated for $C_{22}H_{20}N_6O_2$: 400.16, found (**41**) 401.14 $[M + H]^+$, 399.38 $[M - H]^-$, 421.08 $[M + Na - 2H]^-$, and (**42**) 402.18 $[M + H]^+$. High resolution FAB mass spectroscopy in a glycerol matrix confirms (**42**) as $C_{22}H_{19}O_3N_5$.



Compound **44**. Potassium cyanate (26 mg, 0.32 mmol, 1 equiv) was added to a solution of (**43**) (50 mg, 0.32 mmol, 1 equiv) in glacial acetic acid (5 mL). The reaction was allowed to stir under argon for 2 hours until TLC (20% methanol in chloroform) showed all starting material was consumed. One complete, the reaction was concentrated in vacuo. Flash chromatography (4-8% methanol in chloroform) yielded the desired product as a white solid (43.8 mg, 0.22 mmol, 69%). ^{1}H NMR (400 MHz, D_2O) δ 8.10 (d, 1H), 8.02 (d, 1H), 7.83 (d, 1H), 7.61 (dd, 1H), 7.35 (d, 1H), 4.87 (s, 11H), 3.31 (m, 5H), 2.65 (s, 4H), 2.15 (s, 3H). Exact mass calculated for $C_{11}H_{11}N_3O$: 201.09, found 202.08 $[M + H]^+$.

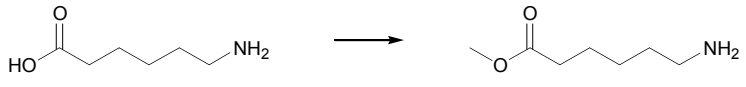


Compound **45**. Triphosgene (47 mg, 0.16 mmol, 1 equiv) was dissolved in 1,4-dioxane (0.2 mL) and added drop wise to a solution of (**43**) (150 mg, 0.95 mmol, 6 equiv) in glacial acetic acid. A white precipitate formed in 30 minutes. The reaction was allowed to stir under argon for 2 hours until TLC (1% ammonium hydroxide and 10% methanol in chloroform) analysis showed all starting material (**43**) was consumed. Ethyl ether (25 mL) was added to the reaction and the white solid was filtered, washed with ethyl ether (100 mL), and dissolved into hot water (50 mL). The solution was alkalinized with 35% aqueous sodium hydroxide and the desired product precipitated out as a free base which was then dried well. The solid was then dissolved in boiling *n*-butanol, and precipitated out as the solution was cooled to 0^{o}C. Reverse phase C-18 flash chromatography (0-20% Acetonitrile (0.1% TFA) in water (0.1% TFA)) afforded the desired product as a yellow solid (26 mg, 0.076 mmol, 48%). ^{1}H NMR (400 MHz, D_2O) δ 8.86 (d, 1H), 8.56 (d, 1H), 8.19 (dd, 1H), 8.08 (d, 1H), 7.84 (d, 1H), 4.93 (s, 8H), 3.31 (m, 4H), 2.95 (s, 3H), 2.03 (s, 2H). Exact mass calculated for $C_{21}H_{18}N_4O$: 342.15, found 343.11 $[M + H]^+$, 706.89 $[2M + Na]^+$.

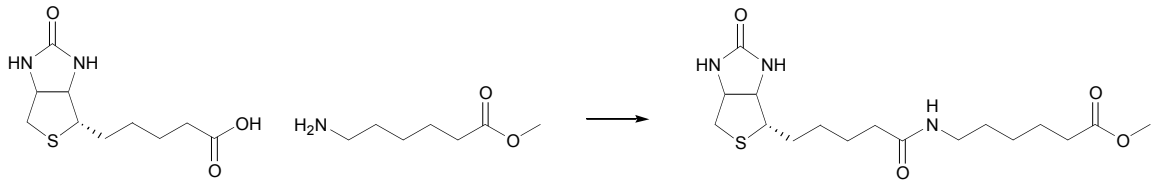


Compound **46**. Oxalyl chloride (26 μL, 0.301 mmol, 1 equiv) was slowly added to a solution of (**43**) (100 mg, 0.632 mmol, 2.1 equiv) in glacial acetic acid (10 mL). A heavy white/brown precipitate formed almost very quickly and the reaction was allowed to stir for two hours. Ethyl ether (25 mL) was added to the reaction and the precipitate was filtered, washed with ethyl ether (50 mL) and dissolved in hot water (65 mL). Alkalinization of the solution with 35% aqueous sodium hydroxide precipitated the desired product as a free base. Reverse phase C-18 flash chromatography (0-20% Acetonitrile (0.1% TFA) in water (0.1% TFA)) and subsequent recrystalization from methanol afforded the desired product as a yellow, crystalline solid (54 mg, 146 μmol, 49%). ^{1}H NMR (400 MHz, D_2O) δ 8.86 (d, 1H), 8.84 (d, 1H), 8.40 (dd, 1H), 8.14 (d, 1H), 7.85 (d, 1H), 4.89 (s, 11H), 3.31 (m, 12H), 2.95 (s, 3H). Exact mass calculated for $C_{22}H_{18}N_4O_2$: 370.14, found 371.12 $[M + H]^+$.

4.3 Synthesis of Biotinylated Neomycin and Guanidino-Neomycin

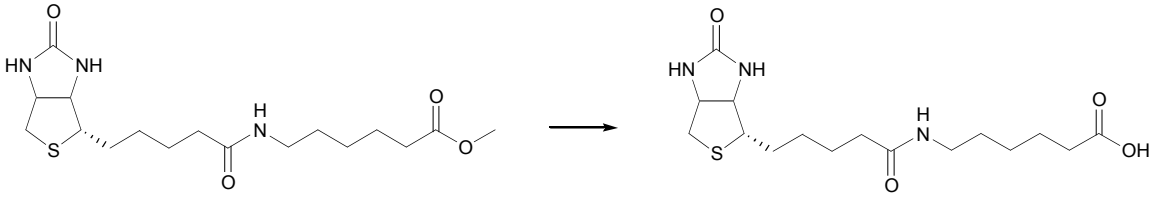


Compound **48**. Thionyl Chloride (2 mL) was slowly added methanol (15 mL at -10°C) and allowed to stir for 10 minutes at -10°C. 6-Aminohexanoic Acid (1g, 7.6 mmol) was added and the reaction was allowed to equilibrate to room temperature and stirred for 16 hours. All volatiles were removed in vacuo and the residue was then dissolved in methanol (7.5 mL). Ethyl ether (30 mL) was then added, precipitating the desired product as white crystals which were then filtered and washed with ethyl ether (10 mL) (1.097 g, 7.5 mmol, 99%). ^{1}H NMR (400 MHz, DMSO-d_6) δ 8.11 (s, 2H), 3.56 (s, 2H), 3.13 (s, 1H), 2.70 (t, 2H), 2.49 (s, 0.2H), 2.28 (t, 2H), 1.52 (m, 4H), 1.28 (m, 2H). Exact mass calculated for $C_7H_{15}NO_2$: 145.11, found 145.98 $[M + H]^+$.



