

G-quadruplex aptamers for cervical cancer therapy

Josué Leandro de Oliveira e Carvalho

Tese para obtenção do Grau de Doutor em
Biomedicina
(3^o ciclo de estudos)

Orientador: Prof.^a Doutora Carla Patrícia Alves Freire Madeira Cruz
Co-orientador: Prof. Doutor João António Queiroz
Co-orientador: Prof. Doutor Gilmar Fernandes Salgado

dezembro de 2020

Provas de doutoramento em Biomedicina realizadas a 10 de dezembro de 2020, sendo o júri das referidas provas constituído por:

- Prof. Doutor Miguel Castelo Branco, Professor Catedrático da Universidade da Beira Interior
- Doutora Sara Richter, Professora Catedrática da Universidade de Padova (Itália)
- Doutor David Monchaud, Investigador Sénior do CNRS da Universidade de Burgogne CNR – ICMUB – UBFC (França)
- Doutor Rui Manuel de Medeiros Melo Silva, Professor Associado Convidado da Faculdade de Ciências da Saúde da Universidade Fernando Pessoa
- Doutora Maria Alexandra da Silva Paulo, Professora Auxiliar da Faculdade de Farmácia da Universidade de Lisboa
- Doutora Cândida Ascensão Teixeira Tomaz, Professora Associada da Universidade da Beira Interior
- Doutora Fani Pereira de Sousa, Professora Auxiliar da Universidade da Beira Interior
- Professora Carla Patrícia Alves Freire Madeira da Cruz, Professora Auxiliar Convidada da Universidade da Beira Interior

Aknowledgments

Em primeiro lugar, quero agradecer à minha orientadora, Professora Carla Cruz, por me ter lançado este desafio há vários anos atrás e pelo incessante apoio e disponibilidade que me ofereceu durante os mesmos. Os ensinamentos que me passou vão além da investigação e contribuíram fortemente para o meu desenvolvimento profissional e pessoal. Obrigado pela amizade e confiança que depositou em mim.

Ao Professor Gilmar Salgado, que me abriu as portas de um país e uma realidade diferentes e me acolheu na sua equipa fazendo-me sentir em casa. Obrigado por estimular o meu sentido crítico, pela disponibilidade em ajudar e partilhar o seu conhecimento comigo.

Ao Professor João Queiroz, pelo seu apoio, disponibilidade e pelos preciosos conselhos, que fruto da sua vasta experiência, ajudaram a derrubar barreiras.

To Doctor Jean-Louis Mergny, for sharing his immense expertise, scientific guidance and for his invaluable constant support. Thank you for hosting me in your research group, it was an immense privilege to have worked with you. Also, to Aurore Guédin and Julien Marquevielle, who assisted me during all my visits to Bordeaux, thank you for your patience.

À Universidade da Beira Interior, particularmente ao Centro de Investigação em Ciências da Saúde, e ao Institut Européen de Chimie et Biologie da Universidade de Bordéus, por me darem todas as condições necessárias ao desenvolvimento deste trabalho. Um agradecimento especial à Sofia Duarte e Margarida Carrilho pela paciência, ajuda e momentos de descontração.

À Fundação para a Ciência e Tecnologia (FCT), por financiar a minha Bolsa de Doutoramento (SFRH/BD/122953/2016).

A todos os colegas de laboratório e amigos que me acompanharam nesta cruzada, João Silva, Tiago Santos, André Esgalhado, Jéssica Nunes, Joana Figueiredo, Daniela Alexandre e muitos outros cujos nomes não caberiam numa página apenas. Saibam que vos estimo muito e agradeço profundamente a vossa amizade, apoio, paciência e colaboração. Aos meus amigos mais fit, João, Vanessa, Jorge e Marta, obrigado por manterem a minha sanidade em dia com o vosso humor constante e apoio incondicional.

À Cátia, minha companheira, obrigado pelo apoio incondicional e por seres um exemplo de força que me manteve à tona quando as águas pareciam mais agitadas. Obrigado pelas colossais paciência e compreensão, e pelo teu carinho e amor. Espero que partilhemos muitas mais conquistas e sonhos.

Por último, quero agradecer do fundo do coração à minha família, ao meu Pai, um exemplo de humildade e empenho, à minha Mãe, um exemplo de genuinidade e perseverança, e aos meus queridos irmãos Carlitos e Filó, pelo vosso amor, pela excelente educação que me deram e por estarem comigo em todos os momentos da minha vida.

Preface

*Neste circo que é o mundo,
É no trapézio o meu lugar.*

Napoleão Mira

Cervical cancer due to Human Papillomavirus (HPV) infection is a serious threat to women's health and lives, and worldwide, one woman dies of cervical cancer every two minutes. Cervical cancer is a public health concern, mainly in low-and middle-income countries where the burden of cervical cancer falls on the women who lack access to health services. No woman should die from cervical cancer. Currently, we have the knowledge and the medical tools to prevent it effectively and to eliminate it as a public health issue which is within reach for all countries. This Thesis is the result of a collaborative effort of several international R&D institutions, particularly the Health Sciences Research Centre (CICS-UBI, Portugal), the European Institute of Chemistry and Biology (IECB, France), the Centre for Nuclear Sciences and Technologies (C2TN, Portugal) and Penn State College of Medicine (PA, USA) who tackled the problem of HPV-induced cervical cancer and the development of novel therapeutic strategies.

In this Thesis we have addressed the potential of targeting the G-quadruplex (G4)-forming sequences within the genomes of both low- and high-risk HPVs and cervical cancer cells with small compounds with the ability of recognizing and stabilizing these structures, yielding a dual antiviral and anticancer effect. Furthermore, we have focused on the development of aptamer-based drug delivery systems. In **Chapters 1** and **2** we introduce cervical cancer, the HPV life cycle, and the attractiveness of G4s for the development of antiviral and anticancer therapeutic strategies. **Chapter 3** establishes the specific aims of the Thesis and its outline. **Chapters 4-6** present three different classes of quadruplex-binding molecules and the study of their interaction with biologically relevant G4s such as telomeric and oncogene promoter sequences. The selected compounds in terms of G4 binding and stabilization were then studied in **Chapter 7** for their antiviral potential in HPV16 and HPV18 infected organotypic cultures. **Chapters 8-10** describe the use of aptamer AS1411 and its derivatives as drug delivery systems of acridine-based G4 compounds to cervical cancer cell lines. In **Chapter 11**, we describe the development of drug-loaded AS1411-functionalized nanoaggregates, and their local administration in vaginal tissue using gel formulations. Finally, **Chapter 12** presents a general discussion of the main findings of this Thesis and their impact on future perspectives in the context of HPV-related diseases and their treatment.

Resumo

Anualmente, mais de 500 mil novos casos de cancro do colo do útero são diagnosticados, com uma taxa de mortalidade a rondar os 60%. Em Portugal, cerca de 750 novos casos são diagnosticados anualmente, que em 2018 resultaram em 340 mortes. Os vírus do papiloma humano (VPHs) são a principal causa do desenvolvimento de cancro do colo do útero. Os VPHs de alto-risco do tipo 16 e 18 são responsáveis por 70% de todos os casos. A infeção persistente por VPH leva à hiperproliferação das células infetadas, através da inibição da diferenciação epitelial, do desenvolvimento epidérmico e da resposta imune. O potencial oncogénico dos VPHs de alto risco é atribuído à expressão das oncoproteínas E6 e E7, que inativam o gene supressor de tumores p53 e membros da família do retinoblastoma, levando ao desenvolvimento de carcinomas. Apesar dos recentes progressos no diagnóstico e vacinação contra o VPH, que reduziram a incidência de cancro do colo do útero, ainda não existe tratamento para mulheres infetadas com VPH. Além disso, as terapias existentes para o cancro do colo do útero tais como quimioterapia, radioterapia e cirurgia, têm eficácia e especificidade limitadas, causando sofrimento aos pacientes, e não beneficiam da investigação sobre os padrões de interação hospedeiro-vírus durante a oncogénese.

O genoma do VPH está dividido em 3 regiões: uma região denominada RLC que contém elementos *cis* necessários à replicação e transcrição do ADN viral; uma região prematura composta por seis fases de leitura aberta que codificam as proteínas E1 - E7; e uma região tardia com duas fases de leitura aberta codificando as proteínas estruturais da cápside, L1 e L2. Recentemente, foram encontradas várias regiões ricas em guaninas no genoma de dez tipos de VPH e o potencial destas sequências formarem estrutura estáveis de G-quadruplex (G4) foi demonstrado. Essas sequências ricas em guanina estão presentes nas regiões RLC, E1, E4 e na sequência codificante da proteína L2, que estão envolvidas na transcrição, replicação e produção de proteínas virais. Assim, o desenvolvimento de moléculas capazes de reconhecer e ligar essas regiões ricas em guaninas do genoma do VPH pode servir como uma estratégia antiviral, dado o seu potencial de regular a expressão de genes virais.

A nucleolina é uma proteína sobreexpressa na superfície das células cancerosas do colo do útero e que demonstrou estar envolvida na ativação da transcrição de oncogenes do VPH18 e também na manutenção da estabilidade do genoma do VPH16. A nucleolina está presente no núcleo, no citoplasma e na membrana plasmática e foi descrita por estar

envolvida na carcinogénese, angiogénese e em vias de sinalização do cancro. Por isso, esta proteína é um alvo atraente para o desenvolvimento de terapias direcionadas para o cancro do colo do útero e a infeção por VPH, devido à sua elevada expressão, múltiplos efeitos na oncogénese e virulência e a sua presença seletiva na membrana plasmática das células cancerosas, mas não em células normais. O AS1411 é um aptamero G4 direcionado à nucleolina, que apresenta efeito antitumoral em vários cancros, cuja aplicação se estende também à entrega específica de agentes terapêuticos e de diagnóstico a tumores.

Esta tese inova ao propor o desenvolvimento de transportadores de ligandos de G4 baseados em aptameros, como uma nova estratégia terapêutica para infeção por VPH e cancro do colo do útero. Inicialmente, vários ligandos específicos de G4 foram sintetizados, pertencentes a diferentes grupos funcionais aromáticos, como indol, fenantrolina e acridina. Depois avaliou-se a sua ligação a estruturas G4, encontradas no genoma das células cancerosas (promotor de oncogenes e G4s teloméricos) ou no genoma do VPH. O seu efeito inibitório de processos dependentes de G4 foi demonstrado pelo potencial para inibir as helicase Pif1 e a polimerase Taq, assim como para regular negativamente a expressão do oncogene c-MYC. Os ligandos demonstraram efeitos antitumorais promissores em linhas celulares do cancro do colo do útero, particularmente o ligando C₈, que também inibiu a replicação e encapsidação do VPH18 em culturas organotípicas, inibindo a proliferação do tecido infetado.

Tendo estabelecido a capacidade dos ligandos como agentes anticancerígenos/antivirais, iniciou-se o desenvolvimento de um sistema de entrega de ligandos baseado em aptameros. O aptamero AS1411 e seus derivados AT11, AT11-Lo e AT11-Bo foram usados para entregar seletivamente os ligandos de G4 a células cancerosas do colo do útero, utilizando uma estratégia simples de conjugação não covalente, mediada pelo reconhecimento da nucleolina na superfície das células. Essa estratégia foi eficiente na acumulação dos ligandos nas células do cancro do colo do útero, melhorando o seu efeito antitumoral e limitando sua toxicidade para as células normais. Uma vez que os aptameros usados também são sequências formadoras de G4, a presença dos ligandos melhorou as propriedades dos aptameros, tais como a penetração celular e resistência às nucleases séricas, sem afetar o reconhecimento de nucleolina.

Finalmente, foram desenvolvidos sistemas mais avançados, com o objetivo de melhorar a aplicabilidade clínica dos agentes anticancerígenos/antivirais. Nanoagregados funcionalizadas com AS1411 foram sintetizadas por forma a encapsular os ligandos G4, permitindo sua libertação gradual e direcionada nas células positivas para VPH em fases pré-cancerosas e cancerosas. Os nanoagregados apresentaram tamanhos pequenos (escala

nano), carga superficial negativa, estabilidade e perfil adequado de libertação dos ligandos. A presença de AS1411 na superfície do nanoagregados como agente de direcionamento tumoral, promoveu a acumulação nas células cancerosas do colo do útero, inibindo o seu crescimento, com pouco efeito nas células normais. Notavelmente, os nanoagregados também foram capazes de penetrar biópsias do cérvix uterino de pacientes com infecção por VPH em estádios pré-cancerosos, demonstrando o potencial deste sistema para tratar o VPH numa fase inicial da infecção, travando a oncogénese e progressão do cancro do colo do útero. A via de administração dos nanossistemas foi também abordada. Os nanoagregados foram formuladas como um gel semissólido para aplicação local no trato genital feminino, especificamente tecido com lesões pré-cancerosas e/ou infecção por VPH. Estudos preliminares de permeação *ex vivo* usando tecido vaginal porcino demonstraram a eficácia da formulação em penetrar e acumular os nanossistemas nas células alvo.

No geral, a investigação realizada ao longo desta tese de doutoramento definiu melhorias no desenho de compostos específicos para G4 com efeitos anticancerígenos/antivirais e levou à descoberta de possíveis sistemas de entrega baseados em aptameros para ligandos de G4. Os resultados obtidos forneceram um conjunto de conhecimentos sobre a aplicabilidade dos ácidos nucleicos G4 para o desenvolvimento de terapias específicas e eficazes contra a virulência e oncogénese do VPH. Além disso, a encorajadora avaliação pré-clínica realizada nas nanopartículas desenvolvidas demonstrou a necessidade de continuar a investigação, particularmente através da realização de estudos *in vivo* para verificar a aplicabilidade clínica desses sistemas. No futuro, uma estratégia semelhante pode ser alargada ao tratamento de outros tipos de cancro (por exemplo, cancros da orofaringe induzidos pelo HPV) e outras infeções virais, como HIV, vírus da hepatite e SARS-CoV-2.

Palavras-chave

Cancro do Colo do Útero;Vírus do Papiloma Humano;Sistema de Entrega de fármacos;Ácidos nucleicos ADN G-quadruplex;Compostos;Aptameros;Nucleolina;Nanoagregados;Formulações vaginais

Resumo Alargado

Mundialmente, o cancro do colo do útero é o terceiro tipo de cancro mais comum nas mulheres, com cerca de 500 mil novos casos diagnosticados anualmente e cerca de 300 mil mortes em 2018. Em Portugal especificamente são diagnosticados 750 casos anualmente e em 2018 foram registadas 340 mortes. O cancro do colo do útero pode ser classificado histologicamente em carcinoma pavimentoso, adenocarcinomas e carcinomas adenoescamosos, sendo que o primeiro é a forma mais comum de cancro do colo do útero. Praticamente todos os casos de cancro do colo do útero são responsáveis por uma infeção pelo Vírus do Papiloma Humano (VPH), uma das infeções sexualmente transmissíveis mais comuns do mundo, uma vez que maioria dos homens e mulheres sexualmente ativos a adquirem no decorrer das suas vidas. Os VPH dividem-se em dois tipos, baixo-risco e alto-risco. Os VPH de baixo-risco são responsáveis pelo desenvolvimento de lesões benignas como verrugas genitais, e são eliminados pelo sistema imune após 1 a 2 anos da infeção. No caso dos VPH de alto-risco, estes são responsáveis pelo desenvolvimento de lesões pré-cancerosas denominadas lesões pavimentosas intraepiteliais de baixo grau (LSIL), que apresentam displasia celular ligeira. Na maioria dos casos, estas lesões desaparecem por si com o tempo, mesmo sem tratamento. No entanto, por vezes podem evoluir para uma lesão de alto grau (HSIL) aumentando o risco de degenerar em cancro do colo do útero. As medidas preventivas para cancro do colo do útero têm passado primariamente e secundariamente pelo uso de vacinas profiláticas para os VPHs de alto risco e medidas de rastreio generalizado recorrendo a citologias e testes de deteção de VPH, respetivamente. Quando lesões pré-cancerosas são detetadas e é necessário tratamento, as opções mais comuns são a excisão e ablação do tecido afetado. No caso de as lesões terem evoluído para cancro, o tipo de tratamento depende do estadió da doença, mas as opções típicas são a cirurgia, radioterapia e quimioterapia. Apesar dos avanços no estabelecimento dos programas de vacinação e rastreio, a taxa de mortalidade feminina devido ao cancro do colo do útero ainda é elevada, particularmente nos países menos desenvolvidos onde os meios financeiros para estabelecer estes programas são escassos. Adicionalmente, as vacinas profiláticas não eliminam uma infeção pré-existente e não existe ainda um antiviral eficiente a eliminar o VPH. Os tratamentos atuais para o cancro do colo do útero são também ineficientes, têm efeitos secundários e o risco de recidiva é elevado.

O genoma do VPH está dividido em 3 regiões: uma região longa de controlo (RLC) que contém elementos *cis* necessários à replicação e transcrição do ADN viral; uma região

prematura composta por seis fases de leitura aberta que codificam as proteínas E1 - E7; e uma região tardia com duas fases de leitura aberta codificando as proteínas estruturais da cápside, L1 e L2. O potencial oncogénico dos VPHs de alto risco é atribuído à expressão das oncoproteínas E6 e E7, que inativam o gene supressor de tumores p53 e membros da família do retinoblastoma, levando ao desenvolvimento de carcinomas. Recentemente, foram encontradas várias regiões ricas em guaninas no genoma de dez tipos de VPH e foi demonstrado o potencial destas sequências formarem estrutura estáveis de G-quadruplex (G4). G4s são estruturas tetraédricas que diferem do emparelhamento canónico Watson-Crick (forma B do ADN) por serem formadas pelo emparelhamento de quatro guaninas via pontes de hidrogénio Hoogsteen em estruturas planares denominadas G-tétradas ou G-quartetos. Estas estruturas são estabilizadas por catiões monovalentes como sódio ou potássio e apresentam uma variedade de topologias mediante a estequiometria e orientação das cadeias de ADN que as compõem, assim como a composição de nucleótidos. Os G4 estão envolvidos na regulação de processos celulares chave como transcrição, tradução e replicação do ADN. No genoma dos VPH, as estruturas G4 estão presentes nas regiões RLC, E1, E4 e na sequência codificante da proteína L2, que estão envolvidas na transcrição, replicação e produção de proteínas virais. Assim, o desenvolvimento de moléculas capazes de reconhecer e estabilizar os G4 do genoma do VPH pode servir como uma estratégia antiviral dado o seu potencial de regular a expressão de genes virais. Esta estratégia foi já aplicada com sucesso em outras infeções virais como o herpes, ébola ou vírus da imunodeficiência humana (VIH) onde a utilização de compostos específicos de G4 (ligandos de G4) levou à inibição da replicação do ADN/ARN viral e expressão de proteínas virais, resultante num potente efeito antiviral. De igual forma, uma vez que os promotores de oncogenes e regiões teloméricas das células cancerosas estão também enriquecidas com estruturas G4, a mesma estratégia tem sido aplicada no desenvolvimento de agentes antitumorais.

A nucleolina é uma proteína multifuncional ubiquamente expressa nas células humanas estando presente no núcleo, no citoplasma e na membrana plasmática. A proteína aparece sobreexpressa na superfície das células cancerosas do colo do útero e foi demonstrado anteriormente estar envolvida na ativação da transcrição de oncogenes do VPH18 e também na manutenção da estabilidade do genoma do VPH16. A nucleolina também foi descrita por estar envolvida na carcinogénese, angiogénese e em vias de sinalização de vários cancros tais como o cancro do colo do útero. Devido à sua elevada expressão, múltiplos efeitos na oncogénese e virulência do VPH e sua presença seletiva na membrana plasmática das células cancerosas, mas não em células normais, a nucleolina é um alvo para o desenvolvimento de terapias direcionadas para o cancro do colo do útero e a

infecção por VPH. O AS1411 é um aptamero G4 direcionado à nucleolina que chegou a ensaios clínicos de Fase II para o tratamento de carcinoma renal metastático e leucemia mieloide aguda. Aptameros são pequenos oligonucleotídeos que reconhecem e ligam alvos celulares com elevada afinidade de forma análoga a anticorpos. Quando comparados a estes, os aptameros têm algumas vantagens tais como estabilidade, custo de produção, não produzem resposta imune do hospedeiro e têm uma captação celular aumentada. O aptamero AS1411, apesar de não ter sido aprovado para aplicação clínica devido a propriedades farmacológicas inadequadas (que podem resultar da sua estrutura polimórfica uma vez que foi demonstrado que o aptamero forma pelo menos oito estruturas de G4 distintas), apresenta um perfil de segurança ótimo, tem efeito antitumoral contra vários cânceros, e pode ser usado também para a entrega específica de agentes terapêuticos e de diagnóstico a tumores, dado o perfil seletivo de células cancerosas e o seu alvo ser a nucleolina. Recentemente foram também desenvolvidos alguns derivados do AS1411 tais como AT11, AT11-Lo e AT11-Bo que contêm alterações de nucleótidos favorecendo a formação de uma estrutura única de G4 que pode ser vantajoso para as suas propriedades farmacológicas.

Esta tese propõe o desenvolvimento de compostos capazes de ligar e estabilizar estruturas G4 presentes nos promotores de oncogenes e regiões teloméricas das células cancerosas do cancro do colo do útero, bem como estruturas G4 encontradas no genoma do VPH, com efeito duplo na carcinogénese e infecção, respetivamente. Adicionalmente, usando aptameros G4 direcionados à nucleolina, tais como o AS1411 e seus derivados, este projeto teve como objetivo desenvolver uma estratégia de entrega específica dos ligandos de G4 para a sua acumulação em células cancerosas do cancro do colo do útero positivas para VPH. Assim, este projeto gerou candidatos promissores a fármacos com efeito anticancerígeno e antiviral para tratar infecções persistentes e a carcinogenicidade do VPH, enquanto abre caminho para que outras doenças virais sejam abordadas usando uma estratégia similar. Este projeto também permitiu entender a função ainda desconhecida das estruturas de G4 no ciclo de vida do VPH e sua relação com o desenvolvimento do cancro do colo do útero. No final, a forma de administração destes potenciais fármacos, foi também alvo de estudo, com o desenvolvimento de formulações vaginais para a aplicação tópica e localizada dos compostos no tecido infetado e/ou em fases pré-cancerosas e cancerosas.

Inicialmente, foram sintetizados vários compostos específicos de G4 pertencentes a diferentes grupos funcionais aromáticos, como indol, fenantrolina e acridina. As estruturas destes compostos contêm as características desejáveis para moléculas

específicas de estruturas G4: uma base aromática larga, carga positiva e braços laterais funcionalizados. Depois avaliou-se capacidade de estes ligarem estruturas de G4 encontradas no genoma das células cancerosas como os promotores de oncogenes *c-MYC*, *KRAS*, *c-KIT*, *VEGF* e G4s teloméricos. Os compostos demonstraram ter elevada afinidade para as estruturas G4 (K_D na ordem dos micromolar) e capacidade de estabilizar termodinamicamente as estruturas (ΔT_m superior a 20 °C em alguns casos). O seu efeito inibitório de processos celulares dependentes de G4 tais como a replicação e transcrição foi demonstrado pelo seu potencial para inibir as helicase Pif1 e a polimerase Taq, assim como para regular negativamente a expressão de oncogenes como o *c-MYC*. Os compostos demonstraram também efeitos antitumorais promissores em linhas celulares do cancro do colo do útero. Os compostos mais promissores foram depois testados também com estruturas G4 do VPH encontradas nas regiões RLC, E1, E4 e na sequência codificante da proteína L2, que estão envolvidas na transcrição, replicação e produção de proteínas virais. Os compostos foram capazes de estabilizar com elevada afinidade esses G4s, com efeitos moduladores na topologia destas estruturas. A realçar, o derivado de acridina C₈, inibiu a replicação e encapsidação do VPH de alto-risco tipo 18 em culturas organotípicas infetadas. O decréscimo da carga viral foi acompanhado por um efeito inibitório na proliferação do tecido infetado, vincando o seu potencial antiviral.

Tendo-se identificado os compostos com potencial anticancerígeno e antiviral, iniciou-se o desenvolvimento de um sistema de entrega de compostos a células positivas para VPH baseado em aptameros. O aptamero AS1411 e seus derivados AT11, AT11-Lo e AT11-Bo foram usados para entregar seletivamente os ligandos de G4 derivados de acridina a células cancerosas do colo do útero, utilizando uma estratégia simples de conjugação não covalente, mediada pelo reconhecimento da nucleolina na superfície das células. Uma vez que estes aptameros são capazes de formar estruturas G4, a interação com os compostos por forma a transportá-los para as células alvo foi facilitada pelo mesmo tipo de interações que estes efetuam com os seus alvos terapêuticos. Esta interação melhorou as propriedades dos aptameros, tais como a penetração e acumulação celular e resistência a nucleases séricas, sem afetar o reconhecimento de nucleolina. A força da interação foi selecionada para ser suficiente para o complexo suportar o tráfego intracelular, mas não comprometer a dissociação do ligando na célula alvo por forma a manifestar o seu efeito antiproliferativo. Esta estratégia foi eficiente na acumulação dos ligandos nas células do cancro do colo do útero, melhorando o seu efeito antitumoral e limitando a sua toxicidade para as células normais.

Finalmente, foram desenvolvidos sistemas recorrendo à nanotecnologia, com o objetivo de melhorar a aplicabilidade clínica dos ligandos de G4 com propriedades anticancerígenas e

antivirais. Para isso, foram sintetizados aptameros AS1411 funcionalizados com um lípido octadecil, capazes de formar nanoagregados espontaneamente em solução contendo íons de magnésio. Os nanoagregados funcionalizados com AS1411 foram depois usadas para encapsular os ligandos de G4, por forma a permitir sua libertação gradual e direcionada em linhas celulares do cancro do colo do útero. Os nanoagregados desenvolvidas apresentaram tamanhos pequenos (cerca de 100 nanómetros), carga superficial negativa, índice de polidispersão adequado a aplicações farmacêuticas, elevada estabilidade e perfil adequado de libertação dos ligandos. A presença de AS1411 na superfície do nanoagregado como agente de direcionamento tumoral, promoveu a acumulação seletiva dos ligandos de G4 nas células cancerígenas do colo do útero, inibindo o seu crescimento, com pouco efeito nas células normais. Notavelmente, os nanoagregados também foram capazes de penetrar biópsias do cérvix uterino de pacientes com infeção por VPH em estadios pré-cancerosos, demonstrando o potencial deste sistema para tratar o VPH numa fase inicial da infeção, travando a oncogénese e progressão do cancro do colo do útero. A utilização de biópsias de pacientes em estadios pré-cancerosos foi também importante pois permitiu avaliar a expressão já aumentada de nucleolina na superfície destas células, mesmo antes de desenvolverem carcinoma. Por fim, estudou-se uma possível via de administração dos nanossistemas através do desenvolvimento de um gel contendo uma dispersão de nanoagregados carregados de ligandos de G4, para aplicação local no trato genital feminino com lesões pré-cancerosas e/ou cancro do colo do útero por infeção de VPH. Estudos preliminares de permeação *ex vivo* utilizando tecido vaginal porcino demonstraram a eficácia da formulação em penetrar e acumular os nanossistemas nas células alvo, mostrando o potencial de aplicação desta estratégia no combate às doenças provadas pelo VPH.

Estes resultados revelaram o potencial de compostos baseados nos ácidos nucleicos G4 para futura aplicação clínica ao nível da eliminação de infeções persistentes de VPH, do tratamento do cancro do colo do útero, como possivelmente de outras patologias que beneficiem da entrega seletiva de compostos direcionados a estruturas G4.

Abstract

Annually, more than 500,000 new cervical cancer cases are diagnosed worldwide with a 60% mortality rate. In Portugal, about 750 new cases are annually diagnosed which in 2018 resulted in 340 deaths. Human papillomaviruses (HPVs) are the principal cause for the development of cervical cancer. High-risk HPV types 16 and 18 are responsible for 70% of all cases. Persistent HPV infection leads to the hyperproliferation of the infected cells by downregulating epithelial differentiation, epidermal development, and innate immune responses. The oncogenic potential of high-risk HPV types is attributed to the expression of oncoproteins, E6 and E7, which inactivate tumour suppressor p53 and members of retinoblastoma family, leading to carcinoma development. Despite recent progresses in screening and HPV vaccination that reduced cervical cancer incidence, treatment options for women with persistent HPV infection or in its latent form are still missing. Furthermore, the available therapeutic options for cervical cancer have limited effectiveness and specificity, causing great pain to patients, and do not benefit from the research addressing the host-HPV interaction patterns during malignant progression.

The genome of HPV is divided into 3 regions: a long control region (LCR) that contains cis-elements necessary for the replication and transcription of viral DNA; the early region, composed of six open reading frames (ORFs) labelled E1–E7, and a late region with two ORFs coding for capsid structural proteins L1 and L2. Recently, several guanine-rich regions have been found in the genome of ten HPV types (both high- and low-risk), and the potential of the sequences to fold into stable G-quadruplex (G4) structures was demonstrated. These guanine-rich regions are present in LCR, E1, E4 regions and L2 protein coding sequence, which are involved in transcription, replication, and viral proteins production. Hence, targeting these G-rich regions of the HPV genome could serve as a potential antiviral strategy, using small ligands that can stabilize the viral G4 regions and regulate gene expression.

Nucleolin is overexpressed in the surface of cervical cancer cells and was shown to be involved in the activation of HPV18 oncogene transcription in cervical cancer and to be related with HPV16 genome stability maintenance. Nucleolin is present in the nucleus, cytoplasm, and cell surface. The roles of cell-surface nucleolin in tumorigenesis, angiogenesis, and cancer signalling pathways have also been described. Therefore, nucleolin is an attractive target for the development of targeted therapies for both cervical cancer and HPV infection given its high abundance, its multifaceted influence on

oncogenesis and virulence of HPV, and its selective presence on the plasma membrane of cancer cells, but not on untransformed, normal cells. AS1411 is a nucleolin-targeted G4 aptamer with antitumoral effect in many cancers that has also been extensively used for the tumour-specific delivery of both therapeutic and imaging agents.

This thesis innovates by proposing the development of aptamer-based carriers of antiviral/anticancer G4 ligands as a new therapeutic strategy for HPV infection and cervical cancer. Firstly, several G4-specific ligands belonging to different aromatic scaffolds such as indole, phenanthroline and acridine, were synthesized and characterized for their ability to bind G4 structures, either found in the cancer cell genome (oncogene promoter and telomeric G4s) or the HPV genome. Their effect in inhibiting key G4-dependent processes was demonstrated by their potential to inhibit Pif1 helicase and Taq polymerase enzymes, and to downregulate *c-MYC* oncogene expression. The ligands demonstrated interesting antitumoral effects in cervical cancer cell lines, particularly ligand C₈, which was also found to inhibit HPV18 replication and encapsidation in organotypic raft cultures, with resulting antiproliferative effects in infected tissue.

Having established the ability of the ligands as anticancer/antiviral agents, the development of an aptamer-based delivery system was envisioned. AS1411 aptamer and its derivatives AT11, AT11-Lo and AT11-Bo were used to selectively convey the G4 ligands to HPV-positive cervical cancer cells, with a simple non-covalent conjugation strategy and mediated by cell-surface nucleolin recognition. This strategy proved to be efficient at accumulating the ligands in cervical cancer cells, improving their antitumoral effect while limiting their off-target toxicity towards normal cells. Interestingly, because the used aptamers are also G4-forming sequences, the presence of the G4 ligands improved the aptamers' properties such as cellular penetration and serum nuclease resistance, without affecting nucleolin recognition.

Finally, advanced and intricate drug delivery systems were developed aiming at enhancing the clinical applicability of the anticancer/antiviral agents. AS1411-decorated nanoaggregates were synthesized to encapsulate the G4 ligands, allowing their sustained and targeted release within HPV-positive cells in precancerous and cancerous stages. The nanoaggregates presented small sizes (nanoscale), negative surface charge, stability, and a suitable drug release profile. The presence of AS1411 on the nanoaggregates surface as a cancer-specific guiding agent promoted the specific accumulation of the G4 ligands in cervical cancer cells which inhibited their growth, with little effect on non-malignant normal cells. The nanoaggregates were also able to penetrate cervix tissue biopsies of

patients with HPV infection in precancerous stages, demonstrating the potential of this system to treat HPV infection in its latent stage, preventing oncogenesis and cervical cancer progression. The administration route of the drug-loaded nanosystems was also addressed. The C₈-loaded nanoaggregates were formulated as a gel for local application in the female genital tract, specifically tissue with precancerous lesions and/or HPV infection. Preliminary *ex vivo* permeation studies using porcine vaginal tissue demonstrated the efficacy of the formulation in penetrating and accumulating the nanosystems in the target cells.

Overall, the research performed throughout this doctoral thesis defined improvements in the design of G4-specific compounds with dual anticancer/antiviral effects and led to the discovery of potential aptamer-based delivery systems for G4 ligands. The original results obtained during the past years provide important knowledge in the applicability of G4 nucleic acids for the development of specific and effective therapeutics against HPV virulence and oncogenesis. Furthermore, the encouraging pre-clinical evaluation performed on the newly developed nanomedicines demonstrated the need of continuing the research initiated by this project, particularly by pursuing the *in vivo* studies to verify the clinical applicability of these systems. In the future, a similar strategy could be extended to the treatment of other types of cancers (for instance HPV-induced oropharyngeal cancers) and other viral infections such as HIV, hepatitis virus and SARS-CoV-2.

Keywords

Cervical Cancer; Human Papillomavirus; Drug Delivery Systems; DNA G-quadruplex Nucleic Acids; Compounds; Aptamers; Nucleolin; Nanoaggregates; Vaginal Formulations

Table of Contents

Chapter 1.....	1
Cervical Cancer and Human Papillomavirus	3
Cervical Cancer	3
Human Papillomavirus	6
Treatment of HPV-induced disease	9
Novel therapeutic prospects	12
References.....	15
Chapter 2.....	27
G-quadruplex, Friend or Foe: The Role of the G-quartet in Anticancer Strategies	27
Abstract.....	29
Role of the G-quadruplex in anticancer therapeutics.....	29
The G-quartet	30
Regulatory roles of G4s	31
G4s as therapeutic targets	33
Telomeric G4 structures	34
Oncogene promoter G4 structures	35
Clinical applicability of G4 ligands.....	37
Perspectives for the development of G4 ligands as antitumor agents	38
G4s as therapeutic agents.....	40
G4 aptamers as anticancer agents.....	41
G4 aptamers as drug delivery systems.....	43
Concluding Remarks	44
Acknowledgments.....	44
References.....	45
Chapter 3.....	57
Aims and Outline of the Thesis	57
Chapter 4	63
Study of the interaction between indole-based compounds and biologically relevant G-quadruplexes	63
Abstract.....	65
Introduction.....	65

Materials and methods	66
Oligonucleotides	66
Ligand Synthesis.....	67
Circular Dichroism	68
Fluorescence Resonance Energy Transfer Melting Assay	69
Steady-State Fluorescence Measurements	69
NMR Spectroscopy	70
Molecular Docking	70
Results.....	71
Design and synthesis of indole-based ligands.....	71
Ligands binding and stabilization of G-quadruplexes	72
Characterization of ligand binding to G-quadruplexes.....	74
NMR studies of c-MYC/ligand interaction	75
Molecular docking simulations	78
Discussion	79
Conclusions	81
Acknowledgments.....	82
References	82
Chapter 5.....	89
Phenanthroline Polyazamacrocycles as G-quadruplex DNA Binders.....	89
Abstract	91
Introduction	91
Experimental Section.....	93
Oligonucleotides and ligands.....	93
Fluorescence Resonance Energy Transfer melting assays	94
Circular Dichroism studies.....	94
Fluorometric titrations	95
TO displacement assay	96
Job plot analysis by conductivity	96
Fluorescence-based helicase assay.....	96
PCR-Stop Assay	97
Cytotoxicity Assay.....	98
Results and discussion.....	98
Phenanthroline-based ligands as potential G4 binders	98

FRET melting studies	99
Circular dichroism experiments	101
Fluorimetric titration studies for binding analysis	103
Fluorescent intercalator displacement assay	105
Inhibition of helicase and polymerase activities	105
PCR stop assay	106
Cytotoxicity assay	107
Conclusions	108
Acknowledgments	109
References	109
Chapter 6	115
Fluorescent Light-Up Acridine Orange Derivatives Bind and Stabilize	
KRAS-22RT G-Quadruplex	115
Abstract	117
Introduction	117
Materials and methods	119
Oligonucleotides and acridine orange derivatives	119
Circular Dichroism Assays	120
Fluorescence Binding Studies	121
Nuclear Magnetic Resonance Spectroscopy	122
Cell Viability Assay	122
Confocal Fluorescence Microscopy	122
Results	123
Acridine orange ligands design	123
Binding and stabilization of KRAS-22RT G-quadruplex	124
Fluorescence binding studies with KRAS-22RT G4	126
NMR studies with KRAS-22RT G4	129
Cell studies: cytotoxicity and intracellular localization	131
Conclusions	132
Acknowledgments	133
References	134
Chapter 7	139
Human Papillomavirus G-Rich Regions as Potential Antiviral Drug Targets	
.....	139

Abstract	141
Introduction	141
Material and methods	144
Oligonucleotides and compounds	144
Fluorescence resonance energy transfer melting	145
Circular dichroism	145
Fluorescent Intercalator Displacement Assay	146
Nuclear Magnetic Resonance	146
Keratinocyte and organotypic raft cultures	147
Histology of Tissue Sections	147
Virus Harvest and Isolation	147
Viral Titers	147
Results and discussion	148
Conclusion	160
Acknowledgments	160
References	160
Chapter 8	167
Aptamer-Based Targeted Delivery of a G-Quadruplex Ligand in Cervical Cancer Cells	167
Abstract	169
Introduction	169
Materials and Methods	171
Oligonucleotides and compounds	171
Thermal Difference Spectra	171
Circular dichroism	171
Fluorescence studies	172
Cell Viability Assay	172
Cell membrane damage	173
Flow cytometry analysis	173
RNA isolation and RT-qPCR for gene expression analysis	174
Fluorescence resonance energy transfer melting	174
Confocal fluorescence microscopy imaging	175
Statistical analysis	175
Results and Discussion	175

Conclusion	186
Acknowledgments.....	187
References.....	187
Chapter 9	193
Aptamer-guided Acridine Derivatives for Cervical Cancer	193
Abstract.....	195
Introduction.....	195
Material and Methods	196
Oligonucleotides and ligands	196
Circular dichroism spectroscopy	197
Fluorescence Titrations.....	197
NMR spectroscopy.....	198
Serum stability assay	199
Cytotoxicity Assay	199
Confocal fluorescence Microscopy	199
Results and Discussion.....	200
Conclusion	210
Acknowledgments.....	211
References.....	211
Chapter 10.....	215
AS1411 Derivatives as Carriers of G-Quadruplex Ligands for Cervical Cancer Cells	215
Abstract.....	217
Introduction.....	217
Materials and Methods.....	219
Acridine Orange Derivatives and Oligonucleotides	219
Thermal Difference Spectra	219
Circular Dichroism Spectroscopy	219
Fluorescence Binding Studies	220
Serum Stability Assay	221
Cell Viability Assay	221
Confocal Fluorescence Microscopy	222
Flow cytometry.....	222
G4 model structures	223

Molecular Dynamics	223
Results and Discussion	224
Conclusions	236
Acknowledgments	236
Chapter 11	241
Nanoaggregate-Forming Lipid-Conjugated AS1411 Aptamer as a Tumor-targeted Delivery System of Anticancer Agents for Cervical Cancer Therapy	241
Abstract	243
Introduction	243
Materials and Methods	245
DNA synthesis and purification	245
Nanoaggregates formation	246
Agarose gel electrophoresis	246
Characterization of nanoaggregates	246
Circular dichroism	247
Nuclear Magnetic Resonance	247
Critical micellar concentration (CMC) determination	247
Drug release	247
Nanoaggregate cellular uptake	248
Nanoaggregate tissue penetration	248
Flow cytometry	249
<i>In vitro</i> antitumoral effect	249
Gel formulation preparation	249
<i>Ex vivo</i> vaginal permeation assay	250
Chromatographic method development and validation	250
Analysis of tissue permeation/penetration	250
Statistical analysis	251
Results and Discussion	251
Conclusions	261
Acknowledgments	261
References	262
Chapter 12	269
General Conclusions and Future Perspectives	269

General Conclusions	271
Future Perspectives	274

List of Figures

Figure 1.1 – Age-standardised incidence rates of cervical cancer in the World (estimates for 2018).....	3
Figure 1.2 – Pathogenesis of HPV infection in cervical cancer. The schematic drawing of the HPV genome organization before and after genome integration is shown at the bottom. The open reading frames encode the six early proteins E1, E2, E4, E5, E6, and E7 and two late structural proteins L1 and L2.....	7
Figure 1.3 – HPV life cycle organization. The HPV life cycle is dependent on the terminal differentiation of the keratinocytes and the expression of both early and late protein.....	8
Figure 1.4 – Conventional and emerging modalities for treatment of both benign and malignant cervical lesions. EBRT - External beam radiation, LEEP - loop electrosurgical excision procedure, TCA – trichloroacetic acid.....	10
Figure 1.5 – HPV genome organization and the positions where G4-forming sequences were found in the respective HPVs. On the right, the NMR solution structure of a G4 found in the HPV52 genome (PDB: 5O4D). Nucleotide residues are shown in green while the DNA backbone is depicted in orange.....	12
Figure 2.1 - G4 structure. (A) Chemical structure of the G-quartet showing the Hoogsteen hydrogen-bonding between guanine pairs and a central monovalent cation that stabilizes the 3D structure. (B) Side and top view of a model G4 structure found for the human telomeric sequence. The backbone is depicted in green, while guanine bases are represented in blue (loop bases were hidden for clarity). (C), (D) and (E) are representative topologies commonly found in G4 structures.....	31
Figure 2.2 - Biological functions of G4 nucleic acids. (A) In the absence of any specialized helicase, G4s cause the arrest of DNA polymerase (DNAPol) potentially leading to DNA single- or double-strand breaks causing genetic instability. (B) G4s on the 3'-overhang of chromosome ends interfere with telomere elongation by telomerase and affect telomere maintenance. (C) Promoter G4s interfere with transcription factor (TF) binding and RNA polymerase (RNAPol) activity, thus altering gene expression at transcriptional level. The	

complementary strand may form other non-canonical nucleic acid structures, known as i-motifs. (D) Gene expression is also regulated by G4 formation in mRNA by inhibiting ribosome scanning or altering splicing, and mRNA maturation and localization.....32

Figure 2.3 - General workflow of SELEX protocol. During SELEX, a single-strand (ss)DNA/ssRNA library (typically a mix of 10^{15} sequences) is incubated with the target protein (1). The unbound sequences are partitioned from the bound sequences (2). At this point the bound sequences may undergo a negative selection procedure by incubating them with negative protein targets (3). The sequences that bound the positive (true) target but failed to bind the negative control target are then amplified by PCR or RT-PCR (if its RNA aptamers) to enrich the pool (4). The enriched pool is finally sequenced and subjected to the next SELEX round.....41

Figure 4.1 - Structures of the synthesized compounds 1-3. Ligands contain an indole aromatic core with polyamine side chains of different lengths.....71

Figure 4.2 - CD melting curves of (A) c-MYC (5 μ M in 30 mM potassium phosphate buffer pH 7.1 and 100 mM LiCl), (B) hybrid-1 22AG (10 μ M in 30 mM potassium phosphate buffer pH 7.1 and 10 mM KCl and 90 mM LiCl) and (C) antiparallel 22AG (10 μ M in 30 mM sodium phosphate buffer pH 7.1 and 100 mM NaCl) in the absence and presence of ligands 1-3.....72

Figure 4.3 - Fluorescence emission spectra of compound 3 with successive additions of (A) c-MYC and (B) hybrid-1 22AG. Insets: Stern-Volmer plots obtained by plotting F_0/F as a function of DNA concentration and fitting to the adequate equation.....74

Figure 4.4 - Imino and aromatic regions of c-MYC ^1H spectra upon titration of compound 3 at 298K. Top tetrad guanines are colored in blue, middle in red and bottom tetrad in green. On the right a schematic representation of c-MYC G4 structure – Reprinted (adapted) with permission from “Ambrus, A.; Chen, D.; Dai, J.; Jones, R. A.; Yang, D. Biochemistry 2005, 44 (6), 2048–2058”. Copyright 2005 American Chemical Society.....75

Figure 4.5 - Mapping of the chemical shift perturbation on the NMR model of c-MYC (PDB ID:1XAV) with residues colored based on the higher $\Delta\delta$ relative to the average $\Delta\delta$ (average $\Delta\delta=0.07$) where: yellow is for $\leq 25\%$, orange for 25-75% and red for $\geq 75\%$ 76

Figure 4.6 - (A) The expanded H8/H6-H1' region of the 2D-NOESY spectrum of the ligand:DNA 2:1 complex at 310 K. The sequential assignment pathway is represented. (B) Expanded region of the 2D-NOESY spectrum showing the intramolecular NOEs within the aromatic core of compound 3 (bottom) and some intermolecular NOEs with the DNA (top). The cross-peaks of compound 3 are broader when compared to those from DNA. Each cross-peak is labeled as F2/F1 dimensions.....77

Figure 4.7 - Best models of compound 3/c-MYC G4 complex produced by docking. All the binding poses show the ligand interacting with G11 through the indole aromatic moiety, with additional hydrogen bonding and electrostatic interactions with the loops. G4 structure is represented in green/purple ribbon, ligand as orange sticks and pink lines indicate H-bonds.....78

Figure 5.1 - Structures of phenanthroline-based macrocycles.....93

Figure 5.2 - Normalized CD melting curves of (A) c-MYC G4 (100 mM LiCl), (B) hybrid-type 22AG G4 (10 mM KCl and 90 mM LiCl) and (C) antiparallel 22AG G4 (100 mM NaCl) DNAs in the absence and presence of 5 equivalents of ligands.....102

Figure 5.3 - Fluorescence titrations of the ligand [32]phen₂N₄ with increasing concentration of (A) c-MYC and (B) 22AG G4s [5 μM in 100 mM KCl, pH 7.2, concentrations 0–25 μM]. Insets: Stern-Volmer plots obtained by plotting F₀/F as a function of DNA concentration.....103

Figure 5.4 - Helicase activity against the (A) S-mut, (B) S- cmyc DNA and (C) S-htelo system in buffer containing 1 mM KCl at 20 nM Pif1. Error bars indicate mean deviation between three independent experiments.....106

Figure 5.5 - Polyacrylamide gel electrophoresis of (A) Pu77myc G4 and (B) Pu77-mut in the presence of increased concentrations of the ligands (0 μM, 0.8 μM; 1.6 μM; 3.2 μM and 6.4 μM).....107

Figure 5.6 - Relative cell viability of HeLa (left) and NHDF (right) cells measured by MTT assay after 48 h incubation with the ligands at concentrations ranging from 0.01 to 100 μM. Bars represent mean ± SEM of cell viability relatively to control cells (no treatment).....108

Figure 6.1 - Schematic diagram showing the structures of C₃, C₅ and C₈ and how the ligands stabilize KRAS G4 structure hence inhibiting the expression of KRAS protein. The

acridine moiety interact with the G-tetrads surface through end stacking interactions (red). The alkylamide substituents interact through groove binding and a positively charged nitrogen in the acridine ring is able to occupy a position similar to that of the G-tetrad central ion (blue). The iodobenzene is expected to enhance the cellular uptake while providing additional interactions with the G4 grooves/loop interface (green). Interaction of the acridine orange derivatives with the G4 structure is expected to maintain the "gene-off" state induced by the conversion of the double stranded "gene-on" form to a G4 "gene-off" form (model for the regulation of KRAS gene transcription on the right).....119

Figure 6.2 - CD melting curves of KRAS-22RT G4 at 5 μ M in the absence and presence of 2 molar equivalents of ligands, in 10 mM K-phosphate buffer and 50 mM KCl. Data points were recorded at 263 nm. The Boltzmann curve fitting of the data points is shown. Temperature corresponds to the temperature set by the Peltier system.....124

Figure 6.3 - Fluorescence emission spectra of (A) C₃, (B) C₅ and (C) C₈ with increasing concentration of (I) KRAS-22RT G4 (0.5 – 10 μ M) and (II) DNA ds26 (0.5 – 20 μ M) in 10 mM K-phosphate buffer and 50 mM KCl. Insets: fraction of ligand bound plots fitted to the saturation binding equation.....127

Figure 6.4 - Imino region of KRAS22-RT G4 ¹H spectra upon titration with compound C₈ at 310 K. The assignment of free KRAS22-RT is shown: top 5' G-quartet protons are labelled in green, middle tetrad in blue and bottom 3' tetrad in red.....129

Figure 6.5 - Expanded region of the 2D-NOESY spectrum showing the intermolecular NOEs between the G2, G6 and G11 H1 protons (imino) and ligand C₈, namely the alkylamide chain aliphatic protons of the ligand (inset). KRAS-22RT G4 at 350 μ M in presence of 4 molar equivalents of C₈.....131

Figure 6.6 - Confocal microscopy images of HeLa cells treated for 1 h with 10 μ M C₈ (middle). Nuclear probe Hoechst 33342 (left) was used for co-localization (merge, right). Scale bar is 20 μ m. C₃ and C₅ shown similar results.....132

Figure 7.1 - Structure of the G4 ligands used in this study. Ligands belong to different starting moieties, particularly phenanthrolines PhenDC3 and [16]phenN₂, acridines BRACO-19 and C₈, pyridines Pyridostatin (PDS) and 360A, and porphyrin TMPyP4.....144

Figure 7.2 - CD titrations of (A) HPV16 and (B) HPV18 G4s with PhenDC3. Upon ligand binding, PhenDC3 seems to convert the multiple, noncanonical structures of HPV16 into an antiparallel G4 while in the case of HPV18, the same overall parallel G4 topology is maintained. The buffer used was 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl.....150

Figure 7.3 - FID results obtained for HPV G4 sequences in the presence of 5 equivalents of ligands. The heat map representation allows the visual comparison of the TO displacement percentages by using the colour code guide provided below. The buffer used was 10 mM K-phosphate buffer, pH 7.2, with 50 mM KCl.....152

Figure 7.4 - ¹H NMR titrations of three different HPV G4s with increasing amounts of ligands. (A) HPV16 G4, (B) HPV57 G4 and (C) HPV58 G4 with increasing ligand:DNA ratios (L:G4) of PhenDC3, C₈ and PDS, respectively. Arrows indicate the set of imino peaks corresponding to the major G4 conformation. The buffer used was 10 mM K-phosphate buffer, pH 7.2, with 50 mM KCl.....154

Figure 7.5 - Effects of ligand PhenDC3 on HCK16-5 and HCK18C organotypic raft cultures, which contain episomal high-risk HPV16 and HPV18 respectively. HCK16-5 rafts were either treated with (A) DMSO solvent control or (B) 0.5 μM PhenDC3. Similarly, HCK18C rafts were treated with (C) DMSO solvent control or (D) 0.5 μM PhenDC3. Treatments were made every other day with feeding for 20 days. Tissue sections were cut and stained with hematoxylin and eosin (H&E).....156

Figure 7.6 - Quantitative RT-PCR infectivity assay of HPV18 organotypic rafts incubated with PhenDC3. (A) total viral genomes and (B) viral titer. The total number of viral genomes were quantified amplifying the E2 open reading frame and measured against a standard curve. Virus titers were measured by qPCR, in which homogenized samples were treated with benzonase to remove any non-protected HPV genomes. Data represent mean ± SD.....157

Figure 7.7 - Effects of compound C₈ on HCK16-5 and HCK18C organotypic raft cultures. HCK16-5 rafts were either treated with (A) DMSO solvent control or (B) 0.1 μM C₈. Similarly, HCK18C rafts were treated with (C) DMSO solvent control or (D) 0.1 μM C₈. Treatments were made every other day with feeding for 20 days. Tissue sections were cut and stained with hematoxylin and eosin (H&E). Of note is the raft thinning upon incubation with C₈ in the case of HPV18 infected cultures.....159

Figure 7.8 - Quantitative RT-PCR infectivity assay of HPV18 organotypic rafts incubated with C₈. (A) total viral genomes and (B) viral titer. The total number of viral genomes were quantified amplifying the E2 open reading frame and measured against a standard curve. Virus titers were measured by qPCR, in which homogenized samples were treated with benzonase to remove any non-protected HPV genomes. No tissue was recovered at 0.5 μ M compound concentration. Data represent mean $\pm$ SD; ***p* < 0.01.....159

Figure 8.1 - Schematic representation of proposed hypothesis of using AS1411 and its derivatives to carry ligand C₈ into cancer cells overexpressing nucleolin.....176

Figure 8.2 - CD spectra (A) and thermal difference spectra (B) of AS1411 aptamer and its derivatives LNA-AS1411 and U-AS1411. Spectra acquisition was performed in 30 mM potassium phosphate buffer containing 100 mM KCl.....178

Figure 8.3 - CD spectra of (A) AS1411, (B) LNA-AS1411 and (C) U-AS1411 aptamers in the presence of increasing amounts of ligand C₈. Spectra acquisition was performed in 30 mM potassium phosphate buffer containing 100 mM KCl.....178

Figure 8.4 - Fluorescence spectroscopy titrations of ligand C₈ with increasing amounts of (A) AS1411, (B) LNA-AS1411 and (C) U-AS1411 DNA. Spectra acquisition was performed in 30 mM of potassium phosphate buffer containing 100 mM KCl, with DNA concentration ranging from 0.1 μ M to 12.5 μ M. Insets: saturation binding plots fitted to Hill model.....179

Figure 8.5 - Relative cell viability of HeLa and NHDF cells incubated for 7 days with free C₈ or the pre-formed aptamer-ligand complex at a C₈ concentration of 1 μ M using aptamer (A) AS1411, (B) LNA-AS1411 and (C) U-AS1411. **p* < 0.05, ***p* < 0.01.....180

Figure 8.6 - Relative LDH activity of HeLa, PC-3 and A549 cells incubated for 7 days with the pre-formed AS1411-ligand complex at a C₈ concentration of 1 μ M.....182

Figure 8.7 - Flow cytometry analysis of free C₈ (blue) and AS1411-C₈ complex (red) in (A) HeLa cancer cells and (B) NHDF non-malignant cells and respective histograms (C and D). Cells were incubated for 24 h with the preformed aptamer-ligand complex at a C₈ concentration of 1 μ M.....182

Figure 8.8 - Saturation binding plots fitted to Michaelis-Menten model. Fluorescence intensity of Cy-AS1411 was recorded after the addition of increasing amounts of nucleolin in the (A) absence and (B) presence of ligand C₈.....183

Figure 8.9 - Effect of AS1411-C₈ complex treatment on c-MYC and nucleolin gene expression measured by RT-qPCR. Data represent the average of three experiments $\pm$ SEM; * $p < 0.05$, ** $p < 0.01$ 184

Figure 8.10 - Confocal microscopy images of HeLa cells incubated with Cy5-AS1411 and Cy5-AS1411-C₈ complex for 7 days. Cell nuclei are stained with Hoechst 33342 (blue), C₈ emits green fluorescence and Cy5-AS1411 is shown in red. Overlapping of the two stains, observed as yellow regions, can be seen in the merged images of the cells. Control cells and C₈ single stain control images are shown on the right. Brightness was adjusted to facilitate analysis.....185

Figure 8.11 - Confocal microscopy images of NHDF cells incubated with Cy5-AS1411 and Cy5-AS1411-C₈ complex for 6 days followed by 1 day incubation in fresh media. Cell nuclei are stained with Hoechst 33342 (blue), C₈ emits green fluorescence while Cy5-AS1411 fluorescence is absent. Control cells and C₈ single stain control images are shown on the right. Brightness was adjusted to facilitate analysis.....186

Figure 9.1 - Chemical structures of the acridine ligands C₃, C₅ and C₈.....200

Figure 9.2 - CD spectra of AT11-Lo in the presence of increasing amounts of ligands (A) C₃ (B) C₅ and (C) C₈. Spectra acquisition was performed in 20 mM potassium phosphate buffer containing 65 mM KCl.....201

Figure 9.3 - Fluorescence titration spectra of ligands (A) C₃, (B) C₅ and (C) C₈ with AT11-Lo in 20 mM potassium phosphate buffer containing 65 mM KCl, and relative enhancement of the fluorescence intensities of ligands vs. the increasing concentration of AT11-Lo G4; λ_{ex} : 498 nm.....202

Figure 9.4 - A) Sequences of the AT11 and AT11-Lo oligonucleotides. (B) Schematic representation of AT11-Lo G4 showing the G-tetrad alignments and folding topology. Red arrows denote thymine loops while the green arrow highlights the linker between the two subunits where a thymine residue was deleted to derive AT11-Lo from AT11. Adapted with permission from reference 7. C) Imino H1 and aromatic H8/H6 region of the ¹H NMR spectra of AT11-Lo, including the spectra obtained upon titration with ligand C₈ at

increasing ligand:DNA ratios (R). Labelled residues were indicated with their corresponding residue number. Top tetrad residues are shown in purple, middle tetrads in red and blue, and bottom tetrad in green. C₈ protons from the acridine moiety and NH group are shown on the top spectrum as broad signals. These were assigned based on NOESY and TOCSY spectra. (D) Chemical structure of ligand C₈ with respective protons indicated by letters.....203

Figure 9.5 - Expanded region of the 2D NOESY spectrum showing the intermolecular NOEs (highlighted with red arrows) between the AT11-Lo imino protons (not assigned due to spectral overlap) and ligand C₈, namely the alkylamide chain aliphatic protons i, h and f of the ligand, and the iodobenzene k protons (inset) 205

Figure 9.6 - Serum stability of AT11-Lo G4 and AT11-Lo G4-C₈ in RPMI medium containing (i) 10 % FBS and (ii) 50% human serum and resolved in a 1% agarose gel electrophoresis. In ii) the large smearing seen on the top of each lane corresponds to the serum proteins aggregation. On the right, control lane (C⁺) of AT11-Lo G4-C₈ complex in the absence of human serum.....206

Figure 9.7 - Percentage of viable HeLa and NHDF cells after incubation with free AT11-Lo aptamer (15 µM) or conjugated with (A) C₃ at 2 µM, (B) C₅ at 2 µM and (C) C₈ at 1 µM for 7 days. ***p<0.001.....207

Figure 9.8 - Saturation binding plots obtained by fluorescence spectroscopy and fitted to Michaelis-Menten model. Fluorescence intensity of (A) free AT11-Lo G4 labelled with Cy5 or complexed with 2 eq of (B) C₃, (C) C₅ and (D) C₈ obtained after addition of increasing amounts of nucleolin.....208

Figure 9.9 - Confocal fluorescence images of (A) HeLa and (B) NHDF cells incubated with Cy5-AT11-Lo-C₈ after 2h at 37°C. Nucleolin is stained with AlexaFluor 546®, C₈ emits green fluorescence and Cy5- AT11-Lo is shown in red. Scale bar: 25 µm. Brightness was adjusted for all images equally.....209

Figure 9.10 - Confocal microscopy images of (A) HeLa cells and (B) NHDF cells incubated with Cy5-AT11-Lo-C₈ complex for 5 days. Cell nuclei are stained with Hoechst 33342 (blue), C₈ emits green fluorescence and Cy5-AT11-Lo is shown in red. Overlapping of the two stains, observed as yellow/orange regions, can be seen in the merged images of the cells. Scale bar: 25 µm.....210

Figure 10.1 - (A) Sequences of AS1411 and its derivatives AT11 and AT11-Bo (modifications shown in red). (B) Chemical structure of the acridine orange ligands used in this study.....	224
Figure 10.2 - TDS spectra obtained by subtracting the spectrum at 20 °C from the spectrum at 90 °C for (A) AT11 and (B) AT11-Bo. Spectra acquired in 20 mM potassium phosphate buffer containing 65 mM KCl, pH 7.....	225
Figure 10.3 - CD spectra of (A) AT11 G4 and (B) AT11-Bo G4 in the absence and presence of 4 molar eq. of each of the acridine orange ligand. Spectra acquired in 20 mM potassium phosphate buffer containing 65 mM KCl, pH 7.....	225
Figure 10.4 - CD melting curves of (A) AT11 G4 and (B) AT11-Bo G4 in the absence and presence of 4 molar eq. of each of the acridine orange ligands. Data points were recorded at 260 nm. The Boltzmann curve fitting of the data points is shown.....	226
Figure 10.5 - Fluorescence emission spectra of (A, D) C ₃ , (B, E) C ₅ and (C, F) C ₈ with increasing concentration of AT11 G4 (top) or AT11-Bo G4 (bottom). Insets: fraction of ligand bound plots fitted to the Hill saturation binding model. Spectra acquired in 20 mM potassium phosphate buffer containing 65 mM KCl, pH 7.....	227
Figure 10.6 - Serum stability of AT11 G4, AT11 G4-C ₈ , AT11-Bo G4 and AT11-Bo G4-C ₈ in 10% FBS at 37 °C for 0, 7, 24 and 48 h (A, B, C and D) and 50% FBS at 37 °C for 0, 7, 24, 48 and 72 h (E, F, G and H). The samples were resolved in a 1% agarose gel, stained with 0.01% Xpert Green DNA and visualized under UV light.....	229
Figure 10.7 - Saturation binding plots obtained by fluorescence titrations with nucleolin and fitted to Michaelis-Menten model for free Cy5-AT11 G4 and Cy5-AT11-Bo G4 (A and B, respectively), or for Cy5-AT11 and Cy5-AT11-Bo in the presence of ligand C ₈ (C and D, respectively).....	230
Figure 10.8 - Final snapshots of 50 ns MD simulations of the complexes formed by (A) AT11 G4- RBD _{1,2} and (B) AT11-Bo G4- RBD _{1,2} . G4 structures are depicted in green with loop thymines responsible for binding highlighted in orange while RBD _{1,2} is represented in blue ribbon (except for atoms participating in H-bonding). H-bonds are shown as cyan lines.....	231

Figure 10.9 - Confocal fluorescence images of HeLa cells incubated with (A) Cy5-AT11-C₈ and (B) Cy5-AT11-Bo-C₈ after 2 h at 37°C. Nucleolin is stained with AlexaFluor 546®, C₈ emits green fluorescence, Cy5-AT11 and Cy5-AT11-Bo are shown in red. Scale bar: 25 µm. Brightness was adjusted for all images equally.....232

Figure 10.10 - Confocal microscopy images of HeLa cells incubated with (A) Cy5-AT11 G4-C₈ and (B) Cy5-AT11-Bo G4-C₈ for 3 days. Cell nuclei are stained with Hoechst 33342 (blue), C₈ emits green fluorescence and Cy5- AT11 G4 or Cy5-AT11-Bo G4 are shown in red. Overlapping of the two stains, observed as yellow/orange regions, can be seen in the merged images of the cells. Scale bar: 25 µm.....233

Figure 10.11 - Confocal microscopy images of NHDF cells incubated for 3 days with (A) Cy5-AT11 G4-C₈ and (B) Cy5-AT11-Bo G4-C₈. Cell nuclei are stained with Hoechst 33342 (blue), C₈ emits green fluorescence and Cy5- AT11 G4 or -AT11-Bo G4 are shown in red. The combination of the two stains, observed as yellow/orange regions, can be seen in the merged images of the cells. Scale bar: 25 µm.....233

Figure 10.12 - Flow cytometry analysis of HeLa cells incubated with (A) free C₈, (B) AT11 G4-C₈ complex and (C) AT11-Bo G4-C₈ complex. Cells were incubated for 24 h with the preformed aptamer-ligand complex at a C₈ concentration of 1 µM.....234

Figure 10.13 - Flow cytometry analysis of NHDF cells incubated with (A) free C₈, (B) AT11 G4-C₈ complex and (C) AT11-Bo G4-C₈ complex. Cells were incubated for 24 h with the preformed aptamer-ligand complex at a C₈ concentration of 1 µM.....234

Figure 10.14 - Percentage of viable HeLa and NHDF cells after incubation with free C₃ (2 µM), C₅ (2 µM) and C₈ (1 µM), or the AT11 G4 -ligands (pre-formed AT11-ligand conjugates by 10 min incubation) for 7 days. Results are presented as mean ± SEM.....235

Figure 10.15 - Percentage of viable HeLa and NHDF cells after incubation with free C₃ (2 µM), C₅ (2 µM) and C₈ (1 µM), or the AT11-Bo G4-ligands (pre-formed AT11-Bo-ligand conjugates by 10 min incubation) for 7 days. Results are presented as mean ± SEM.....235

Figure 11.1 - (A) Lipid-modified AS1411 oligonucleotides self-assemble in nanoaggregates decorated with G4 structures; (B) Chemical structure of the octadecyl lipidic

phosphoramidite used and the sequence of AS1411; (C) Chemical structure of the acridine orange derivative C₈ used in this study.....251

Figure 11.2 - Native agarose gel electrophoresis of 5'-lipid-AS1411 in different ionic conditions or in the absence of any added salt.....252

Figure 11.3 - (A) CD spectrum of 5'-lipid-AS1411 G4 structure showing the typical bands of a parallel stranded G4; (B) CD melting profile of 5'-lipid-AS1411 in 1× buffer obtained by monitoring the ellipticity at 260 nm at increasing temperatures.....253

Figure 11.4 - (A) DLS and (B) TEM analysis (12,000× magnification) of nanoaggregates formed by 5'-lipid-AS1411 in 1× buffer; (C) DLS and (D) TEM (12,000× magnification) analysis of 5'-lipid-AS1411 nanoaggregates encapsulating the acridine orange derivative C₈.....255

Figure 11.5 - Cumulative release profile of C₈ from the AS1411-decorated nanoaggregates in 1× buffer for 72 h. Values are presented as mean ± SD.....256

Figure 11.6 - Relative cell viability of HeLa and NHDF cells incubated for 72 h with either blank nanoaggregates or C₈-loaded nanoaggregates at a drug concentration of 0.5 μM. Values are presented as mean ± SEM.....257

Figure 11.7 - Cellular internalization of the C₈-loaded nanoaggregates determined by flow cytometry (A and B for HeLa and NHDF, respectively) and (C) fluorescence confocal microscopy. The experiments were performed by incubating the cells for 24 h with C₈-loaded nanoaggregates prior to data acquisition.....258

Figure 11.8 - Fluorescence confocal microscopy images of tissue biopsies of CIN1 patients incubated with C₈-loaded nanoaggregates for 3 h. Cell nuclei are stained with Hoechst 33342 in the blue channel, C₈ emits green fluorescence and nucleolin was stained with Alexa Fluor™ 647 antibody.....259

List of Tables

Table 1.1 - Histological classification of pre-cancerous lesions.....	5
Table 2.1 - List of G4 ligands discussed in this review.....	38
Table 2.2 - List of G4 aptamers discussed in this review.....	43
Table 4.1 - Ligand-induced thermal stabilization measured by CD and FRET melting experiments.....	73
Table 4.2 - Chemical shifts and chemical shift changes of selected ¹ H NMR signals.....	76
Table 5.1 - Oligonucleotide sequences used in the helicase assay.....	97
Table 5.2 - Oligonucleotide sequences used in the PCR stop assay.....	97
Table 6.1 - Ligand-induced thermal stabilization of various G4s measured by CD melting experiments in the absence and presence of ligands.....	125
Table 6.2 - Apparent dissociation constants (K_D) obtained from fluorescence titration experiments.....	127
Table 7.1 - HPV oligonucleotide sequences used in this study.....	143
Table 7.2 - Ligand-induced thermal stabilization (ΔT_m) measured by FRET melting experiments.....	149
Table 7.3 - Ligand-induced thermal stabilization (ΔT_m) measured by CD melting experiments.....	151
Table 8.1 - Nucleotide sequence of the aptamers used in this study and their spectroscopic profile characterization.....	176
Table 8.2 - Ligand induced thermal stabilization of AS1411 aptamer in the presence and absence of nucleolin.....	184

Table 10.1 - Ligand-induced thermal stabilization of the AT11 or AT11-Bo measured by CD melting experiments.....	227
Table 10.2 - Apparent dissociation constants (K_D) obtained from fluorescence titration experiments.....	228
Table 11.1 - Characterization of AS1411 nanoaggregates in terms of size and surface charge.....	254
Table 11.2 - HPLC quantitation of C ₈ in the samples recovered from Franz cells and porcine vaginal tissue.....	260

List of Abbreviations

^{13}C	carbon
^1H	hydrogen
5'-UTR	5'-untranslated regions
5-FU	5-fluorouracil
CD	circular dichroism
CH_2Cl_2	dichloromethane
CH_3CN	acetonitrile
CIN1	cervical intraepithelial neoplasia grade 1
CIN2	cervical intraepithelial neoplasia grade 2
CIN3	cervical intraepithelial neoplasia grade 3
CMC	critical micellar concentration
D_2O	deuterium oxide
DAPI	4,6-diamidino-2-phenylindole
DLS	dynamic light scattering
DMEM	Dulbecco's modified Eagle medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DPFGSE	double pulse-field gradient spin echo
ds	double-strand
EBV	Ebola virus
EGFR	epidermal growth factor receptor
EMT	epithelial to mesenchymal transition
ESI	electrospray ionization
EtOH	ethanol
FAM	6-carboxyfluorescein
FBS	fetal bovine serum
FDA	US Food and Drug Administration
FID	fluorescent intercalator displacement
FITC	fluorescein isothiocyanate
FRET	Förster resonance energy transfer
G4	G-quadruplex
G4 ChIP-seq	G4 chromatin immunoprecipitation sequencing
GAPDH	glyceraldehyde 3-phosphate dehydrogenase

H&E	haematoxylin and eosin
H ₂ O	water
HBV	hepatitis B virus
HCl	hydrochloric acid
HIV	immunodeficiency virus
HPV	human papilloma virus
HRMS	high-resolution mass spectrometry
HSIL	high-grade squamous intraepithelial lesion
HSPG	heparan sulfate proteoglycans
HSV	herpesvirus
IC ₅₀	half maximal inhibitory concentration
IR	ionizing radiation
K ⁺	potassium
K ₂ HPO ₄	potassium hydrogenphosphate
KCl	potassium chloride
K _D	dissociation constant
KH ₂ PO ₄	potassium dihydrogen phosphate
KRAS	Kirsten rat sarcoma viral oncogene homolog
K _{SV}	Stern–Volmer constant
LCR	long control region
LDH	lactate dehydrogenase
Li ⁺	lithium
LiCaco	lithium cacodylate
LiCl	lithium chloride
LNA	locked nucleic acids
LSIL	low-grade squamous intraepithelial lesion
MD	molecular dynamics
Mg(OAc) ₂	magnesium acetate
MgCl ₂	magnesium chloride
MTT	thiazolyl blue tetrazolium bromide
Na ⁺	sodium
Na ₂ HPO ₄	sodium hydrogenphosphate
NaBH ₄	sodium borohydride
NaCl	sodium chloride
NaH ₂ PO ₄	sodium dihydrogen phosphate
NCR	noncoding region
NHDF	normal human fibroblasts

NHE	nuclease hypersensitive element
NMR	nuclear magnetic resonance
NSCLC	non-small cell lung cancer
OD	optical density
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PD1	programmed cell death protein 1
PDB	protein data bank
PDGFA	platelet derived growth factor subunit A
PDI	polydispersity index
PD-L1	programmed death-ligand 1
PDS	pyridostatin
PME	Particle mesh Ewald
POT1	protection of telomeres protein 1
pRB	retinoblastoma
RBD	RNA-binding domain
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
RT-PCR	real-time polymerase chain reaction
SARS-CoV	severe acute respiratory syndrome coronavirus
SELEX	systematic evolution of ligands by exponential enrichment
ss	single-strand
TAMRA	6-carboxytetramethylrhodamine
TBA	thrombin-binding aptamer
TDS	thermal difference spectra
TEM	transmission electron microscopy
T _m	melting temperature
TO	thiazole orange
TPP1	tripeptidyl peptidase 1
TRF2	telomeric repeat-binding factor 2
TSP	trimethylsilyl propionate
TXI	inverse triple resonance
URR	upstream regulatory region
VEGF	vascular endothelial growth factor
WHO	World Health Organization

List of Scientific Publications

Articles included in the Thesis:

Carvalho, J., Mergny, J.-L., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2020) G-quadruplex, Friend or Foe: The Role of the G-quartet in Anticancer Strategies. *Trends in Molecular Medicine*, In press. DOI: 10.1016/j.molmed.2020.05.002. IF: 11.028

Carvalho, J., Nottelet, P., Mergny, J.-L., Queiroz, J. A., Salgado, G. F., and Cruz, C. (2017) Study of the interaction between indole-based compounds and biologically relevant G-quadruplexes. *Biochimie* 135, 186–195. DOI: 10.1016/j.biochi.2017.02.005. IF: 3.362

Carvalho, J., Quintela, T., Gueddouda, N. M., Bourdoncle, A., Mergny, J.-L., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2018) Phenanthroline polyazamacrocycles as G-quadruplex DNA binders. *Organic and Biomolecular Chemistry* 16, 2776–2786. DOI: 10.1039/c8ob00247a. IF: 3.490

Carvalho, J., Pereira, E., Marquevielle, J., Campello, M. P. C., Mergny, J.-L., Paulo, A., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2018) Fluorescent light-up acridine orange derivatives bind and stabilize KRAS-22RT G-quadruplex. *Biochimie* 144, 144–152. DOI: 10.1016/j.biochi.2017.11.004. IF: 3.362

Carvalho, J., Lopes-Nunes, J., Campello, M. P. C., Paulo, A., Milici, J., Meyers, C., Mergny, J.-L., Salgado, G. F., Queiroz, J. A., and Cruz, C. Human papillomavirus G-rich regions as potential antiviral drug targets. *Nucleic Acid Therapeutics, Ahead of Print*. DOI: 10.1089/nat.2020.0869. IF: 5.000

Carvalho, J., Paiva, A., Campello, M. P. C., Paulo, A., Mergny, J.-L., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2019) Aptamer-based Targeted Delivery of a G-quadruplex Ligand in Cervical Cancer Cells. *Scientific Reports* 9, 7945. DOI: 10.1038/s41598-019-44388-9. IF: 4.011

Carvalho, J.[†], Lopes-Nunes, J.[†], Lopes, A. C.[†], Campello, M. P. C., Paulo, A., Queiroz, J. A., and Cruz, C. (2019) Aptamer-guided acridine derivatives for cervical cancer. *Organic and Biomolecular Chemistry* 17, 2992–3002. DOI: 10.1039/c9ob00318e. IF: 3.490

Figueiredo, J.[†], Lopes-Nunes, J.[†], **Carvalho, J.**[†], Antunes, F., Ribeiro, M., Campello, M. P. C., Paulo, A., Paiva, A., Salgado, G. F., Queiroz, J. A., Mergny, J.-L., and Cruz, C. (2019) AS1411 derivatives as carriers of G-quadruplex ligands for cervical cancer cells. *International Journal of Pharmaceutics* 568, 118511. DOI: 10.1016/j.ijpharm.2019.118511. IF: 4.213

[†] Equal contribution.