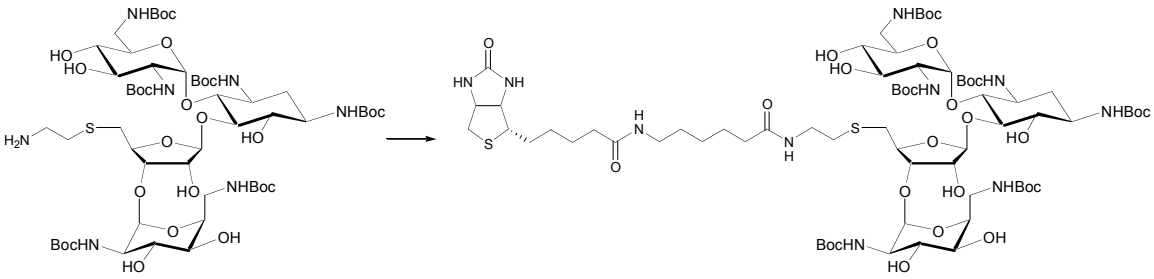
Compound **50**. Biotin (**49**) (225 mg, 920 μmol, 1 equiv) was dissolved in dry DMF (20 mL) which was first passed through silica gel to remove traces of dimethylamine. 1-Hydroxybenzotriazole (HOBT) (190 mg, 1.38 mmol, 1.5 equiv) and 2-(1H-7-

Azabenzotriazol-1-yl)-1,1,3,3-tetramethyl uronium hexafluorophosphate Methanaminium (HATU) (450 mg, 1.38 mmol, 1.5 equiv) were added to the solution of (**49**) in DMF and stirred for 10 minutes until all reagents were dissolved. (**48**) (200 mg, 1.38 mmol, 1.5 equiv) and diisopropylethylamine (1 mL) were added and the reaction was stirred at room temperature for 6 hours. Flash chromatography (2-6% methanol in dichloromethane) afforded the desired product as a white solid (297 mg, 801 μmol, 87%): R_f 0.5 in 10% methanol in dichloromethane, ^{1}H NMR (400 MHz, methanol-d_4) δ 4.87 (s, 3H), 4.48 (q, 1H), 4.31 (q, 1H), 3.72 (q, 3H), 3.64 (s, 2H), 3.60 (q, 1H), 3.33 (s, 2H), 3.22 (m, 6H), 2.91 (dd, 1H), 2.70 (d, 1H), 2.31 (t, 2H), 2.21 (t, 2H), 1.72 (m, 1H), 1.67 (m, 4H), 1.55-1.25 (m, 30H), 1.16 (t, 1H). Exact mass calculated for $C_{17}H_{29}N_3O_4S$: 371.19, found 372.13 $[M + H]^+$, 394.20 $[M + Na]^+$.



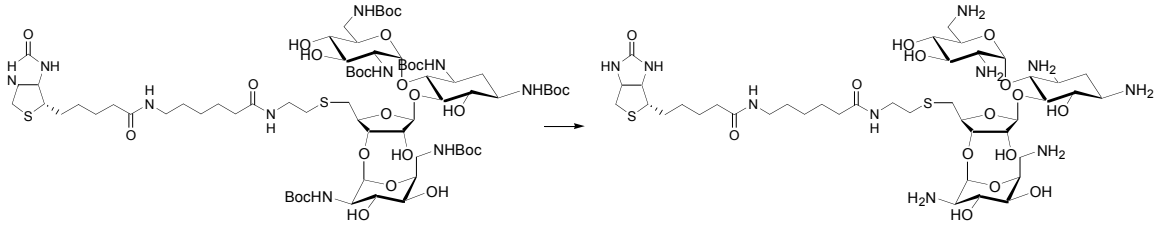
Compound **51**. Sodium hydroxide (25% w/v in water) (1 mL) was added to a solution of (**50**) (250 mg, 674 μmol) in methanol (4 mL). The reaction was stirred at room temperature for 2 hours until TLC analysis (10% methanol in dichloromethane) showed no starting material, and one new product R_f 0.1 in 10% methanol in chloroform. Glacial acetic acid was added dropwise to neutral pH. All volatiles were removed in vacuo and reverse phase C-18 flash chromatography (0-50% Acetonitrile (0.1% TFA)

in water (0.1% TFA)) yielded the desired product as a white solid (107 mg, 299 μmol, 44%). ^{1}H NMR (400 MHz, DMSO-d_6) δ 11.99 (s, 1H), 7.76 (t, 1H), 6.42 (s, 1H), 6.38 (s, 1H), 4.29 (m, 1H), 4.11 (m, 1H), 3.33 (s, 4H), 3.08 (m, 1H), 2.98 (q, 2H), 2.80 (dd, 1H), 2.57 (d, 1H), 2.49 (s, 10H), 2.18 (t, 2H), 2.03 (t, 2H), 1.59 (m, 1H), 1.48 (m, 5H), 1.41-1.18 (m, 7H). Exact mass calculated for $C_{16}H_{27}N_3O_4S$: 357.17, found 356.29 [M – H]$^{-}$.



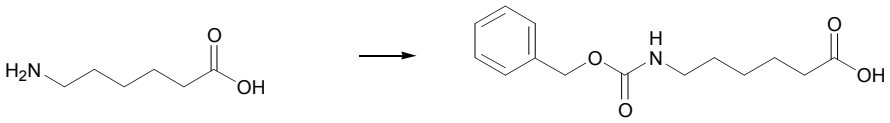
Compound **52**. To a solution of (**51**) (13.5 mg, 33 μmol, 1.1 equiv) in dry DMF (1.5 mL) which was first passed through silica gel to remove traces of dimethylamine, 1-Hydroxybenzotriazole (HOBT) (6 mg, 40.5 μmol, 1.3 equiv) and 2-(1H-7-Azabenzotriazol-1-yl)-1,1,3,3-tetramethyl uronium hexafluorophosphate Methanaminium (HATU) (13.5 mg, 40.5 μmol, 1.3 equiv) were added and the solution was mixed for 10 minutes until all compounds were dissolved. After 10 minutes (**4**) (37.5 mg, 30 μmol, 1 equiv) and triisopropylethylamine (60 μL) were added and the reaction was stirred at room temperature for 10 hours until no starting material was seen and a new spot was seen by TLC R_f 0.6 in 10% methanol in dichloromethane. Flash chromatography (2-10% methanol in dichloromethane) yielded the desired product as a

white solid (34 mg, 21 μmol, 63%). ^{1}H NMR (400 MHz, methanol-d_4) δ 7.78 (dd, 5H), 7.39 (m, 3H), 5.38 (m, 2H), 5.17 (s, 1H) 4.87 (s, 21H), 4.63 (q, 1H), 4.51 (t, 1H), 4.33 (m, 2H), 4.28 (m, 2H), 4.13 (m, 1H), 3.93 (s, 2H), 3.82-3.71 (m, 4H), 3.62-3.12 (m, 30H), 2.92-2.82 (m, 3H), 2.81-2.62 (m, 2H), 2.23 (q, 4H), 1.80-1.23 (m, 89H), 1.22 (m, 4H). Exact mass calculated for $C_{71}H_{124}N_{10}O_{27}S_2$: 1612.81, found 1635.32 $[M + Na]^+$.

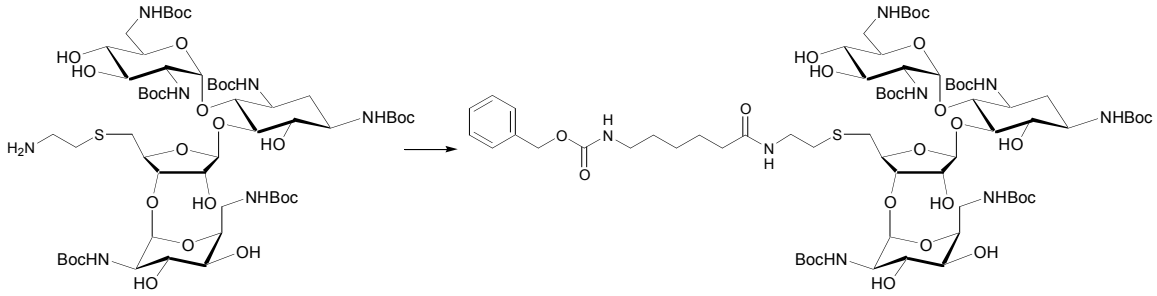


Compound **53**. Trifluoroacetic acid (4 mL) and triisopropylsilane (40 μL) was added to a solution of (**52**) (34 mg, 21 μmol) in dichloromethane (4 mL). This reaction was stirred vigorously at room temperature for 3 hours until TLC (10% methanol in chloroform) showed all starting material consumed and one baseline spot. Water (30 mL) was added and the reaction was washed with chloroform (2 × 20 mL) and ethyl ether (2 × 20 mL). The aqueous layer was evaporated in vacuo and the resulting solid (45 mg) was dissolved in 0.1% TFA in water (1 mL) and lyophilized to dryness. Analytical reverse phase HPLC (C-18, 5-15% acetonitrile (0.1% TFA) in water (0.1% TFA) over 30 minutes showed one peak at 21.03 minutes). HPLC purification (C-18, 5% isocratic acetonitrile (0.1% TFA) for 10 minutes then 5-12% acetonitrile (0.1% TFA) in water over 14 more minutes) gave one major peak at 9.13 minutes. ^{1}H NMR (400 MHz, methanol-d_4) δ 5.93 (d, 1H), 5.30 (s, 1H), 5.19 (s, 1H), 4.67 (s, 10H), 4.31

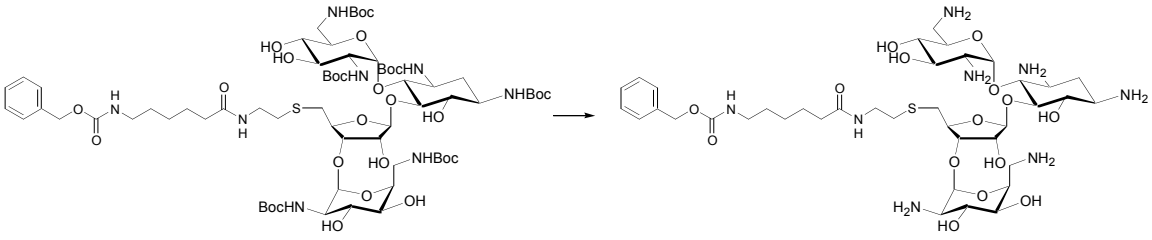
(s, 2H), 4.20 (m, 2H), 4.12 (s, 1H), 3.97 (t, 1H), 3.88 (t, 1H), 3.80 (q, 2H) 3.71 (s, 1H), 3.58 (t, 1H), 3.47 (s, 1H), 3.46-3.22 (m, 12H), 3.21 (s, 8H), 3.15 (m, 2H), 3.07 (t, 16H), 2.82 (d, 8H), 2.76-2.53 (m, 4H), 2.43 (t, 1H), 2.37 (m, 2H), 2.09 (s, 32H), 1.91 (s, 10H), 1.88 (s, 1H), 1.79 (q, 1H), 1.21 (d, 2H), 1.15 (t, 22H). Exact mass calculated for $C_{41}H_{76}N_{10}O_{15}S_2$: 1012.49, found 507.25 $[M + 2H]^{2+}$, 1013.19 $[M + H]^+$, 1035.36 $[M + Na]^+$. High resolution mass confirmed 1013.5018 $[M + H]^+$.