Articles not included in the Thesis:

Lopes-Nunes, J. †, **Carvalho, J.** †, Figueiredo, J., Ramos, C. I. V., Lourenço, L. M. O., Tomé, J. P. C., G.P.M.S. Neves, M., Mergny, J.-L., Queiroz, J. A., Salgado, G. F., and Cruz, C. (2020) Phthalocyanines for G-quadruplex aptamers binding. *Bioorganic Chemistry* 100, 103920. DOI: 10.1016/j.bioorg.2020.103920. IF: 3.926

Carvalho, J., and Cruz, C. (2020) Forster resonance energy transfer for studying nucleic acids denaturation: A chemical and biological sciences laboratory experiment. *Biochemistry and Molecular Biology Education*, 1–8. DOI: 10.1002/bmb.21353. IF: 0.947

Carvalho, J., Santos, T., Carrilho, R., Sousa, F., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2020) Ligand screening to pre-miRNA 149 G-quadruplex investigated by molecular dynamics. *Journal of Biomolecular Structure and Dynamics* 38, 2276–2286. DOI: 10.1080/07391102.2019.1632743. IF: 3.310

Eddahmi, M., Moura, N. M. M., Bouissane, L., Faustino, M. A. F., Cavaleiro, J. A. S., Paz, F. A. A., Mendes, R. F., Figueiredo, J., **Carvalho, J.**, Cruz, C., Neves, M. G. P. M. S., and Rakib, E. M. (2019) Synthesis and Biological Evaluation of New Functionalized Nitroindazolylacetonitrile Derivatives. *ChemistrySelect* 4, 14335–14342. DOI: 10.1002/slct.201904344. IF: 1.716

Ramos, C., Almeida, S., Lourenço, L., Pereira, P., Fernandes, R., Faustino, M., Tomé, J., **Carvalho, J.**, Cruz, C., and Neves, M. (2019) Multicharged Phthalocyanines as Selective Ligands for G-Quadruplex DNA Structures. *Molecules* 24, 733. DOI: 10.3390/molecules24040733. IF: 3.060

Carvalho, J., Alves, S., Castro, M. M. C. A., Gerales, C. F. G. C., Queiroz, J. A., Fonseca, C. P., and Cruz, C. (2019) Development of a bioreactor system for cytotoxic evaluation of pharmacological compounds in living cells using NMR spectroscopy. *Journal of Pharmacological and Toxicological Methods* 95, 70–78. DOI: 10.1016/j.vascn.2018.11.004. IF: 2.679

Carvalho, J., Queiroz, J. A., and Cruz, C. (2017) Circular dichroism of G-Quadruplex: A laboratory experiment for the study of topology and ligand binding. *Journal of Chemical Education* 94, 1547–1551. DOI: 10.1021/acs.jchemed.7b00160. IF: 1.763

Cruz, C., Pedro, A. Q., **Carvalho, J.**, Santos, T., Talhada, D., Paiva, A., Queiroz, J. A., Andrade, M., Pinto, M., Montenegro, L., Pereira, P. Nucleolin as a potential biomarker for canine malignant neoplasia. *Research in Veterinary Science, In Press*. DOI: 10.1016/j.rvsc.2020.10.004. IF: 1.892

Miranda, A. †, Santos, T. †, **Carvalho, J.** †, Alexandre, D., Jardim, A., Caneira, C. R. F., Vaz, V., Pereira, B., Godinho, R., Brito, D., Chu, V., Conde, J. P., Cruz, C. Aptamer-based approaches to detect nucleolin in prostate cancer. *Talanta, Accepted for Publication*. IF: 5.339

† Equal contribution.

Chapter 1

Cervical Cancer and Human Papillomavirus

Cervical Cancer and Human Papillomavirus

Cervical Cancer

Cancer of the cervix uteri is the 3rd most common cancer type amongst women worldwide (2nd most common in women aged 15 to 44 years), with more than 500,000 new cases and around 310,000 deaths in 2018, representing 7.5% of all female cancer deaths [1]. In Portugal, about 750 new cases are annually diagnosed with an estimate of 340 deaths in 2018 alone [1]. Despite advances in prevention, screening, diagnosis, and treatment during the past decade, there are still substantial regional and global disparities in cervical cancer (**Figure 1.1**), being that almost 90% of all cervical cancer deaths occurred in low- and middle-income countries where the mortality is 18 times higher than in developed countries [2]. Histopathologically, cervical cancer can be divided into squamous cell carcinomas, adenocarcinomas and adenosquamous carcinomas which account for approximately 70%, 25% and 3% of all cervical cancers, respectively [2,3]. More rarely, histologies include neuroendocrine or small cell carcinomas.

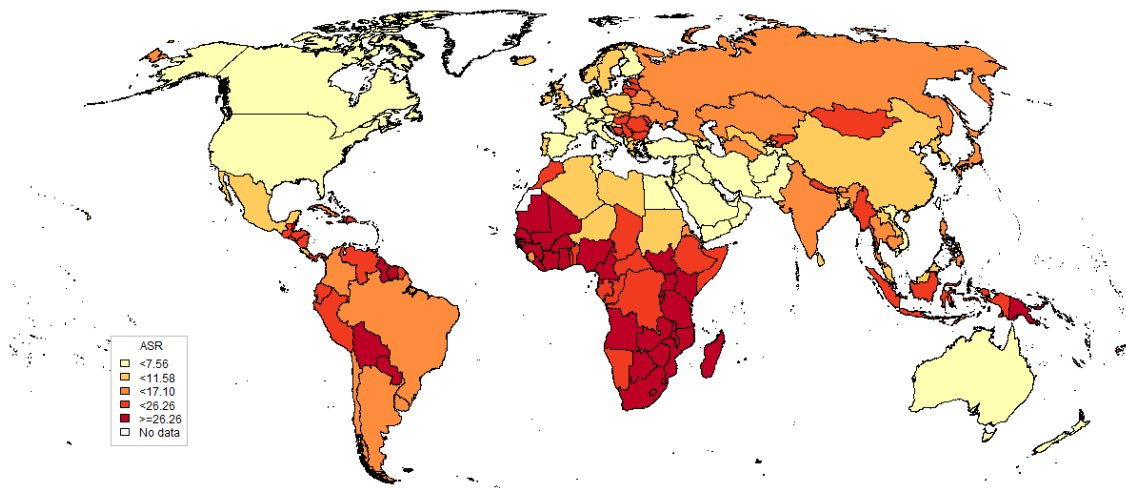


Figure 1.1 - Age-standardised incidence rates of cervical cancer in the World (estimates for 2018). Adapted from [1].

Virtually all cases of cervical cancer are caused by Human Papillomavirus (HPV) persistent infection [4]. Most sexually active women (and men) will acquire a genital HPV infection during their lifetime, despite being generally asymptomatic [5]. Low-risk HPVs induce benign lesions in the mucosal and cutaneous epithelia (most commonly genital warts), and most of the infections are cleared by the immune system within 1-2 years after infection [6]. High-risk HPVs on the other hand are responsible for the development of

pre-cancerous lesions termed low-grade squamous intraepithelial lesions (LSIL), also known as cervical intraepithelial neoplasia grade 1 (CIN 1), which show only mild dysplastic changes [5]. A small fraction of these lesions persists and progress into high-grade squamous intraepithelial lesions (HSIL), also known as CIN 2 (moderate dysplasia) or CIN 3 (severe dysplasia and carcinoma in situ) [5]. If left untreated, HSIL progress into invasive cervical cancer [5]. Risk factors for persistent HPV infection include early age of sexual debut [7], multiple sexual partners [8], immunosuppression [9], history of sexually-transmitted infection or HPV-related vulvar or vaginal dysplasia [2], host genetic susceptibility [10], alcohol consumption [11] and tobacco smoke [12].

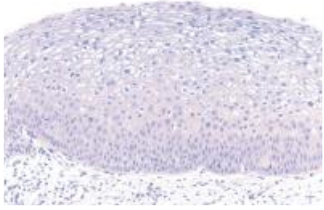
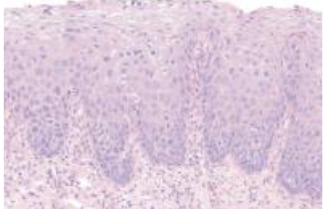
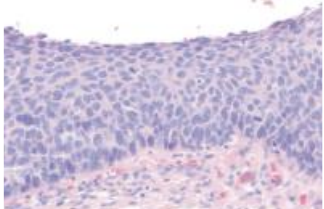
Primary approaches to prevent HPV infection and ultimately cervical cancer include the development of widely available vaccination programmes against primarily high-risk HPV types. Three HPV prophylactic vaccines have been approved for clinical use: the bivalent Cervarix (GSK, UK), the quadrivalent Gardasil and the nonavalent Gardasil9 (both Merck, USA). Cervarix contains HPV16 and HPV18 antigens [13] whereas Gardasil provides protection against HPV6, HPV11, HPV16 and HPV18 [14], both available in 2006. Gardasil9 provides protection against HPV6, HPV11, HPV16, HPV18, HPV31, HPV33, HPV45, HPV52 and HPV58, and became available in 2018 [15]. Vaccination programmes typically target adolescent girls and boys aged 9–13 years [4]. In countries where HPV vaccination programmes have been introduced, a substantial decreases in cervical cancer incidence was observed [15].

Secondary preventive measures include pre-cancer and cervical cancer screening [16]. Screening allows pre-cancerous lesions to be identified at stages when they can easily be treated, preventing up to 80% of cervical cancers while reducing cancer mortality and morbidity [4]. Because pre-cancerous lesions may take several years to develop, screening is recommended for every woman over the age of 30 and regularly afterwards (frequency depends on the screening test used). The most commonly employed alternatives for cervical cancer screening include cytology (also known as Papanicolaou (Pap) smear), HPV-testing (by PCR detection of DNA or RNA of high-risk HPVs 16 and 18) and cytology-HPV co-testing [16]. Despite cytology still being the most widely used, HPV-testing was already shown to be more efficacious at detecting pre-cancers due to the lower sensitivity of the first, and earlier detection of pre-cancers and reduced number of follow-up screens of the latter [4].

When abnormalities are detected in the cytology (LSIL or HSIL) and HPV-testing returns positive (particularly for HPV16 and HPV18), accurate diagnosis generally involving colposcopy with biopsy placement and histological interpretation of the biopsy specimen

is followed [16]. The histopathological assessment allows the classification of the lesions according to the CIN scale together with the Bethesda System (LSIL and HSIL) (**Table 1.1**) and helps to decide how women with HPV-related pre-cancer and cancer should be treated (staging) [17].

Table 1.1 – Histological classification of pre-cancerous lesions.

Histology Grade	Cytology	Description	Example ^a
CIN 1	LSIL	• Mild epithelial dysplasia • Confined to the basal one-third of the epithelium • Typically corresponds to infection with HPV • High rate of regression to normal cells	
CIN 2	HSIL	• Moderate epithelial dysplasia • Confined to the basal two-thirds of the epithelium • Represents a mix of low- and high-grade lesions not easily differentiated by histology	
CIN 3		• Severe epithelial dysplasia • Undifferentiated neoplastic cells that span more than two-thirds of the epithelium • May involve the full thickness • Also referred to as carcinoma in situ	

^a Images adapted from [3].

CIN1 is considered a morphological alteration as consequence of HPV infection and therefore is not treated. CIN2 despite being considered a more severe HPV infection than CIN1, a large proportion still regresses and clinicians may opt for not treating but rather keep under surveillance [4]. CIN3 as a surrogate of pre-cancer is usually treated. In this case, typical options for pre-cancers include excision, ablation (cryotherapy) and possibly immunotherapy [18]. In the case of invasive cervical cancers, the treatment option depends on the stage of the disease. Common options include surgery, external beam radiation, brachytherapy and cisplatin-based chemotherapy [18].

Human Papillomavirus

HPVs are small non-enveloped double-stranded DNA viruses, members of the *Papillomaviridae* family. More than 200 different HPV types were already identified and categorized into five genera [19]. Depending on the genera, HPVs have different propensities to establish in different locations: α -HPVs infect mucosal surfaces and skin, while β -, γ -, μ -, and ν -HPVs infect only cutaneous epithelia [20]. For this reason, the α genus is of the greatest medical importance given they promote cancer particularly at the junction between glandular and squamous epithelia [19]. Based on epidemiological and biological data, HPV types are further classified in high-risk or low-risk types depending on their carcinogenic potential. HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68 and 63 referred to as high-risk HPV types, are responsible for malignant transformations resulting in pre-cancerous and cancerous states, while low-risk HPV types such as HPV6, 11, 42, 43, 44, 54, 61, 70, 72, and 81 are considered non-carcinogenic [6]. HPVs are well established causative agents of cervical cancer (as well as other anogenital and head and neck cancers), being that HPV types 16, 18, and 58 rank as top three high-risk HPV types for invasive carcinoma and account for ~90% of all cervical cancer cases [21]. Low-risk HPV types 6 and 11 are the most common case of benign lesions such as genital warts or condylomata acuminata [6].

HPV genome is approximately 8.0 kb long, circular double-stranded DNA also termed episome (**Figure 1.2**) [22]. The genome organization is shared among all α -HPVs and is divided into three major regions: (i) a LCR, also called the upstream regulatory region (URR) or noncoding region (NCR), located between the L1 and E6 ORFs, which contains the regulatory element and early promoter and involved in viral DNA replication and transcription; (ii) the early region, composed of 6 ORFs which encode the regulatory proteins E1, E2, E4, E5, E6, and E7 involved in viral gene expression, replication, and survival; and (iii) the late region, with 2 ORFs which encode the viral capsid structural proteins L1 and L2 [23]. Only four ORFs (E1, E2, L1, and L2) are necessary to ensure the viral replication and shedding of the virus, and are present in all known HPVs [24].

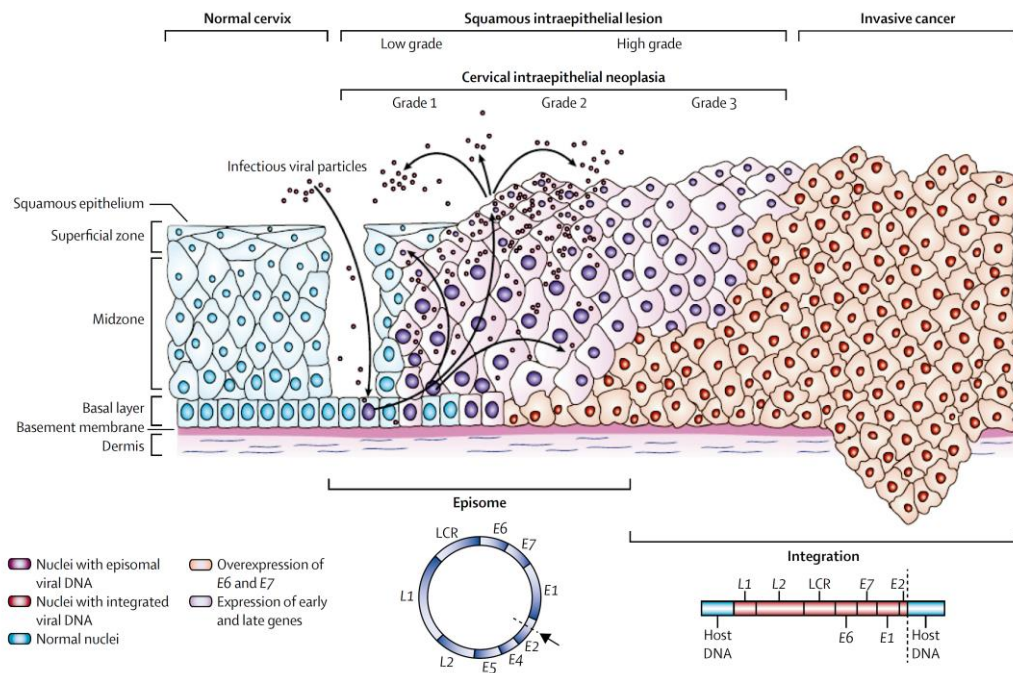


Figure 1.2 – Pathogenesis of HPV infection in cervical cancer. The schematic drawing of the HPV genome organization before and after genome integration is shown at the bottom. The open reading frames encode the six early proteins E1, E2, E4, E5, E6, and E7 and two late structural proteins L1 and L2. Adapted from [2].

HPVs exclusively infect human epithelial cells, more specifically the basal keratinocytes (**Figure 1.2**). Regardless of the pathway, it has been shown that the infection requires epithelial damage to allow viral access to the basal lamina, where the basal keratinocytes are located [25]. Virus entry into the basal layer keratinocytes is initiated by the L1 (major capsid) and L2 (minor capsid) proteins, by interacting with heparan sulfate proteoglycans (HSPGs) located on the surface of basal cells [26–29]. The virus is endocytosed through mechanism similar to macropinocytosis, which is clathrin, caveolin and lipid raft independent [30]. During endosomal trafficking, virions undergo structural changes that result in viral uncoating [31]. The viral genome is then transported into the host cell nucleus and maintained as episomal DNA, in a process which is thought to be dependent on protein L2 [32].

Within the host cell nucleus, HPV expresses the E1 ATP-dependent helicase to facilitate viral DNA replication, which is recruited by a transcription factor encoded in the E2 gene [20]. The infection can follow one of two pathways, either productive or non-productive [33]. In the non-productive stage, viral episomal genomes are amplified to densities of 50–100 copies per cell in the nucleus of proliferative basal cells, in a host-independent mechanism [34]. During this stage, the viral gene expression is minimal. The infected basal keratinocytes then withdraw from the cell cycle and enter the differentiation process, during which HPV begins its productive pathway [19]. In this stage, viral DNA amplification and gene expression activity is greatly amplified [35]. In order to hijack the

host cells' DNA replication machinery, apart from the E1 protein, HPV also expresses E6 and E7 oncoproteins which delay cell differentiation and maintain the HPV genome in a stable episomal form [36]. Both E6 and E7 are transcriptionally regulated by E2 [37]. E6 oncoprotein forms a complex with tumour suppressor protein p53 and recruits enzymes to degrade p53, preventing cellular apoptosis [38]. E7 protein on the other hand, stimulates cell proliferation by binding to the retinoblastoma (pRB) protein and promotes host genome instability [38]. E6 and E7 oncoprotein expression overrides cell cycle checkpoints, therefore allowing HPV to replicate, and are responsible for cell transformation in cervical cancer due to the inactivation of tumour suppression mechanisms, chromosomal instability and the overactivation of telomerase activity [38–40]. Recent studies have also shown the importance of E5 oncoprotein in modulating cell transformation and the immune system, cooperating with oncoproteins E6 and E7 during cell's malignant progression [41].

The viral cycle completes in the upper layers of the epithelium where the viral genomes number is increased up to thousands of copies per host cell [20]. L1 and L2 viral capsid proteins are expressed and assembled in the terminally differentiated cells to allow genome packaging [42]. Following the encapsulation of the viral genomes and maturation of viral particles [43], HPV is then shed from differentiated infected epithelial cells due to their loss of nuclear and cytoplasmic integrity, assisted by E4 binding to cytokeratin filaments that disrupts their structure [44]. The HPV life cycle is summarized in **Figure 1.3**.

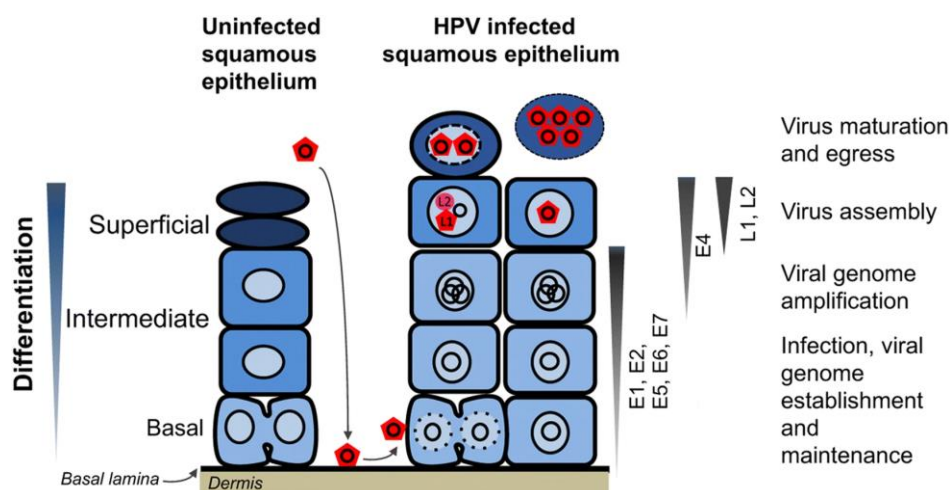


Figure 1.3 – HPV life cycle organization. The HPV life cycle is dependent on the terminal differentiation of the keratinocytes and the expression of both early and late protein. Adapted from [45].

A key event for the development of HPV-associated tumours is the integration of high-risk HPV genomes (such as HPV16 and 18) into the host genome (**Figure 1.2**) [41]. Once the viral DNA is integrated into the host's genome, viral replication is halted due to disruption of E1 and E2 ORFs [41,46]. The integration is triggered by the genomic instability promoted by E6 and E7, which increases the occurrence of double strand breaks in both the host's and viral genomes [47]. The integration of the viral genome occurs in fragile chromosomal regions or near clusters of microRNAs [48]. During this process the E2 ORF is usually deleted or disrupted which leads to deregulated expression of oncoproteins E6 and E7, ultimately leading to carcinogenesis [49]. Hence, E6 and E7 mRNA expression is usually higher in integrated HPV genomes [20].

Treatment of HPV-induced disease

In the past decades, advances in the understanding of the HPV infection mechanisms and its carcinogenesis have been done. Based on these findings, different strategies have been developed to prevent or treat HPV-induced diseases. Despite safe and effective prophylactic vaccines for the prevention of HPV infection, these are not efficacious at treating subjects which are already infected [50]. Furthermore, the existing vaccines do not protect against all HPV types associated with cervical cancer and other diseases, and these are not universally available as they are expensive for developing countries [50,51]. Currently, there is no effective treatment for HPV persistence [51]. The ideal HPV treatment would be able to clear HPV infection even when the virus is present in its latent form.

Traditional treatment for invasive, advanced-stage or recurrent cervical cancer includes systemic chemotherapy, surgery, radiotherapy and more recently immunotherapy [52] (**Figure 1.4**). These treatments cause severe side effects and insignificant survival rates to the patients and do not benefit from the interplay between the host and HPV during malignant progression [51]. Currently, the treatment of high-grade lesions (CIN2-3) includes electrosurgical excision of the diseased tissue, sometimes at cost of infertility [2]. However, the high risk of incomplete excision may facilitate recurrent neoplasia due to the remaining HPV transformed cells. Radiotherapy and chemotherapy treatments on their turn, are unable to differentiate cancerous from healthy tissue, with significant off-target cytotoxicity that usually leads to the appearance of side effects and patient morbidity. Therefore, continuous studies of the molecular biology of HPV-induced carcinogenesis are necessary to develop improved and targeted therapeutic options for the treatment of persistent HPV infection and cervical cancer.

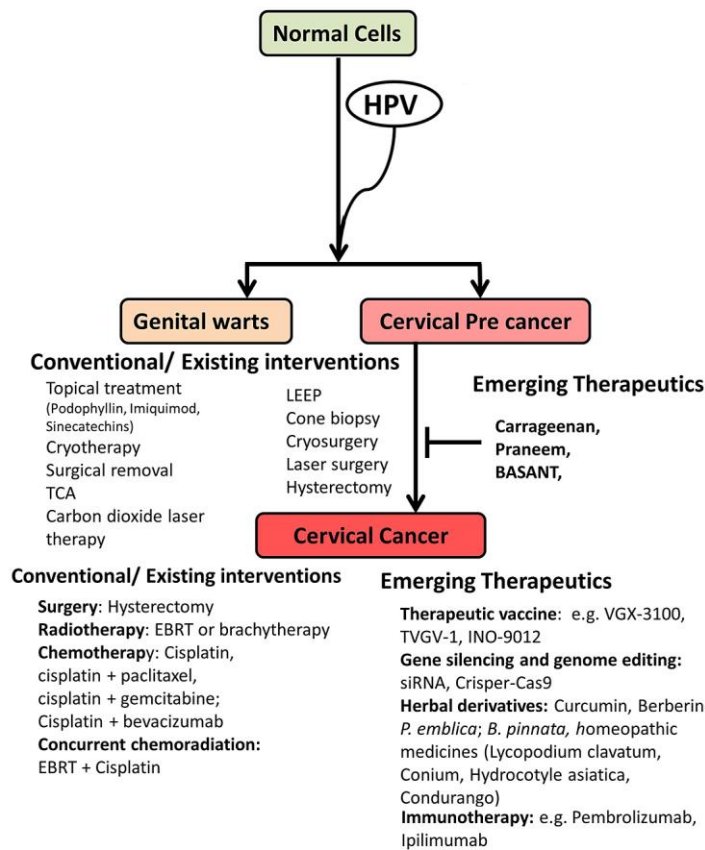


Figure 1.4 – Conventional and emerging modalities for treatment of both benign and malignant cervical lesions. EBRT - External beam radiation, LEEP - loop electrosurgical excision procedure, TCA – trichloroacetic acid. Adapted from [52].

A current approach at clearing HPV infection is via immunotherapy using therapeutic vaccines [53]. The goal of a therapeutic vaccine is to induce a virus-specific T-cell response against the established HPV infection and lesions. This is commonly achieved by using a recombinant DNA vector expressing a viral antigen (commonly E6 or E7 oncoproteins) which are then engulfed, processed and presented by dendritic cells to downstream T-cells initiating the cellular immune response [50,54]. The secreted antigens may also interact with B-cells to initiate a humoral immune response [54]. As a well-documented example, *Lm*-LLO-E7 (also known as ADXS11-001) is a bacterial-based vaccine vector that secretes E7 antigen and was used to treat patients with late-stage metastatic cervical cancer [55]. Currently, *Lm*-LLO-E7 is undergoing phase III clinical trial for high-grade cervical cancer (NCT02853604). Another recent DNA-based vaccine is VGX-3100 which is designed to express E6 and E7 fusion proteins [56]. The E6 and E7 genes within the DNA vaccine were mutated to remove the ability of the oncoproteins to degrade p53 and pRB [56]. GX-188E seems to be a promising therapeutic as it demonstrated ability to induce an E6/E7-specific T-cell immune response in patients with high-grade cervical lesions (NCT03180684) [56]. However, despite the advances made, the numerous therapeutic vaccines developed have

proved to be ineffective clinically as the T-cell responses were neither strong nor broad enough [50]. The use of T-cell immune checkpoint inhibitors has also been developed as an alternative to therapeutic vaccines. These inhibitors target molecules that regulate T-cell immune responses such as the PD-1/PD-L1 axis [57]. Both Pembrolizumab (PD-1 antibody, NCT02628067) [58] and Durvalumab (PD-L1 antibody, NCT01975831) [59] shown great promise in clinical trials in the treatment of advanced cervical cancer by rescuing T-cell function upon blockade of PD-1/PD-L1 interaction. These therapeutic approaches are however not specific of HPV infection as they are not targeted at HPV proteins or genome replication.

Another therapeutic alternative is the targeting of HPV E6/E7 mRNAs with siRNAs that could effectively silence their expression and induce apoptotic cell death in HPV- positive cells [54,60,61]. However, siRNAs were found to only block HPV E6/E7 mRNAs temporarily, and these do not attack HPV DNA in the nucleus which serves as a reservoir of escape mutants that cause resistance to siRNA therapies [51]. As an alternative, several groups have proposed the use of microRNAs to silence the expression of E6 and E7 oncoproteins [54]. It was recently demonstrated that miR-375 was able to directly downregulate E6 and E7 expression in HPV16- and HPV18-positive cervical cancer cell lines, rescuing p53 and pRB pathways [62]. However, several issues such as stability, delivery, and off-target effects remain to be addressed to understand the potential of microRNAs as anti-HPV agents.

The use of small molecule inhibitors and nucleic-acid based systems offer vast opportunities in the development of effective therapeutics against HPV-related diseases, particularly by targeting and interrupting virus-host interactions [50,63]. These small molecules act as inhibitors of the HPV life cycle processes such as DNA replication, episome maintenance, or viral transcription. This leads to the termination of latent infection as the virus is unable to maintain its genome in dividing cells [50]. A recent example encompasses two flavonoid compounds, baicalein and gossypetin, which were identified using an high-throughput screening strategy, that are able to inhibit cell proliferation by interfering with E6-E6AP interaction [64]. E6-Associating Protein (E6AP) is a target of E6 oncoprotein that mediates the upregulation of the p53 tumour suppressor pathway as a response to the E7-mediated degradation of pRB [65]. During HPV infection, E6 forms a complex with E6AP and promotes the degradation of p53 via the ubiquitin-proteasome pathway [65]. Both baicalein and gossypetin were found to inhibit E6-E6AP interaction and protect p53 from degradation in HPV-positive cell lines [64]. Other natural products such as curcumin or berberine have demonstrated promising anti-HPV

activities and can emerge as potential treatments for cervical cancer [52]. Both E1 and E2 proteins, as well as their interaction are also interesting therapeutic targets for HPV-diseases [63,66,67]. A biphenylsulfonacetic acid derivative was shown to effectively inhibit the helicase and ATPase activities of E1 protein of low-risk HPV6 and high-risk HPV18 [68]. Despite recent developments, challenges remain due to various obstacles such as lack of well-defined binding pockets, amino acid divergences between different HPV types, and stage-dependence of the protein targets.

Novel therapeutic prospects

Rather than targeting and disrupting protein-protein interactions [63], an alternative strategy with great potential to persistent HPV infection is the development of small molecules capable of targeting the viral DNA, particularly non-canonical motifs formed within the HPV genome called G-quadruplexes (G4s) [69–71]. G4s are four stranded DNA structures with square planar arrangements of guanine residues stacked on each other and stabilized with cations, particularly K^+ and Na^+ (see **Chapter 2** for more detail on G4s) [72]. Several guanine-rich regions were detected in 10 HPV types (noteworthy in high-risk HPV16 and 18) and are present in LCR, E1, E4 regions and the L2 protein coding sequence (**Figure 1.5, left**), which are involved in transcription, replication and viral proteins production [70,71]. Later on, these guanine-rich regions were found to fold into stable G4 structures (**Figure 1.5, right**) [69–71].

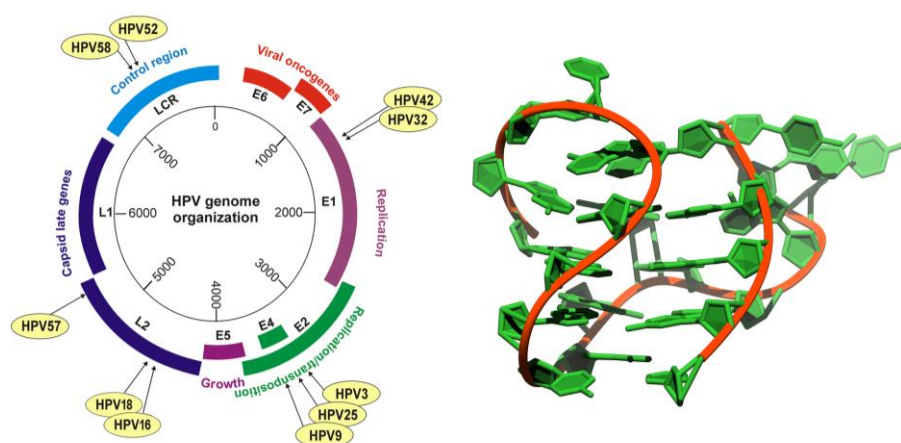


Figure 1.5 – HPV genome organization and the positions where G4-forming sequences were found in the respective HPVs (adapted from [70]). On the right, the NMR solution structure of a G4 found in the HPV52 genome (PDB: 5O4D). Nucleotide residues are shown in green while the DNA backbone is depicted in orange.

Hence, targeting the G4 structures of the HPV genome with small molecule inhibitors could be a potential antiviral strategy, by regulating G4-driven gene expression and DNA

replication. This is a promising, yet unexplored strategy which has already demonstrated to be effective in managing other viral infections (especially in virus with double stranded DNA as HPV) such as herpesvirus (HSV), ebola virus (EBV), hepatitis B virus (HBV), human immunodeficiency virus (HIV) and severe acute respiratory syndrome coronavirus (SARS-CoV) [73–76]. For instance, BRACO-19, a well-studied quadruplex-interacting molecule, was found to reduce the transcriptional activity and DNA replication of EBV by interacting with G4 structures found in the virus RNA [77]. A naphthalene diimide was also found to bind and stabilize various quadruplex-forming sequences of the HSV-1 genome, resulting in strong inhibition of viral DNA replication [78]. Other examples include TMPyP4, a DNA-binding porphyrin which was described to target G4 regions in HIV genome hindering the virus infectivity [79]. More recently, the same compound was shown to block viral replication in two different Jurkat-derived T-cell lines with established HIV latency, and this effect was enhanced by association with DNA damage repair inhibitors [80]. Targeting G4s in the viral genomes of viruses that undergo latency such as HPV have great therapeutic potential and should be pursued. Latent infection is characterized by the maintenance of the virus genome in the human host during its lifetime. The latent virus may reactivate at any time to enter the productive stage. Current therapies that normally target viral proteins as described above fail to remove the latent virus from its host. Selectively targeting the viral genome in a G4-mediated approach would remove both the replicating virus and the latent one, potentially eradicating so far incurable infective agents [73]. Furthermore, by inhibiting HPV DNA replication, this strategy would suppress HPV integration events, which would presumably lead to decreased HPV-induced cervical cancers. More details on G4-targeted drug development is found on **Chapter 2**.

Finally, aiming at interrupting virus-host interactions, targeting host proteins which may regulate viral processes by binding viral DNA emerges as a promising therapeutic strategy for HPV-induced cervical cancer. Due to the limited coding capability of its small genome, HPV heavily relies on host cell proteins for important viral processes such as transcription, replication and segregation [81]. Nucleolin is an abundant, ubiquitously expressed multifunctional protein that is found in the nucleus, cytoplasm and on the cell surface, and regulates many aspects of DNA and RNA metabolism such as ribosome biogenesis, RNA polymerase I transcription and nucleo-cytoplasmic transport [82,83]. Nucleolin is overexpressed in the surface of cervical cancer cells [84] and was shown to be involved in the activation of HPV18 oncogene transcription and to be related with HPV16 genome stability maintenance [85,86]. In HPV18-positive cervical cancer cells, nucleolin selectively controls the proliferation of cancer cells by activating E6 and E7 oncogene

transcription [86]. The protein binds to the HPV18 enhancer in a sequence-specific manner acting as an S phase-specific transcription activator [86]. As to HPV16, nucleolin binds to its genome in a sequence-dependent manner, in a region (nucleotide 531 to 780) known to be the target of many cellular transcription factors [85,87]. Contrarily to HPV18, nucleolin showed no obvious effect on HPV transcription activity of HPV16. Rather, nucleolin may facilitate E1/E2-independent maintenance of the HPV genome as it was able to keep the stability of a plasmid containing an HPV16 insert in human cells, even in the absence of viral E1 and E2 proteins [85]. Given its important role in the HPV life cycle, nucleolin is a potential target for the development of HPV and cervical cancer cell-targeted drugs with therapeutic effect against HPV infection and carcinogenesis. A nucleolin-targeted drug, the aptamer AS1411, reached phase II clinical trials for the treatment of metastatic renal cell carcinoma and acute myeloid leukemia [88,89]. Nucleic acid aptamers are biologically active oligonucleotides whose activity is analogous to antibodies being able to bind proteins via shape-specific recognition while altering the target's function [90]. Aptamers have advantages over monoclonal antibodies, since these are non-immunogenic, thermodynamically and chemically stable, having increased resistance to serum nucleases and enhanced cellular uptake [91]. AS1411 is a 26-mer guanine-rich DNA aptamer that targets cancer cell surface nucleolin with remarkable cancer-selective antiproliferative activity against a variety of cancer cell lines, cervical cancer included [92–94]. Despite the suboptimal pharmacology and low potency demonstrated on clinical trials, its safety profile and ability to induce durable responses suggest that AS1411 still has some potential for clinical application, whether by combining it with other therapeutic agents or by performing some modifications on its structure. Furthermore, its cancer-specific targeting properties make it a potential drug delivery agent for HPV-positive cervical cancer cells. Nucleic acid aptamers with therapeutic antiviral action have been described, particularly against HIV [95–98]. As an example, 93del is a potent anti-HIV agent that targets HIV-1 RNase H and inhibits viral entry, reverse transcription and integration [99,100]. Remarkably, the cellular uptake of the aptamer is enhanced in the presence of viral particles which emphasizes the potential of nucleic acid aptamers as potential antiviral agents [99]. More details on nucleic acid aptamers and the use of AS1411 as a drug delivery agent are provided on **Chapter 2**.

Apart from viral genomes, G4s are also found in the genome of human cells, being particularly enriched in regulatory regions such as gene promoters (notably oncogenes), telomere ends and 5'-untranslated regions (5'-UTR) of mRNAs [101]. Like viral G4s, these structures have been implicated in the regulation of several key processes, such as telomere elongation and maintenance, DNA replication, transcription and translation

[101]. Due to their great implications for various biological events and diseases such as cancer, G4s have attracted vast interest as potential therapeutic targets [102]. For instance, targeting the G4 structures formed within telomeric regions and oncogene promoters with small molecules, prompted strong anticancer effects by interfering with telomerase functions and modulating oncogene expression [102]. This therapeutic strategy has great potential of being applied in the treatment of cervical cancer and is thoroughly reviewed in **Chapter 2** where the clinical applicability of G4s as anticancer targets is discussed. Altogether, these novel targets and strategies offer exciting prospects for the development of HPV-specific therapeutics with potential of clearing viral infection and stopping cervical cancer development and progression and are the main topic of this Thesis.

Having discussed the molecular mechanisms of HPV-induced carcinogenesis (**Chapter 1**) and the potential of G4 nucleic acids for the development of drugs aiming at treating both HPV infection and cervical cancer (**Chapter 2**), the detailed Aims and Outline of this Thesis are presented in **Chapter 3**.

References

- [1] Bruni, L., Albero, G., Serrano, B., Mena, M., Gómez, D., Muñoz, J., Bosch, F. X., and de Sanjosé, S. (2019) Human Papillomavirus and Related Diseases in the World. *ICO/IARC Information Centre on HPV and Cancer (HPV Information Centre)*.
- [2] Cohen, P. A., Jhingran, A., Oaknin, A., and Denny, L. (2019) Cervical cancer. *The Lancet* 393, 169–182. [https://doi.org/10.1016/S0140-6736\(18\)32470-X](https://doi.org/10.1016/S0140-6736(18)32470-X).
- [3] Jenkins, D. (2007) Histopathology and Cytopathology of Cervical Cancer. *Disease Markers* 23, 199–212. <https://doi.org/10.1155/2007/874795>.
- [4] Schiffman, M., Doorbar, J., Wentzensen, N., De Sanjosé, S., Fakhry, C., Monk, B. J., Stanley, M. A., and Franceschi, S. (2016) Carcinogenic human papillomavirus infection. *Nature Reviews Disease Primers* 2, 16086. <https://doi.org/10.1038/nrdp.2016.86>.
- [5] Tsikouras, P., Zervoudis, S., Manav, B., Tomara, E., Iatrakis, G., Romanidis, C., Bothou, A., and Galazios, G. (2016) Cervical cancer: Screening, diagnosis and staging. *Journal of B.U.ON.* 21, 320–325.
- [6] Gheit, T. (2019) Mucosal and Cutaneous Human Papillomavirus Infections and Cancer Biology. *Frontiers in Oncology* 9. <https://doi.org/10.3389/fonc.2019.00355>.

- [7] Louie, K. S., de Sanjose, S., Diaz, M., Castellsagué, X., Herrero, R., Meijer, C. J., Shah, K., Franceschi, S., Muñoz, N., and Bosch, F. X. (2009) Early age at first sexual intercourse and early pregnancy are risk factors for cervical cancer in developing countries. *British Journal of Cancer* 100, 1191–1197. <https://doi.org/10.1038/sj.bjc.6604974>.
- [8] Liu, Z.-C., Liu, W.-D., Liu, Y.-H., Ye, X.-H., and Chen, S.-D. (2015) Multiple Sexual Partners as a Potential Independent Risk Factor for Cervical Cancer: a Meta-analysis of Epidemiological Studies. *Asian Pacific Journal of Cancer Prevention* 16, 3893–3900. <https://doi.org/10.7314/APJCP.2015.16.9.3893>.
- [9] Wieland, U., Kreuter, A., and Pfister, H. (2014) Human Papillomavirus and Immunosuppression, in *Current Problems in Dermatology (Switzerland)*, pp 154–165. <https://doi.org/10.1159/000357907>.
- [10] Nunobiki, O., Ueda, M., Toji, E., Yamamoto, M., Akashi, K., Sato, N., Izuma, S., Torii, K., Tanaka, I., Okamoto, Y., and Noda, S. (2011) Genetic Polymorphism of Cancer Susceptibility Genes and HPV Infection in Cervical Carcinogenesis. *Pathology Research International* 2011, 1–8. <https://doi.org/10.4061/2011/364069>.
- [11] Oh, H. Y., Seo, S.-S., Kim, M. K., Lee, D. O., Chung, Y. K., Lim, M. C., Kim, J.-Y., Lee, C. W., and Park, S.-Y. (2014) Synergistic Effect of Viral Load and Alcohol Consumption on the Risk of Persistent High-Risk Human Papillomavirus Infection. *PLoS ONE* (Liu, X., Ed.) 9, e104374. <https://doi.org/10.1371/journal.pone.0104374>.
- [12] Giuliano, A. R., Sedjo, R. L., Roe, D. J., Harris, R., Baldwin, S., Papenfuss, M. R., Abrahamsen, M., and Inserra, P. (2002) Clearance of oncogenic human papillomavirus (HPV) infection: Effect of smoking. *Cancer Causes and Control*. <https://doi.org/10.1023/A:1020668232219>.
- [13] Wu, T.-C. (2008) Cervarix®: a vaccine for the prevention of HPV 16, 18-associated cervical cancer. *Biologics: Targets & Therapy* 107. <https://doi.org/10.2147/BTT.S1877>.
- [14] Shi, L., Sings, H. L., Bryan, J. T., Wang, B., Wang, Y., Mach, H., Kosinski, M., Washabaugh, M. W., Sitrin, R., and Barr, E. (2007) GARDASIL®: Prophylactic Human Papillomavirus Vaccine Development – From Bench Top to Bed-side. *Clinical Pharmacology & Therapeutics* 81, 259–264. <https://doi.org/10.1038/sj.clpt.6100055>.
- [15] Angioli, R., Lopez, S., Aloisi, A., Terranova, C., De Cicco, C., Scaletta, G., Capriglione, S., Miranda, A., Luvero, D., Ricciardi, R., Montera, R., and Plotti, F. (2016) Ten years of

HPV vaccines: State of art and controversies. *Critical Reviews in Oncology/Hematology* 102, 65–72. <https://doi.org/10.1016/j.critrevonc.2016.03.020>.

[16] Sawaya, G. F., Smith-McCune, K., and Kuppermann, M. (2019) Cervical Cancer Screening. *JAMA* 321, 2018. <https://doi.org/10.1001/jama.2019.4595>.

[17] Waxman, A. G., Chelmow, D., Darragh, T. M., Lawson, H., and Moscicki, A.-B. (2012) Revised Terminology for Cervical Histopathology and Its Implications for Management of High-Grade Squamous Intraepithelial Lesions of the Cervix. *Obstetrics & Gynecology* 120, 1465–1471. <https://doi.org/10.1097/AOG.obo13e31827001d5>.

[18] Dueñas-González, A., and Campbell, S. (2016) Global strategies for the treatment of early-stage and advanced cervical cancer. *Current Opinion in Obstetrics and Gynecology* 28, 11–17. <https://doi.org/10.1097/GCO.0000000000000234>.

[19] Zhou, C., Tuong, Z. K., and Frazer, I. H. (2019) Papillomavirus Immune Evasion Strategies Target the Infected Cell and the Local Immune System. *Frontiers in Oncology* 9. <https://doi.org/10.3389/fonc.2019.00682>.

[20] Harden, M. E., and Munger, K. (2017) Human papillomavirus molecular biology. *Mutation Research - Reviews in Mutation Research* 772, 3–12. <https://doi.org/10.1016/j.mrrev.2016.07.002>.

[21] Liu, Y., Pan, Y., Gao, W., Ke, Y., and Lu, Z. (2017) Whole-Genome Analysis of Human Papillomavirus Types 16, 18, and 58 Isolated from Cervical Precancer and Cancer Samples in Chinese Women. *Scientific Reports* 7, 263. <https://doi.org/10.1038/s41598-017-00364-9>.

[22] Doorbar, J., Egawa, N., Griffin, H., Kranjec, C., and Murakami, I. (2015) Human papillomavirus molecular biology and disease association. *Reviews in Medical Virology* 25, 2–23. <https://doi.org/10.1002/rmv.1822>.

[23] Harari, A., Chen, Z., and Burk, R. D. (2014) Human papillomavirus genomics: Past, present and future. *Current Problems in Dermatology (Switzerland)* 45, 1–18. <https://doi.org/10.1159/000355952>.

[24] Van Doorslaer, K., and McBride, A. A. (2016) Molecular archeological evidence in support of the repeated loss of a papillomavirus gene. *Scientific Reports* 6, 33028. <https://doi.org/10.1038/srep33028>.

[25] Egawa, N., Egawa, K., Griffin, H., and Doorbar, J. (2015) Human Papillomaviruses;

Epithelial Tropisms, and the Development of Neoplasia. *Viruses* 7, 3863–3890. <https://doi.org/10.3390/v7072802>.

[26] Finnen, R. L., Erickson, K. D., Chen, X. S., and Garcea, R. L. (2003) Interactions between Papillomavirus L1 and L2 Capsid Proteins. *Journal of Virology* 77, 4818–4826. <https://doi.org/10.1128/JVI.77.8.4818-4826.2003>.

[27] Richards, R. M., Lowy, D. R., Schiller, J. T., and Day, P. M. (2006) Cleavage of the papillomavirus minor capsid protein, L2, at a furin consensus site is necessary for infection. *Proceedings of the National Academy of Sciences* 103, 1522–1527. <https://doi.org/10.1073/pnas.0508815103>.

[28] Joyce, J. G., Tung, J.-S., Przysiecki, C. T., Cook, J. C., Lehman, E. D., Sands, J. A., Jansen, K. U., and Keller, P. M. (1999) The L1 Major Capsid Protein of Human Papillomavirus Type 11 Recombinant Virus-like Particles Interacts with Heparin and Cell-surface Glycosaminoglycans on Human Keratinocytes. *Journal of Biological Chemistry* 274, 5810–5822. <https://doi.org/10.1074/jbc.274.9.5810>.

[29] Schäfer, G., Blumenthal, M., and Katz, A. (2015) Interaction of Human Tumor Viruses with Host Cell Surface Receptors and Cell Entry. *Viruses* 7, 2592–2617. <https://doi.org/10.3390/v7052592>.

[30] Schelhaas, M., Shah, B., Holzer, M., Blattmann, P., Kühling, L., Day, P. M., Schiller, J. T., and Helenius, A. (2012) Entry of Human Papillomavirus Type 16 by Actin-Dependent, Clathrin- and Lipid Raft-Independent Endocytosis. *PLoS Pathogens* (Meyers, C., Ed.) 8, e1002657. <https://doi.org/10.1371/journal.ppat.1002657>.

[31] Day, P. M., and Schelhaas, M. (2014) Concepts of papillomavirus entry into host cells. *Current Opinion in Virology* 4, 24–31. <https://doi.org/10.1016/j.coviro.2013.11.002>.

[32] Pyeon, D., Pearce, S. M., Lank, S. M., Ahlquist, P., and Lambert, P. F. (2009) Establishment of Human Papillomavirus Infection Requires Cell Cycle Progression. *PLoS Pathogens* (Münger, K., Ed.) 5, e1000318. <https://doi.org/10.1371/journal.ppat.1000318>.

[33] Burd, E. M., and Dean, C. L. (2016) Human Papillomavirus. *Microbiology Spectrum* 4, 1–16. <https://doi.org/10.1128/microbiolspec.DMIH2-0001-2015>.

[34] Moody, C. A., and Laimins, L. A. (2010) Human papillomavirus oncoproteins: pathways to transformation. *Nature Reviews Cancer* 10, 550–560. <https://doi.org/10.1038/nrc2886>.

- [35] Graham, S. (2017) Keratinocyte Differentiation-Dependent Human Papillomavirus Gene Regulation. *Viruses* 9, 245. <https://doi.org/10.3390/v9090245>.
- [36] Thomas, J. T., Hubert, W. G., Ruesch, M. N., and Laimins, L. A. (1999) Human papillomavirus type 31 oncoproteins E6 and E7 are required for the maintenance of episomes during the viral life cycle in normal human keratinocytes. *Proceedings of the National Academy of Sciences* 96, 8449–8454. <https://doi.org/10.1073/pnas.96.15.8449>.
- [37] Li, Y., and Xu, C. (2017) Human papillomavirus-related cancers. *Advances in Experimental Medicine and Biology* 1018, 23–34. https://doi.org/10.1007/978-981-10-5765-6_3.
- [38] Yeo-Teh, N., Ito, Y., and Jha, S. (2018) High-Risk Human Papillomaviral Oncogenes E6 and E7 Target Key Cellular Pathways to Achieve Oncogenesis. *International Journal of Molecular Sciences* 19, 1706. <https://doi.org/10.3390/ijms19061706>.
- [39] Coelho, M. A., de Carné Trécesson, S., Rana, S., Zecchin, D., Moore, C., Molina-Arcas, M., East, P., Spencer-Dene, B., Nye, E., Barnouin, K., Snijders, A. P., Lai, W. S., Blackshear, P. J., and Downward, J. (2017) Oncogenic RAS Signaling Promotes Tumor Immuno-resistance by Stabilizing PD-L1 mRNA. *Immunity* 47, 1083-1099.e6. <https://doi.org/10.1016/j.immuni.2017.11.016>.
- [40] Klingelhutz, A. J., Foster, S. A., and McDougall, J. K. (1996) Telomerase activation by the E6 gene product of human papillomavirus type 16. *Nature* 380, 79–82. <https://doi.org/10.1038/380079a0>.
- [41] Estêvão, D., Costa, N. R., Gil da Costa, R. M., and Medeiros, R. (2019) Hallmarks of HPV carcinogenesis: The role of E6, E7 and E5 oncoproteins in cellular malignancy. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1862, 153–162. <https://doi.org/10.1016/j.bbagr.2019.01.001>.
- [42] McBride, A. A. (2008) Replication and Partitioning of Papillomavirus Genomes, in *Advances in Virus Research*, pp 155–205. [https://doi.org/10.1016/S0065-3527\(08\)00404-1](https://doi.org/10.1016/S0065-3527(08)00404-1).
- [43] Buck, C. B., Thompson, C. D., Pang, Y.-Y. S., Lowy, D. R., and Schiller, J. T. (2005) Maturation of Papillomavirus Capsids. *Journal of Virology* 79, 2839–2846. <https://doi.org/10.1128/JVI.79.5.2839-2846.2005>.
- [44] Doorbar, J. (2013) The E4 protein; structure, function and patterns of expression.

Virology 445, 80–98. <https://doi.org/10.1016/j.virol.2013.07.008>.

[45] Burley, M., Roberts, S., and Parish, J. L. (2020) Epigenetic regulation of human papillomavirus transcription in the productive virus life cycle. *Seminars in Immunopathology* 42, 159–171. <https://doi.org/10.1007/s00281-019-00773-0>.

[46] Weitzman, M. D., and Fradet-Turcotte, A. (2018) Virus DNA Replication and the Host DNA Damage Response. *Annual Review of Virology* 5, 141–164. <https://doi.org/10.1146/annurev-virology-092917-043534>.

[47] Oyervides-Muñoz, M. A., Pérez-Maya, A. A., Rodríguez-Gutiérrez, H. F., Gómez-Macias, G. S., Fajardo-Ramírez, O. R., Treviño, V., Barrera-Saldaña, H. A., and Garza-Rodríguez, M. L. (2018) Understanding the HPV integration and its progression to cervical cancer. *Infection, Genetics and Evolution* 61, 134–144. <https://doi.org/10.1016/j.meegid.2018.03.003>.

[48] Gao, G., Johnson, S. H., Vasmatazis, G., Pauley, C. E., Tombers, N. M., Kasperbauer, J. L., and Smith, D. I. (2017) Common fragile sites (CFS) and extremely large CFS genes are targets for human papillomavirus integrations and chromosome rearrangements in oropharyngeal squamous cell carcinoma. *Genes, Chromosomes and Cancer* 56, 59–74. <https://doi.org/10.1002/gcc.22415>.

[49] Graham, S. V. (2017) The human papillomavirus replication cycle, and its links to cancer progression: a comprehensive review. *Clinical Science* 131, 2201–2221. <https://doi.org/10.1042/CS20160786>.

[50] Shanmugasundaram, S., and You, J. (2017) Targeting persistent human papillomavirus infection. *Viruses* 9, 229. <https://doi.org/10.3390/v9080229>.

[51] Hu, Z., and Ma, D. (2018) The precision prevention and therapy of HPV-related cervical cancer: new concepts and clinical implications. *Cancer Medicine* 7, 5217–5236. <https://doi.org/10.1002/cam4.1501>.

[52] Bharti, A. C., Singh, T., Bhat, A., Pande, D., and Jadli, M. (2018) Therapeutic strategies for human papillomavirus infection and associated cancers. *Frontiers in Bioscience - Elite* 10, 15–73. <https://doi.org/10.2741/e808>.

[53] Garbuglia, A. R., Lapa, D., Sias, C., Capobianchi, M. R., and Del Porto, P. (2020) The Use of Both Therapeutic and Prophylactic Vaccines in the Therapy of Papillomavirus Disease. *Frontiers in Immunology* 11, 1–14. <https://doi.org/10.3389/fimmu.2020.00188>.

- [54] Almeida, A. M., Queiroz, J. A., Sousa, F., and Sousa, Â. (2019) Cervical cancer and HPV infection: ongoing therapeutic research to counteract the action of E6 and E7 oncoproteins. *Drug Discovery Today* 24, 2044–2057. <https://doi.org/10.1016/j.drudis.2019.07.011>.
- [55] Maciag, P. C., Radulovic, S., and Rothman, J. (2009) The first clinical use of a live-attenuated *Listeria monocytogenes* vaccine: A Phase I safety study of Lm-LLO-E7 in patients with advanced carcinoma of the cervix. *Vaccine* 27, 3975–3983. <https://doi.org/10.1016/j.vaccine.2009.04.041>.
- [56] Trimble, C. L., Morrow, M. P., Kraynyak, K. A., Shen, X., Dallas, M., Yan, J., Edwards, L., Parker, R. L., Denny, L., Giffear, M., Brown, A. S., Marcozzi-Pierce, K., Shah, D., Slager, A. M., Sylvester, A. J., Khan, A., Broderick, K. E., Juba, R. J., Herring, T. A., Boyer, J., Lee, J., Sardesai, N. Y., Weiner, D. B., and Bagarazzi, M. L. (2015) Safety, efficacy, and immunogenicity of VGX-3100, a therapeutic synthetic DNA vaccine targeting human papillomavirus 16 and 18 E6 and E7 proteins for cervical intraepithelial neoplasia 2/3: a randomised, double-blind, placebo-controlled phase 2b trial. *The Lancet* 386, 2078–2088. [https://doi.org/10.1016/S0140-6736\(15\)00239-1](https://doi.org/10.1016/S0140-6736(15)00239-1).
- [57] Lee, H., Lee, S., and Heo, Y.-S. (2019) Molecular Interactions of Antibody Drugs Targeting PD-1, PD-L1, and CTLA-4 in Immuno-Oncology. *Molecules* 24, 1190. <https://doi.org/10.3390/molecules24061190>.
- [58] Chung, H. C., Ros, W., Delord, J.-P., Perets, R., Italiano, A., Shapira-Frommer, R., Manzuk, L., Piha-Paul, S. A., Xu, L., Zeigenfuss, S., Pruitt, S. K., and Leary, A. (2019) Efficacy and Safety of Pembrolizumab in Previously Treated Advanced Cervical Cancer: Results From the Phase II KEYNOTE-158 Study. *Journal of Clinical Oncology* 37, 1470–1478. <https://doi.org/10.1200/JCO.18.01265>.
- [59] Rotman, J., Mom, C. H., Jordanova, E. S., de Gruijl, T. D., and Kenter, G. G. (2018) ‘DURVIT’: a phase-I trial of single low-dose durvalumab (Medi4736) IntraTumourally injected in cervical cancer: safety, toxicity and effect on the primary tumour- and lymph node microenvironment. *BMC Cancer* 18, 888. <https://doi.org/10.1186/s12885-018-4764-0>.
- [60] Chang, J. T. C., Kuo, T.-F., Chen, Y.-J., Chiu, C.-C., Lu, Y.-C., Li, H.-F., Shen, C.-R., and Cheng, A.-J. (2010) Highly potent and specific siRNAs against E6 or E7 genes of HPV16- or HPV18-infected cervical cancers. *Cancer Gene Therapy* 17, 827–836. <https://doi.org/10.1038/cgt.2010.38>.

- [61] Butz, K., Ristriani, T., Hengstermann, A., Denk, C., Scheffner, M., and Hoppe-Seyler, F. (2003) siRNA targeting of the viral E6 oncogene efficiently kills human papillomavirus-positive cancer cells. *Oncogene* 22, 5938–5945. <https://doi.org/10.1038/sj.onc.1206894>.
- [62] Jung, H., Phillips, B. L., and Chan, E. K. L. (2014) miR-375 activates p21 and suppresses telomerase activity by coordinately regulating HPV E6/E7, E6AP, CIP2A, and 14-3-3 ζ . *Molecular Cancer* 13, 80. <https://doi.org/10.1186/1476-4598-13-80>.
- [63] D'Abramo, C. M., and Archambault, J. (2011) Small Molecule Inhibitors of Human Papillomavirus Protein - Protein Interactions. *The Open Virology Journal* 5, 80–95. <https://doi.org/10.2174/1874357901105010080>.
- [64] Malecka, K. A., Fera, D., Schultz, D. C., Hodawadekar, S., Reichman, M., Donover, P. S., Murphy, M. E., and Marmorstein, R. (2014) Identification and Characterization of Small Molecule Human Papillomavirus E6 Inhibitors. *ACS Chemical Biology* 9, 1603–1612. <https://doi.org/10.1021/cb500229d>.
- [65] Martinez-Zapien, D., Ruiz, F. X., Poirson, J., Mitschler, A., Ramirez, J., Forster, A., Cousido-Siah, A., Masson, M., Pol, S. Vande, Podjarny, A., Travé, G., and Zanier, K. (2016) Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. *Nature* 529, 541–545. <https://doi.org/10.1038/nature16481>.
- [66] White, P. W., Faucher, A.-M., and Goudreau, N. (2010) Small Molecule Inhibitors of the Human Papillomavirus E1-E2 Interaction, in *Current Topics in Microbiology and Immunology*, pp 61–88. https://doi.org/10.1007/82_2010_92.
- [67] Archambault, J., and Melendy, T. (2013) Targeting human papillomavirus genome replication for antiviral drug discovery. *Antiviral Therapy* 18, 271–283. <https://doi.org/10.3851/IMP2612>.
- [68] White, P. W., Faucher, A.-M., Massariol, M.-J., Welchner, E., Rancourt, J., Cartier, M., and Archambault, J. (2005) Biphenylsulfonacetic Acid Inhibitors of the Human Papillomavirus Type 6 E1 Helicase Inhibit ATP Hydrolysis by an Allosteric Mechanism Involving Tyrosine 486. *Antimicrobial Agents and Chemotherapy* 49, 4834–4842. <https://doi.org/10.1128/AAC.49.12.4834-4842.2005>.
- [69] Marušič, M., and Plavec, J. (2019) Towards Understanding of Polymorphism of the G-rich Region of Human Papillomavirus Type 52. *Molecules* 24, 1294. <https://doi.org/10.3390/molecules24071294>.

- [70] Tlučková, K., Marušič, M., Tóthová, P., Bauer, L., Šket, P., Plavec, J., and Viglasky, V. (2013) Human papillomavirus G-quadruplexes. *Biochemistry* 52, 7207–7216. <https://doi.org/10.1021/bi400897g>.
- [71] Marušič, M., Hošnjak, L., Krafčikova, P., Poljak, M., Viglasky, V., and Plavec, J. (2017) The effect of single nucleotide polymorphisms in G-rich regions of high-risk human papillomaviruses on structural diversity of DNA. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1861, 1229–1236. <https://doi.org/10.1016/j.bbagen.2016.11.007>.
- [72] Carvalho, J., Queiroz, J. A., and Cruz, C. (2017) Circular dichroism of G-Quadruplex: A laboratory experiment for the study of topology and ligand binding. *Journal of Chemical Education* 94, 1547–1551. <https://doi.org/10.1021/acs.jchemed.7b00160>.
- [73] Ruggiero, E., and Richter, S. N. (2018) G-quadruplexes and G-quadruplex ligands: targets and tools in antiviral therapy. *Nucleic Acids Research* 46, 3270–3283. <https://doi.org/10.1093/nar/gky187>.
- [74] Metifiot, M., Amrane, S., Litvak, S., and Andreola, M.-L. (2014) G-quadruplexes in viruses: function and potential therapeutic applications. *Nucleic Acids Research* 42, 12352–12366. <https://doi.org/10.1093/nar/gku999>.
- [75] Lavezzo, E., Berselli, M., Frasson, I., Perrone, R., Palù, G., Brazzale, A. R., Richter, S. N., and Toppo, S. (2018) G-quadruplex forming sequences in the genome of all known human viruses: A comprehensive guide. *PLOS Computational Biology* (Lexa, M., Ed.) 14, e1006675. <https://doi.org/10.1371/journal.pcbi.1006675>.
- [76] Ruggiero, E., and Richter, S. N. (2020) Viral G-quadruplexes: New frontiers in virus pathogenesis and antiviral therapy. *Annual Reports in Medicinal Chemistry*. <https://doi.org/10.1016/bs.armc.2020.04.001>.
- [77] Norseen, J., Johnson, F. B., and Lieberman, P. M. (2009) Role for G-Quadruplex RNA Binding by Epstein-Barr Virus Nuclear Antigen 1 in DNA Replication and Metaphase Chromosome Attachment. *Journal of Virology* 83, 10336–10346. <https://doi.org/10.1128/JVI.00747-09>.
- [78] Callegaro, S., Perrone, R., Scalabrin, M., Doria, F., Palù, G., and Richter, S. N. (2017) A core extended naphthalene diimide G-quadruplex ligand potently inhibits herpes simplex virus 1 replication. *Scientific Reports* 7, 1–9. <https://doi.org/10.1038/s41598-017-02667-3>.

- [79] Perrone, R., Nadai, M., Poe, J. A., Frasson, I., Palumbo, M., Palù, G., Smithgall, T. E., and Richter, S. N. (2013) Formation of a Unique Cluster of G-Quadruplex Structures in the HIV-1 nef Coding Region: Implications for Antiviral Activity. *PLoS ONE* (Qiu, J., Ed.) 8, e73121. <https://doi.org/10.1371/journal.pone.0073121>.
- [80] Piekna-Przybylska, D., Sharma, G., Maggirwar, S. B., and Bambara, R. A. (2017) Deficiency in DNA damage response, a new characteristic of cells infected with latent HIV-1. *Cell Cycle* 16, 968–978. <https://doi.org/10.1080/15384101.2017.1312225>.
- [81] Thierry, F. (2009) Transcriptional regulation of the papillomavirus oncogenes by cellular and viral transcription factors in cervical carcinoma. *Virology* 384, 375–379. <https://doi.org/10.1016/j.virol.2008.11.014>.
- [82] Tajrishi, M. M., Tuteja, R., and Tuteja, N. (2011) Nucleolin. *Communicative & Integrative Biology* 4, 267–275. <https://doi.org/10.4161/cib.4.3.14884>.
- [83] Mongelard, F., and Bouvet, P. (2007) Nucleolin: a multiFACeTed protein. *Trends in Cell Biology* 17, 80–86. <https://doi.org/10.1016/j.tcb.2006.11.010>.
- [84] Berger, C. M., Gaume, X., and Bouvet, P. (2015) The roles of nucleolin subcellular localization in cancer. *Biochimie* 113, 78–85. <https://doi.org/10.1016/j.biochi.2015.03.023>.
- [85] Sato, H., Kusumoto-Matsuo, R., Ishii, Y., Mori, S., Nakahara, T., Shinkai-Ouchi, F., Kawana, K., Fujii, T., Taketani, Y., Kanda, T., and Kukimoto, I. (2009) Identification of nucleolin as a protein that binds to human papillomavirus type 16 DNA. *Biochemical and Biophysical Research Communications* 387, 525–530. <https://doi.org/10.1016/j.bbrc.2009.07.055>.
- [86] Grinstein, E., Wernet, P., Snijders, P. J. F. F., Rösl, F., Weinert, I., Jia, W., Kraft, R., Schewe, C., Schwabe, M., Hauptmann, S., Dietel, M., Meijer, C. J. L. M. L. M., and Royer, H.-D. (2002) Nucleolin as activator of human papillomavirus type 18 oncogene transcription in cervical cancer. *The Journal of Experimental Medicine* 196, 1067–1078. <https://doi.org/10.1084/jem.20011053>.
- [87] Sato, K., Takeuchi, T., Kukimoto, I., Mori, S., Yasugi, T., Yano, T., Taketani, Y., and Kanda, T. (2007) Human papillomavirus type 16 P670 promoter is negatively regulated by CCAAT displacement protein. *Virus Genes* 35, 473–481. <https://doi.org/10.1007/s11262-006-0074-8>.

- [88] Rosenberg, J. E., Bambury, R. M., Van Allen, E. M., Drabkin, H. A., Lara, P. N., Harzstark, A. L., Wagle, N., Figlin, R. A., Smith, G. W., Garraway, L. A., Choueiri, T., Erlandsson, F., and Laber, D. A. (2014) A phase II trial of AS1411 (a novel nucleolin-targeted DNA aptamer) in metastatic renal cell carcinoma. *Investigational New Drugs*. <https://doi.org/10.1007/s10637-013-0045-6>.
- [89] Stuart, R. K., Stockerl-Goldstein, K., Cooper, M., Devetten, M., Herzig, R., Medeiros, B., Schiller, G., Wei, A., Acton, G., and Rizzieri, D. (2009) Randomized phase II trial of the nucleolin targeting aptamer AS1411 combined with high-dose cytarabine in relapsed/refractory acute myeloid leukemia (AML). *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 27, 7019.
- [90] Pofahl, M., Wengel, J., and Mayer, G. (2014) Multifunctional nucleic acids for tumor cell treatment. *Nucleic acid therapeutics* 24, 171–7. <https://doi.org/10.1089/nat.2013.0472>.
- [91] Zhu, G., and Chen, X. (2018) Aptamer-based targeted therapy. *Advanced Drug Delivery Reviews* 134, 65–78. <https://doi.org/10.1016/j.addr.2018.08.005>.
- [92] Bates, P. J., Reyes-Reyes, E. M., Malik, M. T., Murphy, E. M., O'Toole, M. G., and Trent, J. O. (2017) G-quadruplex oligonucleotide AS1411 as a cancer-targeting agent: Uses and mechanisms. *Biochimica et Biophysica Acta - General Subjects* 1861, 1414–1428. <https://doi.org/10.1016/j.bbagen.2016.12.015>.
- [93] Bates, P. J., Laber, D. a., Miller, D. M., Thomas, S. D., and Trent, J. O. (2009) Discovery and development of the G-rich oligonucleotide AS1411 as a novel treatment for cancer. *Experimental and Molecular Pathology* 86, 151–164. <https://doi.org/10.1016/j.yexmp.2009.01.004>.
- [94] Girvan, A. C., Teng, Y., Casson, L. K., Thomas, S. D., Jv°liger, S., Ball, M. W., Klein, J. B., Pierce, W. M., Barve, S. S., and Bates, P. J. (2006) AGRO100 inhibits activation of nuclear factor- κ B (NF- κ B) by forming a complex with NF- κ B essential modulator (NEMO) and nucleolin. *Molecular Cancer Therapeutics* 5, 1790–1799. <https://doi.org/10.1158/1535-7163.mct-05-0361>.
- [95] Do, N. Q., Lim, K. W., Teo, M. H., Heddi, B., and Phan, A. T. (2011) Stacking of G-quadruplexes: NMR structure of a G-rich oligonucleotide with potential anti-HIV and anticancer activity. *Nucleic Acids Research* 39, 9448–9457. <https://doi.org/10.1093/nar/gkr539>.

- [96] Musumeci, D., Riccardi, C., and Montesarchio, D. (2015) G-Quadruplex forming oligonucleotides as Anti-HIV agents. *Molecules* 20, 17511–17532. <https://doi.org/10.3390/molecules200917511>.
- [97] Pedersen, E. B., Nielsen, J. T., Nielsen, C., and Filichev, V. V. (2011) Enhanced anti-HIV-1 activity of G-quadruplexes comprising locked nucleic acids and intercalating nucleic acids. *Nucleic Acids Research* 39, 2470–2481. <https://doi.org/10.1093/nar/gkq1133>.
- [98] D'Onofrio, J., Petraccone, L., Erra, E., Martino, L., Di Fabio, G., De Napoli, L., Giancola, C., and Montesarchio, D. (2007) 5'-Modified G-quadruplex forming oligonucleotides endowed with anti-HIV activity: Synthesis and biophysical properties. *Bioconjugate Chemistry* 18, 1194–1204. <https://doi.org/10.1021/bc070062f>.
- [99] Faure-Perraud, A., Métifiot, M., Reigadas, S., Recordon-Pinson, P., Parissi, V., Ventura, M., and Andréola, M.-L. (2011) The guanine-quadruplex aptamer 93del inhibits HIV-1 replication ex vivo by interfering with viral entry, reverse transcription and integration. *Antiviral Therapy* 16, 383–394. <https://doi.org/10.3851/IMP1756>.
- [100] Andreola, M.-L., Pileur, F., Calmels, C., Ventura, M., Tarrago-Litvak, L., Toulmé, J.-J., and Litvak, S. (2001) DNA Aptamers Selected against the HIV-1 RNase H Display in Vitro Antiviral Activity †. *Biochemistry* 40, 10087–10094. <https://doi.org/10.1021/bio108599>.
- [101] Carvalho, J., Mergny, J.-L., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2020) G-quadruplex, Friend or Foe: The Role of the G-quartet in Anticancer Strategies. *Trends in Molecular Medicine* 1–14. <https://doi.org/10.1016/j.molmed.2020.05.002>.
- [102] Asamitsu, S., Obata, S., Yu, Z., Bando, T., and Sugiyama, H. (2019) Recent progress of targeted G-quadruplex-preferred ligands toward cancer therapy. *Molecules* 24, 429. <https://doi.org/10.3390/molecules24030429>.

Chapter 2

G-quadruplex, Friend or Foe: The Role of the G-quartet in Anticancer Strategies

This chapter was published in:

Carvalho, J., Mergny, J.-L., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2020) G-quadruplex, Friend or Foe: The Role of the G-quartet in Anticancer Strategies. *Trends in Molecular Medicine*, In press. <https://doi.org/10.1016/j.molmed.2020.05.002>.

G-quadruplex, Friend or Foe: The Role of the G-quartet in Anticancer Strategies

Abstract

The clinical applicability of G4s as anticancer drugs is currently being evaluated. Several G4 ligands and aptamers are undergoing clinical trials following the notable examples of quarfloxin and AS1411, respectively. In this review, we summarize the latest achievements and breakthroughs in the use of G4 nucleic acids as both therapeutic tools (*friends*, as healing anticancer drugs) and targets (*foes*, within the harmful cancer cell), particularly using aptamers and quadruplex-targeted ligands, respectively. We explore the recent research on synthetic G4 ligands toward the discovery of anticancer therapeutics and their mechanism of action. Additionally, we highlight recent advances in chemical and structural biology that enable the design of specific G4 aptamers to be used as novel anticancer agents.

Role of the G-quadruplex in anticancer therapeutics

Since the century-old discovery of the ability of guanylic acid to fold into higher-order nucleic acid structures [1], G4s are now recognised for their involvement in several key biological functions [2]. Such findings instigated the quest for understanding how G4-related mechanisms may present opportunities for therapeutic intervention, particularly in cancer. Furthermore, while natural G4 structures are prospective targets for cancer treatment, synthetic G4-forming oligonucleotides aimed at cancer-related protein targets have also been developed as therapeutic agents, with great advantages over antibodies such as ease of production, reproducibility, versatile chemical modification, high stability and lack of immunogenicity. Herein, we discuss the biological relevance of the G4 motif, highlighting the latest key findings regarding its use as both a therapeutic target and a tool in anticancer drug development.

The G-quartet

The G4 structure is the most studied non-canonical nucleic acid conformation with close to 300 structures deposited in the Protein Data Bank (PDB). Through X-ray diffraction studies of fibre formed by guanylic acids, Gellert, Lipsett and Davies found in 1962 that tetrameric guanine residues assembled in vertically stacked hydrogen-bonded guanine tetrads (later termed ‘G-quartet’ or ‘G-tetrad’) [3] (**Figure 2.1A**). Gellert *et al.* also reported the high stability of the structural arrangement provided by Hoogsteen hydrogen-bonding of the four guanine bases, with two bonds between each pair involving the N1, N7, O6 and N2 atoms of each base. The stacking of at least two G-quartets in four-stranded right-handed helical structures results in what are generally designated as G4s (**Figure 2.1B**). These structures are stabilized by cations such as potassium (K⁺) and sodium (Na⁺), which interact with the central electronegative carbonyl O6 atoms of the G-quartet core [2]. G4s display a variety of topologies, owing to possible changes in strand stoichiometry (intramolecular or intermolecular), polarity, and orientation (parallel, antiparallel, or hybrid) (**Figure 2.1C-E**), in addition to loop size and sequence [4,5]. The identity of the cation also contributes to the structural variety of the G4 structure [5]. Only two decades after the first fiber diffraction studies, the formation of G4 structures in guanine-rich sequences of immunoglobulin switch regions and at the ends of eukaryotic chromosomes (telomeres) was suggested [6,7]. Guanine-rich sequences were also identified in the promoter regions of several genes, especially oncogenes [8]. Technological advances in terms of nucleotide sequencing and bioinformatics, together with the availability of the complete human genome sequence led to the identification of new sequences with presumed G4 formation at the DNA and RNA level [2]. Potential G4-forming sequences were initially searched using the sequence motif $G_{3+N_{1-7}}G_{3+N_{1-7}}G_{3+N_{1-7}}G_{3+}$, where ‘G’ represents the guanine residues in each guanine tract, usually directly involved in G-tetrad formation, and N represents any combination of loop residues. More than 300,000 putative G4-forming sequences were found in the human genome using the so-called Quadparser algorithm with the aforementioned sequence motif as a query [9]. Notably, these sequences are non-randomly distributed in key regulatory regions such as telomere ends, gene promoters and 5'-untranslated regions (5'-UTR) [2]. However, the question if all these sequences form stable G4s cannot be answered at the time due to insufficient experimental data on structure and stability. Since then, alternative search tools, such as G4-Hunter [10] have been designed (for a recent review, see [11]). G4 forming sequences were also found and studied in other genomes such as protozoa, bacteria and virus [12–19].

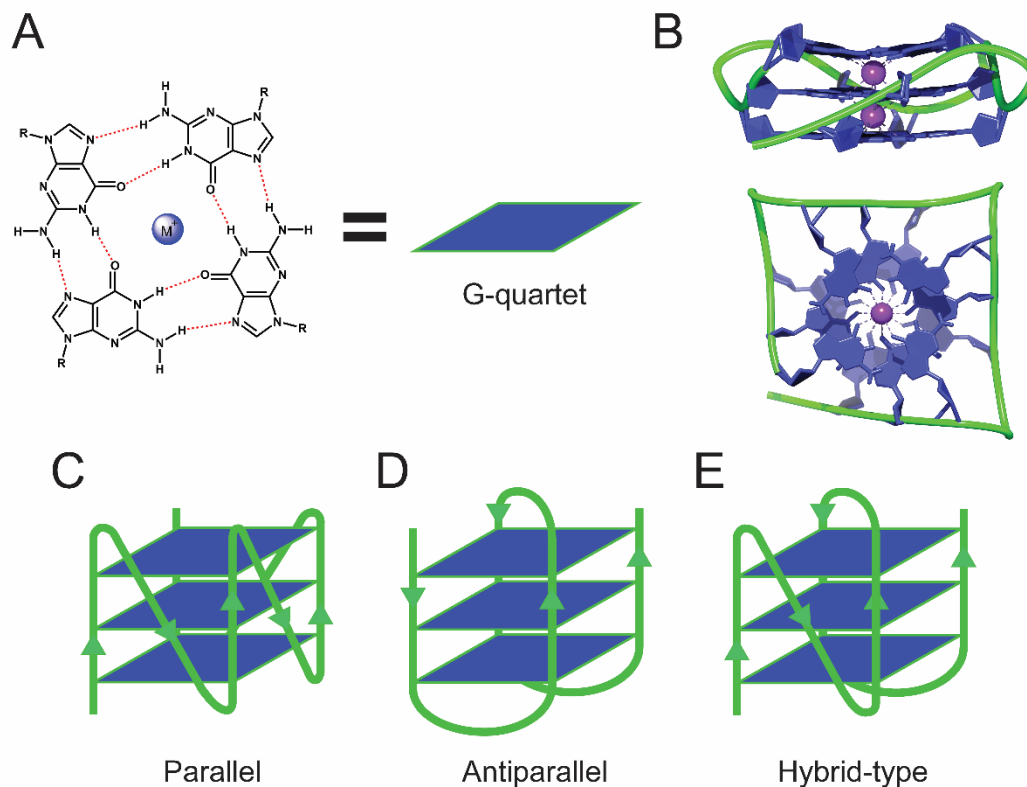


Figure 2.1 - G4 structure. (A) Chemical structure of the G-quartet showing the Hoogsteen hydrogen-bonding between guanine pairs and a central monovalent cation that stabilizes the 3D structure. (B) Side and top view of a model G4 structure found for the human telomeric sequence. The backbone is depicted in green, while guanine bases are represented in blue (loop bases were hidden for clarity). (C), (D) and (E) are representative topologies commonly found in G4 structures.

Regulatory roles of G4s

Evidence for G4 existence *in vivo* has been obtained in living mammalian cells using different strategies [20–24], and these have been implicated in the regulation of several key cellular functions, such as telomere elongation and maintenance [25,26], DNA replication initiation [27] and replication fork progression [28–30], transcription [31,32], translation [33], and genomic instability [29,30] (**Figure 2.2**). Using an antibody-based G4 chromatin immunoprecipitation sequencing (G4 ChIP-seq) approach, the genomic locations involved in the regulation of such cellular processes were shown to be enriched with dynamic G4 structures that strongly depend on chromatin relaxation or cell status [34].

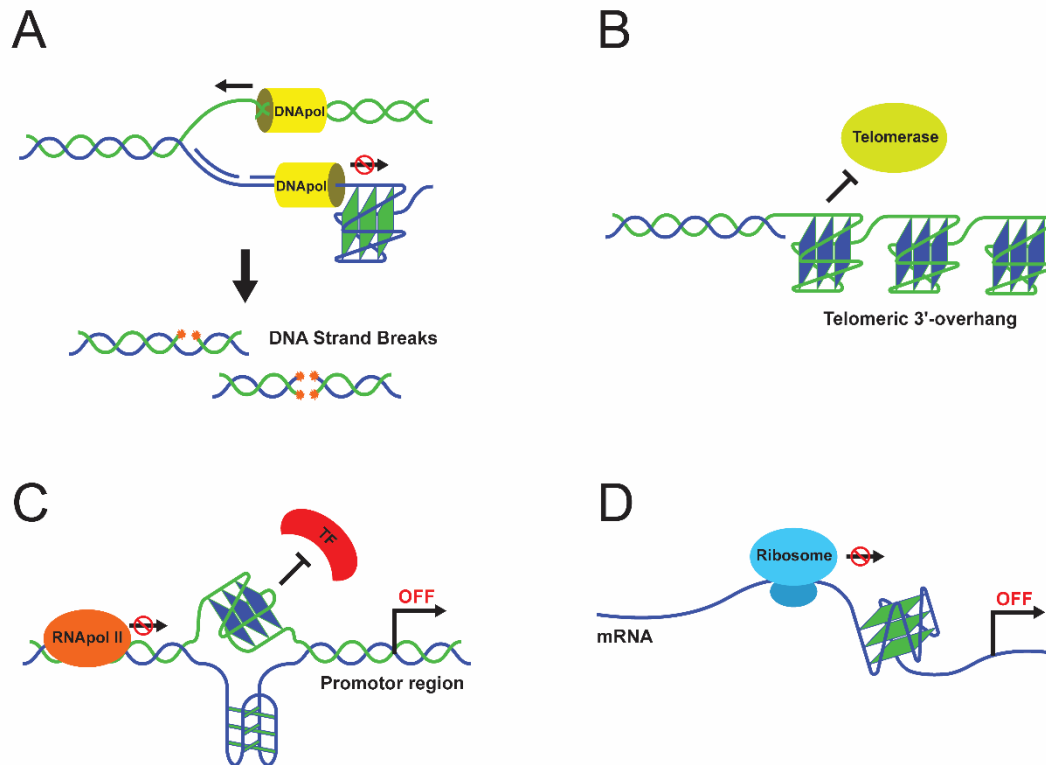


Figure 2.2 - Biological functions of G4 nucleic acids. **(A)** In the absence of any specialized helicase, G4s cause the arrest of DNA polymerase (DNAPol) potentially leading to DNA single- or double-strand breaks causing genetic instability. **(B)** G4s on the 3'-overhang of chromosome ends interfere with telomere elongation by telomerase and affect telomere maintenance. **(C)** Promoter G4s interfere with transcription factor (TF) binding and RNA polymerase (RNAPol) activity, thus altering gene expression at transcriptional level. The complementary strand may form other non-canonical nucleic acid structures, known as i-motifs [35]. **(D)** Gene expression is also regulated by G4 formation in mRNA by inhibiting ribosome scanning or altering splicing, and mRNA maturation and localization [36].

Telomeric G4s were among the first biologically relevant G4 structures to be studied in detail [37]. Human telomeres comprise tandem repeats of the sequence d(TTAGGG), which are able to fold into stable G4 structures [38]. These were already shown to regulate and modulate telomere maintenance [25]. Several proteins involved in the telomere nucleoprotein complex (shelterin) such as POT1 or TPP1, are known to bind and unwind telomeric G4 structures, suggesting the involvement of these structures in the regulation of telomeric functions [39]. Yeast helicase Pif1, which is also able to unwind G4 structures, suppresses both telomere lengthening and G4-associated DNA damage, also demonstrating the association of G4 structures with genomic and epigenetic instability, the latter due to complex genetic-epigenetic events caused by replication impairment leading to gene silencing [40–42].

DNA replication origins, which are found throughout human chromosomes, are known to contain guanine-rich elements in DNA synthesis initiation sites that could form G4

structures [27,43]. Two independent studies showed that either the deletion of these elements in several model origins or the introduction of point mutations that affected G4 stability abolished origin activity in cells suggesting that G4 motifs are required for replication initiation [27,43]. Furthermore, G4-forming oligonucleotides were able to compete for replication factors suggesting a role in the activation of replication by recruiting activating factors such as Sld2, Sld3, GINS, Cdc 45, and Dbp11 [27]. The existence of G4-specialized helicases that unwind these structures *in vivo* during replication fork progression is another example of the role of G4s in replication [44]. These findings suggest an important regulatory function for G4s by backing the replication fork progression process (reviewed in [45]).

As for the role of G4 structures in regulating transcription and translation, G4 ChIP-seq data revealed that the promoters and 5'-UTRs of highly transcribed cancer-related genes are enriched with G4 motifs and that G4 structures in promoters are linked to elevated transcriptional activity in cancer cells, but not in normal cells [34]. The binding sites for transcription factor SP1 [46] and DNA helicases belonging to the RecQ helicase family were also found to contain putative G4 motifs [47]. Mutations causing loss of function in RecQ helicases triggered an altered transcriptional activity, which may result from their failure to properly resolve G4 structures [47]. This suggests that G4 formation impairs the initiation of transcription by inhibiting the binding of transcription factors and RNA polymerase [48]. This mechanism appears to be conserved between vertebrates and microorganisms ranging from yeast to bacteria down to viruses, which emphasizes the importance of G4-mediated regulation of the expression of genes involved in disease progression [17,18,48–50]. Furthermore, G4s also appear to be involved in the regulation of transcriptional activity in non-cancerous cells. Recently, it was proposed that the regulation of mitochondrial nucleic acid homeostasis is G4-dependent [51]. Additionally, using Quadparser and QGRS Mapper algorithms, Cree and Kennedy demonstrated the existence of putative G4 structures in the promoters of genes encoding important metabolic enzymes, receptor proteins, and proteins involved in potassium and sodium channels [52].

G4s as therapeutic targets

Due to their relevant genomic location and the occurrence of several quadruplex-driven genes, G4s have gathered a lot of attention as potential therapeutic targets. Over the past two decades, large amounts of evidence have pointed to the clinical relevance of G4 nucleic acids, particularly in anticancer drug design. Compounds binding to G4s are called

G4 ligands; a list may be found in the G4 ligands database (<http://www.g4ldb.org/ci2/index.php>) [53]. These compounds were initially developed to inhibit telomerase activity, interfere with telomere functions, and/or alter oncogene transcription. Efforts continue to be made to identify new ligands and understand their physico-chemical properties and biological effects (for an overview of recent developments, see [54]). Examples of G4 ligands are provided in **Table 2.1**. Most G4 ligands share defined structural features, such as an extended aromatic core, which favors π - π stacking interactions onto the terminal G-quartets, and one or more flexible sidechains that offer additional binding interactions with the loop-groove interface. Most, but not all, G4 ligands carry one or more positive charges which may be important for electrostatic interactions with the negatively charged DNA/RNA backbone [55], sometimes at the cost of selectivity [56].

Telomeric G4 structures

The first biologically relevant target for G4 studies was telomeric DNA. The structure and stability of telomeres are closely related to cancer due to the upregulation of telomerase expression observed in 80-85% of all human cancers [5]. The rationale behind targeting telomeric G4s lies in the inhibition of telomerase activity thus inducing senescence of tumor cells by stabilizing G4s at the end of chromosomes [2]. The acridine derivative BRACO-19 is an effective and specific ligand for telomeric G4s, and its ability to inhibit telomerase activity was recently shown [57]. Upon incubation in human glioblastoma cells, BRACO-19 triggered a DNA damage response at the telomeres by uncapping the telomeric structures and exposing chromosomal termini to the DNA damage pathway [57]. Telomere uncapping was followed by disassembly of the T-loop structure, particularly by displacing TRF2 and POT1, which resulted in p53 and p21-mediated cell cycle arrest, short-term apoptosis and senescence [57]. The use of natural compounds as G4 ligands or as a prime source of scaffolds for the synthesis of new ligands has been a popular research topic (reviewed in [58]). A recent example is the use of schizocommunin, a fungal natural alkaloid, as starting scaffold for the synthesis of telomeric G4 ligands [59]. One of the synthesized derivatives showed great affinity towards telomeric G4 sequences, with the ability to induce and stabilize the formation of G4 structures in cancer cells as suggested by an increase in quadruplex foci revealed by the BG4 antibody in the nucleus [59]. This triggered a DNA damage response at the telomere level, inducing telomere dysfunction by dissociating telomere binding proteins (TRF2 and POT1), uncapping the telomeres, and producing anaphase bridges [59]. The schizocommunin derivative was later shown to inhibit tumor growth in a cervical squamous cancer xenograft mouse model

with an apparent safe toxicity profile [59]. Telomeric G4 ligands have also been used in combination with other therapeutic methods. RHPS4, which had already been to have strong antitumoral activity in various *in vivo* tumor models, was recently used as a radiosensitizing agent in a glioblastoma multiforme xenograft model [60]. The administration of RHPS4 followed by Ionizing Radiation (IR) exposure was effective in blocking tumor growth for up to 65 days [60]. RHPS4-IR combinatory treatment produced telomere aberration and dysfunctionalities during the first days of treatment which increased the cellular sensitivity to RHPS4 [60]. Interestingly, RHPS4-IR treatment failed to produce similar telomere damage in glioma stem-like cells, but the administration of RHPS4 alone produced a strong antiproliferative effect in these cells, which was suggested to occur due to the induction of replicative stress by reduction of RAD51 and CHK1 levels [60].

Oncogene promoter G4 structures

Another possible therapeutic strategy encompasses the suppression of the expression of oncogenes by stabilizing G4 structures in the promoters of these genes with small ligands. G4 sequences have already been identified in the promoter regions of prominent oncogenes, such as *c-MYC*, *KRAS*, *c-KIT*, *BCL2*, and *VEGF* (see [48] for a review on promoter G4s). In particular, a G4 structure found in the nuclease hypersensitive element III₁ (NHE III₁) of the proto-oncogene *c-MYC* promoter set the paradigm for subsequent studies [48,61]. A recent example is the quinoxaline QN-1, which effectively suppresses tumor growth in a triple-negative breast cancer mouse model [62]. This inhibitory effect was proposed to be due to the stabilization of a G4 structure formed within *c-MYC* promoter, with selectivity over other G4 structures. When compared with other G4-driven oncogenes, such as *BCL2*, *c-KIT*, *VEGF*, and *HRAS*, and the housekeeping gene β -actin, QN-1 was able to selectively downregulate *c-MYC* alone [62]. The *c-MYC* downstream effector Cyclin D1 was also significantly downregulated, thus suggesting that QN-1 affects *c-MYC* in a specific manner [62]. Using an exon-specific assay, *c-MYC* transcription was found to be inhibited by G4-mediated mechanisms, which resulted in cell cycle arrest and apoptosis. Strikingly, QN-1 significantly reduced *in vivo* *c-MYC* expression and tumor growth in 4T1 tumor-bearing mice [62]. Other promising *c-MYC* G4 targeted compounds include the four-leaf clover-like compound IZCZ-3 and thiazole peptide TH3 [63,64]. IZCZ-3, a imidazole/carbazole conjugate, was able to bind and stabilize *c-MYC* G4 with strong affinity with a K_D in the 10^{-7} M range [63]. In cells, IZCZ-3 produced a significant growth inhibitory effect against a variety of cancer cell lines by inducing cell cycle arrest at G₀/G₁ and apoptosis [63]. This effect was later shown to be due to reduced *c-MYC*

expression which resulted in downregulated (Cyclin D1 and CDK6) and upregulated (p15, p27, and PARP) cell cycle and apoptosis regulators, without affecting the expression of β -actin [63]. IZCZ-3 had no effect on the transcription of other G4-driven oncogenes (*VEGF*, *BCL2*, *c-KIT*, *KRAS*, *RET*, *PDGFA*, and *HRAS*). When administered to mice bearing human cervical squamous cancer xenografts, the ligand was equally able to downregulate c-MYC expression with an antitumoral effect comparable with that of doxorubicin [63]. The TH3 thiazole peptide was also found to bind and stabilize c-MYC G4 with a comparable K_D value (10^{-7} M), showing a clear preference towards this structure over other G4s, such as c-KIT or BCL2 [64]. Intriguingly, despite its structure being relatively identical to other groove binders in the literature, TH3 was found to stack on both 5'- and 3'-end G-quartets as suggested by NMR spectroscopy [64]. Using cancer cell lines, TH3 was shown to cross the cell membrane and localize in the nucleus, with significant antiproliferative effect by growth arrest in S phase and induction of apoptosis [64]. Noncancer genes, such as *GAPDH* and 18S rRNA were used as control, as well as G4-driven BCL-2, and the results suggested no significant alteration in their levels. Although these findings suggest that all the aforementioned effects are c-MYC and cancer specific, further *in vivo* evidence may be needed to unambiguously show that the ligands recognize c-MYC G4 in more complex systems.

G4s structures found in oncogene promoters other than c-MYC have also been efficiently targeted by ligands, particularly c-KIT [65], BCL2 [66], and KRAS [67]. The KRAS NHE region contains particular guanine-rich elements capable of folding into stable G4 structures, one of which was recently solved by NMR spectroscopy (KRAS 22RT, PDB accession code: 6T51) [68], and two conformers of a slightly longer version of the KRAS G-rich promoter sequence were identified (PDB accession codes 6T2G and 6SUU). Recent KRAS G4-targeted ligands include porphyrins developed as anticancer agents for pancreatic cancer [69]. Building on the existent knowledge of the G4-binding properties of porphyrins, such as TMPyP4 [70], the ligands tetrakis and octaacetyl were found to strongly stabilize KRAS 32R G4 and to inhibit KRAS expression in PANC-1 and MiaPaCa 2 pancreatic cancer cell lines with strong apoptotic effect [69]. The noncancer gene *GAPDH* was not affected, while other apoptotic response genes such as *Bax* and *P53*, were increased. Although *BCL2* expression was found to be decreased, the authors failed to address whether this could be due to G4-dependent mechanisms because BCL2 is also enriched with G4s in its promoter [66]. Remarkably, both porphyrins were able to hinder metastasis formation by arresting the epithelial to mesenchymal transition (EMT), which is a hallmark of pancreatic cancer [69]. *In vivo* studies further demonstrated the antitumoral effect of these porphyrin derivatives with decreased expression of several

tumor markers and tumor volume upon treatment with nontoxic doses [69]. Other examples of KRAS-targeted ligands include the acridine orange derivative C₈ which was found to bind and stabilize KRAS 22RT G4 with high affinity ($K_D \sim 10^{-7}$ M) [71]. The ligand exhibit strong inhibitory effect against HeLa cervical cancer cells with an IC₅₀ 100-times lower than that of 5-fluorouracil (0.9 - 5.7 μ M) [71], and the ability to decrease oncogene transcription without affecting noncancer genes, such as β -actin [72]. Furthermore, C₈ presented excellent cell-penetrating behavior, localizing particularly in the nucleoli after a short incubation in HeLa cells [71].

Inhibition of oncogene expression can also be achieved by interfering with mRNA translation [36]. For instance, inhibition of RAS proteins was achieved by stabilizing the G4s formed in the NRAS [73] and KRAS [74] mRNAs. The expression of the latter was repressed with an anthrafurandione derivative by stabilizing a G4 structure formed in the 5'-UTR of KRAS mRNA [74]. This resulted in a strong inhibitory effect in Panc-1 pancreatic cancer cells (IC₅₀ = 0.26 μ M). Interestingly, the authors were able to demonstrate that this effect was also replicable using a biotinylated ligand under low-abundance cellular conditions which better simulate *in vivo* conditions compared with the commonly used luciferase assays [74].

Clinical applicability of G4 ligands

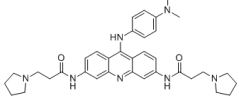
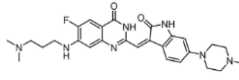
After decades of development, the clinical applicability of these compounds is presently being tested. Some examples encompass CX-3543 [75], a G4 ligand that reached Phase 2 trials as a candidate therapeutic agent against several tumors (NCT00780663), and, more recently, CX-5461 [76] which is currently in advanced Phase 1 clinical trials for patients with BRCA1/2-deficient tumors (NCT02719977). Given that no G4 interacting compound is currently approved by the US Food and Drug Administration (FDA), a large body of evidence should help overcoming the bottlenecks of drug discovery in the G4 field. One of the obstacles hampering clinical applications of G4-interacting compounds appears to be selectivity. Indeed, all the aforementioned antitumoral effects have been linked effects of targeting a particular G4 motif or structure. However, is complete selectivity among G4 structures really needed? While most of the efforts over the past decade focused on designing ligands capable of binding to a particular G4, recent evidence points to multiple G4s being involved in the biological effects of G4 ligands [77]. An interesting example comes from the triarylpyridine 20A, which was shown to bind to telomeric repeats *in vitro* [78] and has recently found to affect several biological pathways dependent on genes in *loci* other than telomeric regions [77]. Transcriptome analysis of 20A-treated HeLa cells

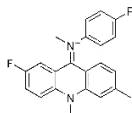
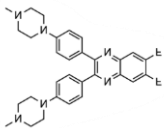
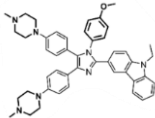
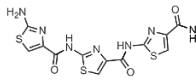
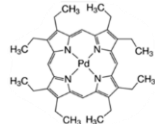
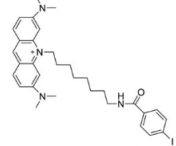
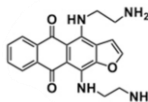
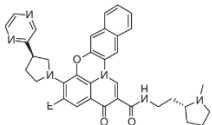
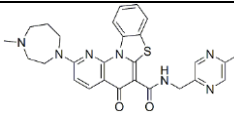
showed the enrichment of pathways related to DNA damage, cell growth, and autophagy, with > 600 genes with up- or downregulated [77]. The downregulated genes were shown to be enriched with G4 motifs compared with unchanged and upregulated genes, which suggests multiple G4-dependent inhibitory effects of 20A [77]. This example deepens the question of whether the selectivity for a particular G4 structure is really needed. Indeed, widely used FDA-approved drugs in clinical practice for cancer therapy are multitargeted compounds, such as the well-known kinase inhibitor crizotinib, which is used for non-small cell lung cancer (NSCLC) treatment [79].

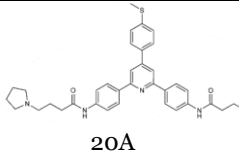
Perspectives for the development of G4 ligands as antitumor agents

An interesting study by Biroccio and colleagues showed that the susceptibility of cancer cells to G4 ligands compared with normal cells is not associated with telomere length or the higher proliferation rate [80]. Rather, it was suggested that cell transformation triggers replication stress and telomere uncapping making cancer cells more prone to DNA damage responses induced by the G4 ligands. More recently, Balasubramanian and colleagues revealed novel disease-related genetic vulnerabilities for two different well-characterized and specific G4-ligands, PhenDC3 and Pyridostatin (PDS) [81]. These screens allowed the identification of several genes involved in a G4 response that can become genetic vulnerabilities to G4 ligands when mutated. The discovery of several genes that, when depleted, enhance cell killing with the G4 ligands opens new therapeutic perspectives for G4 ligands in combinatorial treatments. Another recent interesting observation came from the Capranico group, who showed that G4 ligands increase R-loop levels in human cancer cells while simultaneously stabilizing G4s [82]. R-loop formation is favored by GC skewness and G4 formation is possible in the displaced DNA strand of an R-loop. These observations emphasize the interplay between R-loops and G4s and establish a model of action of G4 ligands in cancer cells that is not yet well understood.

Table 2.1 - List of G4 ligands discussed in this review.

Ligand	Target G4	Tumour type	Effect	Refs
 BRACO-19	Telomere	Glioblastoma	- Telomerase inhibition - Cell cycle arrest and DNA damage - Short-term apoptosis and senescence - Null effect in normal primary astrocytes	[57]
 Schizocommunin	Telomere	Cervical squamous cancer	- Telomere shortening - In-cell G4 formation - Cell cycle arrest and apoptosis - In vivo tumor growth inhibition	[59]

derivative				
 RHSP4	Telomere	Glioblastoma	- Radiosensitizing agent - Antiproliferative effect against glioblastoma stem-like cells - In vivo tumor growth inhibition	[60]
 QN-1	c-MYC	Triple-negative breast cancer	- Downregulation of c-MYC transcription - Metastasis repression - Cell cycle arrest and apoptosis - In vivo tumor growth inhibition	[62]
 IZCZ-3	c-MYC	Cervical squamous cancer	- Downregulation of c-MYC transcription - Cell cycle arrest and apoptosis - In vivo tumor growth inhibition	[63]
 TH3	c-MYC	Cervical and lung cancer	- Downregulation of c-MYC transcription - Cell cycle arrest and apoptosis - Negligible cytotoxicity for normal cells	[64]
 Porphyrin derivatives	KRAS	Pancreatic cancer	- Downregulation of KRAS expression - Metastasis repression - Cell cycle arrest and apoptosis - In vivo tumor growth inhibition	[69]
 C8	KRAS	Cervical cancer	- Downregulation of oncogene expression - Remarkable cytotoxicity for cancer cells - Enhanced cellular uptake	[71]
 Anthrafurandione derivative	KRAS mRNA	Pancreatic cancer	- Downregulation of KRAS translation - KRAS mRNA binding in low-cell conditions - Apoptosis induction in KRAS addicted cells	[74]
 CX-3543	Ribosomal DNA	Carcinoid/ neuroendocrine tumors	- Inhibition of rRNA biogenesis - Apoptosis and DNA damage - In vivo tumor growth inhibition	[75]
 CX-5461	Ribosomal DNA	BRCA1/2 deficient tumors	- Inhibition of rRNA biogenesis - Apoptosis and DNA damage - In vivo tumor growth inhibition	[76]

 20A	Multitargeted	Cervical cancer	- Induction of autophagy - Apoptosis and DNA damage - In vivo tumor growth inhibition	[77]

G4s as therapeutic agents

Apart from drug targets, G4 nucleic acids have also been used as valuable tools in developing cancer therapeutics. Nucleic acid aptamers are single-stranded DNA or RNA sequences that are able to recognize and bind with high affinity cellular targets such as proteins, small molecules, ions, or even whole cells [83]. Remarkably, several aptamers against biologically relevant targets are guanine-rich sequences that fold into stable G4 structures. This may be because G4 nucleic acids are significantly more negatively charge dense than duplex DNA, which favours the interaction with positively charged pockets of proteins. In addition, the G4 element may be seen as a central core presenting specific elements for protein recognition in the loops, bulges and 5' and 3' extremities. G4 aptamers offer several advantages compared with unstructured sequences, such as thermodynamic and chemical stability, lack of immunogenicity, serum nuclease resistance, and enhanced cellular uptake [83]. Aptamers are usually developed by systematic evolution of ligands by exponential enrichment (SELEX) technology, which was introduced three decades ago [84]. A conventional SELEX protocol involves three cycled main steps (**Figure 2.3**): (1) binding (selection), where the target molecule is incubated with a random oligonucleotide library under defined conditions; (2) partitioning, in which the target-bound aptamers are separated from the unbound library; and (3) amplification, where the enriched pool of bound aptamers is amplified either by PCR or reverse-transcription PCR for DNA or RNA oligonucleotides, respectively, before being used in the succeeding round of selection [84]. Ideally, the SELEX protocol should also include a counter-selection step where the selected aptamers are incubated with a negative control target. Several modified and improved SELEX protocols have been developed since and have been reviewed elsewhere [84].

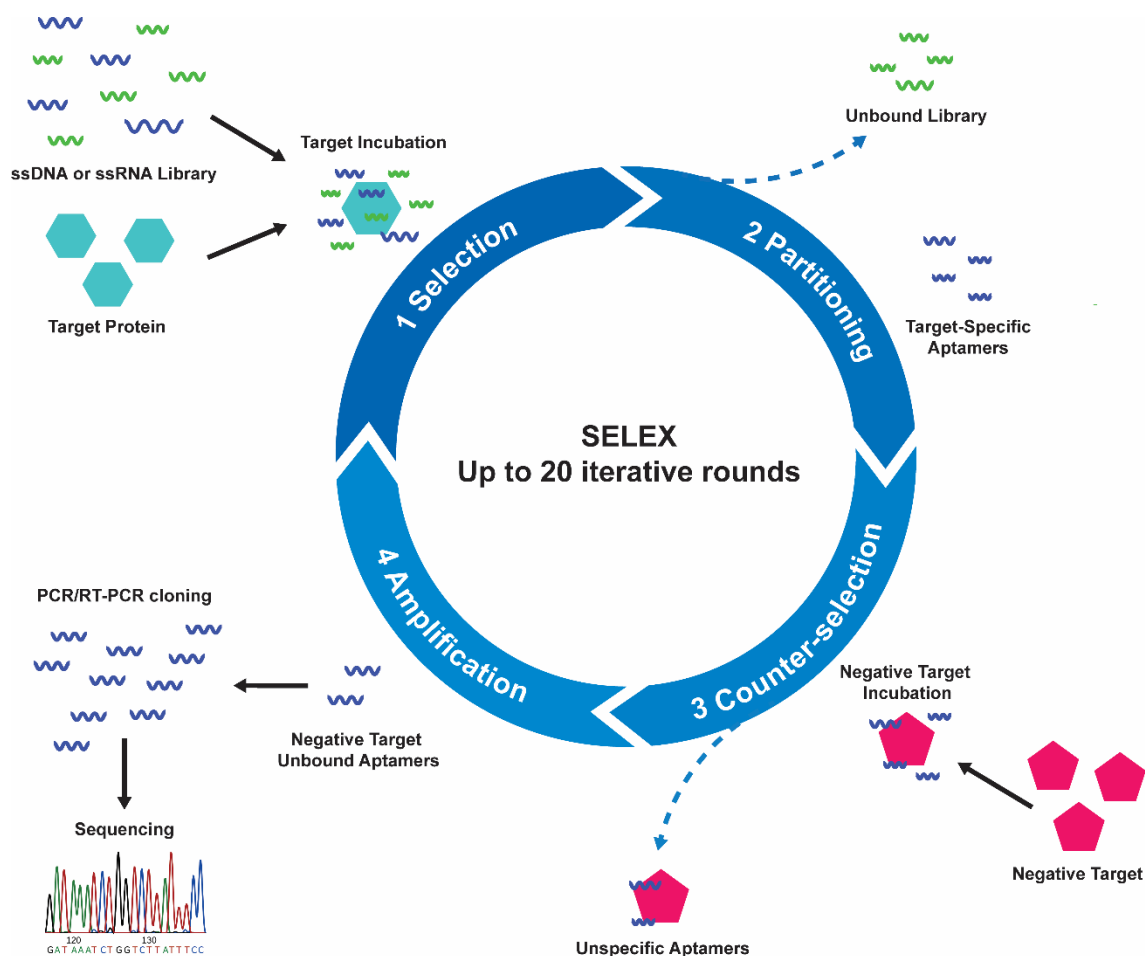


Figure 2.3 - General workflow of SELEX protocol. During SELEX, a single-strand (ss)DNA/ssRNA library (typically a mix of 10^{15} sequences) is incubated with the target protein (1). The unbound sequences are partitioned from the bound sequences (2). At this point the bound sequences may undergo a negative selection procedure by incubating them with negative protein targets (3). The sequences that bound the positive (true) target but failed to bind the negative control target are then amplified by PCR or RT-PCR (if its RNA aptamers) to enrich the pool (4). The enriched pool is finally sequenced and subjected to the next SELEX round.

G4 aptamers as anticancer agents

Over the past two decades, several G4 aptamers have been discovered for their anticancer activity (**Table 2.2**). Indubitably, the most famous example of a G4 aptamer is AS1411 which reached Phase 2 clinical trials (NCT01034410 and NCT00740441) and showed promising activity against metastatic renal cell carcinoma and acute myeloid leukaemia with minimal toxicity [85]. AS1411 is a 26 base-pair guanine-rich aptamer that specifically targets nucleolin, a protein that is overexpressed on the surface of cancer cells, making it a tumour-selective target [86]. Its most recent proposed mechanism of action rests on the hyperstimulation of micropinocytosis by the activation of Rac1, which causes cell death by methuosis, a nonapoptotic cell death pathway characterized by massive vacuolization in

the cytoplasm [87]. More recently, two aptamer sequences derived from the thrombin-binding aptamer (TBA) were developed as anticancer agents against lung cancer cells [88]. Similarly to AS1411, TBA is one of the most investigated therapeutic guanine-rich aptamers, not only with remarkable anticoagulant activity, but also showing some antiproliferative properties as demonstrated in a previous study using HeLa cells [89]. TBA derivatives TBA353 and TBA535 were designed with inverted polarity sites (3'-3' or 5'-5') to introduce different loop arrangements in the G4 structure [88]. TBA535 was shown to inhibit cell proliferation in Calu-6 lung cancer cells, with an activity higher than that observed for TBA after 48 h of treatment [88]. Furthermore, TBA535 inhibited cancer cell motility, suggesting that the aptamer inhibits cell metastasis and/or tumour invasion [88]. The authors did not address whether TBA535 binds thrombin to the same extent as the TBA aptamer, or whether the antiproliferative properties of the TBA derivatives are the result of thrombin recognition. Nonetheless, given the known roles of thrombin in tumour biology and metastasis [90], it is acceptable to speculate that TBA535 mechanism of action is thrombin dependent. This should be addressed in future work to leverage the potential clinical applicability of TBA derivatives. Additional TBA analogs containing L-residues have already been described by the same group to have lower anticoagulant activity but higher antiproliferative properties [91]. Other anticancer G4 aptamers include the kinase MNK-targeted aptamers apMNK2F and apMNK3R [92]. Unlike AS1411 which was not discovered by a SELEX procedure, apMNK2F and apMNK3R were selected using *in vitro* selection. The aptamers presented high binding affinities to the MNK protein ($K_D \sim 10^{-9}$ M) and high stability against nucleases [92]. Both G4 aptamers were shown to colocalize with MNK in the cytoplasm of MDA-MB-231 triple-negative breast cancer cells, inhibiting MNK activity and expression at the translational level (by almost 100% in the case of apMNK2F) [92]. This resulted in the inhibition of tumor cell proliferation and migration *in vitro*, and of colony formation in breast cancer cells [92]. Two G4 DNA aptamers against epidermal growth factor receptor (EGFR)-transfected A549 cells have been developed and their antiproliferative properties tested [93]. Aptamers S13 and S50 formed parallel G4 structures and were swiftly internalized by A549 lung cancer cells [93]. S50 promoted a concentration-dependent inhibitory effect on cancer cell growth without affecting the viability of normal cells [93]. Intriguingly, S50, similar to AS1411 and TBA [72], was shown to be internalized by noncancer cells to a comparable extent as cancer cells, suggesting that the tumor-selective antiproliferative effects of G4 aptamers do not strictly depend on the binding of the G-rich aptamer to the surface of the target cells [93].

Table 2.2 - List of G4 aptamers discussed in this review.

Aptamer	Sequence	Protein Target	Ref.
AS1411	5'-GGTGGTGGTGGTTGTGGTGGTGGTGG-3'	Nucleolin	[85], [86]
TBA	5'-GGTTGGTGTGGTTGG-3'	Thrombin	[94]
TBA353	5'-GG-3'-3'-TTGGTGT-5'-5'-GG-3'-3'-TTGG-5'	Thrombin?	[88]
TBA535	3'-GG-5'-5'-TTGGTGT-3'-3'-GG-5'-5'-TTGG-3'	Thrombin?	[88]
apMNK2F	5'-GCGGATGAAGACTGGTGTGGGGTGGGCGGGC GGGGGTGGGGGTGGTATGGCGCGTTGGCCCTAAATA CGAGCAAC-3'	MAP kinase interacting kinases	[92]
apMNK3R	5'- GTTGCTCGTATTTAGGGCGGGGGTGGGCGGATTTC GCAGGTGGCGTGTACACCACTCTTCATCCGC-3'	MAP kinase interacting kinases	[92]
S13	5'-TGGGTGGGTGGGGGG-3'	Nucleolin	[93]
S50	5'-GGGCGGGGGGTGGGGTGGTTGG-3'	Nucleolin	[93]
AT11	5'-TGGTGGTGGTTGTTGTGGTGGTGGTGGT-3'	Nucleolin	[95]
AT11-Lo	5'-TGGTGGTGGTTGTTGGGTGGTGGTGGT-3'	Nucleolin	[95]
AT11-Bo	5'-TGGTGGTGGTTGGTGGTGGTGGTGGT-3'	Nucleolin	[95]

G4 aptamers as drug delivery systems

G4 aptamers have also been used as drug carriers or targeting agents of drug delivery systems [83]. Using a simple noncovalent approach, AS1411 and its derivatives AT11, AT11-Lo, and AT11-Bo were recently used to selectively accumulate G4 ligands in cancer cells [72,96,97]. AS1411 was shown to form a stable complex with the acridine orange derivative C₈, and to selectively accumulate the compound in HeLa cells compared with normal fibroblasts, limiting the previously reported off-target effects of C₈ [71,72]. The AS1411-C₈ complex inhibited c-MYC transcription with concomitant antiproliferative effects against cervical cancer cells [72]. Interestingly, AS1411 uptake was enhanced upon binding C₈, which may be due to the stabilization of the aptamer G4 structure and/or the observed nucleolin overexpression induced by the AS1411-C₈ complex [72]. Similar results were obtained using AT11, AT11-Lo and AT11-Bo aptamers, emphasizing the potential of using G4 aptamers as drug delivery systems to deliver G4 ligands to cancer cells [96,97]. Due to its tumour-selective properties, AS1411 has also been used as a targeting agent of nanomedicines that target cancer cells specifically. A recent example are AS1411-functionalized polymeric micelles produced to co-deliver a siRNA and doxorubicin in murine tumour models [98]. The micelles were shown to localize and accumulate in tumour sites in a lung tumour mouse model, promoting a rapid and steady intracellular release of drugs and genes in the target tissue [98]. For a detailed analysis of the use of

AS1411 as a targeting agent of cancer-selective drug delivery systems, we refer the reader to an excellent review published in 2017 by Bates and colleagues [86].

Concluding Remarks

G4 nucleic acid structures are remarkable tools for the development of anticancer therapeutics. Excitingly, the first G4-targeted drugs are now reaching and progressing in clinical trials and we hope that in the near future, some will be approved for clinical use. Many questions remain regarding the selectivity, biological evidence, effects, and antitumoral mechanisms of both G4 ligands and G4 aptamers, but the currently available evidence strongly suggests G4 nucleic acids as valuable and innovative tools in the fight against cancer. The G4 field would greatly benefit from *in vivo* evidence and structural data, from both the biologically relevant folding of the target G4 structures and their ligand interactions. Important steps toward this were taken with the development of probes capable of visualizing G4s in living cells, but more efforts are needed. Additionally, further understanding the G4/i-motif/R-loop (and also duplex DNA) interplay could open new prospects for the development of specific ligands capable of controlling gene expression in this context. By contrast, for G4 aptamers, there is still a lack of a clear structure-activity relationship to correlate a given G4 conformation with a specific biological effect. This is particularly true for AS1411, which folds into a variety of G4 structures. This emphasizes the critical need for high-resolution structures of the G4-forming aptamers in complex with their targets, which would provide the necessary scaffolds for the design of new aptamers. Given the abundance of natural G4 motifs in the human genome, particularly in promoters, these could be studied as an alternative to SELEX-derived aptamers in the future.

Acknowledgments

The authors acknowledge their current and past colleagues for stimulating discussion. This work was supported by project “Ações Integradas Luso-Francesas” ref. TC-15/17, “Cooperação Científica e Tecnológica FCT/Acordo Pessoa” project ref. 5079, FCT project ref. IF/00959/2015 financed by Fundo Social Europeu and Programa Operacional Potencial Humano, FCT project ref. UID/Multi/04349/2013, MIT Portugal project BIODEVICE ref. MIT-EXPL/BIO/0008/2017 and UTAustin FCT project DREAM ref. UTAP-EXPL/NTec/0015/2017 (to C.C.) and the SYMBIT project Reg. no. CZ.02.1.01/0.0/0.0/15 003/0000477 financed by the ERDF (to J.L.M.). J.C.

acknowledges the doctoral fellowship grant from FCT - Foundation for Science and Technology ref. SFRH/BD/122953/2016.

References

- [1] Bang, I. (1910) Untersuchungen über die Guanylsäure. *Biochemische Zeitschrift* 26, 293–311.
- [2] Spiegel, J., Adhikari, S., and Balasubramanian, S. (2020) The Structure and Function of DNA G-Quadruplexes. *Trends in Chemistry* 2, 123–136. <https://doi.org/10.1016/j.trechm.2019.07.002>.
- [3] Gellert, M., Lipsett, M., and Davies, D. (1962) Helix formation by guanylic acid. *Proceedings of the National Academy of Sciences of the United States of America* 48, 2013–2018. <https://doi.org/10.1073/pnas.48.12.2013>.
- [4] Neidle, S. (2017) Quadruplex nucleic acids as targets for anticancer therapeutics. *Nature Reviews Chemistry* 1, Article 0041. <https://doi.org/10.1038/s41570-017-0041>.
- [5] Neidle, S. (2016) Quadruplex Nucleic Acids as Novel Therapeutic Targets. *Journal of Medicinal Chemistry* 59, 5987–6011. <https://doi.org/10.1021/acs.jmedchem.5b01835>.
- [6] Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I., and Blackburn, E. H. (1987) Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. *Cell* 51, 899–908.
- [7] Sen, D., and Gilbert, W. (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* 334, 364–366. <https://doi.org/10.1038/334364a0>.
- [8] Simonsson, T., Pecinka, P., and Kubista, M. (1998) DNA tetraplex formation in the control region of c-myc. *Nucleic acids research* 26, 1167–72.
- [9] Chariker, J. H., Miller, D. M., and Rouchka, E. C. (2016) Computational analysis of G-quadruplex forming sequences across chromosomes reveals high density patterns near the terminal ends. *PLoS ONE* 11, e0165101. <https://doi.org/10.1371/journal.pone.0165101>.
- [10] Bedrat, A., Lacroix, L., and Mergny, J.-L. (2016) Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Research* 44, 1746–1759. <https://doi.org/10.1093/nar/gkw006>.

- [11] Puig Lombardi, E., and Londoño-Vallejo, A. (2019) A guide to computational methods for G-quadruplex prediction. *Nucleic Acids Research* 48, 1–15. <https://doi.org/10.1093/nar/gkz1097>.
- [12] Krafčíková, P., Demkovičová, E., and Víglaský, V. (2017) Ebola virus derived G-quadruplexes: Thiazole orange interaction. *Biochimica et Biophysica Acta - General Subjects* 1861, 1321–1328. <https://doi.org/10.1016/j.bbagen.2016.12.009>.
- [13] Fleming, A. M., Ding, Y., Alenko, A., and Burrows, C. J. (2016) Zika Virus Genomic RNA Possesses Conserved G-Quadruplexes Characteristic of the Flaviviridae Family. *ACS Infectious Diseases* 2, 674–681. <https://doi.org/10.1021/acsinfecdis.6b00109>.
- [14] Krafčíková, P., Demkovičová, E., Halaganová, A., and Víglaský, V. (2017) Putative HIV and SIV G-Quadruplex Sequences in Coding and Noncoding Regions Can Form G-Quadruplexes. *Journal of Nucleic Acids* 2017, 1–13. <https://doi.org/10.1155/2017/6513720>.
- [15] Marušič, M., and Plavec, J. (2019) Towards Understanding of Polymorphism of the G-rich Region of Human Papillomavirus Type 52. *Molecules* 24, 1294. <https://doi.org/10.3390/molecules24071294>.
- [16] Marušič, M., Hošnjak, L., Krafčíková, P., Poljak, M., Víglaský, V., and Plavec, J. (2017) The effect of single nucleotide polymorphisms in G-rich regions of high-risk human papillomaviruses on structural diversity of DNA. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1861, 1229–1236. <https://doi.org/10.1016/j.bbagen.2016.11.007>.
- [17] Bartas, M., Cutová, M., Brázda, V., Kaura, P., Št'astný, J., Kolomazník, J., Coufal, J., Goswami, P., Červeň, J., and Pečinka, P. (2019) The presence and localization of G-Quadruplex forming sequences in the domain of bacteria. *Molecules* 24, 1711. <https://doi.org/10.3390/molecules24091711>.
- [18] Ruggiero, E., and Richter, S. N. (2018) G-quadruplexes and G-quadruplex ligands: targets and tools in antiviral therapy. *Nucleic Acids Research* 46, 3270–3283. <https://doi.org/10.1093/nar/gky187>.
- [19] Saad, M., Guédin, A., Amor, S., Bedrat, A., Tourasse, N. J., Fayyad-Kazan, H., Pratviel, G., Lacroix, L., and Mergny, J. L. (2019) Mapping and characterization of G-quadruplexes in the genome of the social amoeba Dictyostelium discoideum. *Nucleic Acids Research* 47, 4363–4374. <https://doi.org/10.1093/nar/gkz196>.

- [20] Biffi, G., Tannahill, D., McCafferty, J., and Balasubramanian, S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nature chemistry* 5, 182–6. <https://doi.org/10.1038/nchem.1548>.
- [21] Zhang, M., Liu, R., and Wang, F. (2019) Telomere and G-Quadruplex Colocalization Analysis by Immunofluorescence Fluorescence In Situ Hybridization (IF-FISH) BT - DNA Repair: Methods and Protocols, in *DNA Repair* (Balakrishnan, L., and Stewart, J. A., Eds.), pp 327–333. Springer New York, New York, NY. https://doi.org/10.1007/978-1-4939-9500-4_23.
- [22] Salgado, G. F., Cazenave, C., Kerkour, A., and Mergny, J.-L. (2015) G-quadruplex DNA and ligand interaction in living cells using NMR spectroscopy. *Chem. Sci.* 6, 3314–3320. <https://doi.org/10.1039/C4SC03853C>.
- [23] Laguerre, A., Hukezalie, K., Winckler, P., Katranji, F., Chanteloup, G., Pirrotta, M., Perrier-Cornet, J. M., Wong, J. M. Y., and Monchaud, D. (2015) Visualization of RNA-Quadruplexes in Live Cells. *Journal of the American Chemical Society* 137, 8521–8525. <https://doi.org/10.1021/jacs.5b03413>.
- [24] Zhang, S., Sun, H., Wang, L., Liu, Y., Chen, H., Li, Q., Guan, A., Liu, M., and Tang, Y. (2018) Real-time monitoring of DNA G-quadruplexes in living cells with a small-molecule fluorescent probe. *Nucleic Acids Research* 46, 7522–7532. <https://doi.org/10.1093/nar/gky665>.
- [25] Moye, A. L., Porter, K. C., Cohen, S. B., Phan, T., Zyner, K. G., Sasaki, N., Lovrecz, G. O., Beck, J. L., and Bryan, T. M. (2015) Telomeric G-quadruplexes are a substrate and site of localization for human telomerase. *Nature Communications* 6, 1–12. <https://doi.org/10.1038/ncomms8643>.
- [26] Wang, Q., Liu, J. Q., Chen, Z., Zheng, K. W., Chen, C. Y., Hao, Y. H., and Tan, Z. (2011) G-quadruplex formation at the 3' end of telomere DNA inhibits its extension by telomerase, polymerase and unwinding by helicase. *Nucleic Acids Research* 39, 6229–6237. <https://doi.org/10.1093/nar/gkr164>.
- [27] Prorok, P., Artufel, M., Aze, A., Coulombe, P., Peiffer, I., Lacroix, L., Guédin, A., Mergny, J.-L., Damaschke, J., Schepers, A., Ballester, B., and Méchali, M. (2019) Involvement of G-quadruplex regions in mammalian replication origin activity. *Nature Communications* 10, 3274. <https://doi.org/10.1038/s41467-019-11104-0>.
- [28] Bryan, T. M. (2019) Mechanisms of DNA replication and repair: Insights from the

study of G-quadruplexes. *Molecules* 24, Article 3439. <https://doi.org/10.3390/molecules24193439>.

[29] Wang, Y., Yang, J., Wild, A. T., Wu, W. H., Shah, R., Danussi, C., Riggins, G. J., Kannan, K., Sulman, E. P., Chan, T. A., and Huse, J. T. (2019) G-quadruplex DNA drives genomic instability and represents a targetable molecular abnormality in ATRX-deficient malignant glioma. *Nature Communications* 10, 913. <https://doi.org/10.1038/s41467-019-08905-8>.

[30] Schiavone, D., Guilbaud, G., Murat, P., Papadopoulou, C., Sarkies, P., Prioleau, M., Balasubramanian, S., and Sale, J. E. (2014) Determinants of G quadruplex-induced epigenetic instability in REV 1-deficient cells . *The EMBO Journal* 33, 2507-2520. <https://doi.org/10.15252/emj.201488398>.

[31] Bay, D. H., Busch, A., Lisdat, F., Iida, K., Ikebukuro, K., Nagasawa, K., Karube, I., and Yoshida, W. (2017) Identification of G-quadruplex structures that possess transcriptional regulating functions in the Dele and Cdc6 CpG islands. *BMC Molecular Biology* 18, 7. <https://doi.org/10.1186/s12867-017-0094-z>.

[32] Wu, G., Xing, Z., Tran, E. J., and Yang, D. (2019) DDX5 helicase resolves G-quadruplex and is involved in MYC gene transcriptional activation. *Proceedings of the National Academy of Sciences of the United States of America* 116, 20453–20461. <https://doi.org/10.1073/pnas.1909047116>.

[33] Jodoin, R., Carrier, J. C., Rivard, N., Bisailon, M., and Perreault, J. P. (2019) G-quadruplex located in the 5'UTR of the BAG-1 mRNA affects both its cap-dependent and cap-independent translation through global secondary structure maintenance. *Nucleic acids research* 47, 10247-10266. <https://doi.org/10.1093/nar/gkz777>.

[34] Hänsel-Hertsch, R., Beraldi, D., Lensing, S. V., Marsico, G., Zyner, K., Parry, A., Di Antonio, M., Pike, J., Kimura, H., Narita, M., Tannahill, D., and Balasubramanian, S. (2016) G-quadruplex structures mark human regulatory chromatin. *Nature Genetics* 48, 1267–1272. <https://doi.org/10.1038/ng.3662>.

[35] Assi, H. A., Garavís, M., González, C., and Damha, M. J. (2018) I-motif DNA: Structural features and significance to cell biology. *Nucleic Acids Research* 46, 8038-8056. <https://doi.org/10.1093/nar/gky735>.

[36] Kharel, P., Balaratnam, S., Beals, N., and Basu, S. (2020) The role of RNA G-quadruplexes in human diseases and therapeutic strategies. *Wiley Interdisciplinary*

Reviews: RNA 11, e1568. <https://doi.org/10.1002/wrna.1568>.

[37] Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I., and Blackburn, E. H. (1987) Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. *Cell* 51, 899–908. [https://doi.org/10.1016/0092-8674\(87\)90577-0](https://doi.org/10.1016/0092-8674(87)90577-0).

[38] Rhodes, D., and Lipps, H. J. (2015) Survey and summary G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Research* 43, 8627–8637. <https://doi.org/10.1093/nar/gkv862>.

[39] Asamitsu, S., Obata, S., Yu, Z., Bando, T., and Sugiyama, H. (2019) Recent progress of targeted G-quadruplex-preferred ligands toward cancer therapy. *Molecules* 24, 429. <https://doi.org/10.3390/molecules24030429>.

[40] Lu, K. Y., Chen, W. F., Rety, S., Liu, N. N., Wu, W. Q., Dai, Y. X., Li, D., Ma, H. Y., Dou, S. X., and Xi, X. G. (2018) Insights into the structural and mechanistic basis of multifunctional *S. Cerevisiae* pif1p helicase. *Nucleic Acids Research* 46, 1486–1500. <https://doi.org/10.1093/nar/gkx1217>.

[41] Paeschke, K., Bochman, M. L., Garcia, P. D., Cejka, P., Friedman, K. L., Kowalczykowski, S. C., and Zakian, V. A. (2013) Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature* 497, 458–462. <https://doi.org/10.1038/nature12149>.

[42] Guilbaud, G., Murat, P., Recolin, B., Campbell, B. C., Maiter, A., Sale, J. E., and Balasubramanian, S. (2017) Local epigenetic reprogramming induced by G-quadruplex ligands. *Nature Chemistry* 9, 1110–1117. <https://doi.org/10.1038/NCHEM.2828>.

[43] Valton, A.-L., Hassan-Zadeh, V., Lema, I., Boggetto, N., Alberti, P., Saintome, C., Riou, J.-F., and Prioleau, M.-N. (2014) G4 motifs affect origin positioning and efficiency in two vertebrate replicators. *The EMBO Journal* 33, 732–746. <https://doi.org/10.1002/embj.201387506>.

[44] Hou, X. M., Wu, W. Q., Duan, X. L., Liu, N. N., Li, H. H., Fu, J., Dou, S. X., Li, M., and Xi, X. G. (2015) Molecular mechanism of G-quadruplex unwinding helicase: Sequential and repetitive unfolding of G-quadruplex by Pif1 helicase. *Biochemical Journal* 466, 189–199. <https://doi.org/10.1042/BJ20140997>.

[45] Lerner, L. K., and Sale, J. E. (2019) Replication of G Quadruplex DNA. *Genes* 10, 95.

<https://doi.org/10.3390/genes10020095>.

[46] Fukuhara, M., Ma, Y., Nagasawa, K., and Toyoshima, F. (2017) A G-quadruplex structure at the 5' end of the H19 coding region regulates H19 transcription. *Scientific Reports* 8, 45815. <https://doi.org/10.1038/srep45815>.

[47] Smestad, J. A., and Maher, L. J. (2015) Relationships between putative G-quadruplex-forming sequences, RecQ helicases, and transcription. *BMC Medical Genetics* 16, 91. <https://doi.org/10.1186/s12881-015-0236-4>.

[48] Rigo, R., Palumbo, M., and Sissi, C. (2017) G-quadruplexes in human promoters: A challenge for therapeutic applications. *Biochimica et Biophysica Acta - General Subjects* 1861, 1399–1413. <https://doi.org/10.1016/j.bbagen.2016.12.024>.

[49] Harris, L. M., and Merrick, C. J. (2015) G-Quadruplexes in Pathogens: A Common Route to Virulence Control? *PLoS Pathogens* 11, 1–15. <https://doi.org/10.1371/journal.ppat.1004562>.

[50] Serber, Z., Keatinge-Clay, a T., Ledwidge, R., Kelly, a E., Miller, S. M., and Dötsch, V. (2001) High-resolution macromolecular NMR spectroscopy inside living cells. *Journal of the American Chemical Society* 123, 2446–2447.

[51] Falabella, M., Kolesar, J. E., Wallace, C., de Jesus, D., Sun, L., Taguchi, Y. V., Wang, C., Wang, T., Xiang, I. M., Alder, J. K., Maheshan, R., Horne, W., Turek-Herman, J., Pagano, P. J., St. Croix, C. M., Sondheimer, N., Yatsunyk, L. A., Johnson, F. B., and Kaufman, B. A. (2019) G-quadruplex dynamics contribute to regulation of mitochondrial gene expression. *Scientific Reports* 9, 1–17. <https://doi.org/10.1038/s41598-019-41464-y>.

[52] Cree, S. L., and Kennedy, M. a. (2014) Relevance of G-quadruplex structures to pharmacogenetics. *Frontiers in Pharmacology* 5, 4–11. <https://doi.org/10.3389/fphar.2014.00160>.

[53] Li, Q., Xiang, J.-F. F., Yang, Q.-F. F., Sun, H.-X. X., Guan, A.-J. J., and Tang, Y.-L. L. (2013) G4LDB: A database for discovering and studying G-quadruplex ligands. *Nucleic Acids Research* 41, D1115–D1123. <https://doi.org/10.1093/nar/gks1101>.

[54] Mergny, J.-L. (2020) Meeting report: Seventh International Meeting on Quadruplex Nucleic Acids (Changchun, P.R. China, September 6–9, 2019). *Biochimie* 168, 100–109. <https://doi.org/10.1016/j.biochi.2019.10.019>.