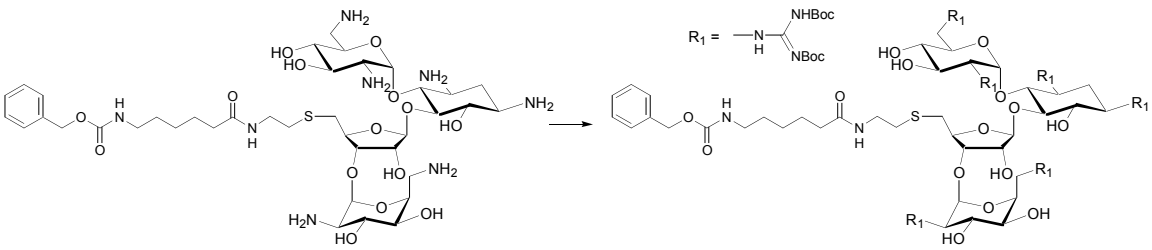
Compound **54**. Sodium carbonate (1.75 g) in water (4 mL) was added to a solution of 6-Aminohexanoic Acid (500 mg, 3.8 mmol) in methanol (45 mL) on ice. The reaction was stirred in an ice water bath for 10 minutes. Benzyl chloroformate (2.5 mL) was slowly added to the reaction which was allowed to equilibrate to room temperature and stirred for 14 hours until all starting material was consumed and a new product was visualized by TLC analysis R_f 0.4 in 8% methanol in dichloromethane. All volatiles were removed in vacuo and flash chromatography (3-6% methanol in dichloromethane) yielded the desired product (585 mg, 2.21 mmol, 58%). ^{1}H NMR (400 MHz, methanol-d_4) δ 7.32 (m, 5H), 6.89 (s, 1H), 5.18 (s, 2H), 5.04 (s, 2H), 3.33 (s, 1H), 3.31 (m, 1H), 3.12 (t, 2H), 2.28 (t, 2H), 1.58 (m, 2H), 1.47 (m, 2H), 1.31 (m, 2H). Exact mass calculated for $C_{14}H_{19}NO_4$: 265.13, found 265.88 $[M + H]^+$, 288.06 $[M + Na]^+$.



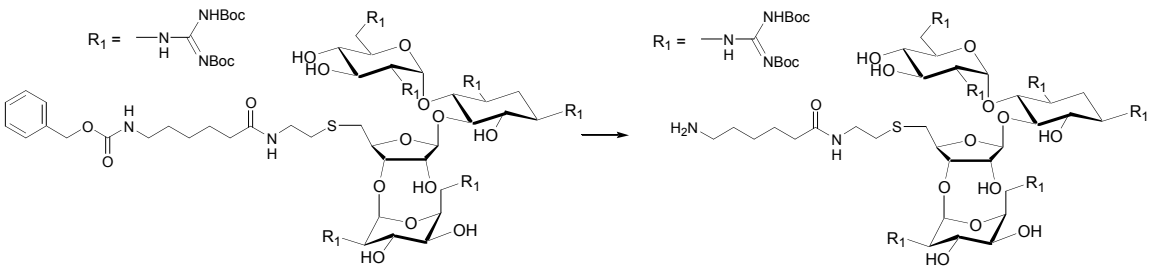
Compound **55**. To a solution of (**54**) (93 mg, 350 μmol, 1.1 equiv) in dry DMF (13 mL) which was first passed through silica gel to remove traces of dimethylamine, 1-Hydroxybenzotriazole (HOBT) (55 mg, 408 μmol, 1.3 equiv) and 2-(1H-7-Azabenzotriazol-1-yl)-1,1,3,3-tetramethyl uronium hexafluorophosphate Methanaminium (HATU) (133 mg, 408 μmol, 1.3 equiv) were added and the reaction was mixed for 10 minutes until all compounds were dissolved. . After 10 minutes (**4**) (400 mg, 314 μmol, 1 equiv) and triisopropylethylamine (600 μL) were added and the reaction was stirred at room temperature for 10 hours until no starting material was seen and a new spot was seen by TLC R_f 0.7 in 10% methanol in dichloromethane. Flash chromatography (2-6% methanol in dichloromethane) yielded the desired product as a white solid (458 mg, 301 μmol, 96%). ^{1}H NMR (400 MHz, methanol-d_4) δ 7.88 (d, 1H), 7.77 (d, 1H), 7.51 (m, 2H), 7.36 (m, 4H), 5.39 (s, 1H), 5.17 (s, 1H), 5.09 (s, 2H), 4.92 (s, 11H), 4.42 (m, 1H), 4.23 (m, 2H), 4.10 (q, 1H), 3.91 (m, 2H), 3.81-3.17 (m, 20H), 3.13 (t, 2H), 2.85 (m, 2H), 2.77 (t, 2H), 2.22 (t, 2H), 1.98 (s, 1H), 1.62 (q, 2H), 1.58-1.25 (m, 55H), 1.21 (t, 4H). Exact mass calculated for $C_{69}H_{116}N_8O_{27}S$: 1520.77, found 1543.28 $[M + Na]^+$.



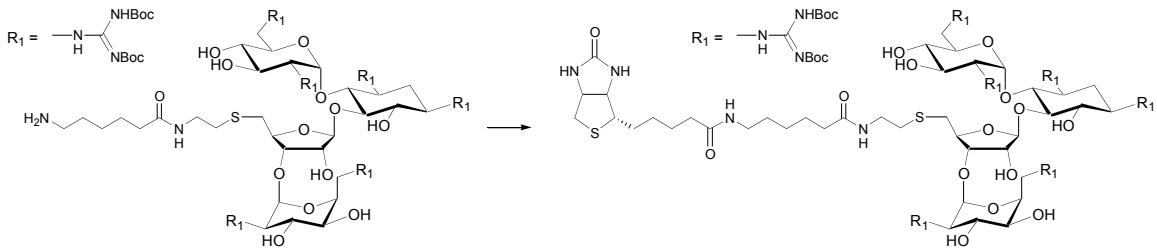
Compound **56**. Trifluoroacetic acid (15 mL) and triisopropylsilane (150 μL) was added to a solution of (**55**) (400 mg, 263μmol) in dichloromethane (4 mL). This reaction was stirred vigorously at room temperature for 3 hours until TLC (10% methanol in chloroform) showed all starting material consumed and one baseline spot. Water (150 mL) was added and the reaction was washed with chloroform (2 × 150 mL) and ethyl ether (2 × 150 mL). The aqueous layer was evaporated in vacuo to yield a the desired product as a white solid (218 mg, 237 μmol, 90%). ^{1}H NMR (400 MHz, methanol-d_4) δ 7.88 (d, 1H), 7.77 (d, 1H), 7.51 (m, 2H), 7.36 (m, 5H), 6.03 (d, 1H), 5.39 (s, 60H), 5.33 (s, 1H), 5.02 (s, 3H), 4.43-4.22 (m, 5H), 4.16 (m, 3H), 4.02 (t, 1H), 3.97 (s, 1H), 3.83 (t, 1H), 3.63 (m, 2H), 3.59 (t, 1H), 3.51-3.02 (m, 21H), 2.90 (t, 1H), 2.70 (m, 4H), 2.45 (d, 1H), 2.20 (t, 2H), 2.08 (m, 1H), 1.61 (t, 3H), 1.50 (t, 2H), 1.34 (t, 8H), 1.22 (m, 2H). Exact mass calculated for $C_{39}H_{68}N_8O_{15}S$: 920.45, found 921.18 $[M + H]^+$, 943.39 $[M + Na]^+$.