[55] Sun, Z. Y., Wang, X. N., Cheng, S. Q., Su, X. X., and Ou, T. M. (2019) Developing

novel G-quadruplex ligands: From interaction with nucleic acids to interfering with nucleic acid–protein interaction. *Molecules* 24, 396. <https://doi.org/10.3390/molecules24030396>.

[56] Zuffo, M., Guédin, A., Leriche, E. D., Doria, F., Pirota, V., Gabelica, V., Mergny, J. L., and Freccero, M. (2018) More is not always better: Finding the right trade-off between affinity and selectivity of a G-quadruplex ligand. *Nucleic Acids Research* 46, e115. <https://doi.org/10.1093/nar/gky607>.

[57] Zhou, G., Liu, X., Li, Y., Xu, S., Ma, C., Wu, X., Cheng, Y., Yu, Z., Zhao, G., and Chen, Y. (2016) Telomere targeting with a novel G-Quadruplex-Interactive ligand BRACO-19 induces T-Loop disassembly and telomerase displacement in human glioblastoma cells. *Oncotarget* 7, 14925–14939. <https://doi.org/10.18632/oncotarget.7483>.

[58] Shan, C., Tan, J.-H., Ou, T.-M., and Huang, Z.-S. (2013) Natural products and their derivatives as G-quadruplex binding ligands. *Science China Chemistry* 56, 1351–1363. <https://doi.org/10.1007/s11426-013-4920-y>.

[59] Che, T., Chen, S. Bin, Tu, J. L., Wang, B., Wang, Y. Q., Zhang, Y., Wang, J., Wang, Z. Q., Zhang, Z. P., Ou, T. M., Zhao, Y., Tan, J. H., and Huang, Z. S. (2018) Discovery of Novel Schizocommunin Derivatives as Telomeric G-Quadruplex Ligands That Trigger Telomere Dysfunction and the Deoxyribonucleic Acid (DNA) Damage Response. *Journal of Medicinal Chemistry* 61, 3436–3453. <https://doi.org/10.1021/acs.jmedchem.7b01615>.

[60] Berardinelli, F., Tanori, M., Muoio, D., Buccarelli, M., Di Masi, A., Leone, S., Ricci-Vitiani, L., Pallini, R., Mancuso, M., and Antocchia, A. (2019) G-quadruplex ligand RHPS4 radiosensitizes glioblastoma xenograft in vivo through a differential targeting of bulky differentiated- and stem-cancer cells. *Journal of Experimental and Clinical Cancer Research* 38, 1–17. <https://doi.org/10.1186/s13046-019-1293-x>.

[61] Yang, D., and Hurley, L. H. (2006) Structure of the biologically relevant G-quadruplex in the c-MYC promoter. *Nucleosides, nucleotides & nucleic acids* 25, 951–968. <https://doi.org/10.1080/15257770600809913>.

[62] Hu, M.-H., Wu, T.-Y., Huang, Q., and Jin, G. (2019) New substituted quinoxalines inhibit triple-negative breast cancer by specifically downregulating the c-MYC transcription. *Nucleic Acids Research* 47, 10529–10542. <https://doi.org/10.1093/nar/gkz835>.

[63] Hu, M. H., Wang, Y. Q., Yu, Z. Y., Hu, L. N., Ou, T. M., Chen, S. Bin, Huang, Z. S., and

Tan, J. H. (2018) Discovery of a New Four-Leaf Clover-Like Ligand as a Potent c-MYC Transcription Inhibitor Specifically Targeting the Promoter G-Quadruplex. *Journal of Medicinal Chemistry* 61, 2447–2459. <https://doi.org/10.1021/acs.jmedchem.7b01697>.

[64] Dutta, D., Debnath, M., Müller, D., Paul, R., Das, T., Bessi, I., Schwalbe, H., and Dash, J. (2018) Cell penetrating thiazole peptides inhibit c-MYC expression via site-specific targeting of c-MYC G-quadruplex. *Nucleic Acids Research* 46, 5355–5365. <https://doi.org/10.1093/nar/gky385>.

[65] Zorzan, E., Da Ros, S., Musetti, C., Shahidian, L. Z., Coelho, N. F. R., Bonsembiante, F., Létard, S., Gelain, M. E., Palumbo, M., Dubreuil, P., Giantin, M., Sissi, C., and Dacasto, M. (2016) Screening of candidate G-quadruplex ligands for the human c-KIT promotorial region and their effects in multiple in-vitro models. *Oncotarget* 7, 21658–21675. <https://doi.org/10.18632/oncotarget.7808>.

[66] Amato, J., Pagano, A., Capasso, D., Di Gaetano, S., Giustiniano, M., Novellino, E., Randazzo, A., and Pagano, B. (2018) Targeting the BCL2 Gene Promoter G-Quadruplex with a New Class of Furopyridazinone-Based Molecules. *ChemMedChem* 13, 406–410. <https://doi.org/10.1002/cmdc.201700749>.

[67] Lavrado, J., Brito, H., Borralho, P. M., Ohnmacht, S. a., Kim, N.-S., Leitão, C., Pisco, S., Gunaratnam, M., Rodrigues, C. M. P., Moreira, R., Neidle, S., and Paulo, A. (2015) KRAS oncogene repression in colon cancer cell lines by G-quadruplex binding indolo[3,2-c]quinolines. *Scientific Reports* 5, 9696. <https://doi.org/10.1038/srep09696>.

[68] Marquevielle, J., Kumar, M. V. V., Mergny, J.-L., and Salgado, G. F. (2018) ¹H,¹³C, and¹⁵N chemical shift assignments of a G-quadruplex forming sequence within the KRAS proto-oncogene promoter region. *Biomolecular NMR Assignments* 12, 123–127. <https://doi.org/10.1007/s12104-017-9793-0>.

[69] Pattanayak, R., Barua, A., Das, A., Chatterjee, T., Pathak, A., Choudhury, P., Sen, S., Saha, P., and Bhattacharyya, M. (2018) Porphyrins to restrict progression of pancreatic cancer by stabilizing KRAS G-quadruplex: In silico, in vitro and in vivo validation of anticancer strategy. *European Journal of Pharmaceutical Sciences* 125, 39–53. <https://doi.org/10.1016/j.ejps.2018.09.011>.

[70] Han, F. X., Wheelhouse, R. T., and Hurley, L. H. (1999) Interactions of TMPyP4 and TMPyP2 with quadruplex DNA. Structural basis for the differential effects on telomerase inhibition. *Journal of the American Chemical Society* 121, 3561–3570. <https://doi.org/10.1021/ja984153m>.

- [71] Carvalho, J., Pereira, E., Marquevielle, J., Campello, M. P. C., Mergny, J.-L., Paulo, A., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2018) Fluorescent light-up acridine orange derivatives bind and stabilize KRAS-22RT G-quadruplex. *Biochimie* 144, 144–152. <https://doi.org/10.1016/j.biochi.2017.11.004>.
- [72] Carvalho, J., Paiva, A., Cabral Campello, M. P., Paulo, A., Mergny, J.-L., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2019) Aptamer-based Targeted Delivery of a G-quadruplex Ligand in Cervical Cancer Cells. *Scientific Reports* 9, 7945. <https://doi.org/10.1038/s41598-019-44388-9>.
- [73] Katsuda, Y., Sato, S., Asano, L., Morimura, Y., Furuta, T., Sugiyama, H., Hagihara, M., and Uesugi, M. (2016) A Small Molecule That Represses Translation of G-Quadruplex-Containing mRNA. *Journal of the American Chemical Society* 138, 9037–9040. <https://doi.org/10.1021/jacs.6b04506>.
- [74] Miglietta, G., Cogoi, S., Marinello, J., Capranico, G., Tikhomirov, A. S., Shchekotikhin, A., and Xodo, L. E. (2017) RNA G-Quadruplexes in Kirsten Ras (KRAS) Oncogene as Targets for Small Molecules Inhibiting Translation. *Journal of Medicinal Chemistry* 60, 9448–9461. <https://doi.org/10.1021/acs.jmedchem.7b00622>.
- [75] Drygin, D., Siddiqui-Jain, A., O'Brien, S., Schwaebe, M., Lin, A., Bliesath, J., Ho, C. B., Proffitt, C., Trent, K., Whitten, J. P., Lim, J. K. C., Von Hoff, D., Anderes, K., and Rice, W. G. (2009) Anticancer activity of CX-3543: A direct inhibitor of rRNA biogenesis. *Cancer Research* 69, 7653–7661. <https://doi.org/10.1158/0008-5472.CAN-09-1304>.
- [76] Xu, H., Di Antonio, M., McKinney, S., Mathew, V., Ho, B., O'Neil, N. J., Santos, N. Dos, Silvester, J., Wei, V., Garcia, J., Kabeer, F., Lai, D., Soriano, P., Banáth, J., Chiu, D. S., Yap, D., Le, D. D., Ye, F. B., Zhang, A., Thu, K., Soong, J., Lin, S. C., Tsai, A. H. C., Osako, T., Algara, T., Saunders, D. N., Wong, J., Xian, J., Bally, M. B., Brenton, J. D., Brown, G. W., Shah, S. P., Cescon, D., Mak, T. W., Caldas, C., Stirling, P. C., Hieter, P., Balasubramanian, S., and Aparicio, S. (2017) CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. *Nature Communications* 8, 14432. <https://doi.org/10.1038/ncomms14432>.
- [77] Beauvarlet, J., Bensadoun, P., Darbo, E., Labrunie, G., Rousseau, B., Richard, E., Draskovic, I., Londono-Vallejo, A., Dupuy, J. W., Nath Das, R., Guédin, A., Robert, G., Orange, F., Croce, S., Valesco, V., Soubeyran, P., Ryan, K. M., Mergny, J. L., and Djavaheri-Mergny, M. (2019) Modulation of the ATM/autophagy pathway by a G-quadruplex ligand tips the balance between senescence and apoptosis in cancer cells.

Nucleic acids research 47, 2739–2756. <https://doi.org/10.1093/nar/gkz095>.

[78] Kerkour, A., Mergny, J.-L., and Salgado, G. F. (2017) NMR based model of human telomeric repeat G-quadruplex in complex with 2,4,6-triarylpyridine family ligand. *Biochimica et Biophysica Acta - General Subjects* 1861, 1293–1302. <https://doi.org/10.1016/j.bbagen.2016.12.016>.

[79] Gentile, C., Martorana, A., Lauria, A., and Bonsignore, R. (2017) Kinase Inhibitors in Multitargeted Cancer Therapy. *Current Medicinal Chemistry* 24, 1671–1686. <https://doi.org/10.2174/0929867324666170112112734>.

[80] Salvati, E., Rizzo, A., Iachettini, S., Zizza, P., Cingolani, C., D'Angelo, C., Porru, M., Mondello, C., Aiello, A., Farsetti, A., Gilson, E., Leonetti, C., and Biroccio, A. (2015) A basal level of DNA damage and telomere deprotection increases the sensitivity of cancer cells to G-quadruplex interactive compounds. *Nucleic Acids Research* 43, 1759–1769. <https://doi.org/10.1093/nar/gkv006>.

[81] Zyner, K. G., Mulhearn, D. S., Adhikari, S., Cuesta, S. M., Di Antonio, M., Erard, N., Hannon, G. J., Tannahill, D., and Balasubramanian, S. (2019) Genetic interactions of G-quadruplexes in humans. *eLife* 8, 1–40. <https://doi.org/10.7554/eLife.46793>.

[82] De Magis, A., Manzo, S. G., Russo, M., Marinello, J., Morigi, R., Sordet, O., and Capranico, G. (2019) DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* 116, 816–825. <https://doi.org/10.1073/pnas.1810409116>.

[83] Roxo, C., Kotkowiak, W., and Pasternak, A. (2019) G-Quadruplex-Forming Aptamers-Characteristics, Applications, and Perspectives. *Molecules (Basel, Switzerland)* 24, 3781. <https://doi.org/10.3390/molecules24203781>.

[84] Bayat, P., Nosrati, R., Alibolandi, M., Rafatpanah, H., Abnous, K., Khedri, M., and Ramezani, M. (2018) SELEX methods on the road to protein targeting with nucleic acid aptamers. *Biochimie* 154, 132–155. <https://doi.org/10.1016/j.biochi.2018.09.001>.

[85] Rosenberg, J., Figlin, R. A., Smith, G. W., Choueiri, T., and Erlandsson, F. (2015) A phase II trial of the nucleolin-targeted DNA aptamer AS1411 in metastatic refractory renal cell carcinoma. *Invest New Drugs* 32, 178–187. <https://doi.org/10.1007/s10637-013-0045-6.A>.

- [86] Bates, P. J., Reyes-Reyes, E. M., Malik, M. T., Murphy, E. M., O'Toole, M. G., and Trent, J. O. (2017) G-quadruplex oligonucleotide AS1411 as a cancer-targeting agent: Uses and mechanisms. *Biochimica et Biophysica Acta - General Subjects* 1861, 1414–1428. <https://doi.org/10.1016/j.bbagen.2016.12.015>.
- [87] Reyes-Reyes, E. M., Šalipur, F. R., Shams, M., Forsthoefel, M. K., and Bates, P. J. (2015) Mechanistic studies of anticancer aptamer AS1411 reveal a novel role for nucleolin in regulating Rac1 activation. *Molecular Oncology* 9, 1392–1405. <https://doi.org/10.1016/j.molonc.2015.03.012>.
- [88] Esposito, V., Russo, A., Vellecco, V., Bucci, M., Russo, G., Mayol, L., Virgilio, A., and Galeone, A. (2018) Thrombin binding aptamer analogues containing inversion of polarity sites endowed with antiproliferative and anti-motility properties against Calu-6 cells. *Biochimica et Biophysica Acta - General Subjects* 1862, 2645–2650. <https://doi.org/10.1016/j.bbagen.2018.07.031>.
- [89] Dapić, V., Abdomerović, V., Marrington, R., Peberdy, J., Rodger, A., Trent, J. O., and Bates, P. J. (2003) Biophysical and biological properties of quadruplex oligodeoxyribonucleotides. *Nucleic Acids Research* 31, 2097–2107. <https://doi.org/10.1093/nar/gkg316>.
- [90] Reddel, C. J., Tan, C. W., and Chen, V. M. (2019) Thrombin generation and cancer: Contributors and consequences. *Cancers* 11, 100. <https://doi.org/10.3390/cancers11010100>.
- [91] Esposito, V., Russo, A., Amato, T., Varra, M., Vellecco, V., Bucci, M., Russo, G., Virgilio, A., and Galeone, A. (2017) Backbone modified TBA analogues endowed with antiproliferative activity. *Biochimica et Biophysica Acta - General Subjects* 1861, 1213–1221. <https://doi.org/10.1016/j.bbagen.2016.09.019>.
- [92] García-Recio, E. M., Pinto-Díez, C., Pérez-Morgado, M. I., García-Hernández, M., Fernández, G., Martín, M. E., and González, V. M. (2016) Characterization of MNK1b DNA Aptamers that inhibit proliferation in MDA-MB231 breast cancer cells. *Molecular Therapy - Nucleic Acids* 5, e275. <https://doi.org/10.1038/mtna.2015.50>.
- [93] Hu, J., Zhao, Z., Liu, Q., Ye, M., Hu, B., Wang, J., and Tan, W. (2015) Study of the Function of G-Rich Aptamers Selected for Lung Adenocarcinoma. *Chemistry - An Asian Journal* 10, 1519–1525. <https://doi.org/10.1002/asia.201500187>.
- [94] Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H., and Toole, J. J. (1992)

Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature* 355, 564–566. <https://doi.org/10.1038/355564a0>.

[95] Do, N. Q., Chung, W. J., Truong, T. H. A., Heddi, B., and Phan, A. T. (2017) G-quadruplex structure of an anti-proliferative DNA sequence. *Nucleic Acids Research* 45, 7487–7493. <https://doi.org/10.1093/nar/gkx274>.

[96] Carvalho, J., Lopes-Nunes, J., Lopes, A. C., Cabral Campello, M. P., Paulo, A., Queiroz, J. A., and Cruz, C. (2019) Aptamer-guided acridine derivatives for cervical cancer. *Organic and Biomolecular Chemistry* 17, 2992–3002. <https://doi.org/10.1039/c9ob00318e>.

[97] Figueiredo, J., Lopes-Nunes, J., Carvalho, J., Antunes, F., Ribeiro, M., Campello, M. P. C., Paulo, A., Paiva, A., Salgado, G. F., Queiroz, J. A., Mergny, J.-L., and Cruz, C. (2019) AS1411 derivatives as carriers of G-quadruplex ligands for cervical cancer cells. *International Journal of Pharmaceutics* 568, 118511. <https://doi.org/10.1016/j.ijpharm.2019.118511>.

[98] Yang, S., Ren, Z., Chen, M., Wang, Y., You, B., Chen, W., Qu, C., Liu, Y., and Zhang, X. (2018) Nucleolin-Targeting AS1411-Aptamer-Modified Graft Polymeric Micelle with Dual pH/Redox Sensitivity Designed to Enhance Tumor Therapy through the Codelivery of Doxorubicin/TLR4 siRNA and Suppression of Invasion. *Molecular Pharmaceutics* 15, 314–325. <https://doi.org/10.1021/acs.molpharmaceut.7b01093>.

[99] Kikin, O., D'Antonio, L., and Bagga, P. S. (2006) QGRS Mapper: A web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Research* 34, W676–W682. <https://doi.org/10.1093/nar/gkl253>.

Chapter 3

Aims and Outline of the Thesis

Aims and Outline of the Thesis

HPVs are well-established causative agents of both benign and malign lesions in the epithelia, particularly cervical cancer. To date, there is no effective cure for an ongoing HPV infection and the therapeutic options available for cervical cancer bring many drawbacks such as lack of targeted behaviour causing nefarious side-effects, multidrug resistance and poor efficacy of chemotherapy agents, and rate of recurrence due to radiation-surviving cancer cells or microscopic cancer cells that evade surgical excision. G4 targeted ligands and G4 aptamers have already been described for their safe and attractive therapeutic action against both viral infections and various cancer types. However, the development of G4-driven therapeutics for HPV-related diseases is still overlooked. Combining cancer-specific agents (G4 aptamers) with antiviral/anticancer drugs (G4 ligands) has a tremendous potential for developing next-generation therapeutics against HPV-diseases, by targeting HPV-positive cancer cells with minimal damage to peripheral healthy tissue.

Thus, the goal of this work is the development of G4 ligands capable of binding and stabilizing oncogene and telomeric G4 structures within cervical cancer cells, as well as G4 structures found in the genome of HPV, with potential dual effect on both carcinogenesis and infection, respectively. Furthermore, using nucleolin-targeted G4 aptamers such as AS1411, this work aims at developing a cancer-specific drug delivery strategy for the accumulation of the developed G4 ligands in HPV-positive cervical cancer cells. This project intends to furnish potential drug candidates to be used both as anticancer and antiviral agents against HPV persistence and carcinogenesis, while paving the way for other viral diseases to be addressed using this same strategy. This work will also shed light and help understand the unclear function of G4 structures in the HPV life cycle and their relationship with cervical cancer development.

The main objectives of this work were:

1. Development of G4-targeted small molecules (G4 ligands) capable of binding and stabilizing G4 structures (human and HPV) with high affinity and specificity over Watson-Crick duplex DNA.
2. Characterization of the antitumoral and antiviral effects of those G4 ligands in cervical cancer cell lines and organotypic raft cultures with replicating HPV, respectively.

3. Study of the feasibility of using AS1411 aptamer or other guanine-rich aptamers with improved pharmacology to be used as drug delivery systems of G4 ligands into HPV-positive cervical cancer cells.
4. Development of drug delivery nanosystems such as AS1411-functionalized nanoaggregates for the topical local delivery of G4 ligands into the cervix.

After the general description of cervical cancer and HPV life cycle (Chapter 1), the revision of G4 structures as both therapeutic targets and agents (Chapter 2), and the establishment of the thesis aims, Chapters 4-11 were organized as follows:

- Chapter 4 describes the development of indole-based G4 ligands and the study of their interaction with biologically relevant G4 structures found in the c-MYC promoter and human telomeres.
- Chapter 5 shows another class of G4 ligands, phenanthroline polyazamacrocycles, and describes their binding, affinity and thermal stabilization towards human telomeric and c-MYC promoter G4s, and their inhibitory effects on the activity of Pif1 helicase, transcription polymerases and cancer cell growth.
- Chapter 6 explores the potential of acridine orange derivatives to bind a G4 structure found in the KRAS oncogene and its cellular uptake and effect on cell viability of cervical cancer cells.
- Using the selected compounds developed in previous chapters in terms of G4 binding and stabilization, together with commercially available G4 ligands, Chapter 7 establishes the anti-HPV potential of the compounds in HPV16 and HPV18 infected organotypic cultures.
- Chapters 8-10 describes the possibility of using of aptamer AS1411 and several modified derivatives as drug delivery systems of acridine orange-based G4 ligands into cervical cancer cells.
- In Chapter 11, the development of nanoparticles comprising drug-loaded AS1411 nanoaggregates, and their formulation as a topical gel for local application in the cervix is described.

- Finally, Chapter 12 presents an integrated view of the main findings of this thesis and discusses their potential impact and the future perspectives in the context of HPV-related diseases and their treatment.

Chapter 4

Study of the interaction between indole-based compounds and biologically relevant G-quadruplexes

This chapter was published in:

Carvalho, J., Nottelet, P., Mergny, J.-L., Queiroz, J. A., Salgado, G. F., and Cruz, C. (2017) Study of the interaction between indole-based compounds and biologically relevant G-quadruplexes. *Biochimie* 135, 186–195. <https://doi.org/10.1016/j.biochi.2017.02.005>.

Study of the interaction between indole-based compounds and biologically relevant G-quadruplexes

Abstract

G4s are high-order secondary structures that modulate several key cell processes such as telomere function, gene expression and DNA replication. In the past years G4s have emerged as promising targets for drug development due to the discovery of small molecules which bind and stabilize these structures. In this work, we report the synthesis of indole-based compounds and the study of their interaction with the biological relevant G4s c-MYC and human telomeric repeat 22AG using several biophysical techniques. The ligands are G4 specific and they can discriminate different G4 structures namely parallel and hybrid-1 topologies. The NMR study of interaction between the most promising indole-derivative (compound **3**) and c-MYC quadruplex suggests that the ligand binds on the external tetrads with additional actions in the loops/grooves in a 2:1 ratio. The molecular docking calculations of compound **3** to c-MYC quadruplex corroborates with the ^1H and NOESY NMR studies.

Overall, the results suggest that indole-based ligands are promising candidates for future lead optimizations in drug development.

Introduction

Apart from the well-known Watson-Crick double helix DNA structure, nucleic acids can adopt other non-canonical conformations under certain conditions [1]. The ability of Guanine to assemble into higher order structures has been described more than a century ago by Ivar Bang [2]. G4s are non-canonical four-stranded DNA structures composed of two or more G-tetrads stacked on top of each other and connected by loops of nucleotides [3]. Each G-tetrad contains four Hoogsteen hydrogen-bound guanine residues, stabilized by central monovalent cations such as K^+ and Na^+ , and constitutes the common structural feature to all G4s [1]. G4s are structurally diverse owing that polymorphism to an array of possible combinations relatively to the strand stoichiometry and orientation, sequence, loop length and the stabilizing metal cation present in solution [4,5].

These DNA structures were found to be involved in important biological processes such as gene transcription, DNA replication and genome stability [3]. Putative G4 forming sequences are spread throughout the human genome, notably in oncogene promoter regions, telomere ends and immunoglobulin switch regions [6,7]. The detection of G4 structures in living cells was already reported [6]. Biffi *et al.* described a monoclonal single chain antibody BG4 with high affinity and specificity for intramolecular G4 structures [8]. This work reported evidence for the formation of G4 DNA within mammalian cells genomes, in a cellular context, the cell-cycle dependence of the structures (the occurrence of G4s was maxed during the S phase) and provided an important basis to understand the biological roles of G4 structures. These findings emphasized the importance of G4s in biological processes making these DNA structures central drug-design targets for the treatment of diseases such as viral infections and cancer [6,9]. Several small specific ligands have been developed with the ability to stabilize the G4 structures, inhibiting telomerase activity or interfere with telomeric functions and modulating the transcription of oncogenes such as c-MYC [10]. For instance cationic porphyrin (TMPyP4), acridine derivatives (BRACO-19) as well as natural macrocycles (telomestatin) have been studied as G4 binders [11–13].

Despite a lot of ligands have been designed and synthesized, few have reached clinical trials, probably due to specificity and/or cytotoxicity problems [10]. Thus, the design and development of new, specific G4 stabilizing agents is a continuous and necessary challenge. The use of the indole moiety for the synthesis of G4 ligands is appealing as its derivatives are well known for their anticancer and anti-inflammatory pharmacological properties [14]. In this work, we report the synthesis of novel G4 indole-based ligand candidates and the study of its interaction with the biologically relevant human G4s 22AG and Pu22. Several biophysical techniques including circular dichroism (CD), fluorescence-based assays and nuclear magnetic resonance (NMR) were employed. Molecular docking calculations were employed in order to predict the binding mode of the ligands.

Materials and methods

Oligonucleotides

All the oligonucleotides were purchased from Eurofins (Germany), Eurogentec (Belgium) or STAB VIDA (Portugal) with HPLC-grade purification. The oligonucleotides sequences used in this work were the following: human telomeric DNA sequences (22AG: 5'-AGGGTTAGGGTTAGGGTTAGGG-3'; 21G: 5'-GGGTTAGGGTTAGGGTTAGGG); promoter

DNA sequence c-MYC Pu22: 5'-TGAGGGTGGGTAGGGTGGGTAA-3') and duplex sequences (dx: d(TATAGCTAT-hexaethyleneglycol-TATAGCTATA); ds-DNA: d(AGAGAG(TTAG)₉). Labelled oligonucleotides for Fluorescence Resonance Energy Transfer (FRET) experiments were modified 5' and 3' with the donor fluorophore FAM (6-carboxyfluorescein) and the acceptor fluorophore TAMRA (6-carboxytetramethylrhodamine), respectively. Stock solutions of approximately 1.5 mM strand concentration were prepared using Milli-Q water and stored at -20 °C until used. For all experiments, G4s were annealed by heating to 95 °C for 10 minutes and cooling in ice for 30 minutes.

Ligand Synthesis

Synthesis of N¹-((1H-indol-3-yl)methyl)-N²-(2-aminoethyl)ethane-1,2-diamine (1):

Indole-3-carboxaldehyde (0.50 g, 3.4×10⁻³ mol) and diethylene triamine (0.35 g, 3.4×10⁻³ mol) were dissolved in 50 mL of dry EtOH-CH₃CN (1:1). The resulting mixture was stirred for 2 hours at room temperature and then the solvent was evaporated. The resultant residue was dissolved in 20 mL EtOH and then NaBH₄ (1.11g, 29.5×10³ mol) was added portion wise. The solution was stirring during 24 hours at room temperature, and then the excess of NaBH₄ was filtered off and the solvent was evaporated to dryness. The resultant solid was treated with deionized water and extracted with CH₂Cl₂ (3×50 mL). The organic phase was evaporated to dryness. The solid was then dissolved in 1.5 mL EtOH and 6 mL dioxane, and precipitated with aqueous HCl 37% to obtain its hydrochloride salt. The precipitate was filtered and recrystallized three times with MeOH to give compound **1** as a pale rose solid (0.245 g, 31%).

Compound 1: M.p.: 239.5 °C. ¹H NMR (D₂O, 400.13 MHz, 25 °C): δ = 3.15 (t, 4H, J = 6.93), 3.24 (t, 4H, J = 6.93), 4.41 (s, 2H), 7.14 (t, 1H, J = 7.14), 7.23 (t, 1H, J = 7.23), 7.47 (d, 1H, J = 2.70), 7.49 (s, 1H), 7.63 (d, 1H, J = 8.40), 10.51 (s, 1H) ppm. ¹³C NMR (D₂O, 100.62 MHz, 25 °C): δ = 41.57, 44.99, 45.76, 105.66, 114.26, 119.86, 122.21, 124.56, 128.15, 130.03, 138.25 ppm. HRMS (ESI): calcd for C₁₃H₂₀N₄ [M + H]⁺ 233.176200; found 233.176073 (Δm -000127 and error -2.7 ppm).

Synthesis of N¹-((1H-indol-3-yl)methyl)-N²-(2-((2-aminoethyl)amino)ethyl)ethane-1,2-diamine (2) and N¹-((1H-indol-3-yl)methyl)-N²-(2-((2-((2-aminoethyl)amino)ethyl)amino)ethyl)ethane-1,2-diamine (3):

The same procedure as described for compound **1** was used with the following amines: triethylene tetraamine (0.5 g, 3.4×10⁻³ mol) and tetraethylene pentamine (0.64 g, 3.4×10⁻³

mol) respectively (0.232 g, 24%; 0.258 g, 24%). Both compounds are yellow and white solids, respectively.

Compound 2: M.p.: 249.0 °C. ¹H NMR (D₂O, 400.13 MHz, 25 °C): δ = 3.28-3.44 (m, 12H), 4.56 (s, 2H), 7.26 (t, 1H, J = 7.34), 7.33 (t, 1H, J = 7.34), 7.58 (s, 1H), 7.59 (d, 1H, J = 2.53), 7.76 (d, 1H, J = 7.84), 10.62 (s, 1H) ppm. ¹³C NMR (D₂O, 100.62 MHz, 25 °C): δ = 45.50, 45.74, 46.69, 47.25, 106.55, 115.25, 120.93, 123.23, 125.52, 129.10, 131.10, 139.12 ppm. HRMS (ESI): calcd for C₁₅H₂₅N₅ [M + H]⁺ 276.218300; found 276.218273 (Δm -0.000027 and error 1.5 ppm).

Compound 3: M.p.: 278.3 °C. ¹H NMR (D₂O, 400.13 MHz, 25 °C): δ = 3.39-3.48 (m, 16H), 4.57 (s, 2H), 7.26 (t, 1H, J = 7.38), 7.33 (t, 1H, J = 7.34), 7.58 (s, 1H), 7.61 (d, 1H, J = 2.48), 7.78 (d, 1H, J = 7.88), 10.62 (s, 1H) ppm. ¹³C NMR (D₂O, 100.62 MHz, 25 °C): δ = 42.20, 42.93, 43.77, 43.92, 44.24, 103.58, 112.32, 118.01, 120.29, 122.62, 126.13, 128.20, 136.19 ppm. HRMS (ESI): calcd for C₁₇H₃₀N₆ [M + H]⁺ 319.260400; found 319.263157 (Δm +0.002757 and error -0.18 ppm).

Circular Dichroism

All CD spectra were recorded using a Jasco J-815 spectropolarimeter equipped with a Peltier-type temperature control system. The measurements were made at 25 °C using an instrument scanning speed of 100 nm/min with a response time of 1 s through wavelengths ranging from 200 to 340 nm. The recording bandwidth was 1 nm with 1 nm step size using a quartz cell with an optical length of 1 mm. Oligonucleotides were used at 6-15 μM strand concentration in 30 mM phosphate buffer (15 mM KH₂PO₄ (or 15 mM NaH₂PO₄), 15 mM K₂HPO₄ (or 15 mM Na₂HPO₄), pH 7.1) containing 100 mM KCl or 100 mM NaCl when appropriate. Ligands were dissolved in water in a 10 mM stock solution and the required volume for the titrations was added directly in the quartz cell. The spectra were signal-averaged over four scans and baseline corrected by subtracting a buffer spectrum.

CD melting experiments were performed in the temperature range 20-100 °C, with a heating rate of 2°C/minute by monitoring the ellipticity at 263 and 290/295 nm for promoter and telomeric G4s, respectively. c-MYC (5 μM in 100 mM LiCl), 22AG (10 μM in 100 mM NaCl or 10 mM KCl + 90 mM LiCl) were annealed as described above. Spectra acquisition was performed in the presence of 8 molar equivalents of ligand. Data was converted into fraction folded (θ) plots using the following equation:

$$\theta = \frac{CD - CD_{\lambda}^{min}}{CD_{\lambda}^{max} - CD_{\lambda}^{min}} \quad (1)$$

where CD is the ellipticity of the monitored wavelength at each temperature and CD_{min} and CD_{max} are the lowest and highest ellipticity, respectively. Data points were then fitted to a Boltzmann distribution (OriginPro 8) and the melting temperatures were determined from the two state transition model used using first derivative method.

Fluorescence Resonance Energy Transfer Melting Assay

FRET melting assays were performed using a Stratagene Mx3005P qPCR instrument. Oligonucleotides were modified with fluorescein (FAM) and rhodamine (TAMRA) in 5' and 3' ends, respectively [15]. The excitation and detection wavelengths were 492 nm and 516 nm, respectively. The assays started with an initial incubation at 25 °C for 5 min, then the temperature was increased by 1 °C every minute until 95 °C was reached. The experiments were performed with samples of 0.2 μM oligonucleotide in 10 mM lithium cacodylate (pH 7.2) buffer, with appropriate concentrations of KCl, NaCl and LiCl (total salt concentration of 100 mM). For measurements in the presence of a ligand, the concentration of ligand was 5 μM. Each experimental condition was tested in duplicate on at least three separate plates. The melting temperature was determined from normalized curves as the mid-transition $T_{1/2}$ temperature.

Steady-State Fluorescence Measurements

Fluorescence titrations were performed at 25 °C using a Horiba FluoroMax4 fluorimeter equipped with a Peltier-type temperature control system. Compound **3** at 5 μM was excited at 268 nm and emission spectra were acquired between 330 and 550 nm in 30 mM phosphate buffer (15 mM KH_2PO_4 , 15 mM K_2HPO_4 , pH 7.1). Excitation and emission slits were both fixed at 5 nm and a quartz cuvette with a path length of 1 cm was used. The association between DNA and ligand was then followed by titrating the oligonucleotides and measuring the change in fluorescence. The titration was performed by adding previously annealed G4 or ds-DNA solution, followed by 3 minutes for equilibration. Data was corrected for the screening effect *i.e.* the amount of excitation light absorbed by DNA at 268 nm [16]. Fluorescence intensity (F/F_0) was plotted against DNA concentration and fitted to Stern-Volmer equation to determine binding constants (K_a) using OriginPro 8. Two equations were used depending on the quenching mechanism and resulting curves [17,18]:

$$\frac{F_0}{F} = 1 + K_{SV}[DNA] \quad (2)$$

$$\frac{F_0}{F} = Ae^{K_{sv}[DNA]} \quad (3)$$

where F and F₀ are the fluorescence intensities of the ligand in the presence and absence of DNA, respectively, A is the amplitude and K_{sv} is the binding constant.

Job-plot method of continuous variation was used in order to determine the binding stoichiometry of compound **3** with c-MYC and 22AG G4s [19]. The analysis was performed through a systematic variation of the ligand/DNA molar fraction while keeping a constant total concentration of 5 μM. The spectra acquisition was carried at 25 °C in the range of 330-550 nm.

NMR Spectroscopy

The ¹H NMR spectra were recorded at temperature of 298K on a Bruker Advance 400 MHz spectrometer equipped with a liquid inverse triple resonance (TXI) 5 mm probe or at 310K on Bruker Advance 700 MHz spectrometer equipped with a liquid TXI ¹H/¹³C/¹⁵N/²H with Z-gradient probe. Final oligonucleotide concentration was typically between 0.5-1.5 mM in 3 mm tubes for a final volume of 200 μL. Samples were annealed in 30 mM K-phosphate buffer pH 6.8 containing 100 mM KCl and supplemented with 10% D₂O. Standard 1D proton NMR spectra were recorded and referenced to external TSP (trimethylsilyl propionate). NMR titration was performed by adding increasing amounts of ligand **3** to the G4 solution at ligand:DNA ratios from 0:1 to 2:1. All spectra were processed with the software Topspin 3.2 and analyzed with MestReNova. Chemical shifts (δ) were measured in ppm. Standard homonuclear 2D-NMR NOESY and ROESY experiments were performed to study complex formation. NOESY spectra were acquired at 310K at ligand:DNA ratio of 2:1 on a 700 MHz spectrometer. Acquisition was performed with 4096 × 144 data points in F2 and F1 respectively, and exchange of longitudinal magnetization occurred during a mixing time of 600 ms. Apodization were applied by multiplying the time-domain, F2 with Gaussian Multiplication and F1 with QSIN functions respectively, data and zero filled to 256 points in F1, with final incremental phase adjustments. No symmetrization was applied to the final spectra.

Molecular Docking

The NMR structure of the 22-mer c-MYC G4 (PDB code 1XAV) was used as target for docking studies with compound **3**. The size of the box was constrained to 85 × 40 × 65 Å along the x, y, and z axes, respectively, with a grid spacing of 0.375 Å. The G4 structure and the ligand were optimized for docking using Dock Prep tool of Chimera 1.11.1 [20]. After assigning polar hydrogens and Gasteiger charges, docking simulations were carried

with AutoDock 4.0 program using Lamarckian genetic algorithm [21]. Grid maps were assigned for each atom type in the ligand using AutoGrid. Each run consisted of an initial population of 150 random individuals, a maximum number of evaluations set to 2.5×10^7 , the rate of mutation and crossover set to 0.02 and 0.8, respectively, and elitism value of 1. The ligand was set with rotatable bonds having full flexibility, while the structure of the G4 was kept rigid. Each run provided 10 different docked conformers. The best binding poses were selected based on the binding free energy and NMR results, and further analyzed with Chimera.

Results

Design and synthesis of indole-based ligands

The majority of the G4 ligands currently being developed share the same structural characteristics: a large aromatic core prone for π stacking interactions with the G-tetrads and a positive charge taking advantage of DNA phosphate backbone [10,22]. In addition, ligand structure may include protonated side arms for the recognition of G4 via actions in the loops/grooves [10,22]. Moreover, compound synthesis should take into account the “Lipinski Rule of Five” which characterizes the drug-like properties of molecules [23]. For this work three derivatives of indole were synthesized and characterized by NMR (Figures S9-S14, **Appendix A**) and high-resolution mass spectrometry (HRMS) (**Figure 4.1**). The ligands structure contains amine side arms for the interaction with the loops/grooves, and the indole moiety for the recognition of the G-tetrads, by π - π stacking and H-bonds between the NH group and O of the deoxyribose. Moreover, previous studies using the indole-containing amino acid L-tryptophan have shown that the indole ring preferentially binds guanine which could provide specificity for the G-tetrads [24].

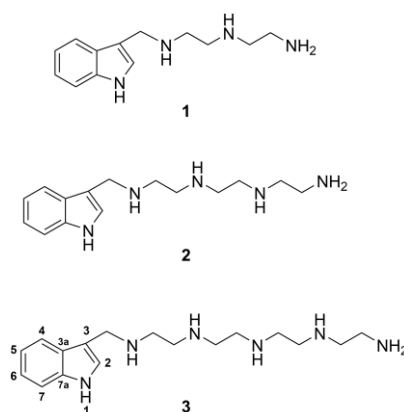


Figure 4.1 - Structures of the synthesized compounds 1-3. Ligands contain an indole aromatic core with polyamine side chains of different lengths.

Ligands binding and stabilization of G-quadruplexes

In order to study the G4 stabilizing properties of the ligands, CD and FRET assays were performed with G4s presenting all the possible folding topologies: parallel, antiparallel and hybrid-type. Human telomeric sequence 22AG, depending on the experimental conditions, is able to form two distinct conformations namely hybrid-1 and antiparallel whether it is under K^+ or Na^+ conditions, respectively [25,26]. Promoter DNA sequence c-MYC on its hand, folds into a parallel G4 conformation [27]. CD titrations were carried in order to assess the binding of the ligands to both G4 sequences and the potential of the compounds to modify the native folding topology of the G4s. Characteristic CD spectra of parallel-stranded G4s shows a positive peak at 260 nm and a negative peak at 240 nm, whereas antiparallel-stranded quadruplexes present positive peaks at 295 and 240 nm, and a negative peak at 260 nm [28]. Hybrid-type 1 spectrum has a positive peak at 290 nm, a shoulder peak at 265 nm and a weak negative peak at 240 nm [29]. Upon addition of the ligands the overall shape of the CD signature remained unchanged, although a slight decrease of ellipticity was observed at higher ligand concentration, thus demonstrating binding of the ligands and preservation of the structure (Figures S1-S3).

The stabilization of the G4s structure was then evaluated by CD and FRET thermal denaturation assays. CD melting curves of the three different G4 conformations in the presence and absence of each ligand were recorded and the ligand-induced change in the melting temperature (T_m) was measured. Thermal denaturation of the quadruplexes was monitored at the wavelength of maximum ellipticity (263, 295 and 290 nm for parallel, antiparallel and hybrid-1, respectively) in the presence of the required amount of K^+ , Na^+ or Li^+ ions (**Figure 4.2**).

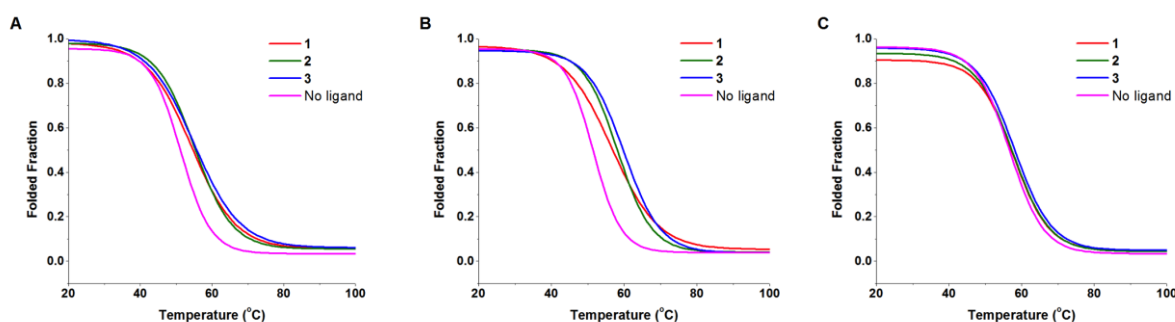


Figure 4.2 - CD melting curves of (A) c-MYC (5 μ M in 30 mM potassium phosphate buffer pH 7.1 and 100 mM LiCl), (B) hybrid-1 22AG (10 μ M in 30 mM potassium phosphate buffer pH 7.1 and 10 mM KCl and 90 mM LiCl) and (C) antiparallel 22AG (10 μ M in 30 mM sodium phosphate buffer pH 7.1 and 100 mM NaCl) in the absence and presence of ligands 1-3.

Upon addition of the ligands the stability of the G4s was enhanced by the ΔT_m values summarized in **Table 4.1**. The results indicate that the compounds significantly increase the T_m of parallel and hybrid-1 G4s ($\Delta T_m \geq 3^\circ\text{C}$) but display a poor stabilization effect in the case of the antiparallel G4 ($\Delta T_m \leq 3^\circ\text{C}$). The higher stabilization effect is observed for compound 3 with ΔT_m values of 3.7, 8.4 and 1.4 $^\circ\text{C}$ for c-MYC, hybrid-1 22AG and antiparallel 22AG, respectively. These results suggest that compound 3 is an effective stabilizer of G4s. Compound 2 moderately increased the T_m values by 3.6, 6.7 and 1.0 for c-MYC, hybrid-1 22AG and antiparallel 22AG, respectively. On its hand, compound 1 showed the lowest stabilization effect with T_m values of 2.9, 5.1 and 0.9 for c-MYC, hybrid-1 22AG and antiparallel 22AG, respectively. In order to assess whether the indole moiety significantly contributes to the thermal stabilization induced by the compounds, a experiment using the amine was performed together with hybrid-1 22AG. In the presence of tetraethylene pentamine, a stabilization was about 2 $^\circ\text{C}$ much lower than the stabilization provided by compound 3 ($\Delta T_m = 8.4^\circ\text{C}$).

Table 4.1 - Ligand-induced thermal stabilization measured by CD and FRET melting experiments.

	$\Delta T_m (^\circ\text{C})^a$ in CD melting experiments			$\Delta T_m (^\circ\text{C})^a$ in FRET melting experiments			
Ligand	c-MYC	22AG-h ^b	22AG-a ^c	FmycT	F21T-h ^b	F21T-a ^c	FdxT
1	2.9±0.1	5.1±0.1	0.9±0.1	0.0±0.3	5.6±1.8	1.3±0.4	0
2	3.6±0.2	6.7±0.1	1.0±0.1	0.7±0.8	8.8±2.1	1.1±0.7	0
3	3.7±0.1	8.4±0.1	1.4±0.1	3.0±0.4	13.7±1.1	5.5±1.9	0

^a ΔT_m represents difference in melting temperature [$\Delta T_m = T_m (\text{DNA} + \text{ligand}) - T_m (\text{DNA})$]. Experiments were performed in triplicate, ΔT_m values are reported as mean $\pm$ SD. The buffer used was 30 mM potassium or sodium phosphate, pH 7.1 for CD melting and 10 mM lithium cacodylate, pH 7.2 for FRET melting. The T_m values are 53.10 $\pm$ 0.1 $^\circ\text{C}$ [c-MYC DNA in 100 mM LiCl]; 52.3 $\pm$ 0.1 $^\circ\text{C}$ [hybrid type telomeric DNA in 10 mM KCl, 90 mM LiCl]; 50.2 $\pm$ 0.1 $^\circ\text{C}$ [antiparallel telomeric DNA in 100 mM NaCl]. All experiments were triplicated, and the values reported are average of 3 measurements with the estimated standard deviation.

^b Hybrid type conformation obtained in 10 mM K⁺.

^c Antiparallel conformation obtained in 100 mM Na⁺.

The thermal stabilization of G4s was further assessed by FRET melting experiments as well as the selectivity of the ligands using labelled G4s FmycT, F21T and double-stranded DNA FdxT. The results of the FRET melting experiments are summarized in **Table 4.1**. The thermal stabilization results obtained are in qualitative agreement with those obtained in CD melting, being compound 3 the ligand with higher ΔT_m values. The differences observed for the ΔT_m values of both techniques are either due to differences in experimental conditions and the presence of fluorophores for the modified sequences used

for FRET melting. Nonetheless, the values are within the same range and the stabilization effect is comparable. Regarding the selectivity of the ligands towards G4 structures over ds-DNA, all the compounds did not improve the thermal stability of FdxT, therefore this suggests that these are selective for G4s.

Characterization of ligand binding to G-quadruplexes

Fluorescence titration experiments were performed in order to quantify the binding affinity of compound 3 for c-MYC and 22AG G4s. Compound 3 has an excitation wavelength of 268 nm and exhibits an emission broad band centered at 398 nm. Upon titrating the G4s in a solution of compound 3, it was possible to observe a quenching of the compound fluorescence intensity (**Figure 4.3**).

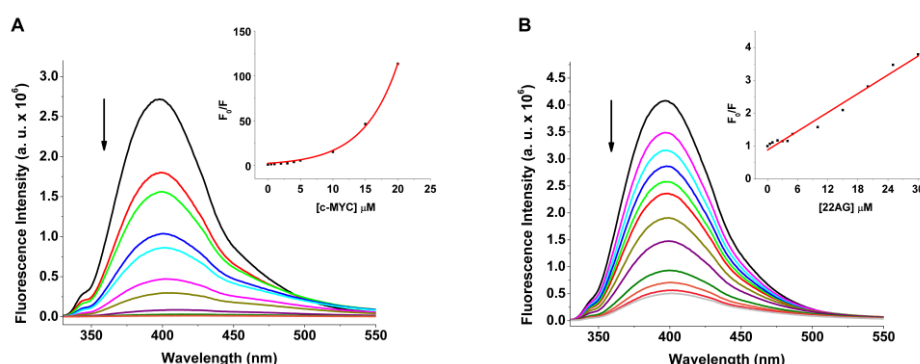


Figure 4.3 - Fluorescence emission spectra of compound 3 with successive additions of (A) c-MYC and (B) hybrid-1 22AG. Insets: Stern-Volmer plots obtained by plotting F_0/F as a function of DNA concentration and fitting to the adequate equation.

The titration data was then fitted to Stern-Volmer models and the binding constants (K_{SV}) were determined (**Figure 4.3**, insets). In the case of 22AG titration a linear model described by Equation 2 was used, whereas in the case of c-MYC a non-linear model was applied due to the resultant upward-curvature plot with exponential dependence described by Equation 3 [17,18]. K_{SV} values obtained were $1.95 \times 10^5 \text{ M}^{-1}$ and $1.0 \times 10^5 \text{ M}^{-1}$ for c-MYC and 22AG, respectively. A similar experiment performed with duplex DNA gives a K_{SV} value of $9.9 \times 10^3 \text{ M}^{-1}$ (Figure S4) i.e., a ≈ 10 -fold lower binding constant.

In order to gain insight on the binding stoichiometry for compound 3 interaction with the tested G4s, Job plot analysis was performed [19]. Through the plot of fluorescence versus the mole fraction of 3 (χ) and fitting of data points both at high and low molar fractions is possible to determine the number of ligand-binding sites (Figure S5). The molar fraction determined by the intersection of the linear regressions was 0.60 and 0.57 for c-MYC and

22AG, respectively, which indicates the possible coexistence of 2:1 and 1:1 (ligand:DNA) stoichiometries [30].

NMR studies of c-MYC/ligand interaction

NMR spectroscopy is a powerful tool for the study of G4s structure and ligand interactions [31,32]. ^1H NMR titration experiments were additionally performed to further study the binding of compound 3 and G4s. Since 22AG ^1H spectrum in K^+ is not easily interpretable due to the coexistence of several topologies, only c-MYC was used for NMR experiments. Moreover, telomeric sequences are known for their aggregation properties which complicate the analysis [33]. c-MYC gives a well resolved imino proton NMR spectra and its resonances were already fully assigned and published [27]. Titrations were carried out by adding compound 3 to a tube containing pre-folded c-MYC and following the changes in chemical shift of the G4 spectra (**Figure 4.4**).

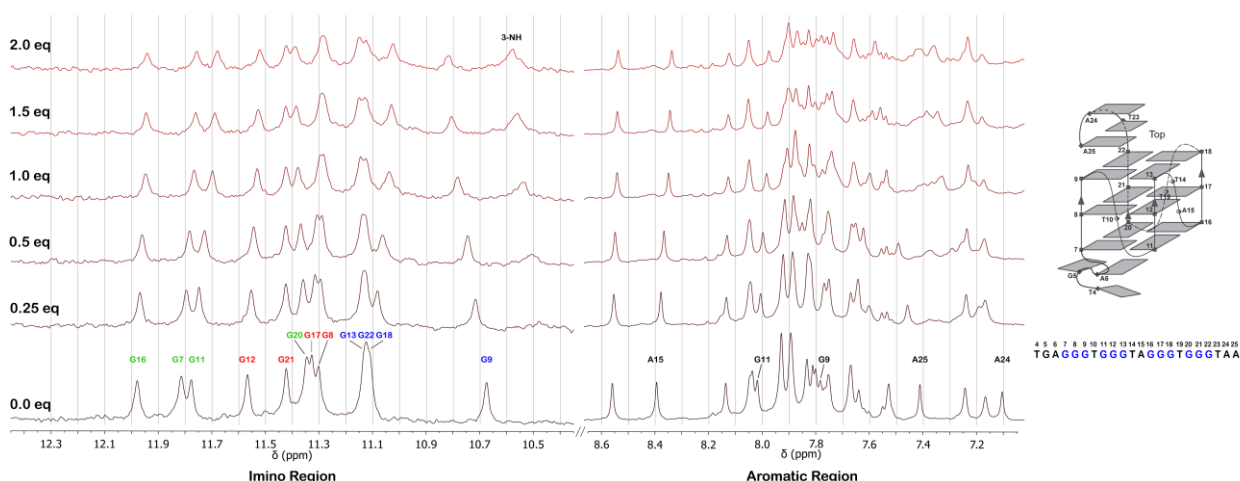


Figure 4.4 - Imino and aromatic regions of c-MYC ^1H spectra upon titration of compound 3 at 298K. Top tetrad guanines are colored in blue, middle in red and bottom tetrad in green. On the right a schematic representation of c-MYC G4 structure – Reprinted (adapted) with permission from “Ambrus, A.; Chen, D.; Dai, J.; Jones, R. A.; Yang, D. *Biochemistry* 2005, 44 (6), 2048–2058”. Copyright 2005 American Chemical Society.

The analysis of the spectrum demonstrates that the G4 signals shifted upon ligand titration, which is indicative of ligand interaction [32]. The major shifts were observed for imino protons belonging to residues G9, G11 and G18, the Me proton of T23, which stacks on the top tetrad, and the aromatic protons of A24 and A25 which stack with residue T23 [27]. Also the anomeric and aromatic proton of residue A15 and T14, respectively, are affected upon ligand binding. The significantly shifted residues and protons are summarized in **Table 4.2**.

Table 4.2 - Chemical shifts and chemical shift changes of selected ¹H NMR signals.

Residue - Atom	Chemical shift perturbation ^a	
	δ	$\Delta\delta$ ^b
G7 - H1	11.71	-0.06
G9 - H1	10.77	0.14
G9 - H8	7.74	-0.05
G11 - H1	11.63	-0.10
G11 - H8	7.98	-0.04
G12 - H1	11.47	-0.04
T14 - H6	7.58	-0.09
A15 - H1'	6.65	-0.07
A15 - H2	8.34	-0.06
G16 - H1	11.89	-0.04
G17 - H1	11.24	-0.04
G18 - H1	10.97	-0.10
T23 - Me	1.61	0.10
A24 - H2	7.42	0.31
A25 - H2	7.58	0.17

^a Measured at 25 °C in ppm (δ) and referenced from external TSP, $R = [\text{ligand}]/[\text{DNA}] = 2$.

^b $\Delta\delta = \delta_{\text{bound}} - \delta_{\text{free}}$

All other protons are almost unchanged or very small shifts were observed ($\Delta\delta \leq 0.03$ ppm). At around 10.5 ppm is possible to detect the appearance of a single broad peak upon titration that is attributed to the indole NH of compound 3. After the addition of 2 equivalents of 3 no further changes in the chemical shifts are observed. A visual representation of the most shifted residues on the c-MYC G4 structure allows a deeper understanding of the possible binding sites based on the chemical shifts (**Figure 4.5**).

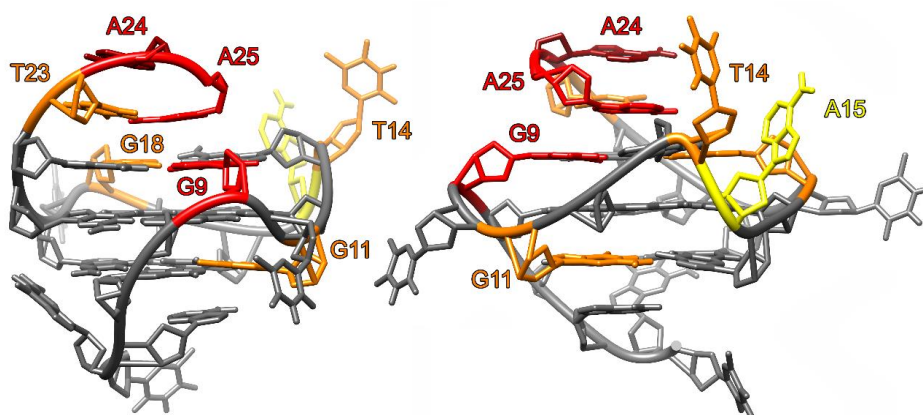


Figure 4.5 - Mapping of the chemical shift perturbation on the NMR model of c-MYC (PDB ID:1XAV) with residues colored based on the higher $\Delta\delta$ relative to the average $\Delta\delta$ (average $\Delta\delta = 0.07$) where: yellow is for $\leq 25\%$, orange for 25-75% and red for $\geq 75\%$.

The analysis of the figure suggests that 3' tetrad may be involved in binding as well as the side loops. Also the 5' tetrad may be involved in complex formation, despite the smaller chemical shift perturbation (**Table 4.2**). The lack of clearly detectable ligand signals on the same NMR timescale indicates that the ligand exchanges between binding sites, consistent with the Job plot results (Figure S5) that suggest the coexistence of 2:1 and 1:1 stoichiometries [29]. NOESY spectra of the ligand:DNA 2:1 mixture was also acquired (**Figures 4.6**, S6-7). The NOESY experiment was performed at 310 K as it favored the formation of a single G4 conformation (Figure S8). At 298 K, it is possible to observe more than one set of peaks in the imino region due to multiple conformations adopted by the 3'-end capping structure formed by T23, A24 and A25 [27,34]. It is important to note that the NMR spectra of the imino region of the ligand:DNA complex is similar at 298 K and 310 K, indicating that the same complex is formed at both temperatures. After the assignment of the guanines H1s (Figure S8), we were able to assign the protons H1-H8/H2 region (Figure S6) and then, using the sequential assignment strategy [27,35], to assign protons H8/H6-H1' (**Figure 4.6**). The assignment of H8/H6s of nucleotide residues other than guanines was assisted by observing the H8/H6-methyl (data not shown) and H1'-H2'/H2'' regions (Figure S7). Due to strong overlap between the signals from both the ligand and G4, it is difficult to correctly identify intermolecular correlation peaks. However, NOE contacts were found between compound 3 aromatic moiety, namely protons H5, H6 and H7, and guanine G11 anomeric protons (**Figure 4.6**), in agreement with the relatively large chemical shift perturbation observed in ^1H spectra for G11.

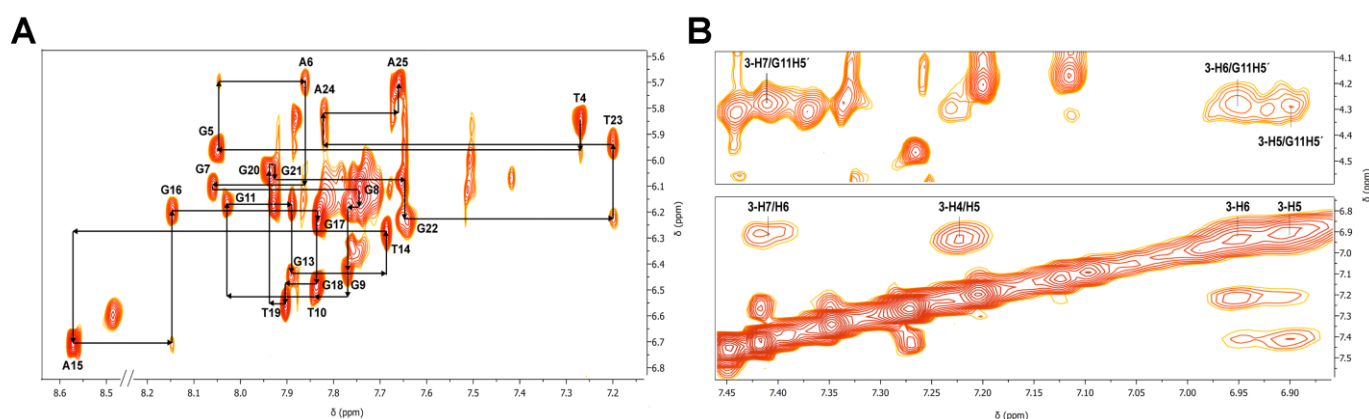


Figure 4.6 - (A) The expanded H8/H6-H1' region of the 2D-NOESY spectrum of the ligand:DNA 2:1 complex at 310 K. The sequential assignment pathway is represented. (B) Expanded region of the 2D-NOESY spectrum showing the intramolecular NOEs within the aromatic core of compound 3 (bottom) and some intermolecular NOEs with the DNA (top). The cross-peaks of compound 3 are broader when compared to those from DNA. Each cross-peak is labeled as F2/F1 dimensions.

Molecular docking simulations

In order to predict the binding mode of the ligand, molecular docking calculations were performed using AutoDock. Such strategy has been widely used to provide insight on the interaction between small ligands and G4s [36]. Compound **3** and c-MYC G4 (PDB entry: 1XAV) were used for the simulations. The best four binding poses are depicted in **Figure 4.7**. The indole aromatic moiety of compound **3** seems to interact with guanine G11 while the amine arm adds additional contacts with the loops/grooves of the G4 structure. The binding free energies of complexes A-D are -6.79, -4.65, -4.94 and -9.75 kcal/mol, respectively. In complex A, three H-bonds are formed, namely between the NH groups of the ligand amine arm and atoms G9O3', G9OP1 and T10OP2. In complex B, four H-bonds are formed between the NH groups of the amine arm and atoms G9O3', T10OP2, G11OP1 and G11OP2. For complex C, three H-bonds are formed also between the NH groups of the amine arm and atoms G11OP2, G12O5' and G12OP2. Finally, in complex D only one H-bond is detected between one NH group of the ligand side arm and atom G13OP2 from the G4 structure.

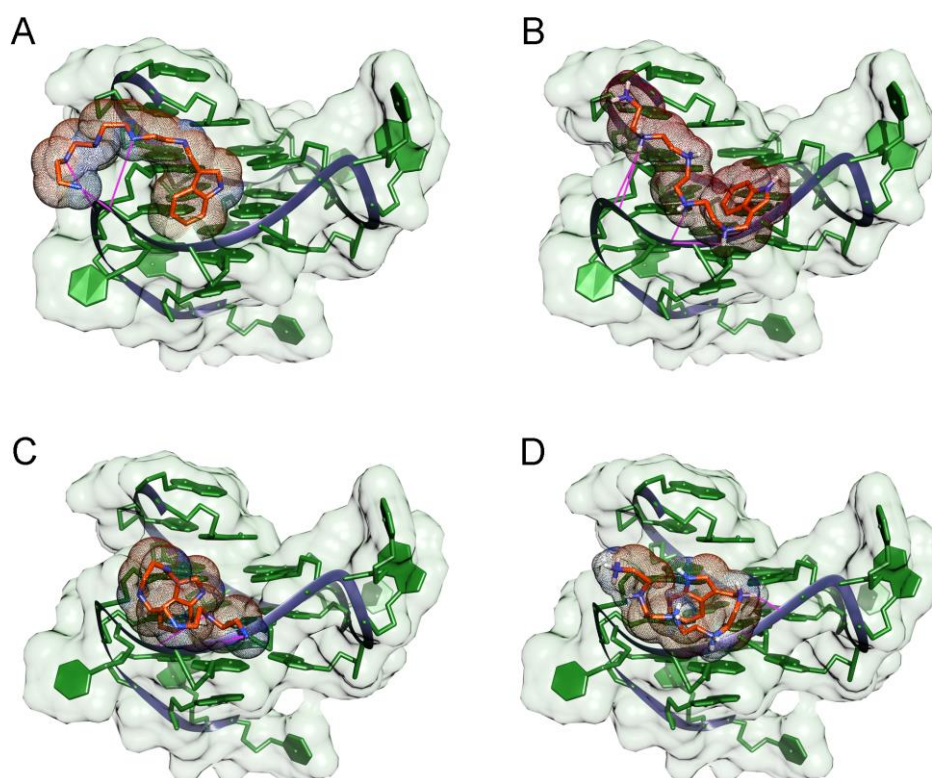


Figure 4.7 - Best models of compound **3**/c-MYC G4 complex produced by docking. All the binding poses show the ligand interacting with G11 through the indole aromatic moiety, with additional hydrogen bonding and electrostatic interactions with the loops. G4 structure is represented in green/purple ribbon, ligand as orange sticks and pink lines indicate H-bonds.

Discussion

G4s are attractive targets for drug development in the field of cancer, viral and neurological diseases [37–39]. Despite the recent progresses in developing efficient and selective G4 ligands, only a few candidate drugs have reached clinical trials such as Quarfloxin which reached Phase II trials [40]. Candidate molecules for drug development should exhibit drug-like properties, high specificity and promote the stabilization of G4 structures among others. In this context, three indole derivatives have been designed in order to provide suitable candidate molecules for drug development (**Figure 4.1**).

The results from CD titrations indicate that the compounds are able to bind the G4 structures tested without disrupting its native topology, even at higher concentrations. In fact, at lower ligand concentration there is an ellipticity increase showing further increase of the fraction of G4 structures (Figures S1-S3). The stabilization effect of the ligands was assessed by CD and FRET melting experiments that shown that compound 3 promoted the higher T_m increase with all the G4s tested, whereas compound 1 was the least stabilizing. A structure-activity relationship based on ΔT_m values suggests that the increase in the NH_2 content on the protonable side arm favors the stabilization of the G4 structure, as seen for compound 3. Therefore, the amine arms seem to play an important role in the interaction between the ligands and the G4s. The presence of the indole moiety is also important for the thermal stabilization as emphasized by the experiment using the amines only, where lower degree of stabilization was observed. The ΔT_m values obtained are similar to other indole-based ligands reported previously [14,41]. The relatively low stabilization for these molecules as compared to the best ligands reported in the literature may be compensated by i) lower molecular weight and ii) better biological and pharmacological properties such as undesired toxicity or cell penetration. The ligands also discriminate between different G4 structures as observed from the almost residual stabilization of antiparallel topology when compared to hybrid-1 telomeric G4. Regarding the specificity of the ligands towards G4 over duplex DNA, the FRET melting results using a ds-DNA sequence indicate that the ligands do not bind and stabilize duplex DNA. Based on these results compound 3 was selected for further studies.

Ligand binding was further characterized by fluorometry. The quenching of fluorescence is described by the Stern-Volmer equation. Linear plots relate to collisional (static) or dynamic (complex formation) quenching whereas systems including both static and dynamic quenching tend to present an upward curvature and are described by non-linear

relationships [17]. These positive deviations of the Stern-Volmer equation are expected when the quenching is greater which is verified for c-MYC G4. Upon G4 titration there was a clear concentration-dependent quenching of the ligand fluorescence intensity. The determined binding affinity constants in the range $K_{SV} \approx 10^5 \text{ M}^{-1}$ are similar to the reported values for moderate ligand-quadruplex binding interactions [42–44]. On the other hand, binding affinity of $K_{SV} \approx 10^3 \text{ M}^{-1}$ was obtained for duplex DNA. The tenfold difference observed for duplex DNA binding constant explains the specificity of compound 3. The data from Job plot analysis suggests a co-existence of 1:1 and 2:1 ligand:G4 complexes for both c-MYC and 22AG [30].

NMR spectroscopy was used in order to learn more about the molecular nature behind the interaction between compound 3 and c-MYC G4 structure. Pu22 is a modified sequence of the c-MYC NHE III₁ regulatory sequence whose structure was already resolved by NMR (PDB: 1XAV) [27]. Pu22 forms a single parallel G4 conformation in physiologically relevant K⁺, with 12-well-resolved imino proton peaks, consistent with 3 tetrads (G7-G11-G16-G20, G8-G12-G17-G21 and G9-G13-G18-G22) and 1:2:1 side loops (**Figure 4.4**). Upon titration of compound 3 some imino protons such as G7, G11, G17 and G18 shifted upfield while others as G20 and G9 shifted downfield. Overall, all imino peaks became broader, but the overall profile, number and peak order remained unchanged. Resonances were assigned using previous published data. Since there were no imino peaks overlapping or disappearance during the titration it was straightforward to follow chemical shifts before and after ligand addition. This indicates that the ligand binds through both stacking and groove binding modes, and that the tetrads organization is maintained [45–47]. The spectra kept constant after 2 equivalents of compound 3 supporting the existence of a 2:1 ligand:DNA complex. By the analysis of the most shifted resonances it is possible to predict the possible binding site of the ligand (**Figure 4.5**). Changes in the chemical shift ($\Delta\delta \geq 0.1 \text{ ppm}$) of imino protons of G9 and G18, and in the aromatic protons of A24 and A25 suggest that the ligand interacts mainly with the top tetrad, possibly through stacking interactions between the indole ring and terminal 3' tetrad bases. Moreover, the smaller chemical shift perturbation observed for the 5' tetrad residues indicate that the ligand may also bind the bottom tetrad, supporting the 2:1 ligand:DNA complex formation. The terminal 5' tetrad imino protons are somewhat shifted ($0.04 \leq \Delta\delta \leq 0.1 \text{ ppm}$), especially G11 which indicates similar involvement of the two external tetrads, despite a clear preference for the top tetrad. The existence of NOE contacts between the aromatic moiety of compound 3 and guanine G11, supports the proposed mode of interaction (**Figure 4.6**). These correlations are consistent with the relatively large chemical shift perturbations of G11 protons (**Table 4.2**). The flanking residues T23, A24 and A25

resonances shifted downfield upon titration which indicates that the binding mode to these residues is different than stacking [48]. Loop residues T14 and A15, which belong to the longest loop, present shifted resonances ($0.07 \leq \Delta\delta \leq 0.09$ ppm) suggesting ligand interaction via action in the loops/grooves interface, possibly with the protonable arms of compound 3. The use of higher temperature for the NOESY experiment allowed us to better observe the major conformation of c-MYC G4 as shown before [27]. Moreover, the minor conformations, when present, do not exchange with the major conformation, being that no exchange NOE cross-peak is expected [27].

In order to further understand the binding mode, unrestrained molecular docking calculations were performed. As anticipated, the docking results show more than one binding site in the surface of the G4 structure as the ligand seems to bind in a 1:1 and 2:1 ligand:G4 mixed mode. The binding site which corroborates with the data derived from the NMR experiments is presented in **Figure 4.7**. The binding mode predictions are in agreement with the NMR results as the ligand establishes contacts with the most shifted residues of the ^1H NMR spectrum. The binding mode seems to be through interactions of the indole aromatic moiety with G11 residue, while the amine arm interacts with the grooves/loops interface, namely the groove formed between G9, T10, G11 and the 3'-cap residues T23, A24 and A25 backbone. Comparing Figures 5 and 7, there seems to be a correlation between the predicted binding sites and the docked pose, which supports the proposed binding mode. The ligand establishes H-bonds and electrostatic interactions with the G4 structure. The H-bonds formed between the NH groups of the ligand amine arm and the residues backbone indicate the potential of the ligand as groove binder.

Conclusions

Compound 3 has shown to be a selective and potent G4 ligand with interesting biophysical properties and seems to bind c-MYC G4 through mixed-mode interactions in a 2:1 ratio. This study has shown that indole-based G4 ligands are promising candidates for future drug-like compounds development as we have provided evidence for its G4 binding and stabilizing properties. In addition, this chemical moiety can be used in ligands that have a larger π -stacking core for improved binding and selectivity towards G4. Also this demonstrates that future lead optimizations of G4 ligands can employ the polyamine side chains in order to improve some drug-like properties such as the number of hydrogen bonds acceptor/donor or water solubility, improving ligands for potential therapeutic use.

Overall, these insights may help in selecting prosperous candidates for development of G4 ligands and to explore its binding properties.

Acknowledgments

This work was supported by FCT - 'Fundação para a Ciência e a Tecnologia' (project FCOMP-01-0124-FEDER-041068 and EXPL/QEQ-MED/1068/2013), FEDER funds through the POCI - COMPETE 2020 - Operational Programme Competitiveness and Internationalization in Axis I - Strengthening research, technological development and innovation (Project POCI-01-0145-FEDER-007491) and National Funds by FCT - 'Fundação para a Ciência e a Tecnologia' (project UID/Multi /00709/2013). It was also supported by grants from Agence Nationale de la Recherche ('Quarupdiem' [ANR-12-BSV8-0008-01] and Conseil Régional d'Aquitaine to J.L.M. G.F.S. is the recipient of a "Chaire Inserm-Université de Bordeaux".

C. Cruz acknowledges post-doctoral grant from FCT SFRH/BPD/100015/2014. J. Carvalho acknowledges a fellowship from UBI-Santander/Totta (Ref BID/ICI-CICS-BST-UBI). We thank the Biophysical & Structural Chemistry platform of the IECB (CNRS UMS3033, Inserm US001, Univ. Bordeaux) for the access to their instruments. We also thank Aurore Guédin for her insight concerning the FRET experiments.

References

- [1] Huppert, J. L. (2008) Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. Chemical Society reviews 37, 1375–1384. <https://doi.org/10.1039/b702491f>.
- [2] Bang, I. (1910) Untersuchungen über die Guanylsäure. Biochemische Zeitschrift 26, 293–311.
- [3] Zhang, S., Wu, Y., and Zhang, W. (2014) G-quadruplex structures and their interaction diversity with ligands. ChemMedChem 9, 899–911. <https://doi.org/10.1002/cmdc.201300566>.
- [4] Phan, A. T. (2010) Human telomeric G-quadruplex: Structures of DNA and RNA sequences. FEBS Journal 277, 1107–1117. <https://doi.org/10.1111/j.1742-4658.2009.07464.x>.
- [5] Phan, A. T., Luu, K. N., and Patel, D. J. (2006) Different loop arrangements of

intramolecular human telomeric (3+1) G-quadruplexes in K⁺ solution. *Nucleic Acids Research* 34, 5715–5719. <https://doi.org/10.1093/nar/gkl726>.

[6] Ou, T. M., Lu, Y. J., Tan, J. H., Huang, Z. S., Wong, K. Y., and Gu, L. Q. (2008) G-quadruplexes: Targets in anticancer drug design. *ChemMedChem* 3, 690–713. <https://doi.org/10.1002/cmdc.200700300>.

[7] Carvalho, J., Ferreira, J., Pereira, P., Coutinho, E., Guédin, A., Nottelet, P., Salgado, G. F., Mergny, J.-L., Queiroz, J. A., Sousa, F., Cabrita, E. J., and Cruz, C. (2016) Stabilization of novel immunoglobulin switch regions G-quadruplexes by naphthalene and quinoline-based ligands. *Tetrahedron* 72, 1229–1237. <https://doi.org/10.1016/j.tet.2016.01.022>.

[8] Biffi, G., Tannahill, D., McCafferty, J., and Balasubramanian, S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nature chemistry* 5, 182–6. <https://doi.org/10.1038/nchem.1548>.

[9] Metifiot, M., Amrane, S., Litvak, S., and Andreola, M.-L. (2014) G-quadruplexes in viruses: function and potential therapeutic applications. *Nucleic Acids Research* 42, 12352–12366. <https://doi.org/10.1093/nar/gku999>.

[10] Collie, G., and Parkinson, G. (2011) The application of DNA and RNA G-quadruplexes to therapeutic medicines. *Chemical Society reviews* 40, 5867–5892. <https://doi.org/10.1039/c1cs15067g>.

[11] Grand, C. L., Han, H., Muñoz, R. M., Weitman, S., Von Hoff, D. D., Hurley, L. H., and Bearss, D. J. (2002) The cationic porphyrin TMPyP4 down-regulates c-MYC and human telomerase reverse transcriptase expression and inhibits tumor growth in vivo. *Molecular cancer therapeutics* 1, 565–573.

[12] Burger, A. M., Dai, F., Schultes, C. M., Reszka, A. P., Moore, M. J., Double, J. A., and Neidle, S. (2005) The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function. *Cancer research* 65, 1489–96. <https://doi.org/10.1158/0008-5472.CAN-04-2910>.

[13] Kim, M. Y., Gleason-Guzman, M., Izbička, E., Nishioka, D., and Hurley, L. H. (2003) The different biological effects of Telomestatin and TMPyP4 can be attributed to their selectivity for interaction with intramolecular or intermolecular G-quadruplex structures. *Cancer Research* 63, 3247–3256.

[14] Nagesh, N., Raju, G., Srinivas, R., Ramesh, P., Reddy, M. D., and Reddy, C. R. (2015)

A dihydroindolizino indole derivative selectively stabilizes G-quadruplex DNA and down-regulates c-MYC expression in human cancer cells. *Biochimica et biophysica acta* 1850, 129–140. <https://doi.org/10.1016/j.bbagen.2014.10.004>.

[15] De Rache, A., and Mergny, J.-L. (2015) Assessment of selectivity of G-quadruplex ligands via an optimised FRET melting assay. *Biochimie* 115, 194–202. <https://doi.org/10.1016/j.biochi.2015.06.002>.

[16] Lakowicz, J. R. (2002) *Topics in Fluorescence Spectroscopy*. Kluwer Academic Publishers. <https://doi.org/10.1007/b115592>.

[17] Lakowicz, J. R. (2006) *Principles of fluorescence spectroscopy* Third Edit. Springer. <https://doi.org/10.1007/978-0-387-46312-4>.

[18] Tansakul, C., Lilie, E., Walter, E. D., Iii, F. R., Wolcott, A., Jin, Z., Millhauser, G. L., Braslau, R., Rivera III, F., Wolcott, A., Jin, Z., Millhauser, G. L., and Braslau, R. (2011) Distance-dependent Fluorescence Quenching and Binding of CdSe Quantum Dots by Functionalized Nitroxide Radicals. *Journal of Physical Chemistry C* 114, 7793–7805. <https://doi.org/10.1021/jp1005023>. Distance-dependent.

[19] Huang, C. Y. (1982) *Enzyme Kinetics and Mechanism - Part C: Intermediates, Stereochemistry, and Rate Studies*. *Methods in Enzymology*. Elsevier. [https://doi.org/10.1016/S0076-6879\(82\)87029-8](https://doi.org/10.1016/S0076-6879(82)87029-8).

[20] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry* 25, 1605–12. <https://doi.org/10.1002/jcc.20084>.

[21] Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of computational chemistry* 30, 2785–91. <https://doi.org/10.1002/jcc.21256>.

[22] Monchaud, D., and Teulade-Fichou, M.-P. (2008) A hitchhiker's guide to G-quadruplex ligands. *Organic & biomolecular chemistry* 6, 627–636. <https://doi.org/10.1039/b714772b>.

[23] Lipinski, C. A. (2004) Lead- and drug-like compounds: the rule-of-five revolution. *Drug discovery today. Technologies* 1, 337–341.

<https://doi.org/10.1016/j.ddtec.2004.11.007>.

[24] Wilson, K. A., Kellie, J. L., and Wetmore, S. D. (2014) DNA-protein π -interactions in nature: Abundance, structure, composition and strength of contacts between aromatic amino acids and DNA nucleobases or deoxyribose sugar. *Nucleic Acids Research* 42, 6726–6741. <https://doi.org/10.1093/nar/gku269>.

[25] Ambrus, A., Chen, D., Dai, J., Bialis, T., Jones, R. a., and Yang, D. (2006) Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Research* 34, 2723–2735. <https://doi.org/10.1093/nar/gkl348>.

[26] Wang, Y., and Patel, D. J. (1993) Solution structure of the human telomeric repeat d[AG₃(T₂AG₃)₃] G-tetraplex. *Structure* (London, England: 1993) 1, 263–282. [https://doi.org/10.1016/0969-2126\(93\)90015-9](https://doi.org/10.1016/0969-2126(93)90015-9).

[27] Ambrus, A., Chen, D., Dai, J., Jones, R. A., and Yang, D. (2005) Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. *Biochemistry* 44, 2048–2058. <https://doi.org/10.1021/bi048242p>.

[28] Diveshkumar, K. V., Sakrikar, S., Harikrishna, S., Dhamodharan, V., and Pradeepkumar, P. I. (2014) Targeting promoter G-quadruplex DNAs by indenopyrimidine-based ligands. *ChemMedChem* 9, 2754–2765. <https://doi.org/10.1002/cmdc.201402394>.

[29] Pagano, B., Amato, J., Iaccarino, N., Cingolani, C., Zizza, P., Biroccio, A., Novellino, E., and Randazzo, A. (2015) Looking for efficient G-quadruplex ligands: Evidence for selective stabilizing properties and telomere damage by drug-like molecules. *ChemMedChem* 10, 640–649. <https://doi.org/10.1002/cmdc.201402552>.

[30] Lavrado, J., Brito, H., Borralho, P. M., Ohnmacht, S. a., Kim, N.-S., Leitão, C., Pisco, S., Gunaratnam, M., Rodrigues, C. M. P., Moreira, R., Neidle, S., and Paulo, A. (2015) KRAS oncogene repression in colon cancer cell lines by G-quadruplex binding indolo[3,2-c]quinolines. *Scientific Reports* 5, 9696. <https://doi.org/10.1038/srep09696>.

[31] Adrian, M., Heddi, B., and Phan, A. T. (2012) NMR spectroscopy of G-quadruplexes. *Methods* 57, 11–24. <https://doi.org/10.1016/j.ymeth.2012.05.003>.

[32] Webba da Silva, M. (2007) NMR methods for studying quadruplex nucleic acids.

Methods 43, 264–277. <https://doi.org/10.1016/j.ymeth.2007.05.007>.

[33] Kato, Y., Ohyama, T., Mita, H., and Yamamoto, Y. (2005) Dynamics and thermodynamics of dimerization of parallel G-quadruplexed DNA formed from d(TTAGn) (n=3-5). *Journal of the American Chemical Society* 127, 9980–9981. <https://doi.org/10.1021/ja050191b>.

[34] Dai, J., Carver, M., Hurley, L. H., and Yang, D. (2011) Solution structure of a 2:1 quindoline-c-MYC G-quadruplex: Insights into G-quadruplex-interactive small molecule drug design. *Journal of the American Chemical Society* 133, 17673–17680. <https://doi.org/10.1021/ja205646q>.

[35] Dai, J., Carver, M., Punchihewa, C., Jones, R. A., and Yang, D. (2007) Structure of the hybrid-2 type intramolecular human telomeric G-quadruplex in K⁺ solution: Insights into structure polymorphism of the human telomeric sequence. *Nucleic Acids Research* 35, 4927–4940. <https://doi.org/10.1093/nar/gkm522>.

[36] Artese, A., Costa, G., Distinto, S., Moraca, F., Ortuso, F., Parrotta, L., and Alcaro, S. (2013) Toward the design of new DNA G-quadruplex ligands through rational analysis of polymorphism and binding data. *European Journal of Medicinal Chemistry* 68, 139–149. <https://doi.org/10.1016/j.ejmech.2013.07.022>.

[37] Patel, D. J., Phan, A. T., and Kuryavyi, V. (2007) Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: Diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Research* 35, 7429–7455. <https://doi.org/10.1093/nar/gkm711>.

[38] Balasubramanian, S., and Neidle, S. (2009) G-quadruplex nucleic acids as therapeutic targets. *Current Opinion in Chemical Biology* 13, 345–353. <https://doi.org/10.1016/j.cbpa.2009.04.637>.

[39] Simone, R., Fratta, P., Neidle, S., Parkinson, G. N., and Isaacs, A. M. (2015) G-quadruplexes: Emerging roles in neurodegenerative diseases and the non-coding transcriptome. *FEBS Letters* 589, 1653–1668. <https://doi.org/10.1016/j.febslet.2015.05.003>.

[40] Balasubramanian, S., Hurley, L. H., and Neidle, S. (2011) Targeting G-quadruplexes in gene promoters : a novel anticancer strategy? *Nature reviews. Drug discovery* 10, 261–275. <https://doi.org/10.1038/nrd3428>.

- [41] Amato, J., Iaccarino, N., Pagano, B., Morigi, R., Locatelli, A., Leoni, A., Rambaldi, M., Zizza, P., Biroccio, A., Novellino, E., and Randazzo, A. (2014) Bis-indole derivatives with antitumor activity turn out to be specific ligands of human telomeric G-quadruplex. *Frontiers in chemistry* 2, 54. <https://doi.org/10.3389/fchem.2014.00054>.
- [42] Ranjan, N., and Arya, D. P. (2013) Targeting C-myc G-quadruplex: Dual recognition by aminosugar- bisbenzimidazoles with varying linker lengths. *Molecules* 18, 14228–14240. <https://doi.org/10.3390/molecules18114228>.
- [43] Martino, L., Virno, A., Pagano, B., Virgilio, A., Di Micco, S., Galeone, A., Giancola, C., Bifulco, G., Mayol, L., and Randazzo, A. (2007) Structural and thermodynamic studies of the interaction of distamycin A with the parallel quadruplex structure [d(TGGGGT)]₄. *Journal of the American Chemical Society* 129, 16048–16056. <https://doi.org/10.1021/ja075710k>.
- [44] White, E. W., Tanious, F., Ismail, M. a., Reszka, A. P., Neidle, S., Boykin, D. W., and Wilson, W. D. (2007) Structure-specific recognition of quadruplex DNA by organic cations: Influence of shape, substituents and charge. *Biophysical Chemistry* 126, 140–153. <https://doi.org/10.1016/j.bpc.2006.06.006>.
- [45] Jaumot, J., and Gargallo, R. (2012) Experimental Methods for Studying the Interactions between G-Quadruplex Structures and Ligands. *Current Pharmaceutical Design* 18, 1900–1916. <https://doi.org/10.2174/138161212799958486>.
- [46] Scaglioni, L., Mondelli, R., Artali, R., Ric-cardi Sirtori, F., Mazzini, S., and Riccardi Sirtori, F. (2016) Nemorubicin and doxorubicin bind the G-quadruplex sequences of the human telomeres and of the c-MYC promoter element Pu22. *BBA - General Subjects* 1860, 1129–1138. <https://doi.org/10.1016/j.bbagen.2016.02.011>.
- [47] Ferreira, R., Artali, R., Benoit, A., Gargallo, R., Eritja, R., Ferguson, D. M., Sham, Y. Y., and Mazzini, S. (2013) Structure and Stability of Human Telomeric G-Quadruplex with Preclinical 9-Amino Acridines. *PLoS ONE* 8, 1–15. <https://doi.org/10.1371/journal.pone.0057701>.
- [48] Li, Q., Xiang, J., Li, X., Chen, L., Xu, X., Tang, Y., Zhou, Q., Li, L., Zhang, H., Sun, H., Guan, A., Yang, Q., Yang, S., and Xu, G. (2009) Stabilizing parallel G-quadruplex DNA by a new class of ligands: Two non-planar alkaloids through interaction in lateral grooves. *Biochimie* 91, 811–819. <https://doi.org/10.1016/j.biochi.2009.03.007>.

Chapter 5

Phenanthroline Polyazamacrocycles as G-quadruplex DNA Binders

This chapter was published in:

Carvalho, J., Quintela, T., Gueddouda, N. M., Bourdoncle, A., Mergny, J.-L., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2018) Phenanthroline polyazamacrocycles as G-quadruplex DNA binders. *Organic and Biomolecular Chemistry* 16, 2776–2786. <https://doi.org/10.1039/c8ob00247a>.

Phenanthroline Polyazamacrocycles as G-quadruplex DNA Binders

Abstract

Targeting quadruplex DNA structures with small molecules is a promising strategy for anti-cancer drug design. Four phenanthroline polyazamacrocycles were studied for their binding affinity, thermal stabilization, inhibitory effect on the activity of helicase towards human telomeric 22AG and oncogene promoter c-MYC G4s, and their ability to inhibit *Taq* polymerase-mediated DNA extension. The FRET melting assay indicates that the melting temperature increases (ΔT_m values) of c-MYC and 22AG G4s are 17.2 and 20.3 °C, respectively, for ligand [32]phen₂N₄ followed by [16]phenN₂ (11.3 and 15.0 °C, for c-MYC and 22AG, respectively). Competitive FRET assays show that [32]phen₂N₄ and [16]phenN₂ exhibit G4 selectivity over duplex DNA. Different G4s were compared; no considerable selectivity of the ligands for a specific G4 was found. CD confirms the formation of G4 structures and the melting experiments show that [16]phenN₂ and [32]phen₂N₄ are the most stabilizing ligands with ΔT_m of 19.3 °C and 15.1 °C, respectively, at 5 molar equivalents for c-MYC G4. The fluorescent intercalator displacement (FID) assay also demonstrates that ligand [32]phen₂N₄ furnishes very low DC₅₀ values (0.87–1.24 μM), indicating high stabilization of c-MYC and 22AG G4s. These results suggest that the hexyl chain in these compounds plays an important role in regulating the stabilization of these G4s. Binding constants, determined by fluorescence titrations, indicate a moderate ligand-G4 binding with K_{SV} between 10⁵–10⁶ M⁻¹ in which [16]phenN₂ has a slightly higher apparent binding constant for telomeric 22AG G4 than for c-MYC G4. The ligands ability to inhibit *Taq* polymerase confirms the biological activity of [16]phenN₂ and [32]phen₂N₄ against the c-MYC G4. In addition, ligands [32]phen₂N₄ and [16]phenN₂ affect the unwinding activity of Pif1 in the presence of DNA systems harboring c-MYC and telomeric G4 motifs.

Introduction

Since their discovery more than 50 years ago, DNA G4s have shown to have a great potential as therapeutic targets, with particular relevance for cancer therapy [1–3].

Targeting these four-stranded DNA structures with small molecules is a promising strategy for anti-cancer drug design, since they might inhibit telomerase activity and other telomere-binding proteins, that could trigger a DNA damage response and eventual cell damages. The formation of G4s at the promoter regions of human oncogenes is related to the regulation of gene transcription and to biological dysfunctions that selectively alter the integrity of cancer cells and block transcription or translation of particular genes by stabilizing G4 DNA. Moreover, by stabilizing G4 structures, ligands inhibit the natural action of helicases that resolve G4 structures leading to genomic instability and replication stalling [2]. As a strategy for G4 ligand design, high selectivity between duplex DNA and G4s DNA is a crucial step, because ligand interaction with duplex DNA may lead to acute toxic and intolerable side effects on normal tissues. Most commonly, the overall chemical construct of G4 ligands is comprised of an extended planar aromatic scaffold, which interacts primarily via π - π stacking to the terminal G-tetrad, and side chains tethered to the aromatic scaffold which interact with the sugar-phosphate backbone, loops or grooves of the G4s. A wide variety of G4 ligand variants exist, such as acridine derivatives, bisquinolinium derivatives, cationic porphyrins, ethidium derivatives, triarylimidazole derivatives and transition metal complexes [4,5].

However, despite promising results in both cell culture and xenograft models, few ligands demonstrated the desirable drug-like character to reach clinical trials, and among the few that did, none succeeded to date [3,6]. An interesting moiety for G4 ligand synthesis is 1,10-phenanthroline, as this has been used before with promising G4 stabilization results [7–14]. In these studies, phenanthroline-based ligands demonstrated high affinity and thermal stabilization of the G4 structure, as well as good specificity over duplex DNA. Phenanthroline-based ligands have also been used for the development of G4-sensing probes, seizing the metal-complexing ability of the phenanthroline moiety [15–18]. As for our research, we have previously described the synthesis of phenanthroline-based ligands with suitable G4 binding properties [12]. These phenanthroline-bis-oxazole ligands, particularly Phen-2, were found to strongly bind ($K = 10^6$ to 10^7 M⁻¹) and stabilize c-MYC and 22AG G4s. This study emphasized the suitability of the phenanthroline moiety and its side chain functionalization to selectively bind and stabilize G4 structures. Herein, we report a group of different phenanthroline-based macrocycles (**Figure 5.1**) and probe their interaction with different G4 structures such as the oncogene promoter c-MYC and a telomeric sequence. The ligands were previously reported as potential cytotoxic agents against cancer cells and as imaging probes [19–22] and could also be interesting for supramolecular chemistry and as ligands for metal ions [21,22]. Their interesting structural features, a dual phenanthroline moiety and polyamine arms, led us to test them as G4 binders. Several biophysical techniques were employed to study the interaction of

these ligands with G4 structures, such as CD, FRET melting and fluorometric titrations. Furthermore, we confirmed the specificity of the ligands as potent molecules to target G4-regulated genes by the PCR stop assay and monitored the effect on G4 unwinding by a helicase enzyme assay.

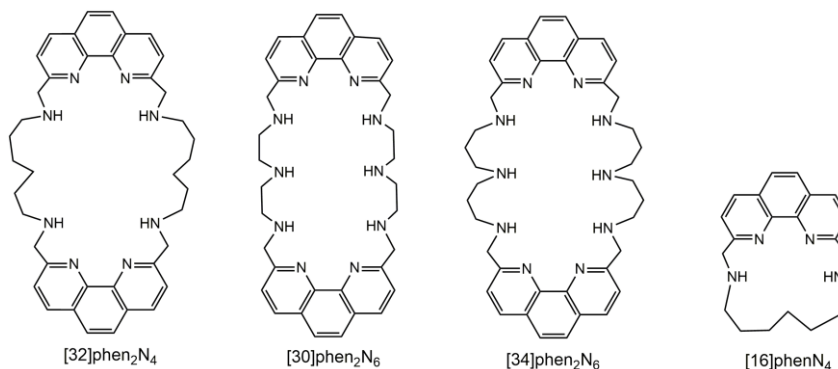


Figure 5.1 - Structures of phenanthroline-based macrocycles.

Experimental Section

Oligonucleotides and ligands

All the oligonucleotides of this study were purchased from Eurofins (Germany) and Eurogentec (Belgium) with HPLC-grade purification. The following oligonucleotide sequences were used: c-MYC: 5'-TGAGGGTGGGTAGGGTGGGTAA-3'; 22AG: 5'-AGGGTTAGGGTTAGGGTTAGGG-3'; 21G: 5'-GGGTTAGGGTTAGGGTTAGGG-3'; c-KIT87: 5'-AGGGAGGGCGCTGGGAGGAGGG-3'; VEGF: CGGGGCGGGCCGGGGGCGGGGT-3'; KRAS21: AGGGCGGTGTGGGAAGAGGGA; dx duplex: 5'-TATAGCTATA-hexaethyleneglycol-TATAGCTATA-3' and ds26: d(5'-CAATCGGATCGAATTCGATCCGATTG-3'). Stock solutions of 0.5 mM were prepared in Milli-Q water and stored at -20 °C. The oligonucleotide concentrations were measured at 260 nm with a UV-Vis spectrophotometer (Thermo Scientific™ Evolution 220) using the molar extinction coefficients (ϵ) provided by the manufacturer. The annealing procedure consisted in heating to 95 °C for 10 min in a heating block, followed by cooling into ice for 30 min.

Ligand [16]phenN₂ was synthesized as previously described [20]. [30]phen₂N₆ and [32]phen₂N₄ were prepared as described in a previous work [21]. [34]phen₂N₆ was prepared as its methylated analogue [21] by using 3,3'-diaminodipropylamine in the first synthetic step. The ligands characterization is presented in Supplementary Information.

Reagents and solvents used were of the purest grade available and were purchased from Sigma-Aldrich (USA) or/and Fischer Chemical (USA).

Electron Spray Ionization (ESI) mass spectra were performed on a Bruker Daltonics Apex-Qe instrument at 300.0 V. The NMR spectra were recorded on a Bruker Avance III 600, operating at 600.10 MHz for protons and 150.90 MHz for carbons equipped with pulse gradient units. The reference used for the ^1H NMR spectra in D_2O was 3-(trimethylsilyl)propanoic acid- d_4 sodium salt. For ^{13}C NMR spectra, dioxane was used as internal reference. All spectra were acquired at temperature of 298 K and using the topspin 2.1 version.

Fluorescence Resonance Energy Transfer melting assays

FRET melting experiments were carried out in 96-well plate using a Stratagene Mx3005P qPCR instrument. The oligonucleotides were purchased labelled with fluorescein (FAM) and rhodamine (TAMRA) at the 5' and 3' ends, respectively. The labelled oligos were annealed before the experiment as described above. The fluorescence intensity of the FAM-labelled oligonucleotide was recorded every 1° C between 25 and 95 °C with a temperature increment of 1 °C/min. The excitation and detection wavelengths were 492 and 516 nm, respectively. The measurements were performed in 10 mM lithium cacodylate (pH 7.2) supplemented with 10 mM KCl and 90 mM LiCl. The concentration of the labelled oligonucleotide was 0.2 μM and the ligands were used at 1 and 2 μM . For competition studies, 3 and 10 μM of double stranded ds26 competitor was added to the wells. Each experimental condition was tested in duplicate on at least three separate plates. The melting temperature was determined from normalized curves as the mid-transition $T_{1/2}$ temperature.

Circular Dichroism studies

CD spectra were done using a Jasco J-815 spectropolarimeter equipped with a Peltier-type temperature control system. Readings were made in a quartz cell with an optical length of 1 mm at 25 °C. CD spectra were recorded using an instrument scanning speed of 200 nm/min with a response time of 1 s through wavelengths ranging from 220 to 340 nm. The recording bandwidth was 1 nm with 1 nm step size. For CD titration, 10 μM strand concentration of oligonucleotides was used in 30 mM phosphate buffer (15 mM KH_2PO_4 , 15 mM K_2HPO_4 , pH 7.1) containing 100 mM KCl or 100 mM NaCl when appropriate. A stock solution of 10 mM of each ligand was prepared. In case of [30]phen $_2\text{N}_6$ and [34]phen $_2\text{N}_6$ ligands were dissolved in water while [16]phen N_2 and [32]phen $_2\text{N}_4$ were

dissolved in DMSO. The volume required for the titrations was added in the quartz cell. Final dilution of the sample at the highest ligand concentration was 5%.

CD melting studies were performed in the temperature range 25 – 95 °C, with a heating rate of 2°C/min by monitoring the ellipticity at 263 nm for c-MYC, 290 nm for hybrid-type 22AG and 295 nm for antiparallel 22AG. Annealing of c-MYC (10 µM in 100 mM LiCl), antiparallel 22AG (10 µM in 100 mM NaCl) and hybrid 22AG (10 µM in 10 mM KCl + 90 mM LiCl) were performed by heating to 95 °C for 10 min followed by cooling in ice for 15 min. Spectra were acquired in a range of 0 to 5 equivalents of ligand. Data were converted into fraction folded (θ) plots using Equation 1, and fitted to a Boltzmann distribution using OriginPro 8.0 and the melting temperature (T_m) obtained.

$$\theta = \frac{CD - CD_{\lambda}^{min}}{CD_{\lambda}^{max} - CD_{\lambda}^{min}} \quad (\text{Equation 1})$$

Where CD is the ellipticity at the monitored wavelength at each temperature and CD^{min} and CD^{max} are the lowest and highest ellipticity, respectively.

Fluorometric titrations

Steady-state fluorescence experiments were carried in a Horiba FluoroMax4 fluorometer equipped with a temperature control system. Ligands (5 µM) were excited at 270 nm and fluorescence emission collected between 285-600 nm in 30 mM potassium phosphate buffer (pH 7.2) containing 100 mM KCl at 25 °C using a 10-mm quartz cuvette. Excitation and emission slits were fixed at 5 nm. In order to study ligand-DNA complex formation, previously annealed oligonucleotides were titrated directly into the cuvette and the spectra acquired after a 3-min equilibration period. Fluorescence spectra reflect the average of three scans. The fluorescence was corrected for the screening effect, *i.e.* the amount of light which is absorbed by DNA at 270 nm [31]. Data points were then fitted to two different Stern-Volmer equations depending on the quenching mechanism and resulting curves [32,33]:

$$\frac{F_0}{F} = 1 + K_{SV}[DNA] \quad (\text{Equation 2})$$

$$\frac{F_0}{F} = A e^{K_{SV}[DNA]} \quad (\text{Equation 3})$$

Where F and F_0 are the fluorescence intensities of the ligand in the presence and absence of DNA, respectively, A is the amplitude and K_{SV} is the Stern-Volmer binding constant.

TO displacement assay

0.25 μM of pre-folded c-MYC and 22AG G4s were mixed with thiazole orange (TO) (0.50 μM TO) in buffer containing 10mM lithium cacodylate and 100 mM KCl or NaCl. Ligands [30]phen₂N₆ and [32]phen₂N₄ were added to the mixture stepwise with a 10 min equilibration period and the fluorescence spectrum was recorded. The emission spectrum was recorded in the range of 505–700 nm with an excitation wavelength of 485 nm. The percentage of TO displacement is calculated by the formula:

$$TO\ Displacement\ (\%) = 100 - ((FA/FA_0) \times 100)$$

where FA_0 is the fluorescence area of DNA-TO complex without ligand and FA is the fluorescence area in the presence of the ligand. FID curves were obtained by plotting the TO displacement percentage versus concentration of ligand used.

Job plot analysis by conductivity

The binding stoichiometry was determined by the Job's continuous variation method using conductivity measurements. Conductivity was measured at 25°C using a Radiometer Danmark CDN210 conductivity meter. The analysis was performed through a systematic variation of the ligand/DNA molar fraction while keeping a constant total concentration of 50 μM , in 20 μM potassium phosphate buffer containing $\approx 100\ \mu\text{M}$ KCl.

Fluorescence-based helicase assay

The fluorescence-based helicase assay was performed using a Tecan Infinite 1000 PRO microplate reader as recently described [30]. All measurements were performed at 25 °C in non-binding 96 wells black plates with flat bottom. The experiment buffer consisted in 20 mM Tris-HCl (pH 7.2) supplemented with 5 mM MgCl₂, 1 mM KCl and 99 mM NaCl. The FAM-Dabcyl systems (S-cmyc, S-htelo, S-mut) were previously annealed in the buffer as described above; sequences are provided in **Table 5.1**. The experiment began by adding 20 nM Pif1 helicase (*S. cerevisiae*), 40 nM FAM-Dabcyl system and 200 nM trap strand into each well. Next, 5 mM ATP was added to each well to activate the helicase and the fluorescence emission was recorded every 10 seconds for 30-45 min (or until reaching a plateau) at 520 nm. The excitation wavelength was set to 492 nm. Once the maximum emission was reached and the signal was stable, 200 nM of complementary sequence was added to provide the maximum intensity of the fluorescence emission. The ligands were tested at 1 μM concentration. Each experiment was replicated on at least three separate plates.

Table 5.1 - Oligonucleotide sequences used in the helicase assay.

System ^a	Sequence
S-cmyc	5'- A ₁₁ -GGGTGGGTAGGGTGGG-TATTCGGTTGAGCAGAG-Dabcyl -3'
S-htelo	5'- A ₁₁ -GGGTAGGGTTAGGGTTAGGG-TATTCGGTTGAGCAGAG-Dabcyl -3'
S-mut	5'- A ₁₁ -TGGTGTGTAGTGTGGTT-TATTCGGTTGAGCAGAG-Dabcyl -3'
C-cmyc	3'- FAM-CTCTGCTCAACGGAA- CCCACCCTACCCACCC-T ₁₁ -5'
C-htelo	3'- FAM-CTCTGCTCAACGGAA-CCCTAACCCTAACCCTAACCC-T ₁₁ -5'
C-mut	3'- FAM-CTCTGCTCAACGGAA- AACCACACTACACACCA-T ₁₁ -5'
Trap	5'- TTCCGTTGAGCAGAG -3'

^a S stands for Pif1 unwinding; C indicates the complementary strand.

PCR-Stop Assay

PCR stop assay was carried out to investigate the stabilization of G4 structures by specific ligands ([16]phenN₂, [30]phen₂N₆, [32]phen₂N₄ and [34]phen₂N₆). The sequences of the test oligomers (Pu77myc; Pu77-mut) and primer (P28) employed in the current assay are presented in **Table 5.2**. Oligonucleotides at a final concentration of 10 µM were annealed in 1× annealing buffer (10 mM Tris·HCl, pH 7.4, 50 mM NaCl, 1 mM EDTA) by heating to 90°C for 5 min, and gradually cooled to room temperature. Reaction mixtures were performed in a final volume of 25 µL containing 30 mM potassium phosphate (pH 7.2), with 100 mM KCl, 1.5 mM Mg(OAc)₂, 0.2 mM dNTPs, 200 nM template DNA and 50 nM primer and further incubated with various concentrations (0.8 µM; 1.6 µM; 3.2 µM and 6.4 µM) of the ligands for 1 h at room temperature. After that time, 1 unit of *Taq* DNA polymerase was added, and PCR reactions were performed with the following cycling conditions: 94°C for 2 min, followed by 30 cycles each comprising 94°C for 30 s, 58°C for 30 s, 72°C for 30 s. Amplified products were resolved on a 15% non-denaturing polyacrylamide gel in 1×TBE and stained with green safe.

Table 5.2 - Oligonucleotide sequences used in the PCR stop assay.

Oligonucleotide	Sequence (from 5' to 3')	Description
Pu77myc	TCCAACATGTATACTGGGGAGGGTGGGGAGGGTGGGGAA GGTTAGCGGCACGCAATTGCTATAGTGAGTCGTATTA	Wild-type <i>c-myc</i> template
Pu77-mut	TCCAACATGTATACTAAGGAAAGTAAGGAAAGTAAGG AAGGTTAGCGGCACGCAATTGCTATAGTGAGTCGTATTA	Mutant <i>c-myc</i> template
P28	TAATACGACTCACTATAGCAATTGCGTG	Primer

Cytotoxicity Assay

HeLa cancer cells were grown in DMEM medium, supplemented with 10% fetal bovine serum (FBS) and 1% streptomycin-penicillin antibiotic. Normal human dermal fibroblasts (NHDF) were grown in RPMI medium, supplemented with 10% FBS, 1% streptomycin-penicillin antibiotic, 0.01 M HEPES, 0.02 M *L*-glutamine and 0.001 sodium pyruvate. Upon confluence, cells were seeded on 48-well culture plates (5×10^4 /well) in 250 μ L volumes. Compounds were added at concentrations ranging from 0.01 μ M to 100 μ M in triplicate. After 48 h of incubation at controlled humidified atmosphere at 37 °C and 5% CO₂, 50 μ L of 1 mg/mL thiazolyl blue tetrazolium bromide (MTT) solution was added to each well and further incubated for 4 h. The resulting formazan crystals were dissolved in 250 μ L of DMSO and the optical density (OD) was recorded at 570 nm. Cell viability was determined using wells without compounds as control.

Results and discussion

Phenanthroline-based ligands as potential G4 binders

The ligands tested here for their ability to bind and stabilize G4 structures were previously described as potential intracellular probes and ligands of metal ions [19–22]. In general, the ligands structure comprises two phenanthroline moieties (except for [16]phenN₂ which has only one) connected by polyamine chains of different lengths and NH content (**Figure 5.1**). The extended planar aromatic moiety of 1,10-phenanthroline has been used before in G4 ligand design due to its ability to π -stack on the external G-tetrads of the G4 structure with stabilizing effect [7–14]. For instance, the well-studied PhenDC3 ligand is a phenanthroline-derivative with interesting G4 binding properties reported by Teulade-Fichou and collaborators [13]; more recently, our research group has also developed and characterized phenanthroline-based ligands and studied their interaction with promoter and telomeric quadruplexes [12]. Additionally, the ligands contain positively charged polyamine chains at physiological pH which could provide additional interactions with the loops and grooves of the G4 structure, namely through H-bonding or electrostatic interactions with the phosphate backbone of the DNA (pH-dependent) [23].

Furthermore, macrocyclic ligands offer an interesting feature for specific binding to G4 over duplex DNA, as their cyclic conformation hinders intercalation between the base pairs for steric reasons; this binding mode being the most common for DNA B-form [23–26]. Particularly, the dibenzophenanthroline analogue BOQ1 showed strong affinity

towards G4 (10^7 M^{-1}) and significant specificity over duplex DNA [26]. This selective behavior was attributed to its cyclic framework. Altogether, this reasoning led us to assess the G4 binding properties of ligands [16]phenN₂, [30]phen₂N₆, [32]phen₂N₄ and [34]phen₂N₆.

FRET melting studies

The initial screening of the phenanthroline-based ligands was performed using the high-throughput FRET melting assay [27]. The ability of the ligands to bind and stabilize G4 structures was evaluated, together with a duplex sequence to evaluate the selectivity towards G4 DNA over double-stranded DNA. The G4 sequences used correspond to the motif found in the c-MYC promoter and the human telomeric repeats (21G). The differences in melting temperature (ΔT_m) at 1 and 2 μM ligand concentrations are summarized in **Table 5.3**. In general, the ligands stabilized the G4 structures in a concentration-dependent manner as the stabilization at 2 μM is higher. Also, the ligands stabilized the telomeric G4 structure (F21T) to a greater extent than the c-MYC promoter (FmycT).

Table 5.3 - Ligand-induced thermal stabilization measured by FRET melting experiments.

Ligand	ΔT_m (°C) ^a at 1 μM ligand concentration			ΔT_m (°C) ^a at 2 μM ligand concentration		
	FmycT	F21T	FdxT ^b	FdxT ^b	FmycT	F21T
[16]phenN ₂	3.6 $\pm$ 0.6	9.2 $\pm$ 2.0	0.0 $\pm$ 0.0	11.3 $\pm$ 0.3	15.0 $\pm$ 1.5	0.3 $\pm$ 0.1
[30]phen ₂ N ₆	6.0 $\pm$ 0.1	6.0 $\pm$ 0.7	0.0 $\pm$ 0.0	8.5 $\pm$ 0.1	11.2 $\pm$ 0.2	0.0 $\pm$ 0.0
[32]phen ₂ N ₄	9.2 $\pm$ 0.2	14.7 $\pm$ 1.2	0.0 $\pm$ 0.0	17.2 $\pm$ 0.8	20.3 $\pm$ 1.3	3.6 $\pm$ 1.4
[34]phen ₂ N ₆	2.3 $\pm$ 1.2	9.4 $\pm$ 0.1	0.0 $\pm$ 0.0	6.5 $\pm$ 1.1	15.9 $\pm$ 0.8	0.4 $\pm$ 0.7

^a ΔT_m represents difference in melting temperature [$\Delta T_m = T_m (\text{DNA} + \text{ligand}) - T_m (\text{DNA})$]. The buffer used was 10 mM lithium cacodylate, pH 7.2. The T_m values for the DNAs are 67.9 $\pm$ 0.2 °C [c-MYC promoter DNA in 10 mM KCl and 90 mM LiCl]; 53.9 $\pm$ 0.3 °C [21G telomeric DNA in 10 mM KCl and 90 mM LiCl] and 63.6 $\pm$ 1.6 °C [duplex DNA in 100 mM KCl]. All experiments were done in triplicate, and the values reported correspond to the average of three measurements with the estimated standard deviation.

^b Negative ΔT_m values are reported as zero.

Based on the above FRET melting studies, [32]phen₂N₄ shows higher thermal stabilization towards FmycT and F21T at 2 μM (17.2 and 20.3 °C, respectively), while [30]phen₂N₆ and [34]phen₂N₆ were weaker stabilizers. The ligand containing one phenanthroline moiety [16]phenN₂ promotes significant thermal stabilization of both G4 structures tested (11.3 and 15.0 °C for FmycT and F21T, respectively). This indicates that one phenanthroline unit is sufficient to bind and stabilize the G4 structure, in agreement with other ligands in

the literature such as PhenDC3 [13]. Moreover, the smaller size of [16]phenN₂ may enhance its tetrad accessibility hence promoting a higher degree of stabilization. The ligand-induced stabilization of c-MYC G4 followed the trend: [32]phen₂N₄ > [16]phenN₂ > [30]phen₂N₆ > [34]phen₂N₆; in the case of 21G G4 is as follows: [32]phen₂N₄ > [34]phen₂N₆ > [16]phenN₂ > [30]phen₂N₆. Despite no clear rule can be established in terms of structure-activity relationship, the ligands with N₄ tend to promote higher stabilization when compared to N₆. Moreover, as reported previously for other macrocycles [24], the higher degree of stabilization of [32]phen₂N₄ suggests that the number of consecutive methylene groups in the linker separating the two phenanthroline moieties is important for the stabilization. We cannot exclude that hydrophobic force may play a role in the case of [32]phen₂N₄. Comparatively, the dibenzophenanthroline analogue BOQ1 has been shown to increase the melting temperature of F21T by 28 °C at 1 μM ligand concentration [26]. The higher degree of stabilization when compared to the ligands hereby described, for instance [30]phen₂N₆ which has the same linker, may be attributed to the larger aromatic core of the benzophenanthroline moiety.

The results with the FdxT duplex sequence suggest all ligands are selective towards G4 over duplex DNA, as a negligible stabilization of the duplex was observed for all ligands. This selectivity was further confirmed by competition FRET melting assays in the presence of an excess of ds26 duplex competitor. The experiments were performed with the best ligands [32]phen₂N₄ and [16]phenN₂, in the presence of 3 and 10 μM of ds26. The selectivity index (*S*) was determined, which corresponds to the ratio between ΔT_m obtained in the presence and absence of ds26. The presence of the competitor did not strongly affect the ligand-induced stabilization of F21T G4 as *S* values for [32]phen₂N₄ and [16]phenN₂ at 10 μM ds26 were 0.89 and 0.81, respectively (Figure S1, **Appendix B**). It should be emphasized that 10 μM of ds26 corresponds to 50-folds excess when compared to F21T G4 (at 0.2 μM). In order to assess the ligands selectivity towards other G4s, additional sequences were used for FRET-melting experiments, namely FkrasT, FvegT and FkitT. The results are summarized in **Table 5.4**. Similarly to the aforementioned results, the N₄ ligands tend to promote higher stabilization when compared to N₆, [32]phen₂N₄ being the best ligand in terms of G4 stabilization. The ligands showed a preference towards KRAS21 G4 when compared to VEGF and c-KIT87 G4s. However, the results for KRAS21 and c-KIT87 G4s are comparable to those in **Table 5.3** suggesting that the ligands do not exhibit a considerable response over other G4s. Regarding VEGF G4, ΔT_m are more modest indicating a weaker binding.

Table 5.4 - Ligand-induced thermal stabilization measured by FRET melting experiments.

Ligand	ΔT_m (°C) ^a at 1 μ M ligand concentration			ΔT_m (°C) ^a at 2 μ M ligand concentration		
	FkrasT	FvegFT	FkitT	FkrasT	FvegFT	FkitT
[16]phenN₂	12.4 $\pm$ 1.5	3.6 $\pm$ 1.0	8.5 $\pm$ 0.5	24.1 $\pm$ 1.5	9.5 $\pm$ 0.0	17.4 $\pm$ 0.9
[30]phen₂N₆	8.1 $\pm$ 0.4	0.2 $\pm$ 0.5	0.3 $\pm$ 0.0	13.7 $\pm$ 0.5	0.7 $\pm$ 0.7	0.8 $\pm$ 0.4
[32]phen₂N₄	21.4 $\pm$ 0.6	5.1 $\pm$ 0.2	9.8 $\pm$ 2.1	32.6 $\pm$ 0.6	9.6 $\pm$ 0.7	16.8 $\pm$ 0.8
[34]phen₂N₆	11.5 $\pm$ 0.6	0.7 $\pm$ 0.0	4.3 $\pm$ 0.4	17.9 $\pm$ 1.1	1.8 $\pm$ 0.7	8.2 $\pm$ 1.6

^a ΔT_m represents difference in melting temperature [$\Delta T_m = T_m$ (DNA + ligand) – T_m (DNA)]. The buffer used was 10 mM lithium cacodylate, pH 7.2. The T_m values for the DNAs are 49.1 $\pm$ 0.5 °C [KRAS promoter DNA in 10 mM KCl and 90 mM LiCl]; 66.8 $\pm$ 0.4 °C [VEGF promoter DNA in 1 mM KCl and 99 mM LiCl]; 53.1 $\pm$ 0.6 °C [c-KIT promoter DNA in 10 mM KCl and 90 mM LiCl]. All experiments were done in triplicate, and the values reported are average of three measurements with the estimated standard deviation.

Circular dichroism experiments

CD spectra were used to study the effect of ligands on the conformation of G4 DNAs. The parallel c-MYC, hybrid-type and antiparallel 22AG G4s were investigated in 100 mM LiCl, 10 mM KCl + 90 mM LiCl and 100 mM NaCl, respectively. Upon titration of the c-MYC G4 with all ligands, no significant change was observed in the characteristic parallel G4 peaks wavelength at both 260 and 240 nm, even after the addition of 5 equivalents of ligand, suggesting that these ligands do not bring structural change of the parallel G4 structure (Figure S2). There is a marked ellipticity decrease at higher molar equivalents of ligand (panels B-D) which may be attributed to the association between the ligands and the G4 structure [28]. In the case of [16]phenN₂ however (panel A), no significant change is observed; such behavior is expected when end-stacking alone is the preferred mode of interaction [29]. The CD spectrum of the hybrid-type G4 formed by 22AG in potassium presented two positive bands at 290 and 268 nm and a weak negative band at around 240 nm. After addition of the ligands, no major structural changes were observed in the G4 structure as the characteristic bands are preserved (Figure S3). Nonetheless, a slight variation of ellipticity is observed probably due to ligand binding. The CD spectra of antiparallel 22AG G4 upon addition of ligands demonstrated that a small increment was observed at 260 nm in a concentration dependent manner (mainly with [30]phen₂N₆) while minor changes in the positive peak ellipticity were observed at 295 nm for all ligands (Figure S4). These results suggest that the ligands do not bring about major structural changes in the antiparallel conformation of the 22AG G4 structure upon binding. Overall, the observations from CD titration studies suggest that the ligands are able to bind the G4s but do not change the pre-formed topologies of c-MYC and 22AG G4s.

CD melting assays were further performed to assess the thermal stabilization of the four ligands towards G4s 22AG (hybrid-type and antiparallel) and c-MYC. Melting experiments for the c-MYC G4, hybrid-type and antiparallel 22AG G4s were performed at 263 (100 mM LiCl), 290 (10 mM KCl and 90 mM LiCl) and 295 nm (100 mM NaCl), which yielded melting temperatures of 55.4, 55.4 and 56.7°C, respectively. After the addition of 5 equivalents of ligand, there was a considerable increase in the melting temperature of parallel c-MYC with [16]phenN₂ and [32]phen₂N₄ harboring a ΔT_m of 19.3 °C and 15.1 °C, respectively (**Table 5.5, Figure 5.2**). The same ligands showed a moderate increase in the melting temperature (maximum $\Delta T_m \sim 6.5$ °C for [32]phen₂N₄) on hybrid-type 22AG G4 and a slight enhancement in the melting temperature (maximum $\Delta T_m \sim 2.2$ °C for [16]phenN₂) of antiparallel 22AG G4, suggesting weak stabilization of these particular topologies (**Table 5.5, Figure 5.2**). The remaining ligands displayed negligible stabilization of hybrid-type (maximum $\Delta T_m \sim 0.3$ °C) and antiparallel 22AG G4s (**Table 5.5, Figure 5.2**).

Table 5.5 - Ligand-induced thermal stabilization measured by CD melting experiments.

Ligand	ΔT_m (°C) ^a		
	c-MYC	22AG-h ^b	22AG-a ^c
[16]phenN ₂	19.3 ± 0.2	5.1 ± 0.1	2.2 ± 0.1
[30]phen ₂ N ₆	13.8 ± 0.7	0.0 ± 0.0	0.0 ± 0.0
[32]phen ₂ N ₄	15.1 ± 0.2	6.5 ± 0.1	1.5 ± 0.1
[34]phen ₂ N ₆	5.9 ± 0.6	0.0 ± 0.0	0.0 ± 0.0

^a ΔT_m represents difference in melting temperature [$\Delta T_m = T_m$ (DNA + 5 molar eq. ligand) – T_m (DNA)]. The buffer used was 30 mM potassium phosphate, pH 7.2. The T_m values for the DNAs are 55.4 ± 2.3 °C [c-MYC promoter DNA in 100 mM LiCl]; 55.4 ± 1.2 °C [22AG telomeric DNA in 10 mM KCl and 90 mM LiCl]; 56.7 ± 0.7 °C [22AG telomeric DNA in 100 mM NaCl]. Values reported with estimated standard error.

^b Hybrid-type telomeric G4 in K⁺ buffer.

^c Antiparallel telomeric G4 in Na⁺ buffer.

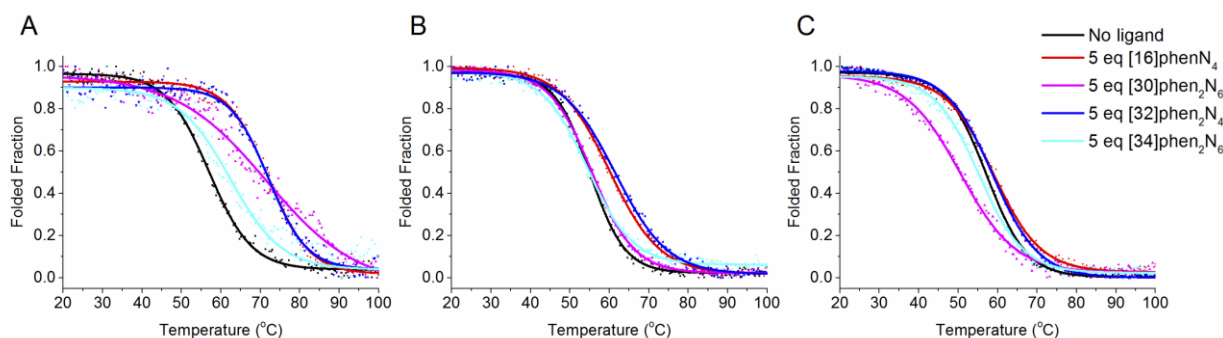


Figure 5.2 - Normalized CD melting curves of (A) c-MYC G4 (100 mM LiCl), (B) hybrid-type 22AG G4 (10 mM KCl and 90 mM LiCl) and (C) antiparallel 22AG G4 (100 mM NaCl) DNAs in the absence and presence of 5 equivalents of ligands.

These results clearly show that all the ligands preferentially stabilize the parallel topology of the promoter c-MYC G4 over the telomeric 22AG G4s. Moreover, ligands [16]phenN₂ and [32]phen₂N₄ with the hexyl side chain are more efficient in achieving the preferential stabilization of parallel topology when compared to the other ligands.

Fluorimetric titration studies for binding analysis

Fluorescence spectroscopy was used to determine the binding affinity of the ligands with c-MYC and 22AG G4s. A concentration-dependent variation of the fluorescence intensity can be used to derive the binding/dissociation constants of the ligand–G4 association. Upon excitation at 271 nm, the ligands emitted fluorescence as a broad band between 375 and 550 nm. Ligands [16]phenN₂, [30]phen₂N₆, [32]phen₂N₄ and [34]phen₂N₆ exhibited emission maxima around 414 nm (**Figure 5.3** and Figures S5-S7). Upon the addition of increasing concentrations of c-MYC and 22AG G4s to ligands into quartz cell, these exhibited an almost complete quenching (see **Figure 5.3** and Figures S5-S7). The titration curves were fitted by two different Stern-Volmer equations (see Experimental section) depending on the quenching mechanism and resulting curves (see **Figure 5.3** and Figures S5-S7).

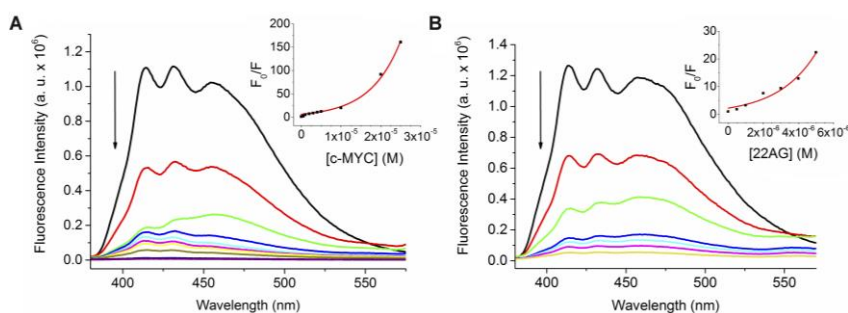


Figure 5.3 - Fluorescence titrations of the ligand [32]phen₂N₄ with increasing concentration of (A) c-MYC and (B) 22AG G4s [5 μM in 100 mM KCl, pH 7.2, concentrations 0–25 μM]. Insets: Stern-Volmer plots obtained by plotting F₀/F as a function of DNA concentration.

The binding affinities of ligands towards c-MYC and 22AG G4s were determined and summarized in **Table 5.6**.

Table 5.6 - Apparent DNA-binding constants (K_{SV} M⁻¹) obtained from fluorescence titration experiments of the ligands.

Ligand	K_{SV} (M ⁻¹)	
	c-MYC	22AG
[16]phenN ₂	1.4×10^5	4.6×10^6
[30]phen ₂ N ₆	2.2×10^5	1.1×10^5
[32]phen ₂ N ₄	1.3×10^5	4.6×10^5
[34]phen ₂ N ₆	1.1×10^5	1.1×10^5

The determined binding affinity constants in the range $K_{SV} = 10^5$ - 10^6 M⁻¹ are typical of moderate-to-good ligand-quadruplex binding. These results show that the shortest macrocycle with the methylene hexyl linker, [16]phenN₂, has a slightly higher apparent binding constant for telomeric 22AG G4 than for c-MYC G4, while macrocycle [32]phen₂N₄ (with the same linker but with two phenanthroline units) presents a lower binding affinity. The remaining ligands displayed lower binding affinity towards antiparallel 22AG G4 and parallel c-MYC G4. Compared to BOQ1 the affinity is slightly lower as the dibenzophenanthroline analogue displayed a binding affinity constant of 10^6 M⁻¹ towards a telomeric G4, measured by surface plasmon resonance (SPR) and fluorimetric titrations [26]. This enhanced behavior was already observed in terms of thermal stabilization.

It should be noted here that, irrespective of the moderate binding preference, ligands show strong stabilization by FRET-melting and CD melting (see above). These results correlate in part with the thermal stabilization obtained in CD melting experiments performed with the same non-labelled oligonucleotide sequences. Ligands [16]phenN₂ and [32]phen₂N₄ exhibit higher affinity towards c-MYC G4 when compared to 22AG and the thermal stabilization for c-MYC correlate with the binding affinity. For the remaining ligands, the affinity towards both c-MYC and 22AG G4s is similar, thus not complying with the thermal stabilization profile.

The continuous variation analysis method was carried out to evaluate the binding stoichiometry by monitoring the conductivity changes of the complex using the continuous variation Job method. The Job plots for [16]phenN₂ and [32]phen₂N₄ displayed a maximum at ≈ 0.67 mole fraction of the ligand indicating 2:1 (ligand:G4) binding complex formation between the ligands and the c-MYC and 22AG quadruplexes (Figure S8).

Fluorescent intercalator displacement assay

To support the findings from the FRET-melting, CD titrations and melting results, FID assay was performed to investigate the binding selectivity of [32]phen₂N₄ and [30]phen₂N₆ towards c-MYC and 22AG G4s (**Table 5.7**). We chose these ligands because their alkyl chain is different, and the hexyl side chain of [32]phen₂N₄ has been found to be more efficient for the stabilization of these topologies. This assay measures the decrease in fluorescence intensity upon the displacement of thiazole-orange (TO) from the G4–TO complex by the ligand and the percentage of displacement can be plotted against the concentration of the ligand to afford the DC₅₀. Low DC₅₀ values represent the high binding affinity of the ligand toward G4.

Table 5.7 - DC₅₀ values for c-MYC and 22AG G4s with ligands [32]phen₂N₄ and [30]phen₂N₆ measured from FID assay.

Ligand	DC ₅₀ (μM)		
	c-MYC	22AG K ⁺	22AG Na ⁺
[32]phen ₂ N ₄	0.87	1.02	1.24
[30]phen ₂ N ₆	0.92	> 2.5	1.55

For c-MYC and 22AG G4s, [32]phen₂N₄ gave low DC₅₀ values (0.87–1.24 μM), indicating high stabilization of these G4s DNA (Table 5) compared to the other ligand as observed by FRET-melting and CD melting experiments. In addition, these ligands showed a slight preference for the c-MYC G4 over the 22AG G4s. The results obtained from the FID assay and the melting experiments show the same trend for the compounds, in which [32]phen₂N₄ has more ability to stabilize G4s DNA and capability to displace TO. The similar DC₅₀ values obtained for both ligands with c-MYC and antiparallel 22AG G4s can be due to the use of TO as a displacement assay, which is different from the other thermal stabilization assays that measure directly the binding affinity.

Inhibition of helicase and polymerase activities

The potential effect of the phenanthroline ligands on biological processes was evaluated *in vitro* using the fluorescence-based helicase assay recently published [30]. The assay is based on the unwinding of G4 structures by the Pif1 helicase and the fact that a ligand, when capable of binding and stabilizing the G4 structure, can inhibit the activity of the enzyme in resolving G4 motifs. The systems used for this helicase assay were those used for the other experiments, the c-MYC (S-cmyc) oncogene promoter and the human telomeric sequence (S-htelo). The S-mut system (substrate containing unstructured

single-stranded DNA in place of the G4 motif) was used as control and to assess the selectivity of the ligands towards G4 structures. The results are presented in **Figure 5.4**.

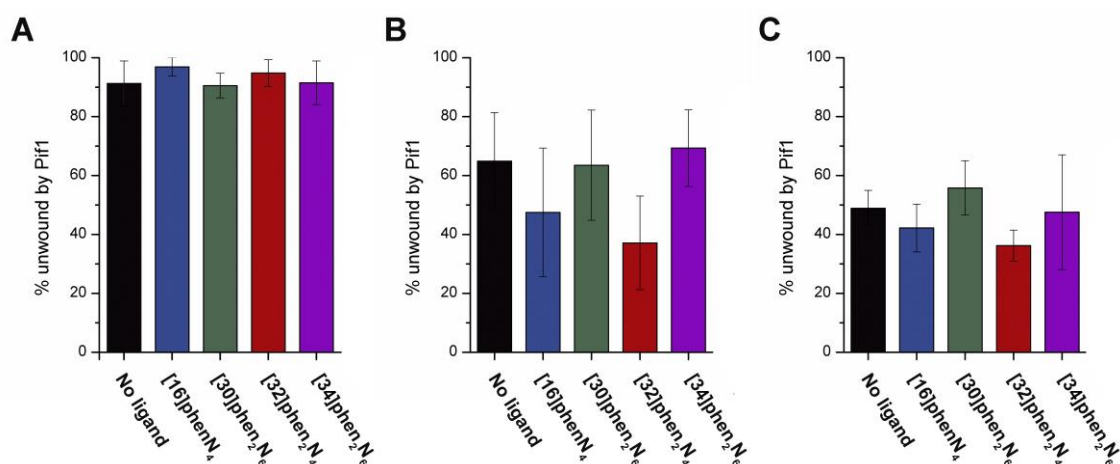


Figure 5.4 - Helicase activity against the (A) S-mut, (B) S-cmyc DNA and (C) S-htelo system in buffer containing 1 mM KCl at 20 nM Pif1. Error bars indicate mean deviation between three independent experiments.

The results suggest that [32]phen₂N₄ and [16]phenN₂ are the best ligands in terms of thermal stabilization and are also capable of interfering with biological processes involving G4 unwinding, because they promote a decrease in unwinding activity of both G4 systems tested. Despite the higher error bars obtained in **Figure 5.4B**, ligands [32]phen₂N₄ and [16]phenN₂ seem to show the most effects on Pif1 activity. A similar trend is observed when compared to **Figure 5.4C** where the error bars are less significant. The other ligands shown a negligible effect on Pif1 activity, which is in agreement with the results from previous experiments. The ligands did not affect the unwinding activity of Pif1 in the presence of the S-mut system suggesting that they are selective binders of these G4 structures.

PCR stop assay

The stabilization of c-MYC G4 by the ligands [16]phenN₂, [30]phen₂N₆, [32]phen₂N₄ and [34]phen₂N₆ was investigated by a PCR-stop assay. Sequences of the test oligonucleotides (Pu77myc and Pu77-mut) and the corresponding complementary sequence (RevPu77) were used in this experiment (**Table 5.2**). In the presence of the ligand, the Pu77myc was stabilized into a G4 that blocked hybridization with a complementary strand overlapping the last G repeat. *Taq* polymerase was unable to recognize the G4 structure and DNA amplification was inhibited, resulting in a reduction/ disappearance of the PCR product. The results of the PCR stop experiment with the ligands are presented in **Figure 5.5**.

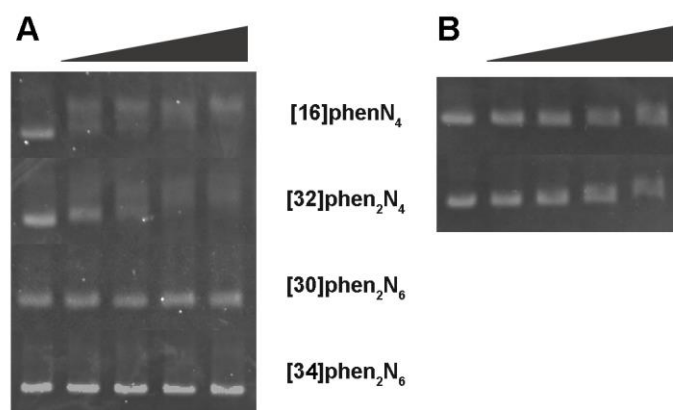


Figure 5.5 - Polyacrylamide gel electrophoresis of (A) Pu77myc G4 and (B) Pu77-mut in the presence of increased concentrations of the ligands (0 μ M, 0.8 μ M; 1.6 μ M; 3.2 μ M and 6.4 μ M).

Inhibition of the DNA amplification was only observed with [16]phenN₂ and [32]phen₂N₄ (**Figure 5.5A**), suggesting that these two ligands stabilize the c-MYC G4. With the addition of 0.8 μ M of [16]phenN₂ and 1.6 μ M of [32]phen₂N₄, the 77-bp double-stranded PCR product completely disappeared. Thus, the concentration of selected ligand that inhibited amplification was found to be higher in [32]phen₂N₄ than in [16]phenN₂, suggesting a stronger inhibitory effect of [16]phenN₂. This result is consistent with competitive FRET-melting and CD melting assays. In contrast, in the presence of [30]phen₂N₆ and [34]phen₂N₆, the final double-stranded PCR product could be still detected with all the concentrations tested, suggesting that these ligands are weak stabilizers and inhibitors of c-MYC G4 (**Figure 5.5A**).

To further demonstrate that inhibition induced by [16]phenN₂ and [32]phen₂N₄ was due to G4 stabilization of the Pu77 oligomer, a parallel PCR stop assay using a mutated Pu-77 strand was included. The positions indicating base modifications in Pu77-mut are shown in **Table 5.2** (Experimental section). By comparison, inhibition of the PCR amplification was not observed with all the concentrations tested, suggesting specific binding to the c-MYC G4 and consequent inhibition of *Taq* polymerase (**Figure 5.5B**).

Cytotoxicity assay

The ligands effect on cell growth was evaluated by MTT assay using HeLa cervical cancer cell line and NHDF. The results are presented in **Figure 5.6**. The compounds did not significantly affect the growth of HeLa cancer cells at low concentrations, only presenting modest toxic effect at 100 μ M (**Figure 5.6**). These results are consistent with those previously reported for cancer cells [20]. [32]phen₂N₄ presented the higher cytotoxicity potential, followed by [34]phen₂N₆ and [16]phenN₂. [30]phen₂N₆ presented a negligible effect on the cell growth being the less toxic ligand. These results partially correlate with

the other techniques, which may indicate that the cytotoxicity of the ligands is related to in cell G4 binding. Regarding NHDF cells (**Figure 5.6**), the results are similar to those obtained for cancer cells, in agreement with previous findings [20]. [32]phen₂N₄ is suggested to be less toxic for normal cells at 100 μ M, as previously reported, while [30]phen₂N₆ seems to have no effect on cell growth. Ligands [34]phen₂N₆ and [16]phenN₂ are the most toxic compounds, lacking any cancer vs normal cells selectivity.

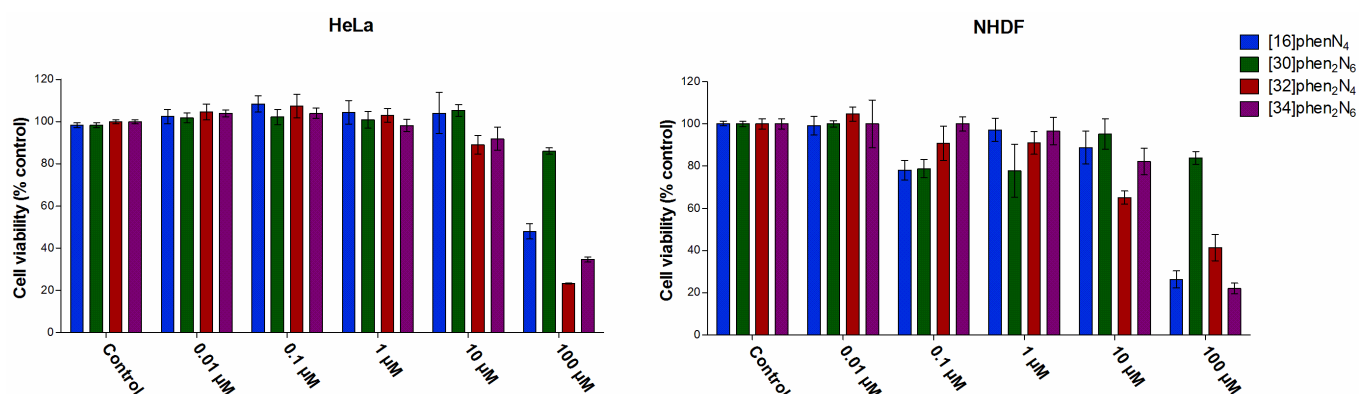


Figure 5.6 - Relative cell viability of HeLa (left) and NHDF (right) cells measured by MTT assay after 48 h incubation with the ligands at concentrations ranging from 0.01 to 100 μ M. Bars represent mean $\pm$ SEM of cell viability relatively to control cells (no treatment).

Conclusions

Four phenanthroline polyazamacrocycles have been investigated for their potential for G4 binding. The hexyl linker in [16]phenN₂ and [32]phen₂N₄ plays an important role in the stabilization ability towards c-MYC and 22AG G4s unlike duplex DNA. Based on the results from FRET-melting, CD melting, FID and *Taq* DNA polymerase stop assays we conclude that ligands [16]phenN₂ and [32]phen₂N₄ can effectively bind and stabilize the quadruplexes formed by telomeric repeats and the c-MYC promoter. Ligand [16]phenN₂ containing a single phenanthroline unit was found to be the most promising c-MYC G4 binding ligand compared to the derivative [32]phen₂N₄ used in this study. This could be due to the G-tetrad accessibility being facilitated by the smaller size of the ligand. Indeed, CD titration results pointed to end-stacking as the main mode of interaction.