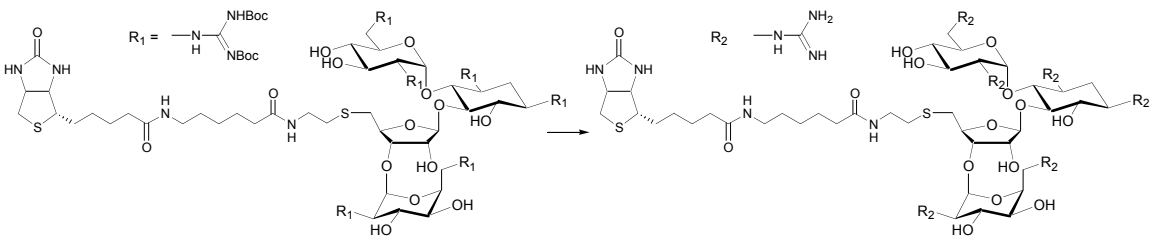
Compound **57**. . Chloroform (4 mL), triethylamine (350 μL) and (**24**) (800 mg, 2.04 mmol, 14 equiv) were added to a solution of (5**6**) (135 mg, 147 μmol) in methanol (1 mL). The reaction was stirred under argon and covered in foil. Monitoring by TLC (4% methanol in dichloromethane) showed multiple new products forming, all with higher R_f than the starting material. Gradually these products all converged to the fastest moving spot. After 48 hours all solvent was removed in vacuo. Flash chromatography (2% methanol in dichloromethane) afforded the desired product as an off white solid (193 mg, 81 μmol, 55%). ^{1}H NMR (400 MHz, methanol-d_4) δ 7.39 (m, 5H), 7.08 (s, 2H), 5.87 (d, 1H), 5.24 (s, 1H), 5.08 (d, 3H), 4.87 (s, 23H), 4.51-4.28 (m, 4H), 4.21 (m, 1H), 4.12 (m, 1H), 4.03 (m, 2H), 3.83 (m, 2H), 3.77-3.59 (m, 4H), 3.47-3.0 (m, 100H), 2.73 (m, 3H). Exact mass calculated for $C_{105}H_{176}N_{20}O_{39}S$: 2373.21, found 791.90 $[M + 3H]^{3+}$, 1187.30 $[M + 2H]^{2+}$.



Compound **58**. Hydrogen gas was bubbled through a solution of (**57**) (180 mg, 76 μmol) and palladium on carbon (100 mg) in methanol (15 mL) for 10 minutes. The reaction was kept under an atmosphere of hydrogen for 18 hours until TLC analysis revealed that all starting material had been consumed and a new, slower moving product appeared R_f 0.3 in 10% methanol in dichloromethane. Flash chromatography (4-10% methanol in dichloromethane) yielded the desired product as an off white solid (146 mg, 65 μmol, 85%). ^{1}H NMR (400 MHz, methanol-d_4) δ 5.92 (d, 1H), 5.50 (s, 2H), 5.08 (d, 2H), 4.87 (s, 40H), 4.59 (m, 4H), 4.34 (m, 4H), 4.30-4.08 (m, 6H), 3.99 (s, 5H), 3.90 (t, 2H), 3.80 (m, 5H), 3.69 (q, 3H), 3.53 (m, 6H), 3.36 (m, 20H), 3.19 (q, 7H), 3.01-2.89 (m, 4H), 2.81-2.67 (m, 8H), 2.38 (t, 5H), 2.19 (s, 3H), 1.90-1.45 (m, 150H), 1.41 (t, 12H), 1.10-0.88 (m, 3H). Exact mass calculated for $C_{97}H_{170}N_{20}O_{37}S$: 2239.18, found 747.30 $[M + 3H]^{3+}$, 1120.32 $[M + 2H]^{2+}$.



Compound **59**. To a solution of Biotin (**49**) (7.3 mg, 30 μmol, 1.2 equiv) in dry DMF (1 mL) which was first passed through silica gel to remove traces of dimethylamine, 1-Hydroxybenzotriazole (HOBT) (5 mg, 33 μmol, 1.3 equiv) and 2-(1H-7-Azabenzotriazol-1-yl)-1,1,3,3-tetramethyl uronium hexafluorophosphate Methanaminium (HATU) (11 mg, 33 μmol, 1.3 equiv) were added and the solution was mixed for 10 minutes until all compounds were dissolved. After 10 minutes (**58**) (55 mg, 25 μmol, 1 equiv) and triisopropylethylamine (40 μL) were added and the reaction was stirred at room temperature for 10 hours until no starting material was seen and a new spot was seen by TLC R_f 0.6 in 10% methanol in dichloromethane. Flash chromatography (2-10% methanol in dichloromethane) yielded the desired product as a white solid (52 mg, 21 μmol, 84%). ^{1}H NMR (400 MHz, methanol-d_4) δ 5.93 (d, 1H), 5.49 (s, 1H), 5.09 (s, 2H), 4.87 (s, 35H), 4.60 (m, 2H), ,4.49 (t, 1H) 4.32 (m, 2H), 4.28-4.10 (m, 3H), 3.99 (s, 2H), 3.90 (t, 1H), 3.85-3.60 (m, 8H), 3.51 (m, 4H), 3.42-3.25 (m, 45H), 3.22 (q, 7H), 2.92 (dd, 2H), 2.77-2.63 (m, 4H), 2.38 (t, 2H), 2.22 (m, 4H), 1.82-1.27 (m, 96H), 1.25-1.10 (m, 3H), 0.90 (m, 2H). Exact mass calculated for $C_{107}H_{184}O_{39}S_2$: 2465.25, found 1233.71 $[M + 2H]^{2+}$.