Additionally, the binding of ligands [16]phenN₂ and [32]phen₂N₄ to Pu77myc G4 blocked the hybridization of the complementary strand and *Taq* DNA polymerase was repressed revealing that these ligands are strong stabilizers and potential transcription inhibitors. The decrease in unwinding of S-cmyc and S-htelo G4s by the Pif1 helicase showed also that [32]phen₂N₄ and [16]phenN₂ are the best ligands in terms of G4 stabilization and are also

capable of interfering with biological processes. Overall polyazamacrocycles [16]phenN₂ and [32]phen₂N₄ interact, stabilize G4 structures and can be exploited *in vitro* as G4 binders and potential transcription inhibitors. This emphasizes the use of the phenanthroline moiety for G4 ligands synthesis and contributes to the macrocyclic ligands field as it introduces important concepts regarding the aromatic surface and side chain size and functionality.

Acknowledgments

Josué Carvalho acknowledges the doctoral fellowship grant from FCT - Foundation for Science and Technology ref. SFRH/BD/122953/2016. This work was supported by project “Ações Integradas Luso-Francesas” ref. TC-15/17 entitled “Aptameros de G-quadruplex para tratamento do cancro do colo do útero”, FCT project “Projeto de Investigação Exploratória” ref. IF/00959/2015 entitled “Nucleolin targeting by G-quadruplex aptamers for cervical cancer therapy” financed by Fundo Social Europeu e Programa Operacional Potencial Humano. This work was also supported by FEDER funds through the POCI - COMPETE 2020 - Operational Programme Competitiveness and Internationalisation in Axis I - Strengthening research, technological development and innovation (Project POCI-01-0145-FEDER-007491), by National Funds by FCT (Project UID/Multi /00709/2013), by Ligue Régionale Contre le Cancer (to G.F.S and J.L.M.) and by the SYMBIT project (reg. no. CZ.02.1.01/0.0/0.0/15_003/0000477) financed by the ERDF.

References

- [1] Neidle, S. (2016) Quadruplex Nucleic Acids as Novel Therapeutic Targets. *Journal of Medicinal Chemistry* 59, 5987–6011. <https://doi.org/10.1021/acs.jmedchem.5b01835>.
- [2] Hänsel-Hertsch, R., Antonio, M. Di, and Balasubramanian, S. (2017) DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nature Reviews Molecular Cell Biology* 18, 279–284. <https://doi.org/10.1038/nrm.2017.3>.
- [3] Neidle, S. (2017) Quadruplex nucleic acids as targets for anticancer therapeutics. *Nature Reviews Chemistry* 1, Article 0041. <https://doi.org/10.1038/s41570-017-0041>.
- [4] Li, Q., Xiang, J.-F. F., Yang, Q.-F. F., Sun, H.-X. X., Guan, A.-J. J., and Tang, Y.-L. L. (2013) G4LDB: A database for discovering and studying G-quadruplex ligands. *Nucleic*

Acids Research 41, D1115–D1123. <https://doi.org/10.1093/nar/gks1101>.

[5] Garci, A., Castor, K. J., Fakhoury, J., Do, J. L., Di Trani, J., Chidchob, P., Stein, R. S., Mittermaier, A. K., Frišćić, T., and Sleiman, H. (2017) Efficient and Rapid Mechanochemical Assembly of Platinum(II) Squares for Guanine Quadruplex Targeting. *Journal of the American Chemical Society* 139, 16913–16922. <https://doi.org/10.1021/jacs.7b09819>.

[6] Müller, S., and Rodriguez, R. (2014) G-quadruplex interacting small molecules and drugs: from bench toward bedside. *Expert Review of Clinical Pharmacology* 7, 663–679. <https://doi.org/10.1586/17512433.2014.945909>.

[7] Dhamodharan, V., Harikrishna, S., Bhasikuttan, A. C., and Pradeepkumar, P. I. (2015) Topology Specific Stabilization of Promoter over Telomeric G-Quadruplex DNAs by Bisbenzimidazole Carboxamide Derivatives. *ACS Chemical Biology* 10, 812–833. <https://doi.org/10.1021/cb5008597>.

[8] Bhasikuttan, A. C., and Mohanty, J. (2015) Targeting G-quadruplex structures with extrinsic fluorogenic dyes: promising fluorescence sensors. *Chem. Commun. Chem. Commun* 51, 7581–7597. <https://doi.org/10.1039/c4cc10030a>.

[9] Mazzolari, A., Maddis, D. De, and Carini, M. (2016) Recent advances in targeting the telomeric G-quadruplex DNA sequence with small molecules as a strategy for anticancer therapies. *Future Medicinal Chemistry* 8, 975–992. <https://doi.org/10.4155/fmc-2016-0017>.

[10] Amato, J., Iaccarino, N., Pagano, B., Morigi, R., Locatelli, A., Leoni, A., Rambaldi, M., Zizza, P., Biroccio, A., Novellino, E., and Randazzo, A. (2014) Bis-indole derivatives with antitumor activity turn out to be specific ligands of human telomeric G-quadruplex. *Frontiers in chemistry* 2, 54. <https://doi.org/10.3389/fchem.2014.00054>.

[11] Zhang, H., Xiang, J., Hu, H., Liu, Y., Yang, F., Shen, G., Tang, Y., and Chen, C. (2015) Selective recognition of specific G-quadruplex vs. duplex DNA by a phenanthroline derivative. *International Journal of Biological Macromolecules* 78, 149–156. <https://doi.org/10.1016/j.ijbiomac.2015.03.034>.

[12] Medeiros-Silva, J., Guédin, A., Salgado, G. F., Mergny, J.-L., Queiroz, J. A., Cabrita, E. J., and Cruz, C. (2017) Phenanthroline-bis-oxazole ligands for binding and stabilization of G-quadruplexes. *Biochimica et Biophysica Acta - General Subjects* 1861, 1281–1292. <https://doi.org/10.1016/j.bbagen.2016.11.024>.

- [13] De Cian, A., DeLemos, E., Mergny, J.-L., Teulade-Fichou, M.-P. P., and Monchaud, D. (2007) Highly efficient G-quadruplex recognition by bisquinolinium compounds. *Journal of the American Chemical Society* 129, 1856–1857. <https://doi.org/10.1021/ja067352b>.
- [14] Beauvineau, C., Guetta, C., Teulade-Fichou, M.-P., and Mahuteau-Betzer, F. (2017) PhenDV, a turn-off fluorescent quadruplex DNA probe for improving the sensitivity of drug screening assays. *Org. Biomol. Chem.* 15, 7117–7121. <https://doi.org/10.1039/C7OB01705G>.
- [15] Wang, M., Wang, W., Kang, T. S., Leung, C. H., and Ma, D. L. (2016) Development of an Iridium(III) Complex as a G-Quadruplex Probe and Its Application for the G-Quadruplex-Based Luminescent Detection of Picomolar Insulin. *Analytical Chemistry* 88, 981–987. <https://doi.org/10.1021/acs.analchem.5b04064>.
- [16] Lin, S., Lu, L., Kang, T. S., Mergny, J.-L., Leung, C. H., and Ma, D. L. (2016) Interaction of an Iridium(III) Complex with G-Quadruplex DNA and Its Application in Luminescent Switch-On Detection of Siglec-5. *Analytical Chemistry* 88, 10290–10295. <https://doi.org/10.1021/acs.analchem.6b03128>.
- [17] Lu, L., Wang, W., Yang, C., Kang, T.-S., Leung, C.-H., and Ma, D.-L. (2016) Iridium(iii) complexes with 1,10-phenanthroline-based N^N ligands as highly selective luminescent G-quadruplex probes and application for switch-on ribonuclease H detection. *Journal of Materials Chemistry B* 4, 6791–6796. <https://doi.org/10.1039/C6TB02316A>.
- [18] Wang, M., Mao, Z., Kang, T.-S., Wong, C.-Y., Mergny, J.-L., Leung, C.-H., and Ma, D.-L. (2016) Conjugating a groove-binding motif to an Ir(iii) complex for the enhancement of G-quadruplex probe behavior. *Chemical Science* 7, 2516–2523. <https://doi.org/10.1039/C6SC00001K>.
- [19] Cruz, C., Cairrão, E., Lourenço, O., Almeida, P., Verde, I., and Queiroz, J. A. (2013) Polyazamacrocycles as potential antitumor agents for human prostate cancer cells. *Chemical biology & drug design* 81, 517–26. <https://doi.org/10.1111/cbdd.12098>.
- [20] Cruz, C., Cairrao, E., Silvestre, S., Breitenfeld, L., Almeida, P., and Queiroz, J. A. (2011) Targeting of mitochondria-endoplasmic reticulum by fluorescent macrocyclic compounds. *PloS one* 6, e27078. <https://doi.org/10.1371/journal.pone.0027078>.
- [21] Cruz, C., Calisto, V., Delgado, R., and Félix, V. (2009) Design of protonated polyazamacrocycles based on phenanthroline motifs for selective uptake of aromatic carboxylate anions and herbicides. *Chemistry - A European Journal* 15, 3277–3289.

<https://doi.org/10.1002/chem.200800993>.

[22] Cruz, C., Delgado, R., Drew, M. G. B., and Félix, V. (2007) Evaluation of the binding ability of a novel dioxatetraazamacrocyclic receptor that contains two phenanthroline units: selective uptake of carboxylate anions. *The Journal of organic chemistry* 72, 4023–34. <https://doi.org/10.1021/jo062653p>.

[23] Monchaud, D., and Teulade-Fichou, M.-P. (2008) A hitchhiker's guide to G-quadruplex ligands. *Organic & biomolecular chemistry* 6, 627–636. <https://doi.org/10.1039/b714772b>.

[24] Marchetti, C., Minarini, A., Tumiatti, V., Moraca, F., Parrotta, L., Alcaro, S., Rigo, R., Sissi, C., Gunaratnam, M., Ohnmacht, S. a, Neidle, S., and Milelli, A. (2015) Macrocyclic Naphthalene Diimides as G-Quadruplex Binders. *Bioorganic & Medicinal Chemistry* 23, 3819–3830. <https://doi.org/10.1016/j.bmc.2015.03.076>.

[25] Kaiser, M., De Cian, A., Sainlos, M., Renner, C., Mergny, J.-L., and Teulade-Fichou, M.-P. (2006) Neomycin-capped aromatic platforms: quadruplex DNA recognition and telomerase inhibition. *Organic & biomolecular chemistry* 4, 1049–1057. <https://doi.org/10.1039/b516378a>.

[26] Teulade-Fichou, M.-P., Carrasco, C., Guittat, L., Bailly, C., Alberti, P., Mergny, J.-L., David, A., Lehn, J. M., and Wilson, W. D. (2003) Selective recognition of G-quadruplex telomeric DNA by a bis(quinacridine) macrocycle. *Journal of the American Chemical Society* 125, 4732–4740. <https://doi.org/10.1021/ja021299j>.

[27] De Cian, A., Guittat, L., Kaiser, M., Saccà, B., Amrane, S., Bourdoncle, A., Alberti, P., Teulade-Fichou, M.-P., Lacroix, L., and Mergny, J.-L. (2007) Fluorescence-based melting assays for studying quadruplex ligands. *Methods* 42, 183–195. <https://doi.org/10.1016/j.ymeth.2006.10.004>.

[28] Jana, J., Mondal, S., Bhattacharjee, P., Sengupta, P., Roychowdhury, T., Saha, P., Kundu, P., and Chatterjee, S. (2017) Chelerythrine down regulates expression of VEGFA, BCL2 and KRAS by arresting G-Quadruplex structures at their promoter regions. *Scientific Reports* 7, 1–14. <https://doi.org/10.1038/srep40706>.

[29] Tripathi, S., Pradeep, T. P., and Barthwal, R. (2016) Molecular Recognition of Parallel DNA Quadruplex d(TTAGGGT)₄ by Mitoxantrone: Binding with 1:2 Stoichiometry Leading to Thermal Stabilization and Telomerase Inhibition. *ChemBioChem* 17, 554–560. <https://doi.org/10.1002/cbic.201500588>.

- [30] Mendoza, O., Gueddouda, N. M., Boule, J.-B., Bourdoncle, A., and Mergny, J.-L. (2015) A fluorescence-based helicase assay: application to the screening of G-quadruplex ligands. *Nucleic Acids Research* 43, e71–e71. <https://doi.org/10.1093/nar/gkv193>.
- [31] Lakowicz, J. R. (2002) Topics in Fluorescence Spectroscopy. Kluwer Academic Publishers. <https://doi.org/10.1007/b115592>.
- [32] Lakowicz, J. R. (2006) Principles of fluorescence spectroscopy Third Edit. Springer. <https://doi.org/10.1007/978-0-387-46312-4>.
- [33] Tansakul, C., Lilie, E., Walter, E. D., Iii, F. R., Wolcott, A., Jin, Z., Millhauser, G. L., Braslau, R., Rivera III, F., Wolcott, A., Jin, Z., Millhauser, G. L., and Braslau, R. (2010) Distance-dependent Fluorescence Quenching and Binding of CdSe Quantum Dots by Functionalized Nitroxide Radicals. *Journal of Physical Chemistry C* 114, 7793–7805. <https://doi.org/10.1021/jp1005023>.

Chapter 6

Fluorescent Light-Up Acridine Orange Derivatives Bind and Stabilize KRAS-22RT G-Quadruplex

This chapter was published in:

Carvalho, J., Pereira, E., Marquevielle, J., Campello, M. P. C., Mergny, J.-L., Paulo, A., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2018) Fluorescent light-up acridine orange derivatives bind and stabilize KRAS-22RT G-quadruplex. *Biochimie* 144, 144–152. <https://doi.org/10.1016/j.biochi.2017.11.004>.

Fluorescent Light-Up Acridine Orange Derivatives Bind and Stabilize KRAS-22RT G-Quadruplex

Abstract

KRAS is often found mutated in lethal cancers and should be an important target for anticancer drug development. However, no effective inhibitor has been reported so far, prompting the scientific community to describe the RAS proteins as nearly “undruggable”. Recent approaches developed to modulate KRAS protein expression comprises the targeting of G4 structures formed within the nuclease hypersensitive element of KRAS promoter region, by designing small and specific ligands to stabilize the tertiary fold and reduce gene expression. In this work, we report *in vitro* and *in silico* studies of novel acridine orange derivatives (C₃-C₈), developed as G4 stabilizing agents. The results show that the ligands bind with high affinity and stabilize KRAS22-RT G4 with modest specificity over duplex DNA. The most promising ligand C₈ stabilizes the structure by ≈ 40 °C. Molecular docking using NMR-derived distance restraints reveal atomic details about the ligand structural features in the interaction with KRAS22-RT G4. In vitro studies with HeLa cells show that the ligands are cytotoxic with IC₅₀ values between 0.9 μ M and 5.7 μ M. Moreover, the ligands tend to localize in the nucleus as shown by confocal fluorescence microscopy. Overall, these results show that the reported acridine orange ligands display favourable properties as G4 ligands and this study provides structural detail for the development of lead KRAS G4 ligands.

Introduction

The KRAS proto-oncogene encodes for a GTPase (p21KRAS) which cycles between GTP-bound active and GDP-inactive states, influencing key signalling pathways related to cell proliferation and apoptosis [1,2]. Due to activating mutations occurring in 30% of all human cancers, the mutated p21KRAS protein remains in activated state leading to increased cell proliferation, anti-apoptotic effects and consequently to cancer development. [2,3]. KRAS mutations are found in a wide variety of cancers such as pancreatic, colorectal, lung and cervical cancers with high incidence in the top-ranked lethal cancers [1,3]. Moreover, p21KRAS is associated with the EGFR and its mutation has

been shown to increase the resistance to anti-EGFR therapies and conventional chemotherapy [3,4]. Therefore, KRAS is an attractive target for cancer therapeutics, despite no effective KRAS inhibitor has been approved thus far. The reported approaches focus on the development of mutated KRAS protein inhibitors, blocking downstream or upstream effectors, and on gene silencing by miRNA or siRNA. However, these strategies failed in achieving a suitable clinical response due to cancerous cell adaption and resistance to inhibitors and inefficient delivery of gene suppressors [1].

One alternative approach to modulate KRAS protein expression relies on the targeting of G4 structures formed in the NHE of KRAS promoter. Those structures were shown to have a regulatory role in the expression of KRAS [2,5,6]. Quadruplexes modulate important mechanisms such as gene transcription, translation and telomere elongation, and are appealing therapeutic targets for the development of specific small molecules that bind and stabilize the G4 structures [7–10]. The KRAS NHE region comprises a particular 32-mer sequence (32R) which contains 6 guanine tracts prone to G4 formation, whose structure was already shown to be highly polymorphic adopting different intramolecular topologies [11]. A smaller 21-mer sequence (21R) adopts the major G4 conformation of 32R. Recently, Salgado and co-workers determined the NMR structure of a 22-mer sequence (22RT) within the NHE which folds into a stable parallel-stranded G4 with 3 G-tetrads [12]. This structure (PDB code: 5I2V) shows unique structural elements such as the presence of a bulge and a 4-nucleotide loop, which can be used for G4-targeted ligands design. Different families of ligands have been synthesized and evaluated in the past decade, among those are porphyrins, acridines, anthraquinones, phenanthrolines, perylenes together with natural compounds such as telomestatin and berberine [13,14]. These compounds interact with the G4 structure through a variety of modes namely external stacking and groove binding [15]. The interest in acridine derivatives is due to its favourable characteristics such as a large planar moiety prone to π -stacking with the tetrads of the G4 structure, possibility of a cationic central ring nitrogen atom, intrinsic fluorescence and cellular uptake with known nucleus penetrating ability [16]. The compound BRACO-19 for instance, is an acridine derivative that binds telomeric G4s specifically over duplex DNA, inhibits telomerase and induces DNA damage with significant *in vivo* anticancer activity in tumour xenografts [17].

In this study, we have evaluated a series of Acridine Orange derivatives (**Figure 6.1**) as a new class of G4 ligands. To tackle this goal, we studied their ability to bind and stabilize the quadruplex found in the KRAS promoter sequence (KRAS-22RT) using spectroscopic techniques such as CD, fluorescence and NMR. Moreover, restrain-assisted molecular

docking simulations using Autodock 4 [18] were performed in order to understand how this family of ligands interact with KRAS-22RT G4. Finally, we have assessed their inhibition properties against cancer cell proliferation using HeLa cell line and normal fibroblasts. These cytotoxicity tests were performed by MTT assays. Moreover, the intracellular localization of the compounds was studied by confocal fluorescence microscopy.

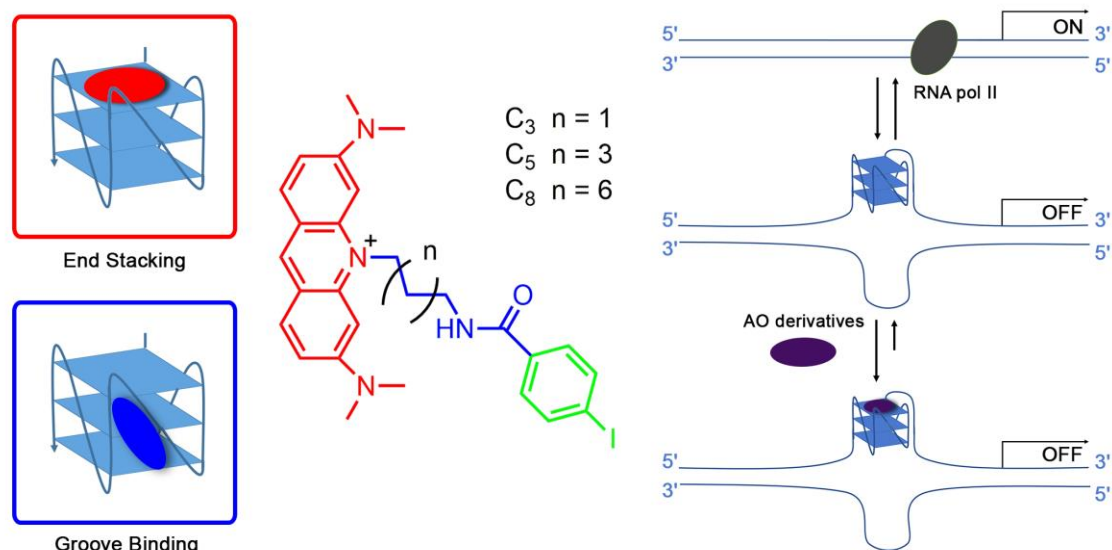


Figure 6.1 - Schematic diagram showing the structures of C_3 , C_5 and C_8 and how the ligands stabilize KRAS G4 structure hence inhibiting the expression of KRAS protein. The acridine moiety interact with the G-tetrads surface through end stacking interactions (red). The alkylamide substituents interact through groove binding and a positively charged nitrogen in the acridine ring is able to occupy a position similar to that of the G-tetrad central ion (blue). The iodobenzene is expected to enhance the cellular uptake while providing additional interactions with the G4 grooves/loop interface (green). Interaction of the acridine orange derivatives with the G4 structure is expected to maintain the "gene-off" state induced by the conversion of the double stranded "gene-on" form to a G4 "gene-off" form (model for the regulation of KRAS gene transcription on the right).

Materials and methods

Oligonucleotides and acridine orange derivatives

All oligonucleotides were obtained from Eurofins MWG Operon (Ebersberg, Germany) or STAB VIDA Genomins Lab (Lisbon, Portugal) with HPLC-grade purification and further purified using Vivaspin® centricons. The following oligonucleotide sequences were used: KRAS-22RT: 5'-AGGGCGGTGTGGAATAGGGAA-3'; c-MYC: 5'-TGAGG GTGGGTAGGGTGGGTAA-3'; 22AG: 5'-AGGGTTAGGGTTAGGGTTAGGG-3' and ds26: d(5'- CAATCGGATCGAATTTCGATCCGATTG-3'). Stock solutions of approximately 500 μ M were prepared using Milli-Q water and stored at -20 °C until used.

The synthesis and purification of 10-(3-(4-iodobenzamide)propyl)-3,6-bis(dimethylamine) acridinium iodide (C₃), 10-(5-(4-iodobenzamide)pentyl)-3,6-bis(dimethylamine) acridinium iodide (C₅) and 10-(8-(4-iodobenzamide)octyl)-3,6-bis(dimethylamine) acridinium iodide (C₈) were performed as previously described [19]. Stock solutions of the compounds were prepared as 10 mM in DMSO. The following dilutions were done using Milli-Q water.

Circular Dichroism Assays

CD spectra were recorded with a Jasco J-815 spectropolarimeter equipped with a temperature control Peltier system. All measurements were made at 25 °C using previously annealed KRAS-22RT G4 (heating to 95 °C for 10 min, followed by cooling in ice) in 10 mM K-phosphate buffer pH 7.0 containing 50 mM K⁺ or in physiologically relevant conditions buffer mimicking the intracellular salt environment (25 mM HEPES (pH = 7.2), 10.5 mM NaCl, 140 mM KCl, 130 nM CaCl₂, 2 mM MgCl₂). Test compounds were dissolved in Milli-Q water at 1 mM stock solution and the titrations performed by adding the required amount directly in a 1 mm path-length quartz cell containing 10 μM G4 DNA. Spectra were collected between 200 and 340 nm using an instrument scanning speed of 100 nm/min, 1 nm bandwidth, 2 s response time and averaged over three scans. All spectra were corrected by subtracting a buffer spectrum. For CD-melting experiments, KRAS22-RT G4 in the absence and presence of the compounds was heated between 20 and 110 °C (temperature set by the Peltier system), with a heating rate of 2 °C/min, while monitoring the ellipticity at 263 nm. Competition assays were performed by adding an excess of ds26 to the cuvette. Additionally, c-MYC and 22AG G4s were tested while monitoring the ellipticity at 263, 290 and 295 nm for parallel, hybrid-1 and antiparallel topologies. For these, buffer conditions were adjusted to favor different topologies and so that the G4s would melt around 50 °C: c-MYC in 100 mM LiCl; 22AG hybrid-1 in 10 mM KCl and 90 mM LiCl; 22AG antiparallel in 100 mM NaCl. Annealing was performed as described above. Data was converted into fraction folded (θ) plots using the following equation:

$$\theta = \frac{CD - CD_{\lambda}^{min}}{CD_{\lambda}^{max} - CD_{\lambda}^{min}} \quad (\text{Equation 1})$$

where CD is the ellipticity of the monitored wavelength at each temperature and $CD^{\min}$ and $CD^{\max}$ are the lowest and highest ellipticity, respectively. Data was fitted to a Boltzmann distribution (OriginPro 8) and the melting temperatures (T_m) determined.

Fluorescence Binding Studies

Fluorescence spectra were recorded with a Horiba FluoroMax4 fluorometer equipped with a temperature control system. Ligands were excited at 498 nm and fluorescence emission collected between 530-700 nm in 10 mM K-phosphate buffer pH 7.0 containing 50 mM KCl at 25 °C using a quartz cuvette with a path length of 1 cm. Similar experiment was conducted using physiologically relevant buffers as described above. Excitation and emission slits were fixed at 2 nm. The association between G4 or duplex DNA and ligands was assessed by titrating the oligonucleotides and measuring the change in fluorescence. The titration was performed by adding previously annealed G4 solution or duplex DNA, followed by 3-min equilibration time. Data was converted into fraction of bound ligand (α) plots using the following equation:

$$\alpha = \frac{I - I_{\lambda}^{free}}{I_{\lambda}^{bound} - I_{\lambda}^{free}} \quad (\text{Equation 2})$$

where I is the fluorescence intensity at 530 nm at each ligand:DNA ratio and I_{free} and I_{bound} are the fluorescence intensity of the free and fully bound ligand, respectively. Data points were then fitted to a hyperbolic function (OriginPro 8) and K_D values were determined from the following saturation binding model [20]:

$$\alpha = \frac{[DNA]}{K_D + [DNA]} \quad (\text{Equation 3})$$

where α is the fraction of ligand bound and $[DNA]$ is the concentration of the DNA. When required, the saturation binding Hill slope model was used [21]:

$$\alpha = \frac{[DNA]^h}{K_D + [DNA]^h} \quad (\text{Equation 4})$$

where n is the Hill constant which describes cooperativity of ligand binding.

The Job plot analysis of ligand C_8 was performed by continuous variation of the ligand/DNA molar fraction while keeping a constant total (DNA + ligand) concentration of

1.5 μM . The excitation wavelength was set at 498 nm and the emission fluorescence recorded at 530 nm.

Nuclear Magnetic Resonance Spectroscopy

^1H NMR spectra were recorded at temperature of 310 K on a Bruker Avance III 700 MHz spectrometer equipped with a liquid TXI $^1\text{H}/^{13}\text{C}/^{15}\text{N}/^2\text{H}$ with Z-gradient probe. Oligonucleotide concentration was typically 0.5 mM in 3 mm tubes with a total volume of 200 μL . G4 samples were annealed as described above in 20 mM K-phosphate buffer pH 6.5 containing 70 mM KCl and supplemented with 10% D_2O . NMR titration was performed by adding increasing amounts of ligand to the G4 solution.

2D ^1H - ^1H NOESY and 1D ^1H - ^{15}N HSQC spectra NMR spectra were recorded in 3 mm tubes at 37 °C unless otherwise stated. Usually the samples were prepared in K-phosphate NMR buffer (20 mM $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$, 70 mM KCl, pH 6.5, 10% D_2O). The concentrations of oligonucleotide were close to 1 mM. Water suppression was achieved with double pulse-field gradient spin echo (DPFGSE) methods. All spectra were processed with the software Topspin 3.2 and analyzed with MestReNova. Chemical shifts (δ) were measured in ppm.

Cell Viability Assay

NHDF and human cervical cancer cells (HeLa) were grown in DMEM supplemented with 10% fetal bovine serum, and 1% antibiotic/antimycotic (DMEM/F-12 was used for NHDF cells). Cultures were maintained at 37 °C in a humidified atmosphere containing 5% CO_2 . Cells were seeded in 48-well plates (1×10^4 cells/well) and after 24 h these were incubated with the compounds at seven different concentrations ranging from 50 nM to 100 μM for 24 h. Wells containing untreated cells were used as control, as well as vehicle control (DMSO) and a common cytotoxic agent 5-fluorouracil (5-FU). At the end of incubation, the media was replaced with fresh media containing MTT and further incubated at 37 °C for 2 h. Finally, MTT containing media was removed, formazan crystals were dissolved in DMSO and absorbance was recorded in a Bio-Rad xMarkTM microplate reader at 570 nm. Cell viability data was expressed as mean $\pm$ SEM from at least three different experiments in comparison with control cells. IC₅₀ values were determined using GraphPad Prism 6.

Confocal Fluorescence Microscopy

HeLa cells were cultured in glass bottom culture dish coated with poly-D-lysine (1×10^5 cells/dish) for 37 °C in a humidified atmosphere containing 5% CO_2 . After growth, cells were incubated in the dark with 10 μM of compound at different times ranging from 5 min

to 24 h. At the end of incubation, the excess of probe was washed off by rinsing with PBS three times. For the co-localization studies, cells were further treated with 1 μ M nuclear probe Hoechst 33342 for 15 min. Prior to visualization, excess probe was washed off by rinsing in PBS three times. Cells were observed using a Zeiss AxioObserver LSM 710 microscope with 405, 488 and 514 nm laser excitation for the 4,6-diamidino-2-phenylindole (DAPI), fluorescein isothiocyanate (FITC) and rhodamine channels, respectively. Appropriate emission bands were selected for the two fluorescent channels, taking into account the spectral overlap and potential bleed-through between both channels. Images were processed with Zeiss ZEN software.

Results

Acridine orange ligands design

Most of the G4 ligands described in the literature share the same characteristics such as a large aromatic core suitable for π - π stacking interaction with the external G-tetrads and side chains enable to establish additional interactions with the grooves/loops of the G4 structure which confer selectivity over duplex DNA. Herein, three acridine orange derivatives with alkylamide side arms containing a iodobenzene substituent were synthesized and evaluated for their G4 binding ability (**Figure 6.1**). These ligands were introduced originally by Paulo and collaborators who also reported the synthesis of their radioiodinated counterparts [19]. Interestingly, the possibility of obtaining the radioiodinated congeners might allow to speed-up the preclinical evaluation of this class of compounds, through biodistribution and in vivo stability/metabolism studies based on quantitative gamma-counting measurements and/or nuclear imaging techniques [19,22]. The planar acridine chromophore was already shown to be an interesting moiety for G4 ligands in the synthesis of 3,6,9-trisubstituted acridine BRACO-19 [17]. The ability of acridine to bind G4s is in part attributed to the fact that it is protonated at physiological pH being able to take advantage of the negative electrostatic potential of the G-tetrads central carbonyl groups [13,23]. The addition of the alkylamide side chains could provide additional interactions between the ligands and the grooves/loops and/or phosphate backbone of the G4 by H-bonding through the amide groups. Moreover, as reported previously, the inclusion of a halogenated substituent such as iodine may increase the binding affinity and stabilization degree of the ligands as well as the cellular uptake [24–26]. This reasoning led us to consider that the acridine orange derivatives presented in **Figure 6.1** would present the structural requisites necessary to act as potent and selective G4-binders.

Binding and stabilization of KRAS-22RT G-quadruplex

CD studies were performed to evaluate the ligands ability to bind and stabilize KRAS-22RT G4 structure. CD spectroscopy is an interesting technique to study the conformation of a G4 structure regarding its strands orientation and to assess the binding potential of G4 ligands [27]. KRAS-22RT adopts a parallel topology in the presence of K^+ cations showing the typical CD spectrum of such conformation (Figure S1, **Appendix C**), with a maximum positive band around 263 nm and a weak minimum negative band around 240 nm. Upon addition of the compounds, only a slight decrease in ellipticity was observed for all ligands, which suggests the preferential binding of the pre-folded parallel topology without disrupting or converting the G4 conformation (Figure S1). The ligand ability to induce thermal stabilization of KRAS-22RT G4 structure was then evaluated by CD melting experiments. The melting temperature (T_m) of the G4 in the absence and presence of 2 molar equivalents of ligands was determined by monitoring the ellipticity at 263 nm and increasing temperature. In the experimental conditions used, KRAS-22RT G4 yielded a melting temperature of about 51 °C. Upon addition of the ligands, the stability of KRAS-22RT was clearly enhanced as denoted by the higher T_m values (**Figure 6.2**, **Table 6.1**).

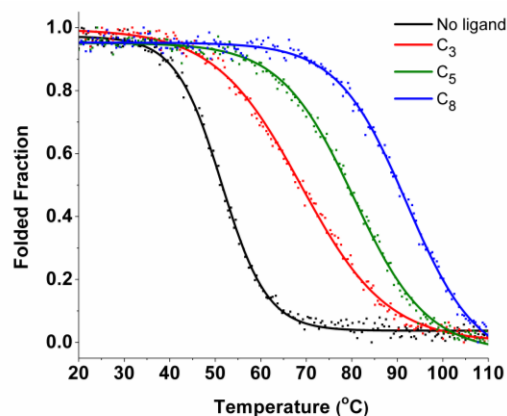


Figure 6.2 - CD melting curves of KRAS-22RT G4 at 5 μ M in the absence and presence of 2 molar equivalents of ligands, in 10 mM K-phosphate buffer and 50 mM KCl. Data points were recorded at 263 nm. The Boltzmann curve fitting of the data points is shown. Temperature corresponds to the temperature set by the Peltier system.

Table 6.1 - Ligand-induced thermal stabilization of various G4s measured by CD melting experiments in the absence and presence of ligands.

Ligand	ΔT_m (°C) ^a			
	KRAS-22RT	c-MYC	22AG-h ^b	22AG-a ^c
C₃	17.8 ± 0.2	18.5 ± 0.3	6.4 ± 0.1	1.5 ± 0.2
C₅	29.0 ± 0.2	22.0 ± 0.2	7.6 ± 0.1	2.7 ± 0.2
C₈	40.7 ± 0.3	33.9 ± 0.7	13.2 ± 0.1	6.8 ± 0.2

^a ΔT_m represents difference in melting temperature [$\Delta T_m = T_m$ (DNA + 2 molar eq. ligand) – T_m (DNA)]. The buffer was 10 mM K-phosphate buffer, pH 7.0. The T_m values are 51.2 ± 0.1 °C [KRAS-22RT (5 μM) in 50 mM KCl]; 51.2 ± 0.2 °C [c-MYC (5 μM) in 100 mM LiCl]; 54.4 ± 0.1 °C [22AG hybrid-1 (5 μM) in 10 mM KCl + 90 mM LiCl]; 56.3 ± 0.2 °C [22AG antiparallel (5 μM) in 100 mM NaCl]. Values reported are average of three measurements with the estimated standard deviation.

^b Hybrid-1 topology.

^c Antiparallel topology.

C₃ promoted an increase in the T_m of 17.8 °C, while C₅ and C₈ stabilized the G4 structure by 29.0 and 40.7 °C, respectively. The ΔT_m value of near 40°C obtained for the ligand C₈ was also corroborated by 1D NMR spectroscopy (Figure S2). The increase in the length of the alkylamide side chain seems to promote a higher thermal stabilization degree as ligand C₈ was the best stabilizer. This feature was studied with more detail by molecular docking studies (see Supporting Information, **Appendix C**). The obtained stabilization values for C₅ and C₈ are higher than those previously reported for other KRAS G4 ligands such as porphyrin TMPyP4 [21,28]. In addition, a competition assay was conducted with the self-complementary DNA sequence ds26, in order to evaluate the specificity of the ligands towards G4 DNA over duplex DNA (Figure S3). Competition melting studies have been used before to assess whether ligands can bind preferentially to G4 structures compared to duplex DNA [28,29]. The ratio of the measured ΔT_m values in the presence and absence of competitor ds26 was used as a measure for specificity (S). In the presence of an excess of ds26 the ΔT_m induced by the ligands was almost unchanged as shown by the S values around 0.9. These results suggest a reasonable specificity of the ligands for KRAS-22RT G4 over duplex DNA.

To assess the selectivity of the ligands towards KRAS-22RT G4, we performed a similar CD melting study with other known G4 structures such as the promoter c-MYC and telomeric sequence 22AG. Promoter sequence c-MYC folds into a parallel G4 conformation while the telomeric sequence 22AG can form two distinct conformations depending on the cation in solution, namely hybrid-1 and antiparallel, in the presence K⁺ or Na⁺, respectively [7]. CD melting curves of the three different G4 conformations in the presence and absence of 2 molar equivalents of ligand were acquired (Figures S4-S6).

Despite using different type of cations, the ionic strength was kept equal for all G4 samples used. Moreover, KRAS-22RT G4 has a very robust and stable conformation even at a KCl concentration 10 times higher than the oligonucleotide concentration and at different pH in the range of 6.0 – 7.0 (checked by NMR, Figures S7-S8). Therefore, the salt conditions used should affect the binding of the acridine orange derivatives in a lesser degree, being the results comparable. The ΔT_m values for each G4 are summarized in **Table 6.1**. The results indicate that the ligands are able to stabilize other G4 structures, despite the lower degree of stabilization observed. A similar trend to that of KRAS-22RT is observed as C₈ promotes the higher stabilization while C₃ is the less effective stabilizer, for all G4 structures. The ligands seem to have a binding preference towards the parallel G4s as the experiments with c-MYC shown the higher ΔT_m values ranging between 18.5 and 33.9 °C, when compared to the hybrid-1 and antiparallel structures of 22AG. The results suggest the order KRAS-22RT > c-MYC > 22AG-h > 22AG-a in terms of thermal stabilization by the ligands.

It is also important to investigate the stabilization effect of the ligands in physiological buffer conditions [30]. Therefore, the CD melting studies (Figure S9) with KRAS-22RT G4 were reconducted in a buffer mimicking the intracellular salt environment (25 mM HEPES (pH = 7.2), 10.5 mM NaCl, 140 mM KCl, 130 nM CaCl₂, 2 mM MgCl₂) [30]. Under these buffer conditions, the T_m of KRAS-22RT G4 structure is 57.9 ± 0.1 °C, which is higher than that obtained previously (51.2 ± 0.2 °C) possibly due to the ionic strength. The obtained ΔT_m values were 9.5 ± 0.2 , 16.4 ± 0.2 and 31.5 ± 0.3 °C for C₃, C₅ and C₈, respectively. Despite a slightly lower stabilization (in terms of relative ΔT_m) as compared to the results shown in **Table 6.1**, the absolute T_m values of KRAS-22RT in the presence of the ligands were roughly the same (68.9 , 80.4 and 92.1 °C *vs* 67.4 , 74.3 and 89.4 °C for C₃, C₅ and C₈, respectively). The difference in the ΔT_m could be a result of the higher T_m value of the G4 structure alone due to the ionic strength.

Fluorescence binding studies with KRAS-22RT G4

The interaction of the ligands with KRAS-22RT G4 was further studied using fluorescence spectroscopy. The ligands excitation at 498 nm produces the emission of light with a broad band centered at 568 nm. Spectrofluorometric titrations were performed with G4 and duplex DNA and the apparent dissociation constants (K_D) determined for the different ligands (**Figure 6.3** and **Table 6.2**) in 10 mM K-phosphate buffer and 50 mM KCl. Upon titration of pre-annealed G4 solution, up to ≈ 10 -fold fluorescence enhancement was observed (**Figure 6.3-I**). The K_D values are presented in **Table 6.2**. The fluorescence titrations were also conducted in a buffer mimicking physiological conditions (same buffer

used in CD) and the fitting of the data points to the saturation binding model are presented in Figure S10 of Supporting Information (**Appendix C**). The K_D values obtained were 3.3×10^{-7} , 1.4×10^{-7} and 2.6×10^{-7} for C_3 , C_5 and C_8 , respectively. The values are similar to those reported in **Table 6.2** and both are indicative of high affinity in the sub-micromolar range, similar to those of potent ligands such as pyridostatin [31].

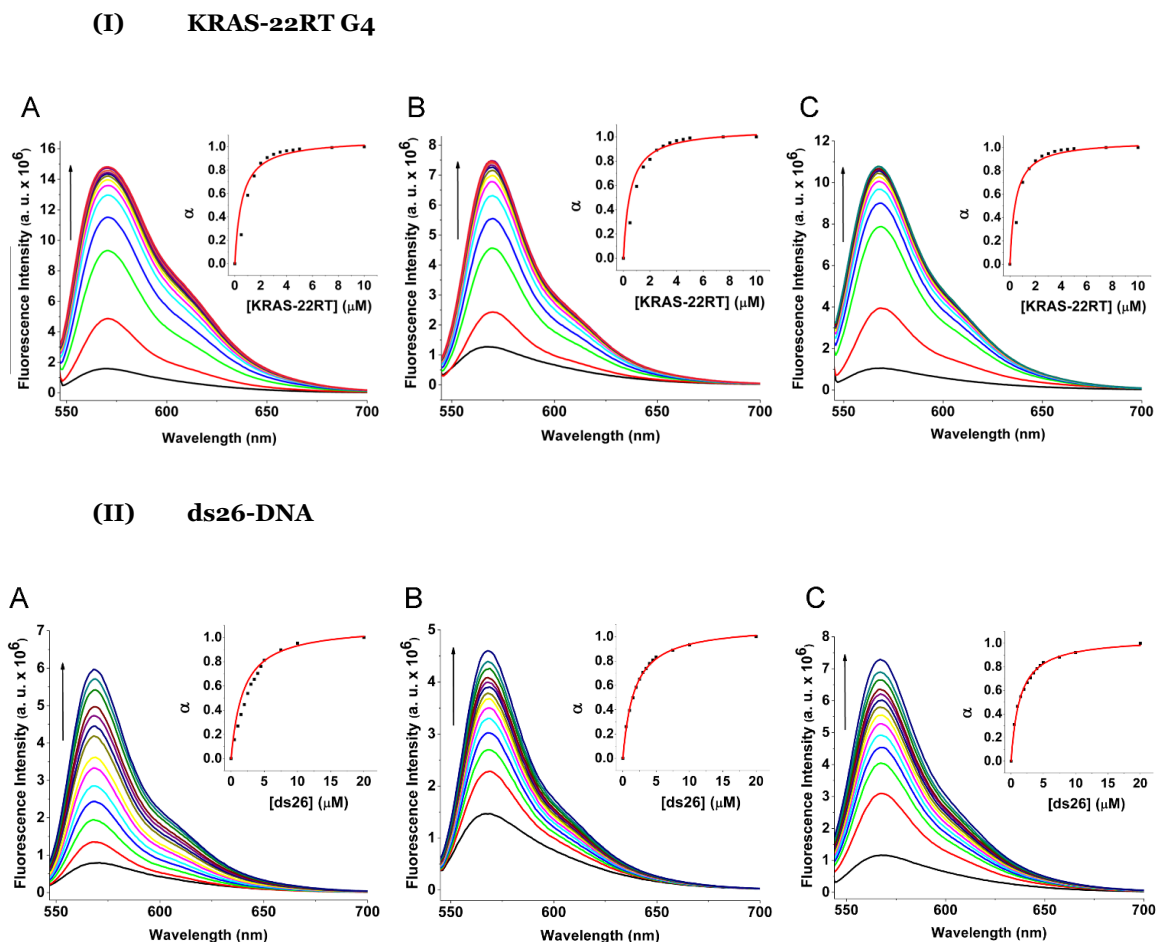


Figure 6.3 - Fluorescence emission spectra of (A) C_3 , (B) C_5 and (C) C_8 with increasing concentration of (I) KRAS-22RT G4 (0.5 – 10 μ M) and (II) DNA ds26 (0.5 – 20 μ M) in 10 mM K-phosphate buffer and 50 mM KCl. Insets: fraction of ligand bound plots fitted to the saturation binding equation.

Table 6.2 - Apparent dissociation constants (K_D) obtained from fluorescence titration experiments.

Ligand	K_D (M) ^a			
	KRAS-22RT	c-MYC	22AG	ds26
C_3	5.5×10^{-7}	1.8×10^{-7}	5.8×10^{-7}	1.9×10^{-6}
C_5	5.1×10^{-7}	2.7×10^{-7}	5.8×10^{-7}	1.7×10^{-6}
C_8	4.4×10^{-7}	7.4×10^{-7}	6.4×10^{-7}	1.4×10^{-6}

^a Determined from fitting the fluorometric titration points to Equations 3 or 4 in the case of c-MYC. The buffer used was 10 mM K-phosphate buffer, pH 7.0. SD $\leq 1.2 \times 10^{-7}$.

In the case of ds26 (**Figure 6.3-II**), the K_D values obtained are about 3.5-fold higher than those obtained for KRAS-22RT G4, indicating weaker binding and modest specificity (both by S and K_D values). Although the specificity was less than that obtained for telomestatin or Phen-DC₃ [32], these results are in line with compounds such as quinacridines (MMQs) [33], indicating that the acridine orange derivatives are interesting scaffolds for further development of G4 specific compounds. For all the complexes formed with G4 and ds26, no variation in the emission wavelength maximum was observed. This suggests a similar chemical environment for the complex formation, although different structures are formed in each complex. However, the fluorescence enhancement upon binding G4 DNA is higher than that observed for ds26 (≈ 3 –7 fold). The varying extent of fluorescence enhancement caused by different DNAs suggests a stronger interaction between the ligands and G4 DNA [34]. These results for ds26 are also in agreement with the CD-melting competition assays and indicate some binding preference of the ligands towards G4 structures.

The K_D values for the interaction of the acridine orange derivatives with c-MYC and 22AG G4s were also determined and are presented in **Table 6.2** and Figures S11 and S12. Intriguingly, the binding curves of the ligands with c-MYC G4 show a sigmoidal rather than hyperbolic shape which may indicate a cooperative mechanism of binding (Figure S11). Therefore, the data points were fitted to Hill equation assuming cooperative binding sites opposingly to that of KRAS-22RT which assumes independent and equivalent binding sites [21]. C₃ has a higher affinity towards c-MYC and the ΔT_m was higher when compared with other G4s; C₈ has lower affinity and ΔT_m towards c-MYC G4 when compared to KRAS-22RT. The Hill constants (h) were greater than 1 for all ligands (1.9, 1.8 and 3.5) which denotes positively cooperative binding, *i.e.* the binding of one ligand increases the affinity of other ligand molecules [21]. In the case of 22AG, the affinity was lower for all ligands when compared to the other G4s studied (**Table 6.2**), which is in agreement with the lower degree of stabilization observed in CD melting studies (Figure S12).

An increase in the alkylamide chain length seems to promote more ligand plasticity, further increasing the binding contacts between the ligands and the G4 structures. This fact may be the explanation for the different ΔT_m values obtained for the ligands since the binding affinities are quite similar between ligands. Furthermore, between different G4 structures, similar K_D values in the order of 10^{-7} M are also observed, indicating that the binding modes are different between the different ligands and the various G4 architectures used.

NMR studies with KRAS-22RT G4

NMR spectroscopy was used to study the complex formation between the acridine orange ligands and KRAS-22RT G4 (**Figures 6.4**, S13 and S14). The imino region of KRAS-22RT ^1H NMR spectrum shows 12 well resolved peaks corresponding to a single G4 conformation with three G-tetrads [12]. Titration of KRAS-22RT G4 with the strongest ligand C₈ is shown in **Figure 6.4**.

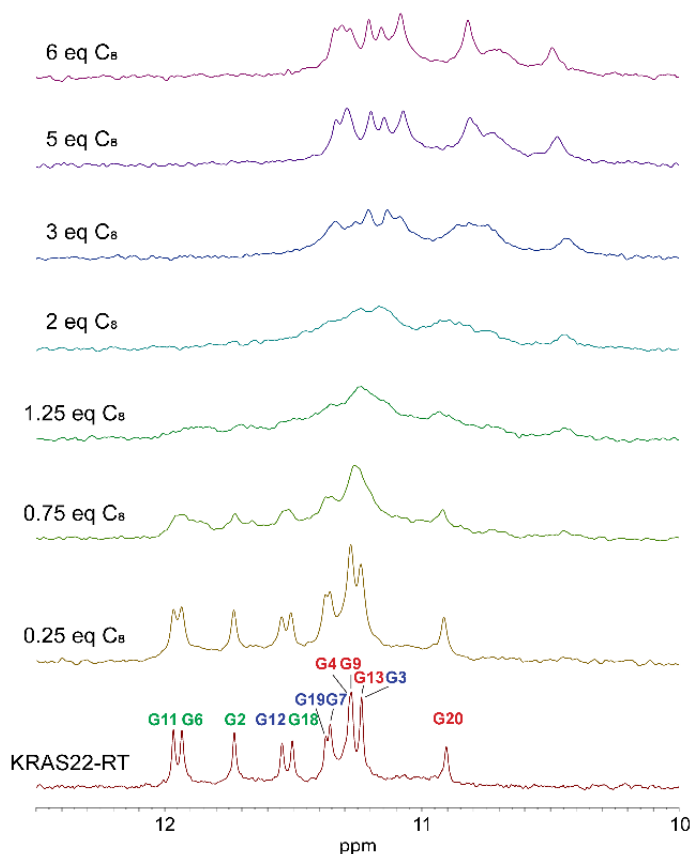


Figure 6.4 - Imino region of KRAS22-RT G4 ^1H spectra upon titration with compound C₈ at 310 K. The assignment of free KRAS22-RT is shown: top 5' G-quartet protons are labelled in green, middle tetrad in blue and bottom 3' tetrad in red.

Upon addition of the ligand, the imino peaks became significantly broader at low ligand concentration. At 2:1 ratio, the imino peaks coalesce and become almost undetectable. This behavior is observed by many different ligands probed against the same oligonucleotide sequence (results not shown), and indicate the formation of higher order complexes with shared ligand molecules which slowly tumble in the liquid state optimal time scale (μs -ps) [35]. In addition, the occurrence of multiple binding modes and conformations in the ensemble, means that one nucleus can have multiple peaks; hence, in the intermediate timescale the averaged NMR signal will be broadened and may

disappear. Finally, at high ligand concentrations, the signals re-appear, indicating the saturation of the binding sites. At a ligand:DNA ratio above 4:1 the imino protons signals did not change significantly. These results suggest a 4:1 ligand:DNA binding ratio. In order to confirm the binding stoichiometry of the C₈ ligand, Job plot analysis was performed using the fluorescence titration data with KRAS-22RT (Figure S15). The molar fraction (χ) of 0.2 was determined for C₈ by the intersection of the linear regressions for high and low molar fractions, which suggests a 4:1 binding ratio. This supports the stoichiometry observed in the ¹H NMR titrations.

The imino peaks were coherently and significantly shifted upfield, which is described as being associated with π - π stacking interactions between ligands and the G-tetrads [36,37]. Such interaction mode is expected with the large aromatic acridine moiety of the ligands. Following the determination of the 4:1 stoichiometry for C₈, the CD melting experiment was carried using 4 molar equivalents of ligand which further increased the KRAS G4 thermal stabilization (> 45 °C). However, the melting temperature was not possible to be determined due to the temperature range limitation. Regarding C₃ and C₅, the analysis of the ¹H NMR titration (Figures S13 and S14) suggest the same ligand:DNA binding stoichiometry of 4:1. The imino protons signals did not change significantly at higher ligand concentrations, suggesting the saturation of the binding sites. Additional ¹H-¹H 2D-NOESY experiments at 37 °C were performed using the most promising ligand C₈ at a ligand:DNA ratio of 4:1. Despite the overlapping between the signals of KRAS22-RT G4 and C₈, several NOE cross-peaks are clearly observed between the ligand and the DNA, namely with the alkylamide chain aliphatic protons of C₈ (**Figure 6.5**). Using ¹⁵N-labelled G4 samples it was possible to assign unambiguously G11. Two other imino-protons (G2 and G6) are overlapped and were assigned to the same chemical shift ($\approx$ 11.95 ppm). In addition, the cross-peaks show almost a two-fold peak volume compared to G11 imino, suggesting that both iminos could participate in the cross peak with aliphatic protons i, g and h. Since G2, G6, G11 and G18 belong to the same tetrad it is reasonable to think that the top 5' tetrad is the principal binding site, although other binding sites are expected due to the 4:1 binding stoichiometry. However, the overlapping of signals volume hinders the unambiguous assignment and analysis of other NOE coming from the 4 ligands bound. Nevertheless, the assigned peaks denote a strong interaction and are expected to correspond to the principal binding site.

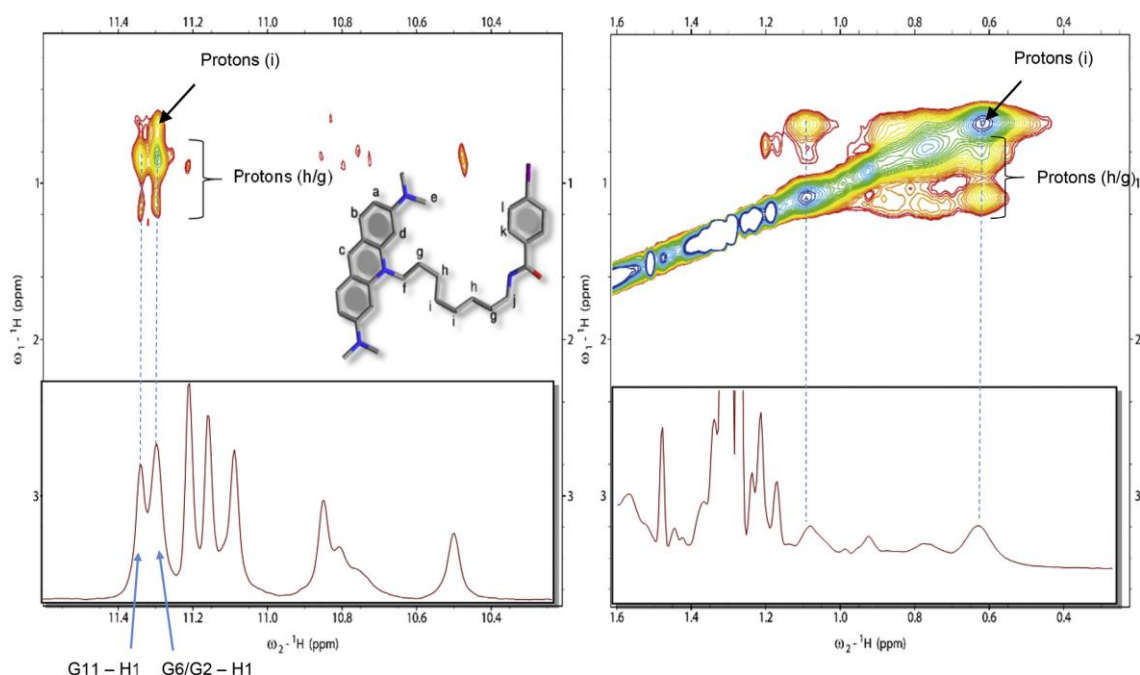


Figure 6.5 - Expanded region of the 2D-NOESY spectrum showing the intermolecular NOEs between the G2, G6 and G11 H1 protons (imino) and ligand C8, namely the alkylamide chain aliphatic protons of the ligand (inset). KRAS-22RT G4 at 350 μ M in presence of 4 molar equivalents of C8.

A deeper insight on the binding mode between ligand C₈ and the G4 structure was obtained by molecular docking studies using the NOE distance restraints obtained in a 250 ms mixing time homonuclear NOESY NMR experiments. The NMR structure of KRAS-22RT (PDB code: 5I2V) was used as docking target for simulation using Autodock 4.2. The results depict the interactions involved in the recognition of the apparent principal binding site, according to the NOE signals observed (see Supporting Information for further discussion, **Appendix C**).

Cell studies: cytotoxicity and intracellular localization

The effect of the ligands on cell proliferation was assessed on a cervical cancer cell line (HeLa) and NHDF by the MTT assay. KRAS gene appears frequently mutated in cervical cancer [38]. The anticancer drug 5-fluorouracil (5-FU) was also used as a control. All the compounds had an effect on the viability of the cancer cells (Figure S16). The concentrations inducing 50% proliferation inhibition (IC₅₀) were remarkably lower for the acridine orange ligands when compared with the reference agent 5-FU. C3 had an IC₅₀ value of 5.7 μ M, while C₅ and C₈ had a lower value of 2.4 and 0.9 μ M, respectively. Notably, 5-FU presented an IC₅₀ value >100 μ M which is >100-fold greater than the value obtained for C₈. As seen previously, the increase in length of the alkylamide chain seems to be important for the activity of the ligand as there is a relationship between higher

stabilization degree, affinity and cytotoxicity, and side chain length. Regrettably, the compounds induced a similar anti-proliferative effect on NHDF cells showing non-selective cytotoxic effect. The IC₅₀ values for NHDF cells were 3.7, 2.5 and 0.7 for C₃, C₅ and C₈, respectively.

Taking advantage of the intrinsic fluorescence of the ligands, confocal fluorescence microscopy was employed to assess their intracellular localization. All the ligands emit green fluorescence upon excitation at 488 nm. After treating the cells with the compounds, the acridine orange derivatives seem to preferentially localize within the nucleus and particularly in the nucleolus, despite some localization on the cytoplasm (**Figure 6.6**). Nucleolus accumulation of ligands often means G4 RNA binding, thus being possible that the ligand binds these types of structures intracellularly as well [39]. The cellular uptake seems to be fast as considerable green fluorescence was observed even at 5 min incubation time of the ligand (Figure S17). The intracellular localization is similar for the three ligands. Acridine orange is known for penetrating the nucleus of cells, thus this aromatic moiety may be responsible for the ligands behavior [40]. The intracellular uptake and nucleus/nucleolus accumulation of the ligands suggest that these compounds may actually reach their intended intracellular target.

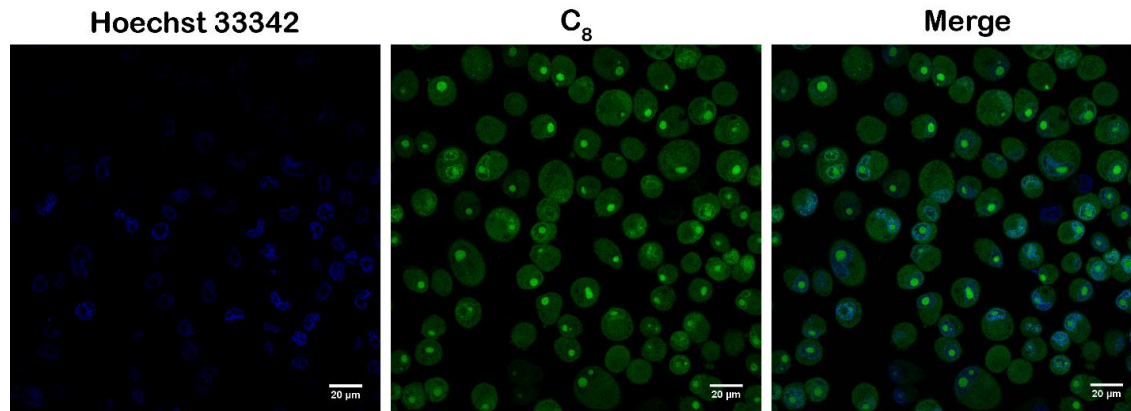


Figure 6.6 - Confocal microscopy images of HeLa cells treated for 1 h with 10 μM C₈ (middle). Nuclear probe Hoechst 33342 (left) was used for co-localization (merge, right). Scale bar is 20 μm. C₃ and C₅ shown similar results.

Conclusions

KRAS plays an important role in the pathogenesis of several types of cancer and is therefore an attractive target for anticancer drugs. Targeting G4 structures formed within KRAS promoter to block transcription and expression has attracted the scientific community in the past years despite no effective suppressor has yet been found [2]. Thus,

the selection of molecules capable of binding to KRAS G4s is still on demand. Overall, this study shows the potential use of acridine orange-based ligands for anticancer drug development based on KRAS oncogene promoter targeting and modulation. Acridine derivatives have been shown already to successfully recognize G4 DNA with antitumor effect, in particular BRACO-19 [17].

The acridine orange moiety together with the alkylamide side chains led to high affinity and reasonable selectivity towards KRAS-22RT G4, promoting a thermal stabilization of about 40 °C for the best ligand C₈. The observed degree of stabilization for C₈ is competitive with other ligands reported for KRAS G4 such as triarylpyridines, indoloquinolines or TMPyP4 [21,28]. Additionally, the *S* values of around 0.9 and 3.5-fold higher *K_D* of the ligand in the presence of duplex DNA shown a modest but significant G4 specificity. The ligands were shown to improve the thermal stability of other known G4 structures but to a lesser extent. Studies using physiological buffers indicate that the ability of the ligands to bind and stabilize G4 structures in a cellular environment is maintained.

The ligands showed high toxicity against HeLa cancer cells; however, the toxicity against non-malignant cells should be improved. To overcome these issues, we plan to introduce additional side chain functionalities in order to increase selectivity towards G4, for instance, by substituting the iodobenzene group with other promising moieties containing fluorine or chloride (work in progress). Another strategy could be to increase the size of the aromatic moiety to limit its access to the smaller surface area of AT and GC base pairs of a typical duplex DNA. Moreover, the nuclear penetration ability and the fluorescence properties of the ligands make them attractive for the development of potential probes for G4 labelling in cellular environment. In brief, the *in silico* and the structural results that we provide in this work will greatly help in designing ligands with better selectivity for KRAS G4 and with promise for G4 DNA based therapeutics.

Acknowledgments

Josué Carvalho acknowledges the doctoral fellowship grant from FCT - Foundation for Science and Technology ref. SFRH/BD/122953/2016. This work was supported by project “Ações Integradas Luso-Francesas” ref. TC-15/17 entitled “Aptameros de G-quadruplex para tratamento do cancro do colo do útero”, FCT project “Projeto de Investigação Exploratória” ref. IF/00959/2015 entitled “Nucleolin targeting by G-quadruplex aptamers for cervical cancer therapy” financed by Fundo Social Europeu e Programa Operacional Potencial Humano. This work was also supported by FEDER funds through the POCI -

COMPETE 2020 - Operational Programme Competitiveness and Internationalisation in Axis I - Strengthening research, technological development and innovation (Project POCI-01-0145-FEDER-007491), by National Funds by FCT (Project UID/Multi /00709/2013), by Conseil régional d'Aquitaine (to J.L.M.) Ligue régionale contre le Cancer (to G.F.S) and by the SYMBIT project (reg. no. CZ.02.1.01/0.0/0.0/15_003/0000477) financed by the ERDF.

References

- [1] Cox, A. D., Fesik, S. W., Kimmelman, A. C., Luo, J., and Der, C. J. (2014) Drugging the undruggable RAS: Mission Possible? *Nature reviews. Drug discovery* 13, 828–51. <https://doi.org/10.1038/nrd4389>.
- [2] Cogoi, S., and Xodo, L. E. (2016) G4 DNA in ras genes and its potential in cancer therapy. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* 1859, 663–674. <https://doi.org/10.1016/j.bbagr.2016.02.002>.
- [3] Chetty, R., and Govender, D. (2013) Gene of the month: KRAS. *Journal of Clinical Pathology* 66, 548–550. <https://doi.org/10.1136/jclinpath-2015-203207>.
- [4] Casadei Gardini, A., Capelli, L., Ulivi, P., Giannini, M., Freier, E., Tamberi, S., Scarpi, E., Passardi, A., Zoli, W., Ragazzini, A., Amadori, D., and Frassinet, G. L. (2014) KRAS, BRAF and PIK3CA status in squamous cell anal carcinoma (SCAC). *PLoS ONE* 9, e92071. <https://doi.org/10.1371/journal.pone.0092071>.
- [5] Paramasivam, M., Cogoi, S., and Xodo, L. E. (2011) Primer extension reactions as a tool to uncover folding motifs within complex G-rich sequences: analysis of the human KRAS NHE. *Chemical Communications* 47, 4965. <https://doi.org/10.1039/c1cc10269a>.
- [6] Paramasivam, M., Membrino, A., Cogoi, S., Fukuda, H., Nakagama, H., and Xodo, L. E. (2009) Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human KRAS promoter: Implications for transcription. *Nucleic Acids Research* 37, 2841–2853. <https://doi.org/10.1093/nar/gkp138>.
- [7] Collie, G., and Parkinson, G. (2011) The application of DNA and RNA G-quadruplexes to therapeutic medicines. *Chemical Society reviews* 40, 5867–5892. <https://doi.org/10.1039/c1cs15067g>.
- [8] Balasubramanian, S., Hurley, L. H., and Neidle, S. (2011) Targeting G-quadruplexes in gene promoters : a novel anticancer strategy? *Nature reviews. Drug discovery* 10, 261–275. <https://doi.org/10.1038/nrd3428>.

- [9] Neidle, S. (2017) Quadruplex nucleic acids as targets for anticancer therapeutics. *Nature Reviews Chemistry* 1, Article 0041. <https://doi.org/10.1038/s41570-017-0041>.
- [10] Neidle, S. (2016) Quadruplex Nucleic Acids as Novel Therapeutic Targets. *Journal of Medicinal Chemistry* 59, 5987–6011. <https://doi.org/10.1021/acs.jmedchem.5b01835>.
- [11] Cogoi, S., Paramasivam, M., Spolaore, B., and Xodo, L. E. (2008) Structural polymorphism within a regulatory element of the human KRAS promoter: Formation of G4-DNA recognized by nuclear proteins. *Nucleic Acids Research* 36, 3765–3780. <https://doi.org/10.1093/nar/gkn120>.
- [12] Kerkour, A., Marquevielle, J., Ivashchenko, S., Yatsunyk, L. A., Mergny, J.-L., and Salgado, G. F. (2017) High-resolution three-dimensional NMR structure of the KRAS proto-oncogene promoter reveals key features of a G-quadruplex involved in transcriptional regulation. *Journal of Biological Chemistry* 292, 8082–8091. <https://doi.org/10.1074/jbc.M117.781906>.
- [13] Monchaud, D., and Teulade-Fichou, M.-P. (2008) A hitchhiker's guide to G-quadruplex ligands. *Organic & biomolecular chemistry* 6, 627–636. <https://doi.org/10.1039/b714772b>.
- [14] Franceschin, M., Rossetti, L., D'Ambrosio, A., Schirripa, S., Bianco, A., Ortaggi, G., Savino, M., Schultes, C., and Neidle, S. (2006) Natural and synthetic G-quadruplex interactive berberine derivatives. *Bioorganic and Medicinal Chemistry Letters* 16, 1707–1711. <https://doi.org/10.1016/j.bmcl.2005.12.001>.
- [15] Zhang, S., Wu, Y., and Zhang, W. (2014) G-quadruplex structures and their interaction diversity with ligands. *ChemMedChem* 9, 899–911. <https://doi.org/10.1002/cmdc.201300566>.
- [16] Debray, J., Zeghida, W., Jourdan, M., Monchaud, D., Dheu-Andries, M.-L., Dumy, P., Teulade-Fichou, M.-P., and Demeunynck, M. (2009) Synthesis and evaluation of fused bispyrimidinoacridines as novel pentacyclic analogues of quadruplex-binder BRACO-19. *Organic & biomolecular chemistry* 7, 5219–28. <https://doi.org/10.1039/b912716j>.
- [17] Burger, A. M., Dai, F., Schultes, C. M., Reszka, A. P., Moore, M. J., Double, J. A., and Neidle, S. (2005) The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function. *Cancer research* 65, 1489–96. <https://doi.org/10.1158/0008-5472.CAN-04-2910>.
- [18] Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., and Olson, A. J. (2009) AutoDock4 and AutoDockTools4: Automated docking with

selective receptor flexibility. *Journal of computational chemistry* 30, 2785–91. <https://doi.org/10.1002/jcc.21256>.

[19] Pereira, E., Do Quental, L., Palma, E., Oliveira, M. C., Mendes, F., Raposinho, P., Correia, I., Lavrado, J., Di Maria, S., Belchior, A., Vaz, P., Santos, I., and Paulo, A. (2017) Evaluation of Acridine Orange Derivatives as DNA-Targeted Radiopharmaceuticals for Auger Therapy: Influence of the Radionuclide and Distance to DNA. *Scientific Reports* 7, 42544. <https://doi.org/10.1038/srep42544>.

[20] Pollard, T. D. (2010) A guide to simple and informative binding assays. *Molecular biology of the cell* 21, 4061–4067. <https://doi.org/10.1091/mbc.E10-08-0683>.

[21] Lavrado, J., Brito, H., Borralho, P. M., Ohnmacht, S. a., Kim, N.-S., Leitão, C., Pisco, S., Gunaratnam, M., Rodrigues, C. M. P., Moreira, R., Neidle, S., and Paulo, A. (2015) KRAS oncogene repression in colon cancer cell lines by G-quadruplex binding indolo[3,2-c]quinolines. *Scientific Reports* 5, 9696. <https://doi.org/10.1038/srep09696>.

[22] Kiess, A. P., Minn, I., Chen, Y., Hobbs, R., Sgouros, G., Mease, R. C., Pullambhatla, M., Shen, C. J., Foss, C. A., and Pomper, M. G. (2015) Auger Radiopharmaceutical Therapy Targeting Prostate-Specific Membrane Antigen. *Journal of Nuclear Medicine* 56, 1401–1407. <https://doi.org/10.2967/jnumed.115.155929>.

[23] Read, M., Harrison, R. J., Romagnoli, B., Tanious, F. A., Gowan, S. H., Reszka, A. P., Wilson, W. D., Kelland, L. R., and Neidle, S. (2001) Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 98, 4844–9. <https://doi.org/10.1073/pnas.081560598>.

[24] Meng, T., Tang, S.-F., Qin, Q.-P., Liang, Y.-L., Wu, C.-X., Wang, C.-Y., Yan, H.-T., Dong, J.-X., and Liu, Y.-C. (2016) Evaluation of the effect of iodine substitution of 8-hydroxyquinoline on its platinum (ii) complex: cytotoxicity, cell apoptosis and telomerase inhibition. *Med. Chem. Commun.* 7, 1802–1811. <https://doi.org/10.1039/C6MD00201C>.

[25] Nanjunda, R., Owens, E. A., Mickelson, L., Dost, T. L., Stroeva, E. M., Huynh, H. T., Germann, M. W., Henary, M. M., and Wilson, W. D. (2017) Selective G-Quadruplex DNA Recognition by a New Class of Designed Cyanines. *Molecules* 18, 13588–13607. <https://doi.org/10.1007/s00210-015-1172-8>.The.

[26] Sidibe, A., Hamon, F., Largy, E., Gomez, D., Teulade-Fichou, M.-P., Trentesaux, C., and Riou, J. F. (2012) Effects of a halogenated G-quadruplex ligand from the pyridine dicarboxamide series on the terminal sequence of XpYp telomere in HT1080 cells. *Biochimie* 94, 2559–2568. <https://doi.org/10.1016/j.biochi.2012.07.003>.

- [27] Paramasivan, S., Rujan, I., and Bolton, P. H. (2007) Circular dichroism of quadruplex DNAs: Applications to structure, cation effects and ligand binding. *Methods* **43**, 324–331. <https://doi.org/10.1016/j.ymeth.2007.02.009>.
- [28] Smith, N. M., Labrunie, G., Corry, B., Tran, P. L. T., Norret, M., Djavaheri-Mergny, M., Raston, C. L., and Mergny, J.-L. (2011) Unraveling the relationship between structure and stabilization of triarylpyridines as G-quadruplex binding ligands. *Organic & biomolecular chemistry* **9**, 6154–6162. <https://doi.org/10.1039/c1ob05560g>.
- [29] Rache, A. De, Gueddouda, N. M., Bourdoncle, A., Hommes, P., Reissig, H.-U., and Mergny, J.-L. (2016) A flexible terpyridine derivative interacts specifically with G-quadruplexes. *Chemistry - A European Journal* **22**, 12651 – 12654. <https://doi.org/10.1002/chem.20160302>.
- [30] Hansel, R., Foldynová-Trantírková, S., Dotsch, V., and Trantírek, L. (2013) Investigation of Quadruplex Structure Under Physiological Conditions Using In-Cell NMR. *Topics in Current Chemistry* **330**, 47–66. <https://doi.org/10.1007/128>.
- [31] Koirala, D., Dhakal, S., Ashbridge, B., Sannohe, Y., Rodriguez, R., Sugiyama, H., Balasubramanian, S., and Mao, H. (2011) A single-molecule platform for investigation of interactions between G-quadruplexes and small-molecule ligands. *Nature chemistry* **3**, 782–7. <https://doi.org/10.1038/nchem.1126>.
- [32] Bončina, M., Podlipnik, Č., Piantanida, I., Eilmes, J., Teulade-Fichou, M. P., Vesnaver, G., and Lah, J. (2015) Thermodynamic fingerprints of ligand binding to human telomeric G-quadruplexes. *Nucleic Acids Research* **43**, 10376–10386. <https://doi.org/10.1093/nar/gkv1167>.
- [33] Hounsou, C., Guittat, L., Monchaud, D., Jourdan, M., Saettel, N., Mergny, J.-L., and Teulade-Fichou, M.-P. (2007) G-quadruplex recognition by quinacridines: A SAR, NMR, and biological study. *ChemMedChem* **2**, 655–666. <https://doi.org/10.1002/cmdc.200600286>.
- [34] Zhang, X., Wei, Y., Bing, T., Liu, X., Zhang, N., Wang, J., He, J., Jin, B., and Shangguan, D. (2017) Development of squaraine based G-quadruplex ligands using click chemistry. *Scientific Reports* **7**, 4766. <https://doi.org/10.1038/s41598-017-04344-x>.
- [35] Palmer, A. G. (2004) NMR Characterization of the Dynamics of Biomacromolecules. *Chemical Reviews* **104**, 3623–3640. <https://doi.org/10.1021/cr030413t>.
- [36] Phan, A. T., Kuryavyi, V., Gaw, H. Y., and Patel, D. J. (2005) Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC

promoter. *Nature Chemical Biology* 1, 167–173. <https://doi.org/10.1038/nchembio723>.

[37] Dhamodharan, V., Harikrishna, S., Bhasikuttan, A. C., and Pradeepkumar, P. I. (2015) Topology Specific Stabilization of Promoter over Telomeric G-Quadruplex DNAs by Bisbenzimidazole Carboxamide Derivatives. *ACS Chemical Biology* 10, 812–833. <https://doi.org/10.1021/cb5008597>.

[38] Wright, A. A., Howitt, B. E., Myers, A. P., Dahlberg, S. E., Palescandolo, E., Van Hummelen, P., MacConaill, L. E., Shoni, M., Wagle, N., Jones, R. T., Quick, C. M., Laury, A., Katz, I. T., Hahn, W. C., Matulonis, U. A., and Hirsch, M. S. (2013) Oncogenic mutations in cervical cancer: genomic differences between adenocarcinomas and squamous cell carcinomas of the cervix. *Cancer* 119, 3776–3783. <https://doi.org/10.1016/j.pestbp.2011.02.012>.

[39] Drygin, D., Siddiqui-Jain, A., O'Brien, S., Schwaebe, M., Lin, A., Bliesath, J., Ho, C. B., Proffitt, C., Trent, K., Whitten, J. P., Lim, J. K. C., Von Hoff, D., Anderes, K., and Rice, W. G. (2009) Anticancer activity of CX-3543: A direct inhibitor of rRNA biogenesis. *Cancer Research* 69, 7653–7661. <https://doi.org/10.1158/0008-5472.CAN-09-1304>.

[40] Agorastos, N., Borsig, L., Renard, A., Antoni, P., Viola, G., Spingler, B., Kurz, P., and Alberto, R. (2007) Cell-Specific and Nuclear Targeting with $[M(CO)_3]^+$ ($M=^{99m}Tc$, Re)-Based Complexes Conjugated to Acridine Orange and Bombesin. *Chemistry - A European Journal* 13, 3842–3852. <https://doi.org/10.1002/chem.200700031>.

Chapter 7

Human Papillomavirus G-Rich Regions as Potential Antiviral Drug Targets

This chapter was published in:

Carvalho, J., Lopes-Nunes, J., Campello, M. P. C., Paulo, A., Milici, J., Meyers, C., Mergny, J.-L., Salgado, G. F., Queiroz, J. A., and Cruz, C. Human papillomavirus G-rich regions as potential antiviral drug targets. *Nucleic Acid Therapeutics*, Ahead of Print. <https://doi.org/10.1089/nat.2020.0869>.

Human Papillomavirus G-Rich Regions as Potential Antiviral Drug Targets

Abstract

Herein we report for the first time the screening of several ligands in terms of their ability to bind and stabilize G4s found in seven HPV genomes. Using a variety of biophysical assays, HPV G4s were shown to possess a high degree of structural polymorphism upon ligand binding which may have an impact on transcription, replication, and viral protein production. A sequence found in high-risk HPV16 genotype folds into multiple non-canonical DNA structures; it was converted into a major G4 conformation upon interaction with a well-characterized highly selective G4-ligand, PhenDC₃, which may have an impact on the viral infection. Likewise, HPV57 and 58, which fold into multiple G4 structures, were found to form single stable complexes in the presence of two other G4-ligands, C₈ and Pyridostatin, respectively. Additionally, one of the selected compounds, the acridine derivative C₈, demonstrated a significant antiviral effect in HPV18-infected organotypic raft cultures. Altogether, these results indicate that targeting HPV G4s may be an alternative route for the development of novel antiviral therapies.

Introduction

HPV infection is the major causative agent of cervical cancer, which is the fourth most frequent cancer in women with an estimated 570,000 new cases per year (data from the World Health Organization, GLOBOCAN 2018). To date, about 200 genotypes of HPV have been identified and divided into two groups, low-risk and high-risk HPV genotypes. Low-risk HPVs cause benign lesions such as condylomas or warts while high-risk types, notably HPV16 and 18, can lead to precancerous lesions and cervical cancer [1]. Around 99% of precancerous cervical lesions and cervical cancer are associated with infection by high-risk HPV genotypes [1]. The carcinogenic mechanisms of HPV infection consist of deregulation of cell cycle and chromosomal instability caused by viral oncoproteins E6 and E7, chronic viral infection, and the integration of HPV DNA in the host genome [1]. Regardless of recent efforts in prevention, early diagnosis, effective screening, follow-up and vaccination programs for HPV, which reduced cancer rate and mortality, the need for treatment of already infected patients is still a priority. The ideal HPV therapeutic strategy

would be able to clear the viral infection during the HPV latent form and to generate an efficient virus-specific response [2,3].

The genome of HPV is divided into 3 regions: a LCR that contains the necessary cis-elements for the replication and transcription of viral DNA; the early region, composed of 6 ORFs labelled E1–E7; and a late region with two ORFs coding for viral structural proteins L1 and L2 [4]. G4s are secondary higher-order nucleic acid structures based on the stacking of two or more planar layers of four guanines (called G-quartets or G-tetrads) held together by Hoogsteen hydrogen bonds and a central cation (particularly K⁺ and Na⁺) [5,6]. The structural complexity and diversity of the G4s contrasts with the relative uniformity of DNA duplexes, as G4 topology depends on the number and orientation of strands, loop length and sequence, molecular crowding, and the monovalent ions present [6]. In humans, G-rich motifs with the potential of forming stable G4s are predominantly found in regulatory regions such as telomeres, oncogene promoters and the 5'-untranslated regions of many mRNAs [6]. This suggests that G4s are likely involved in the regulation of key cellular processes such as replication, transcription, translation and genomic stability. Thus, these motifs are attractive targets for the development of G4-targeted compounds, capable of stabilizing these structures, affecting the binding of G4-binding proteins, such as nucleolin [7,8], and interfering with the regulation of gene expression as described for a variety of cellular targets and viruses such as HIV or EBV [5,9,10]. Indeed, G4 targeting in viruses is increasingly gathering attention as a potential treatment and diagnosis tool. G4s were discovered in HIV, HSV, Epstein–Barr virus and EBV, among others [10–14]. Similarly, to humans, these G4s were found in regulatory regions involved in viral DNA replication and viral proteins expression, and several reports describe the use of G4 compounds with antiviral activity in infected cells [10,11]. For instance, PhenDC3 inhibits KSHV viral replication by blocking the progression of replication forks [15], showing the potential of using G4 compounds to control viral infections.

The genome of HPV displays several G-rich sequences with the potential to form highly stable G4s found in only 10 out of 120 identified HPV types (notably high-risk HPVs 16, 18, 52 and 58) [2,16,17]. These G-rich regions are present in LCR, E1, E4 regions and L2 protein-coding sequences, which are involved in transcription, replication, and viral proteins production. Despite an increased interest in developing G4-mediated antiviral drugs to manage viral infection [10,11,18], no HPV G4 binding compound has ever been reported. Given that vaccination does not treat a pre-existing infection, combined with the lack of effective anti-HPV treatments [19], there is an urgent need for the development of

new therapeutic approaches for the management of HPV diseases. To this end, G4 compounds may provide an attractive option with their innovative mechanisms of action.

Herein, we have employed a combination of biophysical techniques such as CD and NMR to evaluate the binding properties between known G4 compounds and G4 structures of seven different HPV genotypes, namely high and low risk HPV types 9, 16, 18, 32, 52, 57 and 58 (the corresponding sequences are listed in **Table 7.1**). The sequences were chosen based on the ^1H NMR profiles previously shown by Tlučková *et al.* and Marušič *et al.* [2,16], the actual detection of G4 folding *in vitro* and giving priority to high-risk HPVs. Additionally, the effect of a selected ligand on HPV16 and HPV18 replication and encapsidation was evaluated using organotypic raft cultures with episomal viral genomes. The ligands used in this study were two phenanthroline derivatives (PhenDC3 [20] and [16]phenN₂ [21]) two acridines (BRACO-19 [22] and C₈ [23]), two pyridines (Pyridostatin (PDS) [24] and 360A [25]) and one porphyrin TMPyP4 [26] (**Figure 7.1**). These ligands have been previously shown to bind to a variety of G4 structures with high affinity – although with variable specificity - and repress protein transcription or induce cell death in different cell types [20–26].

Table 7.1 - HPV oligonucleotide sequences used in this study.

Virus type	Position ^a	Region (Regulatory or Coding) ^b	Sequence ^c	G4-Hunter score [44]	Molar extinction coefficient (l.mol ⁻¹ .cm ⁻¹)
HPV9	3645	E4 (Coding)	5'-GTGGGAGCGGGAACGGGAACGGGA-3'	1.46	249000
HPV16	4422	L2 (Coding)	5'-AGGGTCGGGTACAGGCGGACGCACTGGGT-3'	1.03	283900
HPV18	4430	L2 (Coding)	5'-AGTGGTACAGGGGGTCGTACAGGGTACATTCATTGGGTGGGCG-3'	1.02	434400
HPV32	1321	E1 (Coding)	5'-GGGAGTATGGGTAACGGGGGGGG-3'	2.17	238400
HPV52	7445	LCR (Regulatory)	5'-GGGTAGGGCAGGGGACACAGGGT-3'	1.74	235700
HPV57	5464	L2 (Coding)	5'-GGGAAAGGGTACCTCGAGGGGCCGCGGG-3'	1.45	285800
HPV58	7362	LCR (Regulatory)	5'-GGGCAGGGTAGGGCAATTTAGGG-3'	1.48	234400

^a Position of the G-rich sequence in the respective HPV genome

^b Indicated in parentheses is the location of the G4 sequence into a regulatory region (i.e. promoters or enhancers) or into a coding region according to the NCBI Genbank database (<https://www.ncbi.nlm.nih.gov/genbank/>)

^c Runs of two or more guanines are shown in bold characters. The propensity to fold into a G4 structure was evaluated by G4 hunter algorithm using the previously described protocol [44]

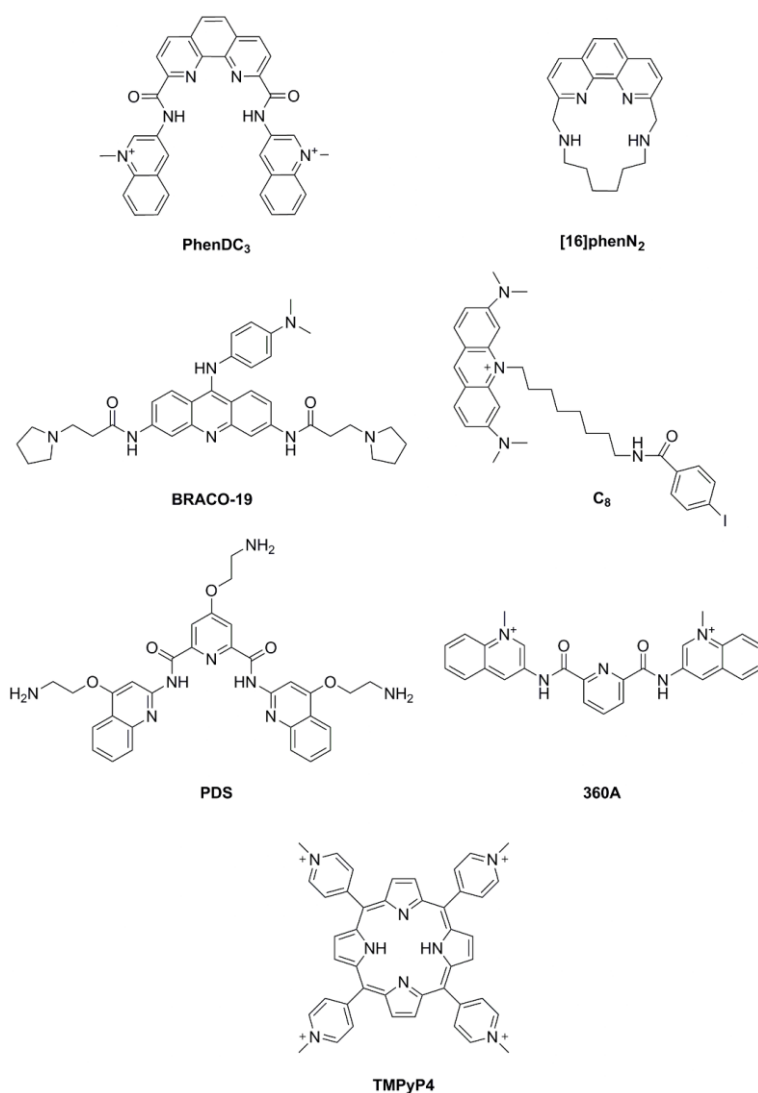


Figure 7.1 - Structure of the G4 ligands used in this study. Ligands belong to different starting moieties, particularly phenanthrolines PhenDC3 and [16]phenN₂, acridines BRACO-19 and C₈, pyridines Pyridostatin (PDS) and 360A, and porphyrin TMPyP4.

Material and methods

Oligonucleotides and compounds

The HPV sequences used are listed in **Table 7.1** and were purchased from Eurogentec (Belgium) with HPLC-grade purification and used without further treatment. Stock solutions were prepared in Milli-Q water and the concentrations determined using the molar extinction coefficients provided by the manufacturer (**Table 7.1**) by measuring the UV absorbance at 260 nm. For all experiments, the oligos were annealed by heating to 95 °C for 10 min and slowly cooled to room temperature until used. PhenDC3 (CAS: 929895-45-4) was purchased from Polysciences, Inc. (USA). Pyridostatin (CAS: 1085412-37-8) was purchased from Sigma-Aldrich (USA). [16]phenN₂ and C₈ were synthesized as previously

described [21,23]. TMPyP4 (CAS: 36951-72-1) was purchased from Tokyo Chemical Industry Co., Ltd. (Japan). 360A (CAS: 794458-56-3) and BRACO-19 (CAS: 1177798-88-7) were purchased from Sigma-Aldrich (USA). The compounds were dissolved either in DMSO or milli-Q water to a 10 mM stock solution and further diluted in Milli-Q water. TO (CAS: 107091-89-4) was bought from Sigma-Aldrich (USA).

Evaluation of G4 propensity (G4 Hunter)

G4 propensity of each sequence was calculated with G4-Hunter [27]: higher score increases the formation of a stable quadruplex. Previous experimental validation on a large set of sequences ([44] and unpublished data) indicates that nearly all (>95%) sequences with G4 Hunter score above 1.5 form a quadruplex under near-physiological conditions, while a G4 formation is likely for motifs with scores between 1.2 and 1.5, and unlikely for scores below 1.

Fluorescence resonance energy transfer melting

FRET melting assays were performed on a CFX Connect™ Real-Time PCR Detection System (Bio-Rad, USA). The HPV sequences were purchased with double labelling by fluorescein (FAM) and carboxytetramethylrhodamine (TAMRA) at the 5' and 3' ends, respectively. The measurements were performed in 10 mM lithium cacodylate supplemented with 10 mM KCl and 90 mM LiCl. The concentration of the labelled oligonucleotide was 0.2 μ M and ligands were used at 0.2 and 1 μ M. The buffer conditions were selected so that the quadruplexes would melt around 50 °C in the absence of ligands. These conditions ensure sufficient plateau phases (folded *vs* unfolded) to determine optimal baselines and allow accurate ΔT_m determination [27]. The fluorescence intensity of FAM was recorded in the 25-95 °C temperature range, with a temperature increment of 1 °C/min. The excitation and detection wavelengths were 492 and 516 nm, respectively. Each experimental condition was tested in triplicate. The melting temperatures were determined from normalized curves as the temperature for which the normalized emission was equal to 0.5.

Circular dichroism

CD spectra were collected on a Jasco J-815 CD spectrometer (France) equipped with a Peltier-type temperature controller (model CDF-426S/15). Samples were prepared in 1-mm quartz cuvette at a concentration of 10 μ M in 10 mM lithium cacodylate supplemented with 10 mM KCl and 90 mM LiCl. Additional experiments were performed at 50 mM KCl concentration to assess its effect on the G4 folding (Figure S3). CD spectra

were acquired in the 220-340 nm range, with a scan speed of 100 nm/min, 1 nm bandwidth and 1 s integration time. The individual spectra were averaged over 4 accumulations. During titrations, the required volume of ligand was added to the quartz cell. For CD melting studies, DNA sequences were used at a concentration of 1 μ M. The buffer conditions were selected to adjust the melting temperatures around 50 °C. These conditions ensure sufficient plateau phases (folded *vs* unfolded) to determine optimal baselines and allows accurate ΔT_m determination [27]. Melting curves were obtained by monitoring ellipticity at a single wavelength, either 265 nm or 290 nm, as a function of temperature in the 20-100 °C range with a heating rate of 1 °C/min. Data were converted into fraction folded (θ) plots, fitted to a Boltzmann distribution using OriginPro 2016 and the melting temperatures (T_m) obtained.

Fluorescent Intercalator Displacement Assay

The experiments were performed in Greiner 96-well opaque clear-bottom microplates. 0.5 μ M of pre-folded HPV G4s were mixed with 1 μ M of TO in 10 mM K-phosphate buffer, pH 7.2, with 50 mM KCl. Five equivalents of each ligand (2.5 μ M) were added for each HPV sequence. Wells with ligands only were included to account the ligand's natural fluorescence; control wells with ligands and TO were also included to screen for probes interaction. TO fluorescence was measured at 25 °C in a Spectramax Gemini EM spectrofluorometer (Molecular Devices LLC, USA) with the excitation at 485 nm and the emission collected at 520 nm. Each condition was tested in duplicate in 2 separate plates in a volume of 100 μ L per sample. The percentage of TO displacement was calculated by the formula:

$$TO\ Displacement\ (\%) = 100 - ((FA/FA_0) \times 100) \quad (1)$$

where the initial fluorescence with and without ligand (FI and FI_0 , respectively) is subtracted by the background fluorescence F_b : $FA = FI - F_b$ or $FA_0 = FI_0 - F_b$. In order to test possible rigidification/conformational change of the HPV G4s upon interacting with TO and ligands, CD spectra of the G4s in the presence of 2 molar equivalents TO were acquired, followed by the addition of 5 molar equivalents of ligand.

Nuclear Magnetic Resonance

Standard ^1H NMR spectra were recorded on a Bruker Avance III 600 MHz spectrometer equipped with a QCI CryoProbe at 25 °C. The zgpg30 pulse sequence was used to suppress the water signal. Oligonucleotide concentration was typically 200 μ M in 3 mm tubes with a total volume of 200 μ L. Samples were annealed as described above in 10 mM K-

phosphate buffer, pH 7.2 containing 50 mM KCl and supplemented with 10% D₂O. NMR titrations were performed by adding increasing amounts of ligands to the 3 mm tube. Titrations were stopped once no imino signals were observed or the chemical shifts remained unchanged. Chemical shifts (δ) were measured in ppm. All spectra were acquired and processed with the software Topspin 3.1.

Keratinocyte and organotypic raft cultures

Human cervical keratinocyte lines stably maintaining HPV DNA (HCK16-5 and HCK18C) were grown in monolayer culture using E medium in the presence of mitomycin C-treated J2 3T3 feeder cells. J2 3T3 feeder cells were grown in DMEM medium supplemented with 10% FBS and gentamycin. HPV organotypic raft cultures were established as previously described [28–30]. In brief, cell lines were seeded onto rat-tail type 1 collagen matrices containing J2 3T3 feeder cells. Following cell attachment and growth to confluence, the matrices were lifted onto stainless steel grids. The raft cultures were then fed by diffusion from below with epithelium medium (E-medium) supplemented with 100, 250 and 500 nM ligand or DMSO alone at the higher concentration used for the ligands as a solvent control. The raft cultures were allowed to stratify and differentiate for 20 days. Over this growth period, cells were fed with fresh media containing the compounds every other day.

Histology of Tissue Sections

Raft tissues were fixed in 10% neutral buffered formalin. Tissue was embedded in paraffin, cut into sections and stained with hematoxylin and eosin (H&E) as previously described [30]. Images were adjusted for brightness and contrast identically.

Virus Harvest and Isolation

Viruses were isolated by Dounce homogenization of tissue in sodium phosphate buffer (50 mM sodium phosphate supplemented with 1 M NaCl, pH 8.0). Post homogenization, rafts were either left untreated to quantify total genomes copies or were treated with 2 mM MgCl₂ and 375 U of benzonase for 1 h at 37 °C to remove unprotected genomes. Samples were then adjusted to 1 M NaCl and centrifuged for 10 min at 4 °C and 10,500 rpm to remove cellular debris. Viral preparations were stored at -20 °C for short term until further analysis.

Viral Titers

Viral titers were determined as previously described [30]. In brief, encapsidated viral genomes were released by resuspending 10 μ L of benzonase-treated virus preparation in

166 μ L of Hirt DNA extraction buffer (400 mM NaCl/10 mM Tris-HCl, pH 7.4/10 mM EDTA, pH 8.0), 2 μ L 20 mg/mL proteinase K, and 10 μ L 10% SDS for 2 h at 37 °C. Following extraction, the DNA was purified by adding an equal amount of phenol–chloroform–isoamyl alcohol (25:24:1) to the mixture and extracting the aqueous phase. An equal amount of chloroform was added, and the aqueous phase extracted again. DNA was then precipitated overnight at 20 °C by adding 2.5 volumes of 100% ethanol and 0.1 volumes of 3 M sodium acetate. The DNA was pelleted by centrifugation at 14,000 rpm for 10 min and then washed with 70% ethanol and resuspended in 20 μ L Tris-EDTA. To quantify the viral genomes, a QIAGEN Quantitect SYBR Green PCR kit (Qiagen, USA) was used. The 5' and 3' primers used for HPV16 titrating were 5'-CCATATAGACTATTGGAAACACATGCGCC-3' and 5'-CGTTAGTTGCAGTTCAATTGCTTGTAAATGC-3', respectively. For HPV18 the primers used were 5'-CAGTATTAACCACCAGGTGGTGGTG-3' and 5'-GTTCAAGACTTGTTTGCTGCATTGTCC-3'. A Bio-Rad CFX-96 Real-Time qPCR machine (Biorad, USA) and corresponding software were used for data collection and subsequent analysis.

Results and discussion

G4 sequences were recently found in the genome of both low- and high-risk HPV genotypes and are potential drug targets to control the viral life cycle [2,16,17]. For this study, seven different G-rich sequences were selected from HPV types 9, 16, 18, 32, 52, 57 and 58 and are listed in **Table 7.1**. The sequences were chosen based on the ^1H NMR profiles previously shown by Tlučková et al. and Marušič et al. [2,16] and the detection of stable G4 structures in vitro. The G4 sequences were found in the LCR (HPVs 52 and 58), E1 (HPV32), E4 (HPV9) and L2 (HPVs 16, 18 and 57) regions. For this study, the G4 folding propensity was re-evaluated using the G4-Hunter algorithm [31]. All sequences presented a G4-Hunter score above 1 which is indicative of the likelihood of the sequences to fold into stable G4 structures, in agreement with the previous studies by Tlučková et al. and Marušič et al. [2,16]. The sequences derived HPV16 and HPV18 genomes presented the lower G4-Hunter scores (1.03 and 1.02); indeed, HPV16 was already suggested to form other non-B-DNA structures such as hairpin, while HPV18 requires two DNA strands to fold into a G4 structure (see below for further analysis) [2]. The interaction of the HPV G4 structures with known ligands was then assessed using a variety of biophysical assays.

In order to assess the binding and stabilizing capability of the ligands towards HPV G4 structures, we initially performed FRET melting experiments using fluorescent-labelled

oligonucleotides mimicking motifs found in the HPV genome. The FHPV9T, FHPV16T, FHPV18T, FHPV32T, FHPV52T, FHPV57T and FHPV58T oligonucleotides were tested in the presence of each ligand at 0.2 and 1 μM concentration (1 and 5 molar equivalents, respectively). Among the tested ligands, PhenDC3 and TMPyP4 were the stronger stabilizers of all HPV G4 structures studied while [16]phenN₂ was the ligand which presented lower ΔT_m values (**Table 7.2**, Figure S1 **Appendix D**). When comparing different HPV sequences, HPV16 was the least prone to ligand-induced stabilization while HPV52 and HPV58 were the most stabilized structures. Marušič *et al.* reported that HPV16 could form multiple unstable G4 structures that coexist in solution with a hairpin-like structure, hence the lower stabilization values [2]. These authors also found HPV52 and HPV58 to be the most suitable targets for small-molecule mediated stabilization. Indeed, these sequences were the most prone to ligand stabilization. Furthermore, TMPyP4, which was already shown to possess a poor G4/duplex selectivity [32], was able to stabilize HPV16 by more than 30 °C, suggesting that indeed non-G4 structures may be formed by this sequence. The results at 0.2 μM ligand concentration are shown in Table S1.

Table 7.2 - Ligand-induced thermal stabilization (ΔT_m) measured by FRET melting experiments.

Ligand	ΔT_m (°C) ^a at 1 μM ligand concentration						
	FHPV9T	FHPV16T	FHPV18T	FHPV32T	FHPV52T	FHPV57T	FHPV58T
PhenDC3	29.2 $\pm$ 1.3	21.0 $\pm$ 2.9	> 30	26.8 $\pm$ 2.6	> 30	24.9 $\pm$ 1.9	> 30
[16]phenN₂	4.1 $\pm$ 0.2	4.5 $\pm$ 0.8	4.3 $\pm$ 0.9	8.5 $\pm$ 0.7	3.1 $\pm$ 0.4	7.7 $\pm$ 0.4	5.0 $\pm$ 0.6
BRACO-19	14.8 $\pm$ 1.3	10.8 $\pm$ 0.5	21.9 $\pm$ 2.8	26.5 $\pm$ 0.5	28.4 $\pm$ 0.4	13.8 $\pm$ 1.1	> 30
C₈	2.9 $\pm$ 0.6	0.1 $\pm$ 0.2	1.4 $\pm$ 0.3	19.5 $\pm$ 0.6	11.9 $\pm$ 2.2	4.1 $\pm$ 0.5	15.6 $\pm$ 1.5
PDS	16.1 $\pm$ 2.6	10.8 $\pm$ 0.9	24.4 $\pm$ 2.3	26.3 $\pm$ 1.3	29.9 $\pm$ 2.1	18.3 $\pm$ 2.8	> 30
360A	10.0 $\pm$ 1.5	4.6 $\pm$ 1.0	20.3 $\pm$ 2.7	23.9 $\pm$ 1.5	> 30	9.5 $\pm$ 1.1	> 30
TMPyP4	> 30	> 30	> 30	> 30	> 30	> 30	> 30

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_m$ (DNA + ligand) – T_m (DNA)]. The buffer used was 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl. The T_m values for the HPV G4 DNAs are 52.9 $\pm$ 0.5 °C for FHPV9T, 52.8 $\pm$ 0.2 °C for FHPV16T, 39.0 $\pm$ 0.4 °C for FHPV18T, 53.7 $\pm$ 0.9 °C for FHPV32T, 49.9 $\pm$ 0.5 °C for FHPV52T, 52.3 $\pm$ 0.5 °C for FHPV57T and 50.3 $\pm$ 0.4 °C for FHPV58T. All experiments were done a minimum of three times, and the values reported correspond to the average of three measurements with the estimated standard deviation given as the error.

Next, we used CD spectroscopy to evaluate the conformations of HPV G4s and the effect of ligand binding on the HPV G4 structures. We first evaluated the folding of the different HPV G4s at different KCl concentrations, specifically 10 and 50 mM and potassium physiological conditions, 140 mM. HPV9, HPV18, HPV32, HPV57 fold into parallel G4s, and HPV52 and HPV58 fold into hybrid (3+1) G4s while HPV16 folds into non-canonical

multiple DNA structures (Figure S2, S3), as previously described [2,16,17]. As seen by the characteristic CD bands of parallel (≈ 264 nm max, ≈ 245 nm min), antiparallel (≈ 295 max, ≈ 260 min) or hybrid-type topologies (≈ 290 max, ≈ 260 max, ≈ 245 min) [33] all G4s seem to adopt the same fold in 10 and 50 mM KCl, with minor differences in ellipticity due to changes in ionic strength. Furthermore, the CD experiments under potassium physiological conditions suggest that these may be the biologically relevant folding conformations as the overall signature of the CD spectra remains unchanged at 140 mM KCl. We then titrated a 10 μ M solution of each HPV sequence in a potassium-containing buffer with increasing amounts of ligands (**Figures 7.2** and S4-10). The CD spectra of all HPV G4s showed to be very polymorphic upon ligand binding, except for HPV18 G4 (**Figure 7.2B**). Changes in the characteristic signatures of HPV G4s were found, namely, the conversion of HPV9 parallel G4 into either a hybrid-type topology in the presence of 5 molar eq. of C₈ and TMPyP4 (Figures S7A and S10A), or into an antiparallel topology in the case of 360A and PhenDC3 (Figures S9A and S4A) at 2 molar eq. of ligand. Interestingly, spectral changes occurred in the spectra of HPV16 upon ligand binding, as the characteristic bands of antiparallel G4 in presence of 2 eq. of PhenDC3 and 360A (Figures 2A and S9B) or hybrid-type in presence of 5 eq. of C₈ (Figure S1B) were observed.

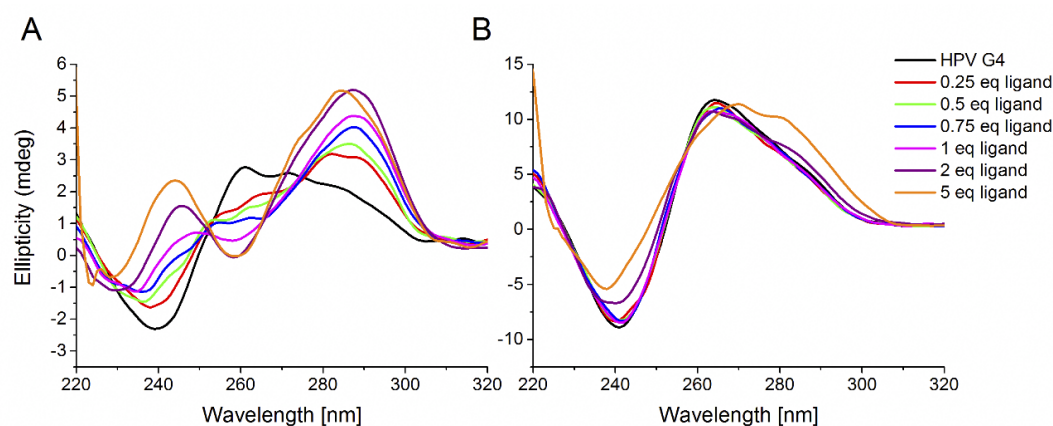


Figure 7.2. - CD titrations of (A) HPV16 and (B) HPV18 G4s with PhenDC3. Upon ligand binding, PhenDC3 seems to convert the multiple, noncanonical structures of HPV16 into an antiparallel G4 while in the case of HPV18, the same overall parallel G4 topology is maintained. The buffer used was 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl.

The CD data indicated that HPV16 G4 structures were favoured over hairpin upon ligand binding. Additionally, HPV52 and HPV57 were converted into antiparallel G4s upon BRACO-19 and 360A binding (Figures S4, S6 and S9). Titration of hybrid-type HPV58 G4 with 360A and TMPyP4 showed a change in topology to parallel G4 (Figures S9G and S10G) while in the case of PhenDC3, BRACO-19 and PDS a different hybrid-type topology similar to that of human telomeric G4s in K⁺ seems preferred [33] (Figures S4G, S6G and

S8G). Details of the CD structural changes and topological characteristic bands are provided in the Supporting Information in Table S2 (**Appendix D**). In addition, CD melting curves were acquired in the presence of the different ligands (Figure S11). The results summarized in **Table 7.3** indicate that PhenDC3 strongly stabilized most of the HPV G4s while [16]phenN₂ is the least stabilizing, in qualitative agreement with the FRET-melting results. High-risk HPVs16 and 18 quadruplexes are strongly stabilized by PhenDC3 which may be due to the formation of stable complexes as further demonstrated by NMR experiments (please see below for discussion). The remaining ligands showed a modest increase in the melting temperature of these particular HPV G4s. Regarding HPV32, the fitting of the data was not possible for most ligands due to the absence of a baseline at high temperatures; nonetheless, the observation of the CD melting curves (Figure S11D) suggests that most ligands stabilize the structure by more than 20 °C, in agreement with FRET melting results. The CD data obtained for HPV9 and HPV52 are in strong agreement with the FRET melting results as [16]phenN₂ and C₈ are the least stabilizing for both sequences. TMPyP4, despite showing the ability to stabilize all structures in FRET melting experiments, failed to replicate the CD melting results. The analysis of the melting curves profile suggests that aggregation and/or precipitation may be occurring. The differences observed may be due to the use of unlabeled sequences in CD experiments, unlike the 5'-FAM and 3'-TAMRA labelled oligos used in FRET, and that the ligands may interact with the FRET dyes.

Table 7.3 - Ligand-induced thermal stabilization (ΔT_m) measured by CD melting experiments.

	ΔT_m (°C) ^a						
Ligand	HPV9	HPV16	HPV18	HPV32	HPV52	HPV57	HPV58
PhenDC3	> 30	>30	>30	n.d.	>30	> 30	>30
[16]phenN₂	4.2 ± 1.1	2.0 ± 0.7	4.8 ± 0.3	5.0 ± 1.2	7.9 ± 0.5	0.0 ± 0.0	5.3 ± 0.3
BRACO-19	24.0 ± 1.1	0.0 ± 0.0	9.8 ± 0.6	n.d.	15.3 ± 0.4	1.8 ± 0.9	2.2 ± 0.5
C₈	15.5 ± 1.1	3.0 ± 1.7	5.8 ± 0.1	14.7 ± 1.2	4.7 ± 0.8	4.6 ± 0.4	10.1 ± 0.4
PDS	> 30	16.0 ± 0.9	9.0 ± 0.8	n.d.	15.0 ± 1.6	11.2 ± 0.8	19.6 ± 0.3 ^c
360A	14.6 ± 1.0	4.1 ± 1.7	6.6 ± 0.7	n.d.	13.9 ± 0.6	5.6 ± 0.7	13.2 ± 0.5
TMPyP4	4.7 ± 0.2	n.d. ^b	n.d. ^b	n.d. ^b	n.d. ^b	n.d. ^b	5.6 ± 3.3

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_m(\text{DNA} + 5 \text{ eq. of ligand}) - T_m(\text{DNA})$]. The buffer used was 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl. The T_m values for the HPV DNAs are 42.4 ± 1.7 °C for HPV9, 55.0 ± 2.2 °C for HPV16, 45.1 ± 2.5 °C for HPV18, 50.9 ± 3.5 °C for HPV32, 45.5 ± 0.5 °C for HPV52, 61.7 ± 0.8 °C for HPV57 and 43.0 ± 0.3 °C for HPV58. Values are reported as $\Delta T_m \pm \text{SE}$.

^b not determined, fitting of normalized data was not possible.

^c value determined at 2 molar eq. as no CD signal was detected at 5 eq.

In order to evaluate the binding strength between the ligands and the HPV G4s, a FID assay was performed using TO as a probe (**Figure 7.3**). TO is a well-described probe for screening G4-binding ligands [35]. The oligos were first tested for their suitability to be used in FID assay, namely by assessing the fluorescence enhancement of TO in the presence of the DNAs. At 2 equivalents of TO relatively to DNA, the observed fluorescence intensity was >2500 a.u. in all cases, which is considered acceptable for the accurate determination of the displacement efficiency [35]. In agreement with FRET-melting results, PhenDC3, BRACO-19, 360A and TMPyP4, presented the higher affinity towards HPV sequences with TO displacement percentages above 50% at 5 molar equivalents.

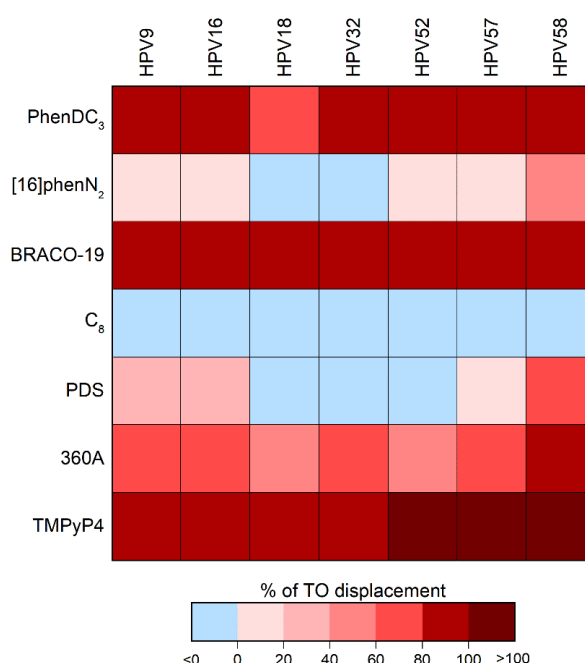


Figure 7.3 - FID results obtained for HPV G4 sequences in the presence of 5 equivalents of ligands. The heat map representation allows the visual comparison of the TO displacement percentages by using the colour code guide provided below. The buffer used was 10 mM K-phosphate buffer, pH 7.2, with 50 mM KCl.

For ligand C₈, despite FRET results suggested its ability to stabilize the HPV G4s, the obtained TO displacement values were below 0% for all G4s. This owes to the fact that C₈ fluorescence is enhanced upon interacting with the HPV G4s and that its emission spectrum is similar to that of TO, thus impairing the determination of the correct displacement values. Among other compounds, [16]phenN₂ promoted the lowest TO displacement, also in agreement with the thermal stabilization studies. PDS, despite the high ΔT_m values observed, performed poorly in the TO displacement assays. This may be attributed to the ligand binding mode being different from that of TO, therefore resulting in an indirect competition, or the formation of a ternary complex thus not displacing TO [35]. Furthermore, ligand binding may promote the rigidification/conformational change

of the G4 structure, thus altering TO properties [36]. To further elucidate this assumption, we acquired CD spectra of the oligos with 2 molar equivalents of TO, followed by the addition of PDS. The results presented in Figure S12 showed that the addition of PDS promotes changes in the CD bands of the HPV G4s. Furthermore, TO itself seems to change the G4 folding, particularly for HPV16 and HPV57 (Figures S8B and S8F). In contrast, TO apparent displacement above 100% was observed in the FID assay for TMPyP4 in the presence of HPV52, 57 and 58 (**Figure 7.3**). This may be explained by the significant absorbance of TMPyP4 between 480 and 540 nm, which promotes a strong decrease of TO fluorescence due to an inner filter effect interfering with the probe's absorption [36]. This effect was minimized by using 485 nm as TO excitation wavelength rather than 501 nm in the case of the other HPV G4s but remained a hurdle to obtain true values for HPV52, 57 and 58.

^1H NMR titrations were then performed to study the effect of ligand binding on the topology of the HPV G4s, particularly high-risk genotypes HPV16, HPV18, HPV52 and HPV58. The analysis of the guanines H1 imino signals (typically between 10.5 and 12.5 ppm) in the presence of a ligand may help to understand the structural arrangement of the interaction between the molecules [37,38]. HPV9 exhibits a set of 12 sharp imino signals indicative of a single G4 structure in solution (Figure S13). Upon ligand titration the imino signals of the G4 structure were found to broaden and shift upfield, which is indicative of ligand binding, particularly through end-stacking [39,40]. At higher ligand:DNA ratios, a significant coalescence of the imino peaks was observed due to the formation of higher order complexes with shared ligand molecules with intermediate exchange in the NMR time scale [38]. Similar behaviour was observed for all ligands, except for TMPyP4 whose spectrum shows an almost complete loss of the imino signals. Regarding HPV16, we investigated if the ligands could favour the formation of G4 structures over hairpins, as indicated by CD titrations. The ^1H NMR spectrum of HPV16 confirmed the existence of multiple non-canonical DNA structures as seen by the existence of signals between 10.5 and 12 ppm, characteristic of G4 structures, and signals between 12 and 14 ppm, characteristic of Watson-Crick base pairing due to the formation of a hairpin structure [2,38]. The analysis of Figure S14 revealed that upon addition of increasing amounts of ligand, the signals corresponding to the hairpin structure started to disappear in the case of PhenDC3, C₈, and 360A, while the G4 imino signals were significantly shifted upfield, as expected for end-stacking binding interactions. Particularly for PhenDC3 (**Figure 7.4A**), the hairpin signals completely disappeared at a 2:1 ligand:DNA ratio and the G4 folding seemed to prevail. Furthermore, despite the multiple imino signals observed, a major set of 12 well-distinguished imino peaks (highlighted by arrows) were identifiable,

possibly belonging to the 12 guanines involved in the formation of the antiparallel G4 conformation seen by CD experiments (**Figure 7.2**).

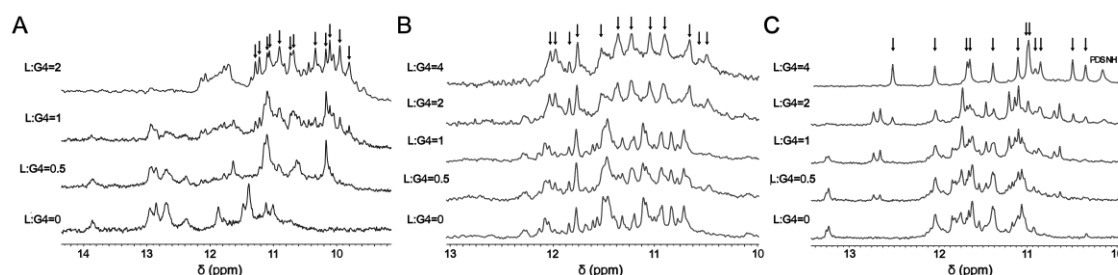


Figure 7.4 - ^1H NMR titrations of three different HPV G4s with increasing amounts of ligands. (A) HPV16 G4, (B) HPV57 G4 and (C) HPV58 G4 with increasing ligand:DNA ratios (L:G4) of PhenDC3, C₈ and PDS, respectively. Arrows indicate the set of imino peaks corresponding to the major G4 conformation. The buffer used was 10 mM K-phosphate buffer, pH 7.2, with 50 mM KCl.

The ^1H NMR titration spectra of HPV18 are shown in Figure S15. In the experimental conditions used, HPV18 showed a well-defined set of 6 imino signals (guanines G36-G42) consistent with a bimolecular G4 structure [2]. Upon addition of ligands, broadening and upfield shift of the signals were observed, characteristic of ligand binding. In the case of PDS, 360A and TMPyP4 a single broad peak was observed at around 10.5-11 ppm which might indicate aggregation or a binding process with kinetics on the NMR time scale. C₈ and PhenDC3 promoted major changes in the imino region with the overall signals shifted upfield, denoting strong interaction.

The ^1H NMR spectrum of HPV32 G4 is consistent with the reported propensity to form aggregates [16] as indicated by the broad signal at around 10.7 ppm (Figure S16). The most interesting result was obtained with PhenDC3 which seemed to be able to resolve the aggregates with the appearance of several imino signals, possibly belonging to multiple G4 conformations.

HPV52 gave a single set of well-resolved 12 signals corresponding to a single G4 conformation. This G4 structure was recently determined by NMR spectroscopy and was shown to adopt a three G-quartet snap back (3 + 1) type scaffold with two edgewise loops, a no-residue V loop and a propeller-type loop [17]. As shown in Figure S17, and unlike the other HPV G4 structures described here, the majority of the imino signals remained well resolved after ligand titration, with moderate shifts, possibly indicating the formation of a single stable complex in solution. The only exception was PhenDC3, where the imino

signals were almost undetectable possibly due to aggregation, which was already suggested by the intense decrease in ellipticity observed in CD experiments (Figure S4E).

HPV57 folds into multiple G4 conformations as suggested by the multiple imino signals observed in the ^1H NMR spectrum (Figure S18). Upon ligand titrations, the overall signals shifted upfield for the majority of ligands. Interestingly, at a 4:1 ligand:DNA ratio, **C₈** promoted the appearance of a major set of 12 well-defined peaks (**Figure 7.4B**) suggesting the formation of a single conformer in opposition to the multiple G4s in the free state [38]. This outcome agrees with the results obtained in CD titrations where a significant variation of the CD bands was observed (Figure S7F), with the appearance of the characteristic CD profile of an antiparallel G4 structure opposingly to the initial typical parallel CD bands.

Finally, the HPV58 imino region was composed of multiple sets of signals and an additional GC base-pair signal can be observed at ≈ 13.1 ppm, showing the existence of multiple G4 structures in solution. Ligand titration had a modest effect on the imino signals for the majority of ligands as observed in Figure S19, as these compounds induced small shifts. Moreover, 360A seems to promote DNA aggregation as suggested by the broad unresolved signal between 11-12 ppm, or the existence of a dynamic binding process in the NMR time scale. The exceptions were PhenDC3 and PDS for which sharp and defined sets of 12 imino signals were observed at a 4:1 ligand:DNA ratio (**Figure 7.4C**) indicating the formation of a single complex formed by HPV58 G4 and each ligand [38].

Finally, we evaluated the antiviral effect of compounds PhenDC3 and **C₈** in organotypic (raft) epithelial cultures infected with episomal high-risk HPVs 16 and 18 [29,30]. Ligands were fed to the cells every other day throughout the raft culture growth and differentiation period (20 days) using DMSO as vehicle control. Given the long incubation period, ligands were used at concentrations that were not toxic to the cells *per se*, allowing the evaluation of the antiviral effects. As observed in **Figures 7.5** and S20, PhenDC3 does not seem to have any effect on infected tissue growth, in high-risk HPVs 16 or 18 rafts, even at the highest concentration tested (0.5 μM). The total number of HPV18 episomal genome copies and viral particle titer, *i.e.*, encapsidated viral genomes resistant to benzonase treatment, were then quantified by qPCR. Benzonase was used to remove un-encapsidated viral genomes thus allowing the evaluation of the compound's effect on genome encapsidation [30]. As indicated by the graphs in **Figure 7.6**, PhenDC3 had no effect on the total genome copies (**Figure 7.6A**), nor on viral encapsidation as indicated by the viral titer quantified after removing non-protected HPV genomes with benzonase which

degrades all forms of DNA and RNA due to its endonuclease activity (**Figure 7.6B**). These results indicate that, despite having a potent stabilizing effect on the HPV G4 structures in vitro as shown by CD and FRET results, the compound failed to reduce viral replication and modulate protein expression as it was able to do in other viruses such as EBV [10]. This may be attributed to a lack of selectivity of the compound towards HPV G4s as PhenDC3 is a potent G4 ligand that was already shown to strongly bind human promoter and telomeric G4 structures. Indeed, a study by A. de Rache *et al.* showed that PhenDC3 presents a selectivity profile towards telomeric G4 structures over other G4s such as myc and kit promoters and TBA thrombin binding aptamer [41]. In a cellular environment, *i.e.* the HPV-infected organotypic raft cultures, one may hypothesize that PhenDC3 preferentially binds the cell's G4 structures rather than the HPV genome thus failing to inhibit viral replication

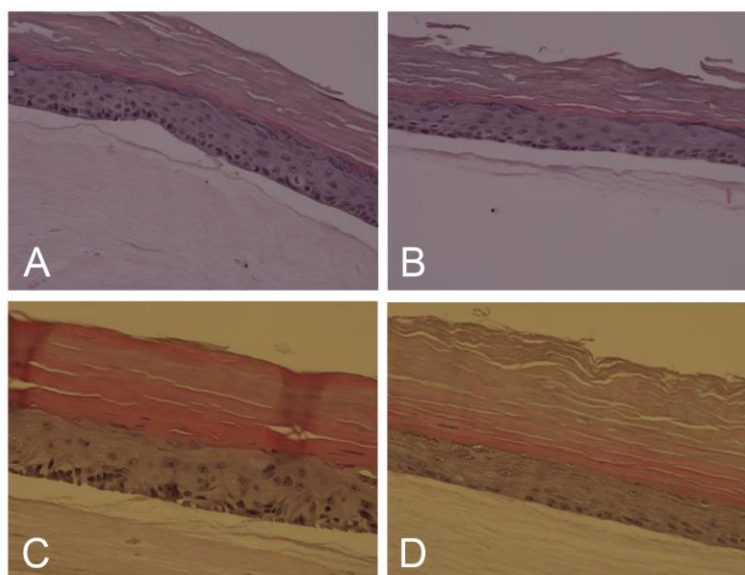


Figure 7.5 - Effects of ligand PhenDC3 on HCK16-5 and HCK18C organotypic raft cultures, which contain episomal high-risk HPV16 and HPV18 respectively. HCK16-5 rafts were either treated with (A) DMSO solvent control or (B) 0.5 μ M PhenDC3. Similarly, HCK18C rafts were treated with (C) DMSO solvent control or (D) 0.5 μ M PhenDC3. Treatments were made every other day with feeding for 20 days. Tissue sections were cut and stained with hematoxylin and eosin (H&E).

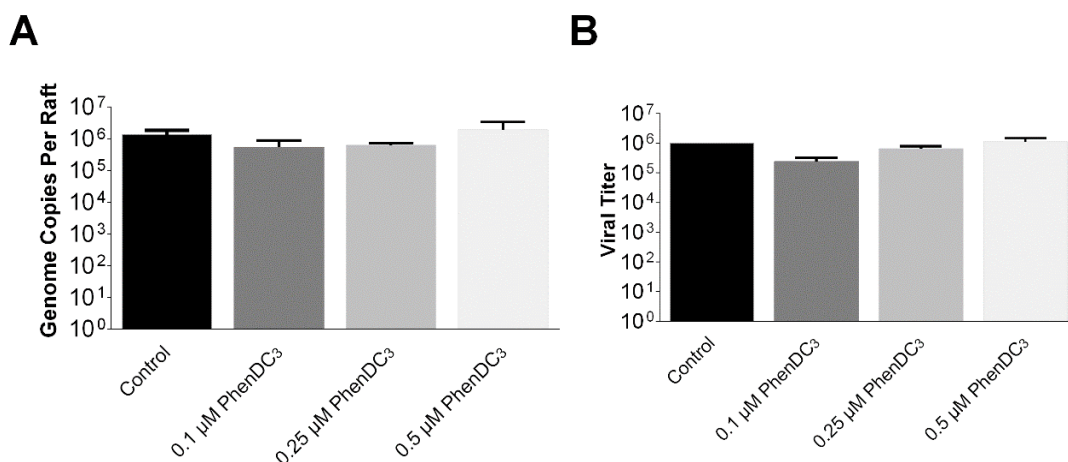


Figure 7.6 - Quantitative RT-PCR infectivity assay of HPV18 organotypic rafts incubated with PhenDC3. (A) total viral genomes and (B) viral titer. The total number of viral genomes were quantified amplifying the E2 open reading frame and measured against a standard curve. Virus titers were measured by qPCR, in which homogenized samples were treated with benzonase to remove any non-protected HPV genomes. Data represent mean $\pm$ SD.

Ligand C₈ was also tested in the rafts as we already demonstrated its significant cell growth inhibitory effect on HPV-positive cell lines such as HeLa cervical cancer cells [23]. As shown in **Figures 7.7** and S21, treatment with increasing concentrations of C₈ did not affect the organotypic raft growth of HPV16 (**Figure 7.7B**), while in the case of HPV18 rafts there is a pronounced thinning of the raft when compared to the control (**Figure 7.7D**) at 0.1 μ M C₈. At higher C₈ concentrations we were not able to retrieve enough tissue for correct histological analysis, probably due to general toxicity of the compound. Similar effects have already been reported for some strong HPV18 inhibitors with other antiviral mechanisms that do not involve G4 binding [40,41].

To assess if the growth inhibition was due to general toxicity of C₈ or to HPV-dependent effects, we quantified the total viral genomes of the tissues. The results plotted in **Figure 7.8A** show a concentration-dependent decreased number of HPV18 genome copies, from 4.1×10^5 to 1.1×10^5 genomes/ μ L at 0.1 μ M of C₈, and to 7.6×10^3 genomes/ μ L at 0.25 μ M compound concentration. The viral titers after benzonase-treatment, *i.e.*, the effect of C₈ on encapsidation, were also determined (**Figure 7.8B**). The exposure of cervical cells to C₈ concentration of 0.25 μ M resulted in a >100-fold decrease in HPV18 viral titer (2.2×10^5 to 5.0×10^2). This result suggests that while C₈ has a modest effect on viral genome amplification (as seen by total genome copies) it plays an important role in viral genome encapsidation. Indeed, the tested HPV18 G4 sequence is found in L2 region which codes for capsid proteins, suggesting that G4-binding may contribute to this reduction in the viral titer. Although biophysical studies suggest modest stabilization of the HPV18 G4 by

C₈, one may not exclude the involvement of multiple G4-mediated mechanisms in the observed results, as suggested for other compounds in the literature such as PDS and PhenDC3 [10]. As an example, PhenDC3 was shown to modulate the synthesis of an EBV genome maintenance protein called EBNA1 by interfering with the interaction of nucleolin and the virus mRNAs [42]. In that study, nucleolin, a G4 interacting protein (which was also described as an activator of HPV18 oncogene transcription in an unrelated study [43]), was prevented from interacting with EBNA1 mRNA hence having antiviral effect by interfering with the mechanism which allows EBV to evade the immune system [42]. Thus, similar mechanisms of antiviral action can be hypothesized for C₈ and HPV18. Notwithstanding C₈ poor selectivity towards HPV18 G4 sequence, most of the G4 ligands found in the literature that were tested in viruses have demonstrated promising antiviral activity, despite having been described to stabilize other G4 structures to a higher extent. In the majority of cases, the antiviral activity was shown to be G4-dependent. A possible explanation for this outcome is the amount of viral G4s, which during infection exceed the cellular G4s, especially during replication, as reported for HSV-1 [46]. Considering that the HPV G4s are located in regulatory regions crucial for the viral life cycle, a combination of the abundance of viral G4s and impairment of key viral functions might explain the observed antiviral effect of C₈. Future studies should be aimed at establishing a direct correlation between G4 stabilization and antiviral effect as it was not clearly demonstrated in this work. Acridine derivatives have been already demonstrated to possess antiviral activity against a variety of RNA and DNA virus [47-50] by other G4-independent mechanisms, mainly by DNA intercalation and interactions with enzymes [49], which may also be ascribed to the effects observed for C₈. Regarding HPV16, PhenDC3 and C₈ showed no significant effect both on total genome copies and viral titer (data not shown).

Altogether these results point for the feasibility and viability of developing G4-targeted antiviral drugs with clinical applicability, which may represent a landmark in the management of viral diseases, particularly in the case of latent infections which is characteristic of HPV. Currently, there is no effective treatment for HPV persistence [51]. The ideal HPV treatment would be able to clear HPV infection even when the virus is present in its latent form. Current therapies which are mainly targeted at viral proteins fail to remove the viral genomes from the host. Ligands that selectively target the G4s in the viral genome have the potential of removing not only the replicating virus but also the latent virus, therefore preventing the development of cervical cancer and other HPV-induced cancers. Recent works have demonstrated that latently infected cells are more susceptible to long-term exposure to G4 ligands [52]. The results obtained herein agree with the aforementioned observations as the organotypic raft cultures were treated for 20

days with C₈, which may be considered a long-term exposure, presenting a large reduction of the viral genomes. It is therefore feasible that G4-targeted therapeutics could be used to clear both productive and latent HPV infection. Furthermore, by inhibiting HPV DNA replication, this strategy could suppress HPV integration events, which would presumably lead to decreased HPV-induced cervical cancers.

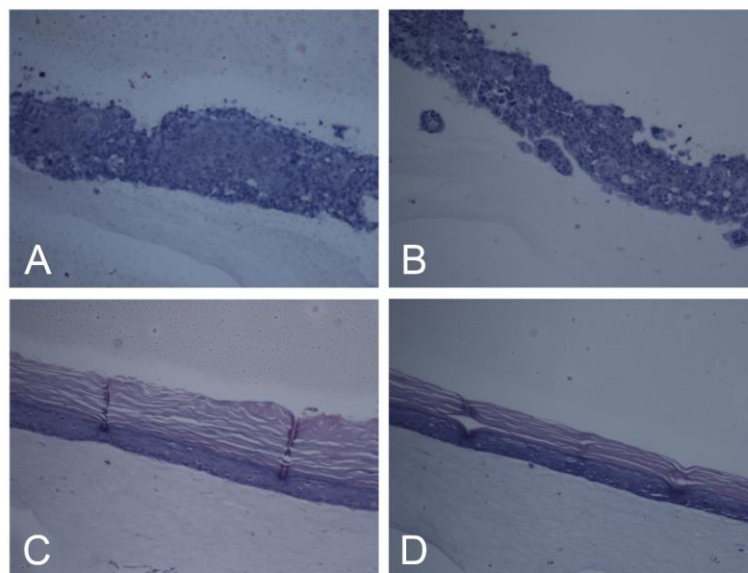


Figure 7.7 - Effects of compound C₈ on HCK16-5 and HCK18C organotypic raft cultures. HCK16-5 rafts were either treated with (A) DMSO solvent control or (B) 0.1 μ M C₈. Similarly, HCK18C rafts were treated with (C) DMSO solvent control or (D) 0.1 μ M C₈. Treatments were made every other day with feeding for 20 days. Tissue sections were cut and stained with hematoxylin and eosin (H&E). Of note is the raft thinning upon incubation with C₈ in the case of HPV18 infected cultures.

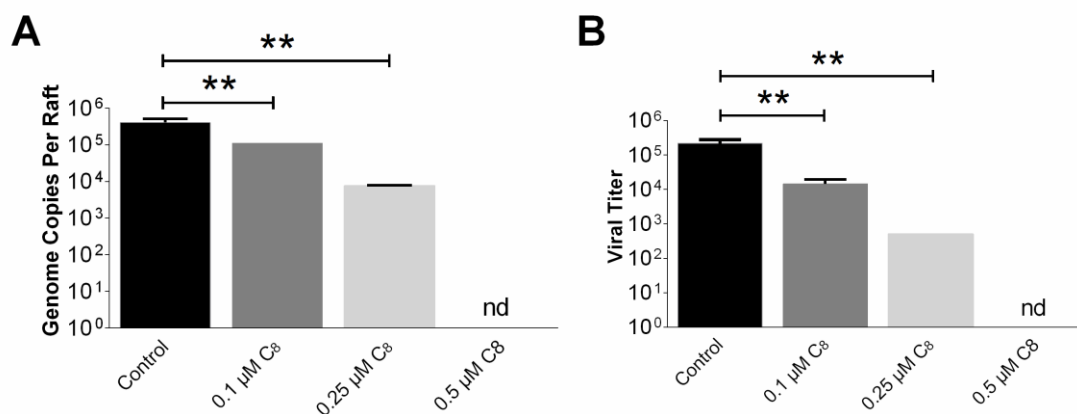


Figure 7.8 - Quantitative RT-PCR infectivity assay of HPV18 organotypic rafts incubated with C₈. (A) total viral genomes and (B) viral titer. The total number of viral genomes were quantified amplifying the E2 open reading frame and measured against a standard curve. Virus titers were measured by qPCR, in which homogenized samples were treated with benzonase to remove any non-protected HPV genomes. No tissue was recovered at 0.5 μ M compound concentration. Data represent mean $\pm$ SD; ** p < 0.01.

Conclusion

Our results demonstrate for the first time that several compounds bind to high- and low-risk HPV G4 structures using a variety of different techniques. Ligand binding promotes a high degree of conformational polymorphism for all G4 structures except those of HPV18 which remained stable. These conformational changes may have an impact on the recognition of these G4 motifs by proteins during the establishment and/or progression of the viral infection. Additionally, most of these G4 structures were found within regions involved in transcription, replication and viral proteins production. Therefore, the stabilization of the G4 structures within these regions may hinder these processes, thus having an antiviral effect. The antiviral effect was confirmed by testing compound C₈ on organotypic raft cultures containing replicating HPV18 and the compound was able to decrease the viral load by several orders of magnitude. Also interesting was the fact that the compounds were able to favour the formation of a quadruplex *vs* hairpin structures in the case of HPV16, as well as the formation of single G4 structures in polymorphic HPVs G4 such as HPV57 and HPV58. Altogether, we provided new insights into G4 possible roles in HPV replication which may represent a significant turning point in the development of HPV G4-mediated antiviral drugs.

Acknowledgments

J. Carvalho acknowledges a doctoral fellowship grant from the FCT – Foundation for Science and Technology ref. SFRH/BD/122953/2016. Jéssica Lopes-Nunes acknowledges the fellowship reference UTAP-EXPL/NTec/0015/2017. This work was supported by project “Ações Integradas Luso-Francesas” ref. TC-15/17, “Cooperação Científica e Tecnológica FCT/Acordo Pessoa” project ref. 5079, FCT project ref. IF/00959/2015 financed by Fundo Social Europeu and Programa Operacional Potencial Humano, FCT project ref. UID/Multi/04349/2013, MIT Portugal project BIODEVICE ref. MIT-EXPL/BIO/0008/2017 and UTAustin FCT project DREAM ref. UTAP-EXPL/NTec/0015/2017. The work was done on NMR spectrometers of CICS-UBI that are part of Rede Nacional de RMN (PTNMR), supported by FCT-MCTES (ROTEIRO/0031/2013 - PINFRA/22161/2016) co-funded by FEDER through COMPETE 2020, POCI, and PORL and FCT through PIDDAC. This work was also supported by the SYMBIT project reg. no. CZ.02.1.01/0.0/0.0/15_003/0000477 financed from the ERDF.

References

- [1] Schiffman, M., Doorbar, J., Wentzensen, N., De Sanjosé, S., Fakhry, C., Monk, B. J., Stanley, M. A., and Franceschi, S. (2016) Carcinogenic human papillomavirus infection. *Nature Reviews Disease Primers* 2, 16086. <https://doi.org/10.1038/nrdp.2016.86>.
- [2] Marušič, M., Hošnjak, L., Krafčikova, P., Poljak, M., Viglasky, V., and Plavec, J. (2017) The effect of single nucleotide polymorphisms in G-rich regions of high-risk human papillomaviruses on structural diversity of DNA. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1861, 1229–1236. <https://doi.org/10.1016/j.bbagen.2016.11.007>.
- [3] Edwards, T. G., Koeller, K. J., Slomczynska, U., Fok, K., Helmus, M., Bashkin, J. K., and Fisher, C. (2011) HPV episome levels are potently decreased by pyrrole–imidazole polyamides. *Antiviral Research* 91, 177–186. <https://doi.org/10.1016/j.antiviral.2011.05.014>.
- [4] Biryukov, J., and Meyers, C. (2015) Papillomavirus infectious pathways: A comparison of systems. *Viruses* 7, 4303–4325. <https://doi.org/10.3390/v7082823>.
- [5] Asamitsu, S., Obata, S., Yu, Z., Bando, T., and Sugiyama, H. (2019) Recent progress of targeted G-quadruplex-preferred ligands toward cancer therapy. *Molecules* 24, 429. <https://doi.org/10.3390/molecules24030429>.
- [6] Neidle, S. (2016) Quadruplex Nucleic Acids as Novel Therapeutic Targets. *Journal of Medicinal Chemistry* 59, 5987–6011. <https://doi.org/10.1021/acs.jmedchem.5b01835>.
- [7] González, V., Guo, K., Hurley, L., and Sun, D. (2009) Identification and Characterization of Nucleolin as a c- myc G-quadruplex-binding Protein. *Journal of Biological Chemistry* 284, 23622–23635. <https://doi.org/10.1074/jbc.M109.018028>.
- [8] Tosoni, E., Frasson, I., Scalabrin, M., Perrone, R., Butovskaya, E., Nadai, M., Palič¹/₂, G., Fabris, D., and Richter, S. N. (2015) Nucleolin stabilizes G-quadruplex structures folded by the LTR promoter and silences HIV-1 viral transcription. *Nucleic Acids Research* 43, 8884–8897. <https://doi.org/10.1093/nar/gkv897>.
- [9] Butovskaya, E., Soldà, P., Scalabrin, M., Nadai, M., and Richter, S. N. (2019) HIV-1 Nucleocapsid Protein Unfolds Stable RNA G-Quadruplexes in the Viral Genome and Is Inhibited by G-Quadruplex Ligands. *ACS Infectious Diseases* 5, 2127–2135. <https://doi.org/10.1021/acsinfecdis.9b00272>.
- [10] Ruggiero, E., and Richter, S. N. (2018) G-quadruplexes and G-quadruplex ligands: targets and tools in antiviral therapy. *Nucleic Acids Research* 46, 3270–3283.