Compound **60**. Trifluoroacetic acid (5 mL) and triisopropylsilane (50 μL) was added to a solution of (**59**) (50 mg, 20 μmol) in dichloromethane (5 mL). This reaction was stirred vigorously at room temperature for 3 hours until TLC (10% methanol in chloroform) showed all starting material consumed and one baseline spot. Water (25 mL) was added and the reaction was washed with chloroform (2 × 15 mL) and ethyl ether (2 × 15 mL). The aqueous layer was evaporated in vacuo to yield a the desired product as a white solid which was dissolved in 0.1% TFA in water (1 mL) and lyophilized to dryness. Analytical reverse phase HPLC (C-18, 5-20% acetonitrile (0.1% TFA) in water (0.1% TFA) over 30 minutes showed one peak at 11.51 minutes). HPLC purification (C-18, 5%-20% acetonitrile (0.1% TFA) in water over 30 minutes) yielded the desired product as a white solid (18 mg, 14 μmol, 70%). ^{1}H NMR (400 MHz, methanol-d_4) δ 5.92 (s, 1H), 5.17 (s, 1H), 5.07 (s, 1H), 4.82 (s, 90H), 4.59 (t, 1H), 4.38 (m, 2H), 4.12 (m, 2H), 3.80 (t, 2H), 3.77-3.68 (m, 2H), 3.65-3.41 (m, 11H), 3.39-3.26 (m, 45H), 3.18 (q, 6H), 2.96 (m, 2H), 2.77 (d, 1H), 2.69 (t, 2H), 2.23 (q, 4H), 2.09 (s, 18H), 1.78-1.53 (m, 6H), 1.50 (t, 2H), 1.39 (m, 2H), 1.38 (m, 7H). Exact mass calculated for $C_{47}H_{88}N_{22}O_{15}S_2$: 1264.62, found 633.47 $[M + 2H]^{2+}$, 1265.40 $[M + H]^+$. High resolution mass confirmed 1265.6293 $[M + H]^+$.

Synthesis of arginine$_9$-alanine-hexanoicacid-biotin peptide. Peptide synthesis was carried out using standard Fmoc/HBTU chemistry on ABI Applied Biosystems 431A peptide synthesizer using protected aminoacids: Fmoc-Arg-(Pbf)-OH, Fmoc-Ala-OH, and Fmoc-6-aminohexanoic acid. Upon completion of automated synthesis, resin was washed with DMF and dissolved in a solution of biotin (244 mg, 1 mmol), , 1-Hydroxybenzotriazole (HOBT) (130 mg, 858 μmol) and 2-(1H-7-Azabenzotriazol-1-yl)-1,1,3,3-tetramethyl uronium hexafluorophosphate Methanaminium (HATU) (308 mg, 924 μmol) in 5 mL DMF/DMSO (1/1). The reaction was stirred for 18 hours, and upon completion, the resin was washed with dichloromethane and suspended in trifluoroacetic acid/water (19/1) (10 mL), and stirred for 2 hours to cleave the peptide from resin. The resin was then filtered and water (20 mL) was added to the filtrate. All volatiles were removed in vacuo and HPLC purification (C-18, 5%-20% acetonitrile (0.1% TFA) in water over 30 minutes) yielded the desired product as a white solid. ^{1}H NMR (400 MHz, D_2O) δ 4.71 (s, 24H), 4.38 (t, 1H), 4.18 (t, 1H), 4.08 (m, 2H), 3.97 (m, 1H), 3.12 (s, 2H), 3.11 (m, 1H), 2.95 (s, 7H), 2.74 (dd, 1H), 2.51 (d, 1H), 2.04 (m, 3H), 1.82 (s, 2H), 1.67-1.48 (m, 7H), 1.48-1.33 (m, 7H), 1.29 (t, 1H), 1.20-0.99 (m, 4H). Exact mass calculated for $C_{73}H_{141}N_{41}O_{13}S$: 1832.14, found 612.06 $[M + 3H]^{3+}$, 917.26 $[M + 2H]^{2+}$, 1833.74 $[M + H]^{+}$. High resolution mass confirmed 1833.1409 $[M + H]^{+}$.

4.4 Preparing Cells for Tissue Culture

Cells are stored in liquid nitrogen. To prepare them for tissue culture use they must be thawed and passaged once. The following protocol is a general procedure for preparing cells for use. Once thawed, experiments must be performed before the 25th passage to minimize genetic drift.

HAMS F12 media was warmed in water bath to 37^{o}C. Fetal bovine serum was added to the warm F12 media, bringing the solution to 10% calf serum. Fetal bovine serum provides nutrients for cells to grow in. A solution of penicillin/streptomycin was added to the media to a final concentration of 1000 units/mL. The antibiotics are used as anti-bacterial agents to prevent any prokaryotic cell growth. Once the media is prepared, cells are thawed for use.

Mammalian cells survive best when frozen slowly and thawed quickly. Cells are stored in aliquots of HAMS F12 media, fetal bovine serum and DMSO (1 mL). Stocks of CHO-K1, pgsA-745, pgsD-677 or other cell lines were quickly thawed in a water bath at 37^{o}C and transferred into labeled 15 mL falcon tubes. Cell stock tubes were washed with the prepared F12/calf serum/pen/strep media (1 mL) which was then added to the falcon tubes. An additional aliquot (5 mL) of media was added to the tubes.

Cells were then pelleted via centrifugation at 1800 rpm for 5 minutes. All media was carefully aspirated, leaving cell pellets which were resuspended in media (1 mL). Cells, in media, were added to cell growth plates which already contain media

(10 mL). Plates are gently mixed to evenly spread the cells around, but making sure that no media reaches the lip of the plate.

Plates were placed under a microscope to visualize the cells, ensure that they are in suspension and to make sure that there is a large population in the media. Cells are then placed in an incubator for 24 hours and grown at 37^{o}C, 5% CO_2. Live cells will adhere to the surface of the plate and grow in a monolayer.

To prepare cells in monolayers in 6-well plates for individual experiments, they were released, and added to 6 well plates at a concentration of 5×10^5 cells per well in the prepared HAMS F12 media with serum and antibiotics (3 mL). Colonies were grown to a confluent monolayer at 37^{o}C, 5% CO_2 in 24 hours.

4.5 High Throughput Screening For Heparan Sulfate Binders

To screen large numbers of compounds for their ability to bind to cell surface heparan sulfate proteoglycans, an assay was developed. Cells expressing heparan sulfate are exposed to compounds at different concentrations before being incubated with a toxin which acts through binding to the proteoglycan. If the molecules tested bind heparan sulfate, they will protect the cells from toxins. If not, the toxins will be able to bind to, enter and kill the cells.

On the first day, Chinese hamster ovarian (CHO) cells were seeded at 100 cells per well in 192-well plates in 100 μL of HAMS F12 media with 5% fetal bovine serum. On each plate, compounds were tested in multiple concentrations alongside controls with no tested compound, no toxin or sodium chlorate titration.

The wells containing cells and toxin but with no compounds tested serve as a negative control while the wells containing cells but lacking toxin exhibit maximal growth. In a third control, 4 wells were treated with sodium chlorate (6.25, 12.5, 25 and 50 mM) which removes cell surface proteoglycans and serves to mimic the effects of an inhibitor.

On the second day of the assay, compounds were prepared either in DMSO or water. If in DMSO, the minimal stock concentration must be greater than or equal to 10 mM. Solutions of compounds to be tested were pipetted into 96-well plates and serially diluted so that multiple concentrations of each compound can be tested. Compounds in DMSO were diluted 50 fold in complete growth medium while compounds in water were diluted 10 fold into the medium. 5 μL of each of these dilutions were added to the cells with a maximum of not more than 0.2% DMSO in each well.

The compounds were allowed to incubate with the cells for two days, allowing potential inhibitors to act and pre-existing heparan sulfate to turn over.

Two toxins are used to insure that a positive hit in the assay is not resulting from a compound which acts as a toxin inhibitor. An assayed compound which blocks toxin activity would be indistinguishable from a heparan sulfate binding molecule. FGF-saporin is used at 1 μg/mL in complete growth medium and DT (a different toxin specific for heparan sulfate proteoglycans) is used at 3 ng/mL in complete growth medium.

On the fourth day, after the compounds and cells have been incubating for two days, the toxins were introduced. 5 μL of each was added to the wells so that each compound is separately assayed with both toxins at all tested concentrations.

On the seventh day, Alamar blue was added and the cells were incubated at 37°C (5% CO_2) for 24 hours. After incubation, the plates were read for Alamar blue fluorescence in a fluorescent plate reader (Applied Biosystems CytoFluor® Series 4000 Multi-Well Plate Reader) ($530_{EX}/580_{EM}$).

Cells expressing heparan sulfate die from the toxins and show low fluorescent signals. Cells which have been protected from the toxins are not killed and therefore display increased fluorescent signals. Cells can be protected from the toxins through lack of heparan sulfate or through compounds binding to the heparan sulfate and blocking the interactions which lead to toxin activity on the cell.

4.6 Heparan Sulfate Binding Through Blocking of Fibroblast Growth Factor

The Fibroblast Growth Factor (FGF) family is responsible for maintaining the building blocks of fibrous tissue. Basic FGF (FGF-2) binds specifically to heparan sulfate on the surface of cells, and can be used as a tool to evaluate the binding of ligands to these proteoglycans. In this assay, cells are exposed to a compound and biotinylated FGF-2. Cells are then washed with a streptavidinylated fluorophore which will bind to any cell surface bound FGF-2. Analysis via flow cytometry allows for quantification of bound FGF-2 with which IC_{50} values for inhibitors can be derived.

Wild type Chinese hamster ovarian (CHO-K1) cells were grown to confluency in 6-well tissue culture plates in HAMS F-12 media containing 10% Fetal Calf Serum (FCS) and penicillin/streptomycin (to a final concentration of 1000 until/mL). Fetal calf serum provides nutrients for cells to grow in and the antibiotics are used as anti-bacterial agents to prevent any prokaryotic cell growth.

Once confluency was achieved, all media was gently aspirated and all wells were washed once with Dubelcos phosphate buffered saline (PBS) (2 mL). Cells were then incubated with of EDTA (10 mM in PBS, 1 mL per well) at 37^{o}C, 5% CO_2 for 10 minutes. Cells were then gently pipetted to ensure release from the surface of the tissue culture plates. Each well was then flooded with PBS (2 mL) and all solution, with cells, was then transferred to flow cytometry centrifuge tubes.

Cells were then pelleted via centrifugation at 1500 rpm for 5 minutes and suspended into pre-chilled FACS buffer (0.1% bovine serum albumen (BSA) in PBS) which contains a 1-1000 dilution of FGF-2-biotin and varying concentrations of the potential inhibitor ligand (0.5 mL). The suspension was then incubated at 4^{o}C for one hour.

After one hour, PBS (1 mL) was added to the suspension and cells were pelleted via centrifugation at 1500 rpm for 5 minutes. The supernatant was gently aspirated, leaving small white cell pellets at the bottom of each tube. Cells were resuspended in PBS (1 mL) and the centrifugation/aspiration/wash was repeated twice.

After the third wash, cells were resuspended in FACS buffer (0.5 mL) which contained a 1-1000 dilution of Phycoerythrin-Cy5-Streptavidin (Streptavidin-PE-Cy5) and incubated at room temperature, shielded from light, for 20 minutes. After

incubation, cells were pelleted via centrifugation at 1500 rpm for 5 minutes. The supernatant was gently aspirated, leaving small white cell pellets at the bottom of each tube. Cells were resuspended in PBS (1 mL) and the centrifugation/aspiration/wash was repeated twice.

Cell pellets were then re-suspended in FACS buffer (0.5 mL) and analyzed via flow cytometry. Fluorescence activated cell sorting (FACS) analysis was used to quantify the fluorescence of 25000 cells per sample. Comparison of the mean fluorescent values at multiple ligand concentrations allows for generation of IC_{50} values for FGF-2 inhibition.

4.7 Direct Binding and Uptake Experiments with BODIPY and Biotin labeled compounds

To directly measure compound binding to and internalization into cells, fluorescent tagged molecules were synthesized and incubated with cells. Comparison of the signal seen from different compounds with a single cell line helps to generate an understanding of pharmacophores necessary for interactions with the cell, while comparison of signal from one compound with different cell lines can elucidate the necessary cellular components for binding and uptake.

These experiments are done with either covalently tagged BODIPY conjugates (neomycin, tobramycin, guanidino-neomycin, guanidino-tobramycin and arg_9 peptide) or biotinylated compounds which were conjugated to a streptavidin bound

Phycoerythrin-Cy5 fluorophore (see Chapter 2 for details on BODIPY and Biotin tagged compounds).

BODIPY tagged compounds were stored as dry solid at -78^{o}C and freshly dissolved in water prior to each experiment. All BODIPY-glycoside conjugates are slightly to moderately hygroscopic, therefore concentrations were determined in methanol using ultraviolet absorbance with ε_{502} = 76,000 $cm^{-1}M^{-1}$. Compounds were then diluted to 1X concentrations in HAMS F-12 media. For experiments done at 4^{o}C these solutions were incubated on ice for 30 minutes. Samples for competition experiments where guanidino-neomycin was used to compete off arg_9-BODIPY signal were done where a fixed concentration of arg_9-BODIPY and varying concentrations of guanidino-neomycin were prepared together before being added to cells.

Biotinylated compounds were stored as a solution in water at -20^{o}C. After thawing at room temperature, compounds were diluted into HAMS F-12 media to 1 μM. To this solution, streptavidin-PE-Cy5 was added in a 1-1000 dilution. To ensure completion of the biotin–streptavidin reaction, this solution was gently mixed and allowed to incubate at room temperature, shielded from light, for 30 minutes. Following this incubation, compounds were diluted to 1X concentrations in HAMS F-12 media. For experiments done at 4^{o}C these solutions were incubated on ice for 30 minutes.

Wild type Chinese hamster ovarian (CHO-K1) (or CHO mutant) cells were grown to confluency in 6-well tissue culture plates in HAMS F-12 media containing 10% Fetal Calf Serum (FCS) and penicillin/streptomycin (to a final concentration of 1000 until/mL). For experiments done at 4^{o}C tissue culture plates were placed on ice

during the wash and incubation steps and PBS was pre-chilled on ice for 30 minutes. Following aspiration of the media from each well, cells were washed with PBS (1×). Aspiration of the PBS wash was followed by addition of prepared media with the appropriate compound and concentration (2 mL) to each well.

Cells were then incubated on ice, in a 4^{o}C atmosphere for experiments where endocytosis was excluded or at 37^{o}C in an atmosphere of 5% CO_2 for all other experiments. All incubations were for one hour unless otherwise noted.

After this incubation, media was quickly aspirated from each well, and cells in each well were washed with PBS (5× 2 mL). Washes were done gently to ensure minimal cell loss as strong shearing forces can cause cellular release. After the fifth wash, cells were then incubated with of EDTA (10 mM in PBS, 1 mL per well) at 37^{o}C, 5% CO_2 for 10 minutes. Cells were then gently pipetted to ensure release from the surface of the tissue culture plates. Each well was then flooded with PBS (2 mL) and all solution, with cells, was then transferred to flow cytometry centrifuge tubes.

Cells were pelleted via centrifugation at 1500 rpm for 5 minutes. The supernatant was then gently aspirated, leaving small white cell pellets at the bottom of each tube. Cell pellets were then re-suspended in FACS buffer (0.5 mL) and analyzed via flow cytometry. Fluorescence activated cell sorting (FACS) analysis was used to quantify the fluorescence of 25000 cells per sample.

4.8 Viability Assay with Saporin Conjugated Compounds

To examine the ability of guanidino-neomycin to facilitate the delivery of bioactive molecules into the cell, it was conjugated to saporin and assayed for the ability of this conjugate to kill cells. Wild type and pgsA-745 cells were plated in 96 well plates at 300 cells per well, in 100 μL of HAMS-F12 media with 10% Fetal Calf Serum and 100 U/mL penicillin and 100 μg/mL streptomycin. Cells were allowed to adhere overnight at 37°C in an atmosphere of 5% CO_2.

A conjugate of Saporin-streptavidin or non-conjugated Saporin (without streptavidin) (Advanced Targeting Systems, San Diego, California) was allowed to complex with neomycin-biotin, guanidino-neomycin biotin or Arg_9-biotin in HAMS-F12 media with 10% Fetal Calf Serum for 30 minutes at room temperature. Saporin-streptavidin is a covalent complex of Saporin and streptavidin which allows the toxin to be tethered to biotinylated molecules. As a control, Saporin, which has no inherent affinity for biotin was also used. Molar ratios of biotinylated compound to Saporin were kept constant at 4:1. Cells were washed with phosphate buffered saline, treated with 150 μL of prepared media and incubated for 4 days at 37°C in a 5% CO_2 atmosphere.

On the last day, 50 μL of media was removed from each well, leaving 100 μL remaining. To each well, 20 μL of CellTiter-Blue (promega) was added, and allowed to incubate for 4 hours at 37°C in an atmosphere of 5% CO_2. CellTiter-Blue contains a buffered solution of resazurin, a redox indicator dye that can be added directly to cells in culture. Living cells can convert the oxidized form of the dye (resazurin) into a

reduced form (resorufin) which has a distinct excitation/emission profile. Excitation at 488 nm results in emission at 575 nm. Because only living cells will metabolize resazurin into resorufin, this fluorescence is directly proportional to viable cell numbers. Analysis of fluorescence measurements from each well allow for determination of the relative proportions of living cells.

Wild type and pgsA cells were examined with guanidino-neomycin-biotin streptavidin-saporin at concentrations ranging from 82 pM to 84 nM. Control experiments were performed with guanidino-neomycin-biotin in the absence of saporin and also with guanidino-neomycin-biotin being incubated with saporin which was not streptavidinylated. While guanidino-neomycin shows no inherent toxicity to cells within the concentration range studied, unconjugated saporin does retain the ability to kill cells (figure 25). However, cell populations were only reduced by roughly 50% at the highest saporin concentrations (84 nM).

The conjugated guanidino-neomycin-saporin system, however, induces cell death at much lower concentrations. IC_{50} values for cell death with wild type cells are in the low (1-5) nanomolar range. While the guanidino-neomycin-saporin system does show some cell lysis activity at higher concentrations with pgsA cell, it is roughly equivalent to the activity of unconjugated saporin, indicating that cell death here is a result of the inherent activity of saporin as opposed to delivery from guanidinoglycosides.

4.9 Synthesis and ^{32}P-end labeling of RNA molecules

Modified and biotinylated RNA sequences were purchased from Dharmacon (www.dharmacon.com) and other RNA sequences were prepared using T7 RNA polymerase in "run-off" *in vitro* transcriptions as previously described[1] and purified using denaturing polyacrylamide gel electrophoresis followed by extraction into 200 mM potassium acetate/10 mM EDTA and ethanol precipitation. RNA molecules were quantified using molar extinction coefficients (A = 15,400 cm^{-1} M^{-1}, G = 11,700 cm^{-1} M^{-1}, C = 7,300 cm^{-1} M^{-1}, U = 10,100 cm^{-1} M^{-1}) in water. RNA was first dephosphorylated with Calf Intestinal Phosphotase in water and end labeled with γ-ATP and T4 nucleotide kinase.[2]

4.10 Fluorescent Anisotropy and Solid Phase Peptide Displacement

Fluorescence peptide displacement anisotropy was performed as previously described.[3] The solid phase assay for identification of RRE selectivity was performed as previously described[4] with both the full length RRE-66 and the RRE IIB constructs. Gel shift experiments and direct titrations were performed as described.[5]

4.11 References

(1) Uhlenbec, O C; Milligan, J F *Methods Enzymol*. **1989**, 180, 51-62.

(2) Kirk, S R; Luedtke, Nathan W.; Tor, Y *J. Am. Chem. Soc*. **2000**, 122, 980-981.

(3) Luedtke, N. W.; Tor, Y. *Biopolymers* **2003**, 70, 103-119.

(4) Luedtke, N. W.; Tor, Y. *Angew Chem Int Ed Engl* **2000**, 39, 1788-1790.

(5) Luedtke, N. W. **2003**, *Thesis Dissertation.*