<https://doi.org/10.1093/nar/gky187>.

[11] Metifiot, M., Amrane, S., Litvak, S., and Andreola, M.-L. (2014) G-quadruplexes in viruses: function and potential therapeutic applications. *Nucleic Acids Research* 42, 12352–12366. <https://doi.org/10.1093/nar/gku999>.

[12] Fleming, A. M., Ding, Y., Alenko, A., and Burrows, C. J. (2016) Zika Virus Genomic RNA Possesses Conserved G-Quadruplexes Characteristic of the Flaviviridae Family. *ACS Infectious Diseases* 2, 674–681. <https://doi.org/10.1021/acsinfecdis.6b00109>.

[13] Ruggiero, E., Tassinari, M., Perrone, R., Nadai, M., and Richter, S. N. (2019) Stable and Conserved G-Quadruplexes in the Long Terminal Repeat Promoter of Retroviruses. *ACS Infectious Diseases*. <https://doi.org/10.1021/acsinfecdis.9b00011>.

[14] Biswas, B., Kumari, P., and Vivekanandan, P. (2018) Pac1 Signals of Human Herpesviruses Contain a Highly Conserved G-Quadruplex Motif. *ACS Infectious Diseases* 4, 744–751. <https://doi.org/10.1021/acsinfecdis.7b00279>.

[15] Madireddy, A., Purushothaman, P., Loosbroock, C. P., Robertson, E. S., Schildkraut, C. L., and Verma, S. C. (2016) G-quadruplex-interacting compounds alter latent DNA replication and episomal persistence of KSHV. *Nucleic Acids Research* 44, 3675–3694. <https://doi.org/10.1093/nar/gkw038>.

[16] Tlučková, K., Marušič, M., Tóthová, P., Bauer, L., Šket, P., Plavec, J., and Viglasky, V. (2013) Human papillomavirus G-quadruplexes. *Biochemistry* 52, 7207–7216. <https://doi.org/10.1021/bi400897g>.

[17] Marušič, M., and Plavec, J. (2019) Towards Understanding of Polymorphism of the G-rich Region of Human Papillomavirus Type 52. *Molecules* 24, 1294. <https://doi.org/10.3390/molecules24071294>.

[18] Harris, L. M., and Merrick, C. J. (2015) G-Quadruplexes in Pathogens: A Common Route to Virulence Control? *PLoS Pathogens* 11, 1–15. <https://doi.org/10.1371/journal.ppat.1004562>.

[19] Shanmugasundaram, S., and You, J. (2017) Targeting persistent human papillomavirus infection. *Viruses* 9, 229. <https://doi.org/10.3390/v9080229>.

[20] De Cian, A., DeLemos, E., Mergny, J.-L., Teulade-Fichou, M.-P. P., and Monchaud, D. (2007) Highly efficient G-quadruplex recognition by bisquinolinium compounds. *Journal of the American Chemical Society* 129, 1856–1857. <https://doi.org/10.1021/ja067352b>.

- [21] Carvalho, J., Quintela, T., Gueddouda, N. M., Bourdoncle, A., Mergny, J.-L., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2018) Phenanthroline polyazamacrocycles as G-quadruplex DNA binders. *Organic and Biomolecular Chemistry* 16, 2776–2786. <https://doi.org/10.1039/c8ob00247a>.
- [22] Read, M., Harrison, R. J., Romagnoli, B., Tanious, F. A., Gowan, S. H., Reszka, A. P., Wilson, W. D., Kelland, L. R., and Neidle, S. (2001) Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 98, 4844–9. <https://doi.org/10.1073/pnas.081560598>.
- [23] Carvalho, J., Pereira, E., Marquevielle, J., Campello, M. P. C., Mergny, J.-L., Paulo, A., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2018) Fluorescent light-up acridine orange derivatives bind and stabilize KRAS-22RT G-quadruplex. *Biochimie* 144, 144–152. <https://doi.org/10.1016/j.biochi.2017.11.004>.
- [24] Rodriguez, R., Müller, S., Yeoman, J. A., Trentesaux, C., Riou, J.-F., and Balasubramanian, S. (2008) A Novel Small Molecule That Alters Shelterin Integrity and Triggers a DNA-Damage Response at Telomeres. *Journal of the American Chemical Society* 130, 15758–15759. <https://doi.org/10.1021/ja805615w>.
- [25] Granotier, C., Pennarun, G., Riou, L., Hoffschir, F., Gauthier, L. R., De Cian, A., Gomez, D., Mandine, E., Riou, J.-F. F., Mergny, J.-L. L., Mailliet, P., Dutrillaux, B., and Boussin, F. D. (2005) Preferential binding of a G-quadruplex ligand to human chromosome ends. *Nucleic Acids Research* 33, 4182–4190. <https://doi.org/10.1093/nar/gki722>.
- [26] Han, F. X., Wheelhouse, R. T., and Hurley, L. H. (1999) Interactions of TMPyP4 and TMPyP2 with quadruplex DNA. Structural basis for the differential effects on telomerase inhibition. *Journal of the American Chemical Society* 121, 3561–3570. <https://doi.org/10.1021/ja984153m>.
- [27] Renčiuk, D., Zhou, J., Beaurepaire, L., Guédin, A., Bourdoncle, A., and Mergny, J.-L. (2012) A FRET-based screening assay for nucleic acid ligands. *Methods* 57, 122–128. <https://doi.org/10.1016/j.ymeth.2012.03.020>.
- [28] Meyers, C. (1996) Organotypic (raft) epithelial tissue culture system for the differentiation-dependent replication of papillomavirus. *Methods in Cell Science* 18, 201–210. <https://doi.org/10.1007/BF00132885>.

- [29] Conway, M. J., and Meyers, C. (2009) Replication and assembly of human papillomaviruses. *Journal of Dental Research* 88, 307–317. <https://doi.org/10.1177/0022034509333446>.
- [30] Biryukov, J., Cruz, L., Ryndock, E. J., and Meyers, C. (2015) Native human papillomavirus production, quantification, and infectivity analysis, in *Methods in molecular biology* (Clifton, N.J.), pp 317–331. https://doi.org/10.1007/978-1-4939-2013-6_24.
- [31] Bedrat, A., Lacroix, L., and Mergny, J.-L. (2016) Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Research* 44, 1746–1759. <https://doi.org/10.1093/nar/gkw006>.
- [32] Ruan, T. L., Davis, S. J., Powell, B. M., Harbeck, C. P., Habdas, J., Habdas, P., and Yatsunyk, L. A. (2017) Lowering the overall charge on TMPyP4 improves its selectivity for G-quadruplex DNA. *Biochimie* 132, 121–130. <https://doi.org/10.1016/j.biochi.2016.11.003>.
- [33] del Villar-Guerra, R., Trent, J. O., and Chaires, J. B. (2018) G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy. *Angewandte Chemie - International Edition* 57, 7171–7175. <https://doi.org/10.1002/anie.201709184>.
- [34] Bončina, M., Podlipnik, Č., Piantanida, I., Eilmes, J., Teulade-Fichou, M. P., Vesnaver, G., and Lah, J. (2015) Thermodynamic fingerprints of ligand binding to human telomeric G-quadruplexes. *Nucleic Acids Research* 43, 10376–10386. <https://doi.org/10.1093/nar/gkv1167>.
- [35] Tran, P. L. T., Largy, E., Hamon, F., Teulade-Fichou, M.-P., and Mergny, J.-L. (2011) Fluorescence intercalator displacement assay for screening G4 ligands towards a variety of G-quadruplex structures. *Biochimie* 93, 1288–1296. <https://doi.org/10.1016/j.biochi.2011.05.011>.
- [36] Largy, E., Hamon, F., and Teulade-Fichou, M.-P. (2011) Development of a high-throughput G4-FID assay for screening and evaluation of small molecules binding quadruplex nucleic acid structures. *Analytical and Bioanalytical Chemistry* 400, 3419–3427. <https://doi.org/10.1007/s00216-011-5018-z>.
- [37] Webba da Silva, M. (2007) NMR methods for studying quadruplex nucleic acids. *Methods* 43, 264–277. <https://doi.org/10.1016/j.ymeth.2007.05.007>.

- [38] Bessi, I., Wirmer-Bartoschek, J., Dash, J., and Schwalbe, H. (2017) Targeting G-quadruplex with Small Molecules: An NMR View, in *Modern Magnetic Resonance*, pp 1–22. <https://doi.org/10.1007/978-3-319-28275-6>.
- [39] Dai, J., Carver, M., Hurley, L. H., and Yang, D. (2011) Solution structure of a 2:1 quindoline-c-MYC G-quadruplex: Insights into G-quadruplex-interactive small molecule drug design. *Journal of the American Chemical Society* 133, 17673–17680. <https://doi.org/10.1021/ja205646q>.
- [40] Adrian, M., Heddi, B., and Phan, A. T. (2012) NMR spectroscopy of G-quadruplexes. *Methods* 57, 11–24. <https://doi.org/10.1016/j.ymeth.2012.05.003>.
- [41] De Rache, A., and Mergny, J.-L. (2015) Assessment of selectivity of G-quadruplex ligands via an optimised FRET melting assay. *Biochimie* 115, 194–202. <https://doi.org/10.1016/j.biochi.2015.06.002>.
- [42] Drubin, D. A., McLaughlin-Drubin, M. E., Clawson, G. A., and Meyers, C. (2006) A Protease Inhibitor Specifically Inhibits Growth of HPV-Infected Keratinocytes. *Molecular Therapy* 13, 1142–1148. <https://doi.org/10.1016/j.ymthe.2006.02.005>.
- [43] Sanjib Banerjee, N., Moore, D. W., Broker, T. R., and Chow, L. T. (2018) Vorinostat, a pan-HDAC inhibitor, abrogates productive HPV-18 DNA amplification. *Proceedings of the National Academy of Sciences of the United States of America* 115, E11138–E11147. <https://doi.org/10.1073/pnas.1801156115>.
- [44] Lista, M. J., Martins, R. P., Billant, O., Contesse, M. A., Findakly, S., Pochard, P., Daskalogianni, C., Beauvineau, C., Guetta, C., Jamin, C., Teulade-Fichou, M. P., Fåhræus, R., Voisset, C., and Blondel, M. (2017) Nucleolin directly mediates Epstein-Barr virus immune evasion through binding to G-quadruplexes of EBNA1 mRNA. *Nature Communications* 8, 16043. <https://doi.org/10.1038/ncomms16043>.
- [45] Grinstein, E., Wernet, P., Snijders, P. J. F. F., Rösl, F., Weinert, I., Jia, W., Kraft, R., Schewe, C., Schwabe, M., Hauptmann, S., Dietel, M., Meijer, C. J. L. M. L. M., and Royer, H.-D. (2002) Nucleolin as activator of human papillomavirus type 18 oncogene transcription in cervical cancer. *The Journal of Experimental Medicine* 196, 1067–1078. <https://doi.org/10.1084/jem.20011053>.
- [46] Artusi, S., Perrone, R., Lago, S., Raffa, P., Di Iorio, E., Palù, G., and Richter, S. N. (2016) Visualization of DNA G-quadruplexes in herpes simplex virus 1-infected cells. *Nucleic Acids Research* 44, 10343–10353. <https://doi.org/10.1093/nar/gkw968>.

- [47] El-Sabbagh, O. I., and Rady, H. M. (2009) Synthesis of new acridines and hydrazones derived from cyclic β -diketone for cytotoxic and antiviral evaluation. *European Journal of Medicinal Chemistry* 44, 3680–3686. <https://doi.org/10.1016/j.ejmech.2009.04.001>.
- [48] Goodell, J. R., Madhok, A. A., Hiasa, H., and Ferguson, D. M. (2006) Synthesis and evaluation of acridine- and acridone-based anti-herpes agents with topoisomerase activity. *Bioorganic and Medicinal Chemistry* 14, 5467–5480. <https://doi.org/10.1016/j.bmc.2006.04.044>.
- [49] Rupar, J., Dobričić, V., Aleksić, M., Brborić, J., and Čudina, O. (2018) A review of published data on acridine derivatives with different biological activities. *Kragujevac Journal of Science* 83–101. <https://doi.org/10.5937/kgjsci1840083r>.
- [50] Tonelli, M., Vettoretti, G., Tasso, B., Novelli, F., Boido, V., Sparatore, F., Busonera, B., Ouhtit, A., Farci, P., Blois, S., Giliberti, G., and La Colla, P. (2011) Acridine derivatives as anti-BVDV agents. *Antiviral Research* 91, 133–141. <https://doi.org/10.1016/j.antiviral.2011.05.005>.
- [51] Hu, Z., and Ma, D. (2018) The precision prevention and therapy of HPV-related cervical cancer: new concepts and clinical implications. *Cancer Medicine* 7, 5217–5236. <https://doi.org/10.1002/cam4.1501>.
- [52] Piekna-Przybylska, D., Sharma, G., Maggirwar, S. B., and Bambara, R. A. (2017) Deficiency in DNA damage response, a new characteristic of cells infected with latent HIV-1. *Cell Cycle* 16, 968–978. <https://doi.org/10.1080/15384101.2017.1312225>.

Chapter 8

Aptamer-Based Targeted Delivery of a G-Quadruplex Ligand in Cervical Cancer Cells

This chapter was published in:

Carvalho, J., Paiva, A., Campello, M. P. C., Paulo, A., Mergny, J.-L., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2019) Aptamer-based Targeted Delivery of a G-quadruplex Ligand in Cervical Cancer Cells. *Scientific Reports* 9, 7945. <https://doi.org/10.1038/s41598-019-44388-9>.

Aptamer-Based Targeted Delivery of a G-Quadruplex Ligand in Cervical Cancer Cells

Abstract

AS1411 is a G-rich DNA oligonucleotide that functions as an aptamer of the protein nucleolin, found at high levels on the surface of cancer cells but not on the surface of normal cells. Herein, we have studied AS1411 as a supramolecular carrier for the delivery of an acridine-based G4 ligand, C₈, to HeLa cancer cells. Two AS1411 derivatives, LNA-AS1411 and U-AS1411, were also tested, in an attempt to compare AS1411 pharmacological properties. The results showed that AS1411-C₈ complexation was made with great binding strength and that it lowered the ligand's cytotoxicity towards non-malignant cells. This effect was suggested to be due to a decreased internalization of the complexed versus free C₈ as shown by flow cytometry. The AS1411 derivatives, despite forming a stable complex with C₈, lacked the necessary tumour-selective behaviour. The binding of C₈ to AS1411 G4 structure did not negatively affect the recognition of nucleolin by the aptamer. The AS1411-C₈ repressed c-MYC expression at the transcriptional level, possibly due to C₈ ability to stabilize the c-MYC promoter G4s. Overall, this study demonstrates the usefulness of AS1411 as a supramolecular carrier of the G4 binder C₈ and the potential of using its tumour-selective properties for the delivery of ligands for cancer therapy.

Introduction

AS1411 is a synthetic 26-base DNA aptamer previously discovered by Bates and co-workers as an agent targeting the nucleolin protein with high affinity and specificity, eliciting a potent antiproliferative effect on a variety of cell lines [1,2]. Nucleolin is a multifunctional protein playing essential roles in cell survival, growth and proliferation. The protein is localized primarily in the nucleus of normal cells, while on cancer cells it is also present in the cytoplasm and at the cell surface [3]. This property confers a tumour-selective behaviour to AS1411 which targets preferentially the external domain of surface nucleolin of cancer cells. This oligonucleotide reached Phase 2b clinical trials for acute myeloid leukaemia and renal cell carcinoma; however, despite a good tolerance and safety profile, the trial was terminated due to suboptimal pharmacokinetics (rapid clearance) and low potency [2]. The mechanism of the cytotoxicity of AS1411 is still under debate and

multiple nucleolin-dependent and independent biological effects have been described [2]. Recently, methuosis has been proposed as the likely cell death mechanism induced by AS1411, which is characterized by altered cell morphology and hyperstimulation of macropinocytosis combined with perturbed vesicle trafficking [4]. AS1411 is able to fold into a highly polymorphic G4 structure which makes it more stable against the serum nucleases and pH fluctuations, with increased cellular uptake efficacy [1,2].

Regardless of the clinical application of AS1411 being hindered, its tumour-selective properties have increasingly motivated its study as a drug delivery system of therapeutic agents and imaging probes. The variety of applications range from nanoparticle-linked AS1411 formulations to covalent/noncovalent drug conjugates [2]. The diversity of cargoes that have been efficiently delivered by AS1411 extend to G4 ligands as its own structure is a stable G4. Recently, AS1411 was noncovalently bound to porphyrin TMPyP4, in an attempt to increase its delivery to MCF7 breast cancer cells within a photodynamic therapy approach [5]. This strategy was shown to increase the accumulation of TMPyP4 in MCF7 cancer cells when compared to non-malignant cells. Several examples in the literature evidence that attaching AS1411 to a wide variety of nanoparticles or other entities is an excellent strategy for delivering these cargoes inside of cancer cells. Moreover, improving AS1411 stability and biological properties, mainly through chemical modification on deoxynucleotides at appropriate positions in AS1411, could also be a strategy to improve the therapeutic function of the aptamer [2].

Herein, we have evaluated AS1411 and its modified derivatives designated by LNA-AS1411 and U-AS1411 as drug delivery systems to carry an acridine orange-based ligand (C_8) [6] to HeLa cervical cancer cells. The ligand C_8 was previously published and shown to bind and stabilize promoter and telomeric G4s with high affinity ($K_D \approx 10^{-7}$ M). C_8 showed cytotoxic activity against HeLa cells with an IC_{50} value around 1 μ M at 24h incubation; however, similar values were observed for non-malignant cells. Therefore, targeted carriers that promote the selective accumulation of C_8 in cancer cells, while preventing toxicity in healthy surrounding cells, are of crucial importance to improve its pharmacological profile as a potential drug for anticancer therapies. To address this goal, we have studied the formation of supramolecular complexes between C_8 and AS1411, as well as with its derivatives LNA-AS1411 and U-AS1411. Herein, we also describe the physical properties and biological activities of the resulting C_8 -oligonucleotide supramolecular complexes.

Materials and Methods

Oligonucleotides and compounds

AS1411 (5'-GGTGGTGGTGGTTGTGGTGGTGGTGG-3'), LNA-AS1411 (5'-gGTGGTGGTGGTTGTGGTGGTGGTGG-3', where g stands for LNA-modified guanine), U-AS1411 (5'-GGTGGTGGTGGUUGTGGTGGTGGTGG-3', where U stands for uracil base), 5'-FAM-AS1411-TAMRA-3' and Cy5-AS1411 were purchased from Eurogentec (Belgium) with HPLC-grade purification and used without further treatment. Stock solutions of about 1 mM were prepared using Milli-Q water and stored at -20 °C until used. For all experiments, the oligo was annealed by heating to 95 °C for 10 min and cooling down on ice until used. Hoechst 33342 was purchased from Thermo Scientific (USA). Compound 10-(8-(4-iodobenzamide)octyl)-3,6-bis(dimethylamine) acridinium iodide (C₈) was synthesized as previously described [7]. The compound was dissolved in DMSO to obtain a 10 mM stock solution and further diluted in Milli-Q water.

Thermal Difference Spectra

Thermal difference spectra (TDS) were collected on a Thermo Scientific™ Evolution™ 201 UV-Visible Spectrophotometer (Thermo Fisher Scientific, USA). Spectra were obtained in the 220-335 nm range (scan rate of 200 nm/min and 1 nm data intervals) above and below the melting temperature, at 90 °C and 20 °C, respectively [8]. DNA was used at 3 μM concentration in 30 mM phosphate buffer (15 mM KH₂PO₄, 15 mM K₂HPO₄, pH 7.1) containing 100 mM KCl. The TDS spectrum was calculated by subtracting the spectrum at 20 °C from the spectrum at 90 °C. The following UV TDS factors were used for analysis: $\Delta A_{240}/\Delta A_{295}$, $\Delta A_{255}/\Delta A_{295}$, $\Delta A_{275}/\Delta A_{295}$. The data was normalized relative to the maximum absorbance.

Circular dichroism

CD experiments were performed on a Jasco J-815 CD spectrometer equipped with a Peltier-type temperature controller (model CDF-426S/15). Spectra acquisition was performed in 1- or 10-mm quartz cuvettes at a DNA concentration of 10 μM in 30 mM potassium phosphate buffer containing 100 mM KCl. Spectral width was set to 220-340 nm, with a scan speed of 100 nm/min, 1 nm bandwidth, 1 s integration time over 4 averaged accumulations. During titrations, the required volume of ligand solution was added to the quartz cell.

For CD melting studies, buffer conditions were adjusted so that the oligos would melt around 50 °C, particularly by using 10 mM lithium cacodylate buffer, pH 7.2, containing 10 mM KCl and 90 mM LiCl. Melting curves were obtained by monitoring a single wavelength between 20 and 110 °C with a heating rate of 2 °C/min. Data were converted into fraction folded (θ) plots, fitted to a Boltzmann distribution using OriginPro 2016 and the melting temperatures (T_m) determined from the two-state transition model.

Fluorescence studies

Fluorescence titrations were conducted on a Horiba FluoroMax4 fluorometer (Japan) using a high-precision quartz suprasil cuvette (light path 10 mm $\times$ 4 mm). The fluorescence spectra were obtained between 530-700 nm and averaged over three scans. The excitation wavelength was 498 nm, matching the maximum absorbance of C₈. To study ligand–aptamer complex formation, previously annealed aptamer sequence was titrated into the cuvette containing a solution of 5 μ M C₈ in 30 mM potassium phosphate buffer (pH 7.2) containing 100 mM KCl. Spectra were acquired after a 5 min equilibration period. Fluorescence data was converted into fraction of bound ligand (α) plots, fitted to the saturation binding Hill model (OriginPro 8) and the dissociation constant (K_D) and Hill coefficient (n) values determined.

Fluorescence titrations were also carried with nucleolin protein. In this case, labelled Cy3-AS1411 at 1 μ M, was titrated with nucleolin in the absence and presence of C₈ to evaluate the effect of ligand binding in the affinity of the aptamer towards the protein. The fluorescence emission spectra were obtained between 555 and 700 nm and averaged over three scans. The excitation wavelength was 540 nm, matching the maximum absorbance of Cy3.

Cell Viability Assay

NHDF were grown in RPMI-1640 medium supplemented with 0.01 M HEPES, 0.02 M L-glutamine, 0.001 M sodium pyruvate, 10% fetal bovine serum, and 1% penicillium/streptomycin antibiotic. Human cervical cancer cells (HeLa) were grown in DMEM medium supplemented with 10% fetal bovine serum, and 1% penicillium/streptomycin antibiotic. Cultures were maintained at 37 °C in a humidified atmosphere containing 5% CO₂. For MTT assays, cells were seeded in 48-well plates (1 $\times$ 10³ cells/well) and incubated overnight for cell adhesion. Then, cells were incubated for 7 days with the aptamers at 15 μ M, C₈ at 1 μ M and the preformed ligand-aptamer complex using the same compound/DNA concentrations. Complex formation was performed by

incubating the aptamers with C₈ for 10 min prior to incubation. Wells containing untreated cells were used as control. At the end of incubation, the media was replaced with fresh media containing MTT and further incubated at 37 °C for 1 or 4 h for HeLa and NHDF, respectively. Finally, the formazan crystals were dissolved in DMSO and absorbance was recorded in a Bio-Rad xMark™ microplate reader at 570 nm. Cell viability relatively to control was expressed as mean ± SEM from at least three different experiments. The IC₅₀ values of C₈ at 7 days incubation period were determined by incubating the cells with different compound concentrations ranging from 0.01 to 2 µM. All data treatment was performed using GraphPad Prism 6.

Cell membrane damage

The effect of the ligand-aptamer complex on cell membrane integrity was assessed by the lactate dehydrogenase (LDH) assay using Pierce™ LDH Cytotoxicity Assay Kit (Thermo Fisher Scientific, USA). Cells were seeded in 48-well plates (1×10³ cells/well) and incubated overnight for cell adhesion. Then, cells were incubated for 7 days with the preformed ligand-aptamer complex using the same compound/DNA concentrations used for MTT assay (15 µM aptamer and 1 µM C₈). After exposure, 50 µL/well of supernatant were transferred into optically clear 96-well flat-bottom microplates and mixed with 50 µL/well of reaction mixture and incubated for 30 min at room temperature. The absorbance of the samples was then measured at 490 nm using a Bio-Rad xMark™ microplate reader. The percentage of cytotoxicity is calculated with the following equation [9]:

$$\% \text{ cytotoxicity} = \frac{\text{Experimental LDH release (OD}_{490})}{\text{Maximum LDH release (OD}_{490})} \times 100$$

Cells treated with lysis buffer were used as positive control (maximum LDH release). The experiments were performed in at least two different plates for each cell line, using triplicate wells for each condition.

Flow cytometry analysis

Cells were seeded in 12-well plates (5×10⁵ cells/well) and incubated overnight for cell adhesion. Then, cells were incubated for 24 h with C₈ at 1 µM as control and the preformed ligand-aptamer complex using the same compound/DNA concentrations as used in the MTT assays. After the incubation period, the wells were washed by rinsing with PBS three times. Cells were then trypsinized, resuspended in PBS and analysed in a BD FACSCanto™ II flow cytometry system (BD Life Sciences, US) to evaluate the uptake of the aptamer-ligand complex. Non-specific colouring and debris were excluded by

analysing FSC *vs* SSC density plot as dead cells and debris have lower forward scatter levels.

RNA isolation and RT-qPCR for gene expression analysis

To analyse the effect of the ligand-aptamer complex at genetic level, RT-qPCR analysis of c-MYC and nucleolin expression was performed. Cells were seeded in 12-well plates (1×10^5 cells/well) and incubated overnight. Following cell adhesion, these were incubated for 7 days with free C₈ at 1 μ M and free aptamer at 15 μ M as control, and the preformed ligand-aptamer complex using the same compound/DNA concentrations. Total RNA was extracted from the cells using TRIzol reagent (Invitrogen, USA). 500 ng of total RNA were reverse transcribed using RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, USA), according to the kit instructions. For quantitative analysis of gene expression, RT-qPCR amplification of cDNA was performed using SYBR® Green PCR Master Mix (Thermo Fisher Scientific, USA) on a CFX Connect™ Real-Time PCR Detection System (Bio-Rad, USA). The relative quantification of gene expression was based on the comparative threshold cycle (C_T) method in which the quantity of transcripts is determined as $2^{-(\Delta C_T \text{ target} - \Delta C_T \text{ control})}$, normalized to levels of β -actin and relative to the untreated control cells. The following forward (Fw) and reverse (Rv) primers were used: c-MYC-Fw 5'-TGAGGAGACACCGCCAC-3' and c-MYC-Rv 5'-CAACATCGATTTCTTCCTCATCTTC-3'; NCL-Fw 5'-AGGAGGAGGAAGAAGAGGAG-3' and NCL-Rv 5'-ACAAAGAGATTGAAAGCCGTAG-3'; β -actin-Fw 5'-AAGAGAGGCATCCTCACCCT-3' and β -actin-Rv 5'-TACATGGCTGGGGTGTGAA-3'. The RT-qPCR parameters were the following: an initial denaturation at 95 °C for 3 min, followed by 40 cycles with denaturation at 95 °C for 30 s, annealing at 57 °C for 30 s, and elongation at 72 °C for 30 s, followed by a dissociation stage at 65 °C 5 s. Each sample was run in triplicate from three different experiments and results are expressed in relative levels to respective controls (means $\pm$ SEM).

Fluorescence resonance energy transfer melting

FRET melting experiments were performed on a CFX Connect™ Real-Time PCR Detection System (Bio-Rad, USA). AS1411 was purchased with FAM and TAMRA at the 5' and 3' ends, respectively. The fluorescence measurements were performed in 10 mM lithium cacodylate buffer containing 10 mM KCl and 90 mM LiCl as previously reported [10]. Labelled-AS1411 was used at 0.2 μ M and C₈ at 1 μ M. For competition experiments with nucleolin, the protein was previously incubated with AS1411 at 1 μ M and 2 μ M for 30 min at 37 °C. FAM fluorescence intensity was recorded between 25-95 °C, with a temperature

increment of 1 °C/min. The excitation and detection wavelengths were 492 and 516 nm, respectively. Each experimental condition was tested in triplicate in at least two separate plates. The melting temperatures were determined from the normalized curves as the temperature for which the normalized emission was 0.5.

Confocal fluorescence microscopy imaging

HeLa and NHDF cells were cultured in Ibidi 8-well μ -slides (1×10^4 cells/well) overnight at 37 °C in a humidified atmosphere containing 5% CO₂. After cell adhesion, cells were incubated with 1 μ M Cy5-AS1411 in the presence and absence of 0.5 μ M C₈, for 7 days. At the end of each of the 7 days of incubation, the excess of fluorophores was washed off by rinsing with PBS three times. Cell nucleoli were then stained with 1 μ M nuclear probe Hoechst 33342 for 15 min. Prior to visualization, excess probe was washed off by rinsing with PBS three times. Images were acquired using a Zeiss AxioObserver LSM 710 microscope with 405, 488 and 564 nm laser excitation for Hoechst 33342, C₈ and Cy5-AS1411, respectively. Appropriate emission bands were selected for the three fluorescent channels, considering the spectral overlap and potential bleed-through between channels. Images were processed with Zeiss ZEN software. The quantitative analysis of Cy5-AS1411 internalization in the presence and absence of C₈ was determined as described in literature [11]. In brief, the corrected total cell fluorescence (CTCF) of at least 100 AS1411-positive cells for each experimental condition in three different samples was measured. CTCF can be determined as $CTCF = ID - (A \times MFB)$ where the ID, A, and MFB correspond to integrated density, selected cell area, and mean fluorescence of background readings, respectively.

Statistical analysis

The statistical analysis was performed by using Student's unpaired t test. A *p* value < 0.05 was considered statistically significant. Data analysis was performed in GraphPad Prism 6 (San Diego, CA, USA).

Results and Discussion

AS1411 aptamer is extensively documented as a potential targeting agent for the development of cancer-selective drug delivery systems. It has been shown before that it can deliver a variety of attached cargoes (both covalently and non-covalently) to various cancer cell types, both in vitro and in vivo [2]. Herein, we non-covalently conjugated ligand C₈ to AS1411, LNA-AS1411 and U-AS1411 in an attempt to improve its biological properties (**Table 8.1**). Using a simple supramolecular strategy lacking any linkage

conjugation, we tested a ligand delivery system aimed at HeLa cancer cells as schematically depicted in **Figure 8.1**. The use of locked nucleic acids (LNA) or via thymine-to-uracil substitutions may enhance the binding affinity of the aptamer to its target nucleolin [12]. LNA nucleotides comprise a 2'-O,4'-C-methylene linkage in the sugar moiety, locking it in a *N*-type conformation mimicking RNA nucleotide. LNAs display enhanced thermal and ribozyme stability, and low cytotoxicity for healthy cells [12]. dT→rU substitutions can contribute to the enhancement of stability due to loss of hydration leading to improved stacking of G-tetrads [13].

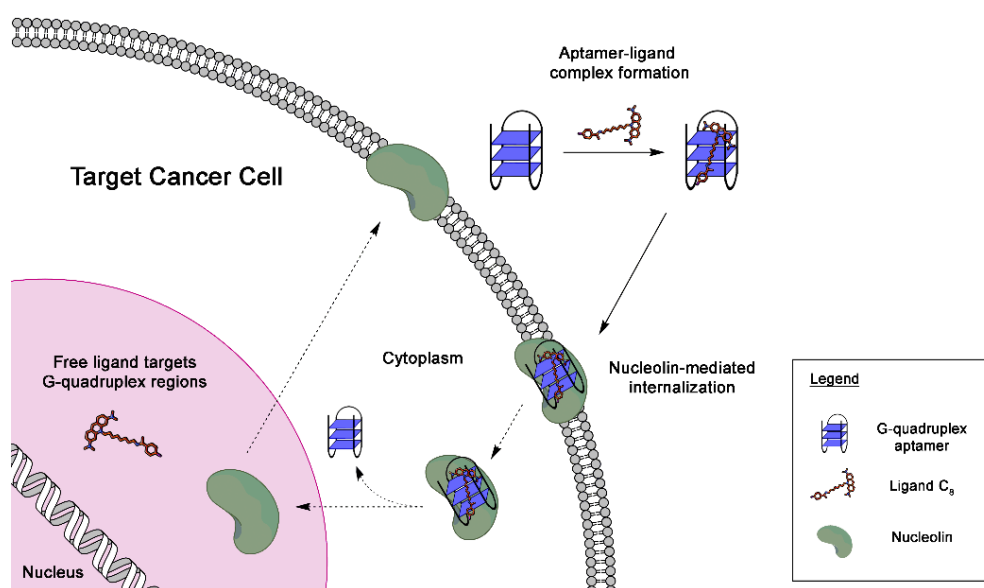


Figure 8.1 - Schematic representation of proposed hypothesis of using AS1411 and its derivatives to carry ligand C₈ into cancer cells overexpressing nucleolin.

Table 8.1 - Nucleotide sequence of the aptamers used in this study and their spectroscopic profile characterization.

Aptamer	Sequence ^a	CD profile ^b		TDS factor ^c		
		240 nm	260 nm	$\frac{\Delta\Delta 240}{\Delta\Delta 295}$	$\frac{\Delta\Delta 255}{\Delta\Delta 295}$	$\frac{\Delta\Delta 275}{\Delta\Delta 295}$
AS1411	5'-GGTGGTGGTGGTGTGGTGGTGGTGG-3'	-	+	2.42	1.51	1.84
LNA-AS1411	5'- g GTGGTGGT g TTGTGGTGGTGGT g -3'	-	+	2.87	3.32	2.90
U-AS1411	5'-GGTGGTGGTGGUUGTGGTGGTGGTGG-3'	-	+	2.17	2.04	2.13

^a Modifications are highlighted in bold, **g** represents LNA guanine.

^b “+” and “-” refer to the presence of a positive or negative peak, respectively, at the indicated wavelength.

^c Ratio between values at two different wavelengths (*e.g.*, 240 and 285 nm for the left column)

The effect of the modifications on the overall G4 structure was evaluated by CD and UV/Vis spectroscopy, namely by the use of TDS. The CD and TDS spectra of AS1411, LNA-AS1411 and U-AS1411 are presented in **Figure 8.2**. The CD spectra of the three sequences,

despite minor changes between them, show the typical CD signature of parallel G4 topologies (**Figure 8.2A**), with a positive peak at around 260 nm and a negative peak at around 240 nm. However, as shown previously by Trent and collaborators, AS1411 forms a mixture of at least eight G4 species in solution and CD spectra could not discriminate among the different species [14]. The TDS spectra confirm the structural polymorphism inherent to the sequence (**Figure 8.2B**). The positive peaks around 243 and 273 nm, and negative peak around 295 nm confirm that a G4 structure is adopted by the modified AS1411 aptamers [8]. Three TDS factors were determined and presented in **Table 8.1**. The TDS factor ($\Delta A_{240}/\Delta A_{295}$) appeared between 2.17 and 2.87, indicating a mixture of G4 topologies as a factor lower than 2 and higher than 4 is characteristic of antiparallel and parallel topologies, respectively [15]. $\Delta A_{255}/\Delta A_{295}$ factor on its turn, was between 1.51 for AS1411 and 3.32 for LNA-AS1411. Likewise, these values demonstrate structural polymorphism as a factor lower than 1.5 and higher than 3.5 is characteristic of antiparallel and parallel topologies, respectively [15]. Finally, $\Delta A_{275}/\Delta A_{295}$ factor was 1.84, 2.13 and 2.90 for AS1411, U-AS1411 and LNA-AS1411 respectively. Factors lower than 2 and higher than 3 are characteristic of antiparallel and parallel topologies, respectively [15]. AS1411 value is indicative of an antiparallel topology, despite the inherent polymorphism. This may be due to the fact that some of the possible topologies predicted for AS1411 are antiparallel dimers [14].

The global analysis of the TDS factors shows that despite the evident structural polymorphism, LNA-AS1411 may fold preferentially into a parallel G4 as the $\Delta A_{255}/\Delta A_{295}$ and $\Delta A_{275}/\Delta A_{295}$ values are higher when compared with the other sequences, and closer to the threshold value for a parallel topology. Indeed, LNA guanosines are incompatible with a *syn* conformation which prevents its assembly into the anti-parallel conformation [16]. It has been proposed that the most stable parallel G- quadruplex is obtained when substituting every guanosine to the LNA except the 5'-terminal one [16]. The observation of the CD spectra shows the loss of a weak negative band at around 280 nm which may be indicative of the structural transitions [17]. Nonetheless, the modifications introduced in AS1411 sequence do not seem to promote the formation of a single G4 structure nor affect G4 formation.

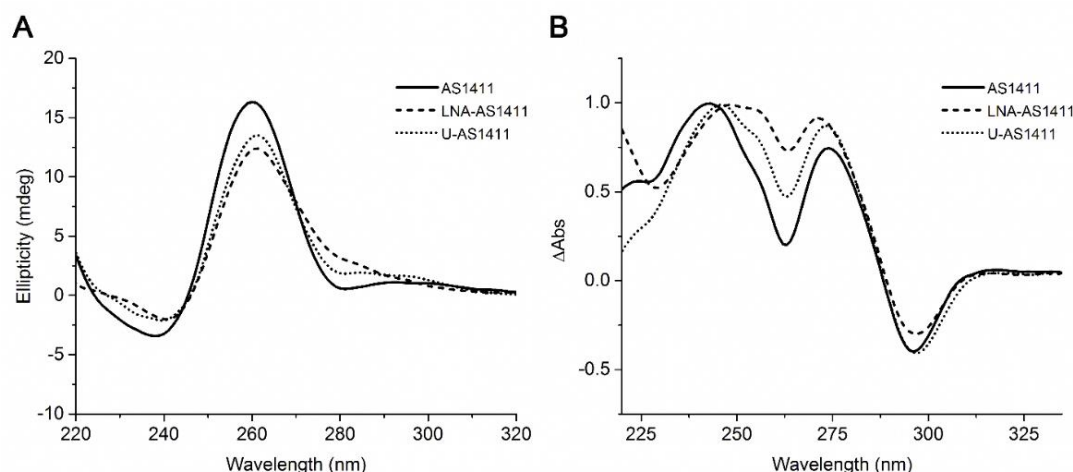


Figure 8.2 - CD spectra (A) and thermal difference spectra (B) of AS1411 aptamer and its derivatives LNA-AS1411 and U-AS1411. Spectra acquisition was performed in 30 mM potassium phosphate buffer containing 100 mM KCl.

The formation of a stable complex between the aptamers and ligand C_8 was assessed and characterized by CD and fluorescence spectroscopy. CD titrations were first performed to assess complex formation and/or structural modification of the G4 structure upon ligand binding. The results are shown in **Figure 8.3**. Upon ligand titration, the overall topology was maintained followed by an increase in ellipticity indicating aptamer-ligand association and stabilization of the preformed G4 structure. A deeper analysis of **Figures 8.3A** and **8.3C** reveals significant changes in the 270-300 nm region. This may be due to the preferential stabilization of one of the possible topologies of AS1411 and U-AS1411. Such effect is not observed for LNA-AS1411 potentially due to the formation of parallel topologies only.

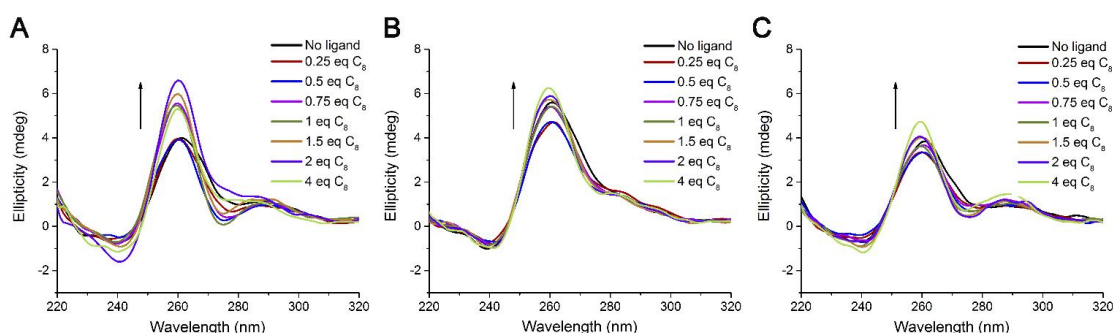


Figure 8.3 - CD spectra of (A) AS1411, (B) LNA-AS1411 and (C) U-AS1411 aptamers in the presence of increasing amounts of ligand C_8 . Spectra acquisition was performed in 30 mM potassium phosphate buffer containing 100 mM KCl.

The binding strength between ligand C_8 and the aptamers was further assessed using fluorimetric titrations. As seen in Figure 4, upon excitation at 498 nm, ligand C_8 emitted fluorescence as a broad band centred at 568 nm as previously reported [6]. After addition

of previously annealed aptamer solution, a ≈ 10 -fold fluorescence enhancement was observed for all sequences denoting a strong interaction between the ligand and the aptamers G4 structure. The fluorescence intensity values were then fitted to a saturation binding model and the apparent dissociation constants (K_D) determined. The K_D values for AS1411, LNA-AS1411 and U-1411 were 1.29×10^{-6} , 1.92×10^{-6} and 1.37×10^{-6} M, respectively ($SD \leq 1 \times 10^{-8}$ M). These values in the low micromolar range are indicative of moderate to high affinity and stable complex formation between the ligand and the aptamers G4 structure [18].

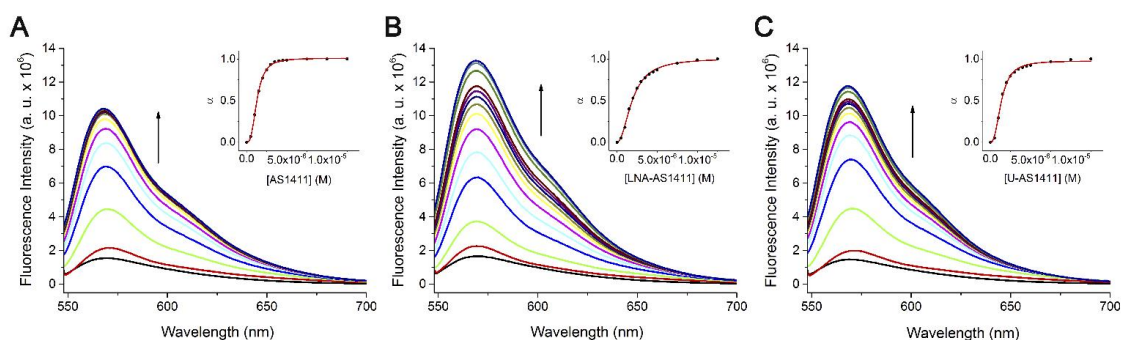


Figure 8.4 - Fluorescence spectroscopy titrations of ligand C_8 with increasing amounts of (A) AS1411, (B) LNA-AS1411 and (C) U-AS1411 DNA. Spectra acquisition was performed in 30 mM of potassium phosphate buffer containing 100 mM KCl, with DNA concentration ranging from 0.1 μ M to 12.5 μ M. Insets: saturation binding plots fitted to Hill model.

The Hill coefficient (n) which can be interpreted as the average number of bound ligands per G4, was 2.8, 2.1 and 2.5 for AS1411, LNA-AS1411 and U-1411, respectively. This suggests a ligand:DNA ratio of 3:1 in the case of AS1411 and U-1411, and 2:1 in the case of LNA-AS1411. Regarding AS1411, as it is assumed to fold into a variety of structures, for instance a dimeric antiparallel G4 [2], the 3:1 ratio may be misestimated. At this point, we have no information on whether ligand C_8 binds preferentially to a specific AS1411 conformation. Additionally, the Hill coefficient is a measure of cooperativity, whereas the n values greater than 1 indicate positive cooperativity meaning that a binding event increases the affinity of additional ligand molecules [6].

The influence of ligand C_8 on the thermal stability of the aptamers was investigated by thermal denaturation. For that, we performed CD melting experiments and assessed the ligand's effect on the aptamers melting temperature (T_m) (Figure S1, **Appendix E**). The ligand clearly increased the T_m values of all the aptamers G4 structure, being the effect more prominent on AS1411, as C_8 promoted a ΔT_m of 28.4 $^{\circ}$ C at 2 molar equivalents. U-AS1411 was stabilized for 24.5 $^{\circ}$ C at the same ligand:G4 ratio, while LNA-AS1411 was the least stabilized structure with a ΔT_m of 17.5 $^{\circ}$ C. The ΔT_m values agree with the K_D values as the AS1411 presented the higher K_D and ΔT_m while on the other end LNA-AS1411 presented

the lowest values. As the G4 folding is necessary for the biological properties of the aptamer, the ligand-induced thermal stability may enhance its nuclease resistance, nucleolin affinity and intracellular trafficking [12,19,20]. The ΔT_m values at different ligand:G4 ratios are summarized in Table S1.

The ability of the aptamers AS1411, LNA-AS1411 and U-AS1411 to accumulate C₈ in the cancer cells and to prevent C₈ accumulation in healthy cells was indirectly assessed by the MTT assay. As previously reported, AS1411 has a weak inhibitory effect in HeLa cancer cells and more than 5 days are needed to observe a slight effect on cell viability, even at 10 μ M aptamer concentration [21,22]. As such, we tested the effect of the aptamer-ligand complexes for 7 days incubation period at 15 μ M aptamer concentration. The free aptamers cytotoxicity towards HeLa and non-malignant NHDF cells is shown in Figure S2. AS1411 showed reduced cytotoxicity towards HeLa cells (79% mean viability) and nearly no toxicity towards non-malignant cells. The LNA-modified aptamer did not have any effect on the cell viability of both cancer and non-malignant cells. Finally, U-AS1411 presented a similar cytotoxicity for HeLa cells (71% mean viability) as AS1411, however in the case of NHDF cells its cytotoxicity was significantly higher (53% mean viability).

Then, the IC₅₀ values of C₈ at 7 days incubation time were determined for HeLa and NHDF cells (Figure S3). The IC₅₀ was 0.15 and 0.48 μ M for HeLa and NHDF, respectively. Based on these values, we decided to use a C₈ concentration of 1 μ M which is high enough to elicit significant cell toxicity and still suitable to observe any changes induced by the ligand complexation to the aptamer. After mixing 15 μ M aptamer with 1 μ M C₈ to allow complex formation, the aptamer-ligand complexes were incubated for 7 days in HeLa and NHDF cells. The results are presented in **Figure 8.5**.

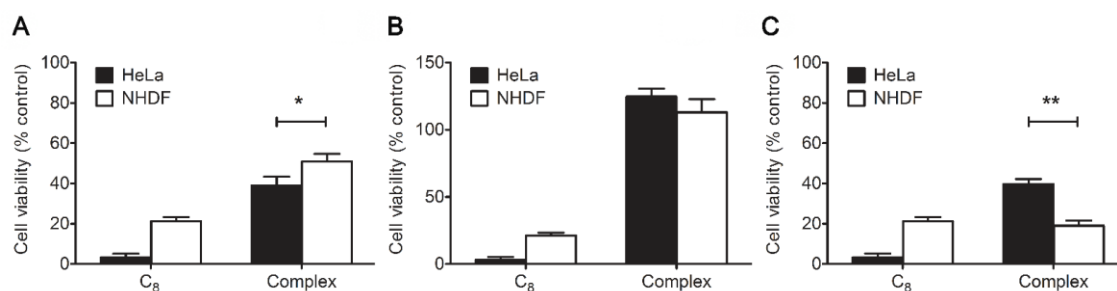


Figure 8.5 - Relative cell viability of HeLa and NHDF cells incubated for 7 days with free C₈ or the pre-formed aptamer-ligand complex at a C₈ concentration of 1 μ M using aptamer (A) AS1411, (B) LNA-AS1411 and (C) U-AS1411. * $p < 0.05$, ** $p < 0.01$.

As seen in **Figure 8.5A**, the C₈ toxicity towards non-malignant cells is significantly attenuated upon complexation with AS1411. The lower cytotoxicity of C₈ towards NHDF

cells when complexed with AS1411 may be due to the aptamer mechanism involved in selective accumulation in cells. It has been proposed that AS1411 is gradually cleared from normal cells by efflux or exocytosis 24-72 h post-treatment [2]. Additionally, normal cells have increased lysosomal activity in comparison with AS1411-treated cancer cells [2]. Therefore, if the aptamer-ligand complex is capable of sustaining cellular trafficking it may be cleared from normal cells by lysosomal degradation, thus clearing C₈ and reducing its effects. Although AS1411-C₈ has a more pronounced effect on HeLa than on NHDF cells, the cytotoxicity towards HeLa cells is also hindered. In this case, the explanation may be the high affinity of C₈ towards the aptamer which may prevent the ligand to detach from AS1411 once inside the cell. LNA-AS1411 complex on its hand, showed no cytotoxic effect on both cell lines with viability values around 100% (**Figure 8.5B**). Indeed, LNA-containing DNA sequences were already shown to be poorly toxic for HeLa cells [23]. This suggests that the modification of the sugar ring may affect the aptamer cellular behaviour, namely the interaction with its target nucleolin. The apparent viability values above 100% observed for LNA-AS1411 treatment may be due to the compounds influencing the cell's metabolism thus resulting in increased metabolization of MTT [24] (as the MTT assay is a measure of the cellular mitochondrial enzymatic activity) and/or cell growth (in response to stress, a phenomena termed hormesis [25]) when compared with the control. Opposingly, U-AS1411 complex showed an undesired increased cytotoxicity towards NHDF cells, significantly higher than that observed for HeLa cells (**Figure 8.5C**).

To show the applicability of the system, the MTT experiments were extended to two additional cancer cell lines using AS1411-C₈ complex, namely PC-3 prostate cancer cell line and A549 lung cancer cell lines (Figure S4). Similarly, to what was observed for HeLa cells, the aptamer cytotoxicity was improved when complexed with C₈. In the case of PC-3 cells the complex promoted a viability decrease of about 85% while in the case of A549 cells a milder effect was observed (47% mean viability). These results suggest that this strategy may be employed in other cancer types, particularly prostate cancer because PC-3 cells showed the most promising results. Additionally, the effect of AS1411-C₈ complex on cell membrane integrity was evaluated in the same cancer cell line panel used in the MTT assays. The results shown in **Figure 8.6** suggest that the AS1411-C₈ complex promoted modest cell membrane damage in the tested cell lines (LDH activity < 50%), despite reducing its viability as indicated by the MTT results. Similar effect was already observed for other nanosystems [26] and may be related to a non-apoptotic cell death mechanism [27]. Furthermore, it should be emphasized that while the LDH assay is a measure of cytotoxicity induced by cell membrane damage, the MTT assay provides data on cell viability, specifically the cell's metabolic activity. A correlation between the results is therefore not always attained.

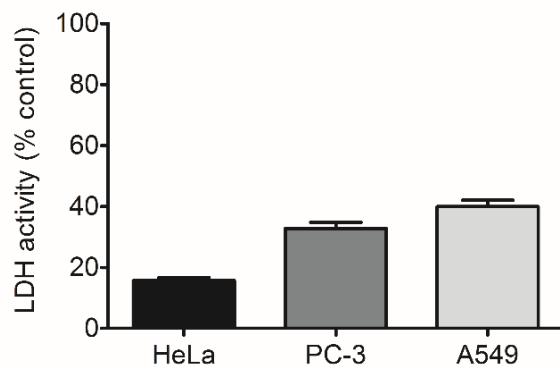


Figure 8.6 - Relative LDH activity of HeLa, PC-3 and A549 cells incubated for 7 days with the pre-formed AS1411-ligand complex at a C₈ concentration of 1 μ M.

Following the MTT and LDH assays, we employed flow cytometry to assess intracellular accumulation of free and AS1411-complexed C₈ in HeLa cancer cells and NHDF non-malignant cells, based on the intrinsic fluorescence of ligand C₈. The results are depicted in **Figure 8.7**.

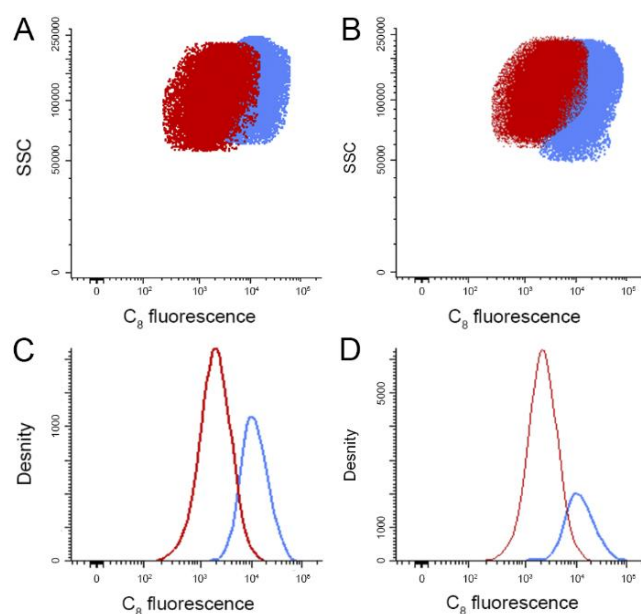


Figure 8.7 - Flow cytometry analysis of free C₈ (blue) and AS1411-C₈ complex (red) in (A) HeLa cancer cells and (B) NHDF non-malignant cells and respective histograms (C and D). Cells were incubated for 24 h with the preformed aptamer-ligand complex at a C₈ concentration of 1 μ M.

Flow cytometry analysis indicated that the AS1411-C₈ complex is less intracellularly accumulated in both cell types when compared to free C₈, as observed by the lower fluorescence intensity emitted from the cells, despite an enhanced fluorescence quantum yield of C₈ when complexed to DNA. These results corroborate the lower cytotoxicity observed for the AS1411-C₈ complex when compared with free C₈.

Then, we assessed if the reduced cellular uptake of the aptamer-ligand complex could be due to a potential interference with nucleolin binding. For that we employed fluorescence spectroscopy in two different approaches. First, using a solution of fluorescently labelled Cy3-AS1411, we titrated increasing amounts of nucleolin in the presence and absence of C₈ and determined the K_D values from the fluorescence intensity changes. The saturation binding plots are shown in **Figure 8.8**. The obtained K_D values for the binding interaction between the free AS1411 aptamer and the AS1411-C₈ complex with nucleolin were 8.63 pM and 6.38 pM, respectively. The obtained K_D values are lower than those reported in the literature, being potentially overestimated by the use of labelled oligonucleotides or the use of a different peptide than that used in those studies [28,29]. Nonetheless, both values are in the same range, being slightly lower in the case of the aptamer-ligand complex which suggests that the ligand does not negatively affect the recognition of the protein by the aptamer.

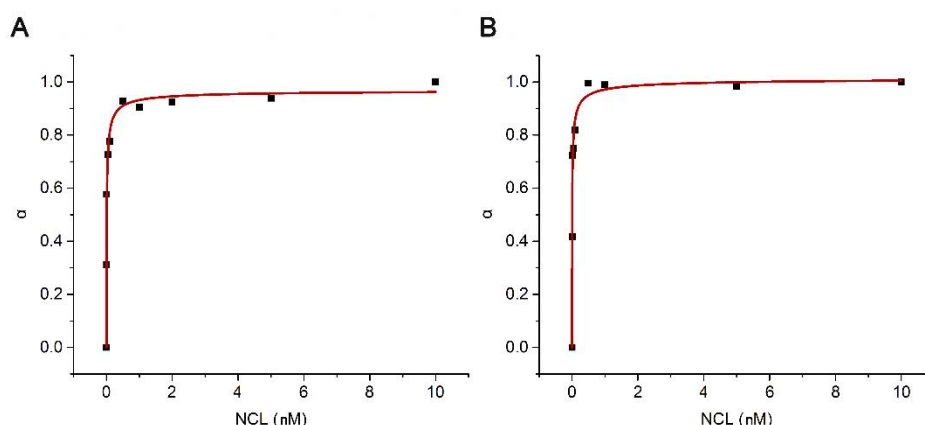


Figure 8.8 - Saturation binding plots fitted to Michaelis-Menten model. Fluorescence intensity of Cy-AS1411 was recorded after the addition of increasing amounts of nucleolin in the (A) absence and (B) presence of ligand C₈.

Additionally, a competition FRET-melting experiment was conducted to support the fluorescence titration findings. The rationale for the assay design was based on the hypothesis that if the ligand-induced stabilization of AS1411 was similar in the presence and absence of nucleolin, ligand binding would not affect protein recognition. The ligand-induced thermal stabilization of AS1411 was 13.1 °C in the presence of 1 μ M C₈ (**Table 8.2**). Upon pre-incubation of AS1411 with equimolar nucleolin, the ΔT_m increased to 20.0 °C. Even with an excess of protein (2 molar eq.) the ΔT_m was 15.1 °C, reflecting a higher stabilization of the aptamer by C₈ in the presence of nucleolin. This indicates that the ligand does not prevent the recognition of the protein by the aptamer, as suggested by the fluorescence titration experiments, and that it may even enhance the stabilization of AS1411 G4 structure. This may be due to the ligand providing additional binding sites for the protein interaction (ternary complex), or since the aptamer G4 structure is necessary

for nucleolin binding, the C₈-induced stabilization of the structure may enhance its binding strength [30].

Table 8.2 - Ligand induced thermal stabilization of AS1411 aptamer in the presence and absence of nucleolin.

	ΔT_m AS1411 (° C)	ΔT_m AS1411 + 1 μ M nucleolin (° C)	ΔT_m AS1411 + 2 μ M nucleolin (° C)
C₈	13.1 $\pm$ 0.8	20.0 $\pm$ 1.4	15.1 $\pm$ 0.7

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_m$ (DNA + 1 μ M of ligand) – T_m (DNA)]. The buffer used was 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl. The T_m value for AS1411 is 46.9 $\pm$ 0.2 °C. Reported values correspond to the average of three measurements with the estimated standard deviation.

To assess the potential effect of the AS1411-C₈ complex on gene expression, RT-qPCR analysis was performed. Oncogene c-MYC and nucleolin mRNA levels were analysed since c-MYC is commonly found upregulated in tumour cells being responsible for cancer cell proliferation [31], and nucleolin is the AS1411's target. The results are depicted in **Figure 8.9**. Upon treatment with AS1411-C₈ complex, the expression of c-MYC mRNA was downregulated by more than 50% ($p < 0.01$) which may inhibit cancer cell progression, eventually leading to cell death [31]. The acridine derivative C₈ was already tested for its ability to stabilize the G4 structure formed by c-MYC NHE element of the promoter region [6]. Therefore, it is reasonable to assume that the decrease in the c-MYC mRNA levels may be due to transcription modulation induced by interaction of C₈ and c-MYC G4 structures upon intracellular dissociation from the AS1411-ligand complex [32]. Opposingly, nucleolin mRNA expression was upregulated by 1.5-fold similarly to what was reported for free AS1411 before [33]. This suggests that the biological effects of AS1411 are not negatively affected by its complexation with ligand C₈. This feedback loop induction may be responsible for increasing the cells sensitivity to AS1411 and/or increasing its intracellular accumulation [33].

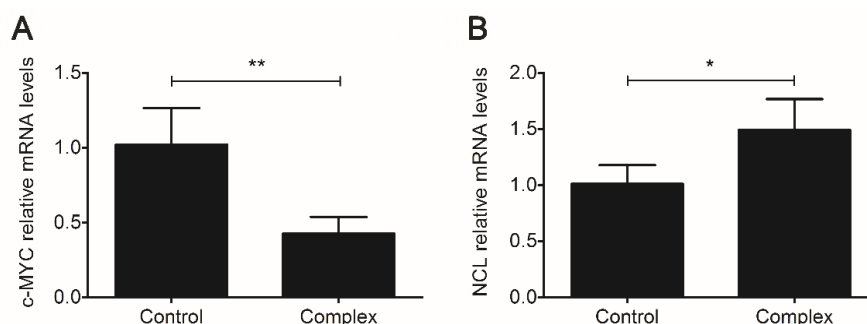


Figure 8.9 - Effect of AS1411-C₈ complex treatment on c-MYC and nucleolin gene expression measured by RT-qPCR. Data represent the average of three experiments $\pm$ SEM; * $p < 0.05$, ** $p < 0.01$.

Finally, the subcellular localization of the AS1411- C_8 complex in HeLa and NHDF cells was evaluated by fluorescence confocal microscopy. As seen in **Figure 8.10**, the complex is able to penetrate cell membrane and localizes in the cytoplasm of HeLa cells. The overlapping of C_8 (green) and AS1411 (red) suggests that the complex is still formed intracellularly (yellow) and maintained during cell internalization and trafficking. C_8 can be seen in the nucleoli in a free state which may suggest decomplexation and localization of the nucleolus as described for the ligand [6]. This may be particularly important to explain the downregulation of c-MYC transcriptional level as showed by the RT-qPCR experiments. A modest increase in the uptake of AS1411 can be observed in the presence of C_8 which may correlate with the proposed increased stability of aptamer upon ligand binding, possibly enhancing its association with nucleolin, and the fact that nucleolin is found overexpressed upon treatment with the complex as suggested by RT-qPCR results. The quantitative analysis of the Cy5-AS1411 fluorescence in the absence and presence of C_8 shown in Figure S5 substantiates an increased uptake of AS1411 in the presence of the ligand.

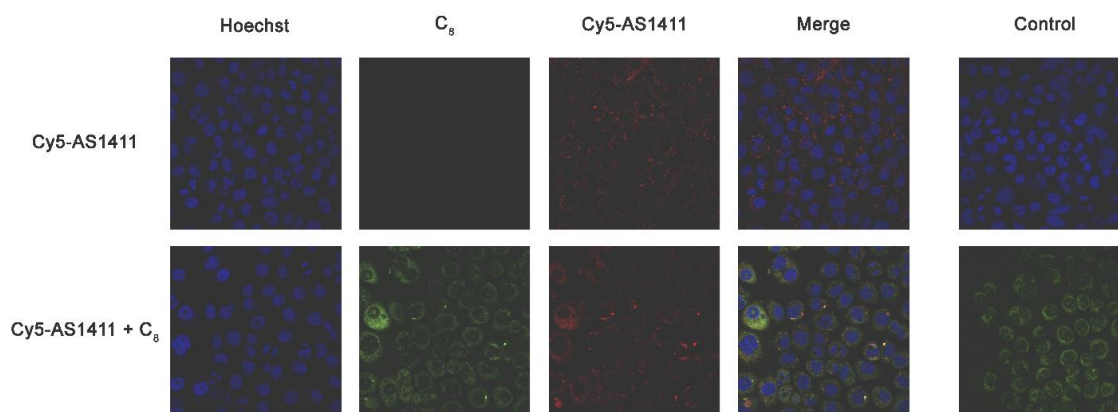


Figure 8.10. Confocal microscopy images of HeLa cells incubated with Cy5-AS1411 and Cy-AS1411- C_8 complex for 7 days. Cell nuclei are stained with Hoechst 33342 (blue), C_8 emits green fluorescence and Cy-AS1411 is shown in red. Overlapping of the two stains, observed as yellow regions, can be seen in the merged images of the cells. Control cells and C_8 single stain control images are shown on the right. Brightness was adjusted to facilitate analysis.

Similar findings were observed for NHDF cells (Figure S6). These results disagree with the selective accumulation described above where it has been shown that AS1411 is gradually cleared from normal cells by efflux 24-72 h post-treatment. The explanation may arise from the fact that cells are constantly incubated in media containing the compounds for 7 days straight. To test this hypothesis, we incubated the cells for 6 days in the presence of the aptamer-ligand complex, followed by 1 day in free media. The results shown in **Figure 8.11** confirm that indeed AS1411 is eliminated from normal cells after one day. C_8 despite

being uptaken complexed with AS1411, remains inside the cells, suggesting that the observed differences in cytotoxicity in MTT assays between free C₈ and AS1411-C₈ complex are mostly due to differences in uptake as shown by flow cytometry. A similar experiment was performed in HeLa cells (Figure S7) and the results suggest that cancer cells are indeed less efficient in eliminating the aptamer in its free or C₈-complexed state as Cy5-AS1411 red staining is still observed after 1-day incubation in free media. These results agree with the proposed mechanism of cancer-selective accumulation of AS1411 where AS1411-conjugates are able to escape lysosomal vesicles with increased uptake [2,34].

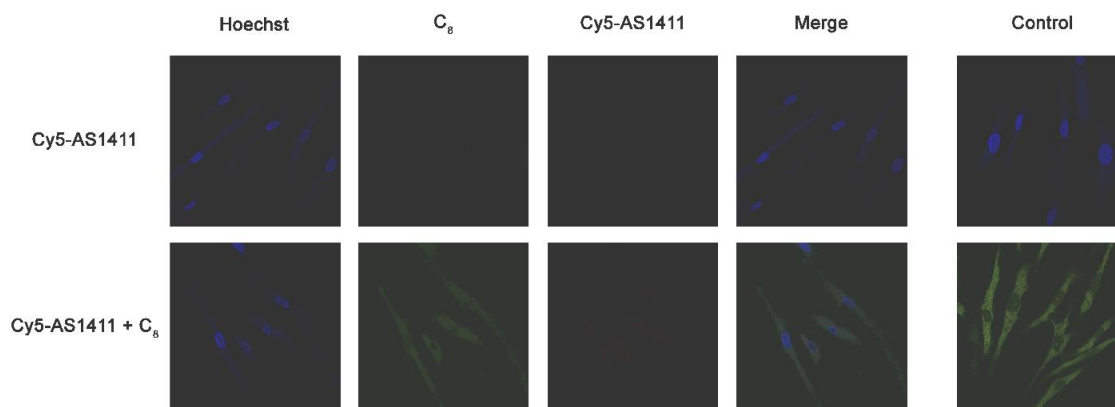


Figure 8.11 - Confocal microscopy images of NHDF cells incubated with Cy5-AS1411 and Cy-AS1411-C₈ complex for 6 days followed by 1-day incubation in fresh media. Cell nuclei are stained with Hoechst 33342 (blue), C₈ emits green fluorescence while Cy5-AS1411 fluorescence is absent. Control cells and C₈ single stain control images are shown on the right. Brightness was adjusted to facilitate analysis.

Conclusion

AS1411 aptamer has been increasingly suggested as an effective targeting system for the selective delivery of a wide range of cargoes to cancer cells. Herein, using a simple supramolecular strategy lacking any chemical modification of the aptamer, we propose a delivery system for conveying a G4 ligand C₈ into cancer cells in a selective manner. AS1411 unmodified aptamer and two derivatives containing LNA and uracil nucleotides were used. Ligand C₈ was shown to strongly associate with the three aptamers with K_D in the 10^{-6} M⁻¹ range. Additionally, C₈ is able to thermally stabilize the aptamers structure, possible enhancing its properties. The MTT assays shown that the complexation of C₈ with AS1411 decreases its cytotoxicity to NHDF cells, which was later revealed to be due to differential uptake of free C₈ vs complexed C₈ by flow cytometry analysis. LNA-AS1411 and U-AS1411 complexes shown non-adequate cytotoxic profile, being that the first is non-toxic for HeLa cells and the latter is highly toxic for non-malignant cells. MTT experiments with prostate and lung cancer cells show the potential applicability of the system in other cancer types. Using fluorescence spectroscopy and FRET-melting we propose that ligand

C₈ complexation with AS1411 does not affect its recognition of protein nucleolin, which is necessary for its biological effects. RT-qPCR experiments demonstrated that AS1411-C₈ complex is able to downregulate c-MYC expression, which may be attributed to the stabilization of c-MYC G4 structures by the ligand. Finally, confocal microscopy indicated that the complex is efficiently taken up into the cells and that AS1411-C₈ is maintained during cell internalization and trafficking. Moreover, AS1411 uptake is increased in the presence of C₈ which may be due to nucleolin overexpression induced by the complex as suggested by RT-qPCR. This study paves the way for the development of cancer-specific AS1411-based delivery systems for the selective accumulation of G4 ligands in the target tissues, in a simple non-covalent strategy.

Acknowledgments

J. Carvalho acknowledges a doctoral fellowship grant from the FCT – Foundation for Science and Technology ref. SFRH/BD/122953/2016. This work was supported by project “Ações Integradas Luso-Francesas” ref. TC-15/17, FCT project ref. IF/00959/2015 financed by Fundo Social Europeu and Programa Operacional Potencial Humano, FCT project ref. UID/Multi/04349/2019, project Fundação Luso-Americana (FLAD) Healthcare 2020 ref. 45/2018, MIT Portugal FCT project BIODEVICE ref. MIT-EXPL/BIO/0008/2017 and UTAustin Portugal Program Exploratory project DREAM ref. UTAP-EXPL/NTec/0015/2017 This work was also supported by POCI – COMPETE 2020 – Operational Programme Competitiveness and Internationalisation in Axis I – Strengthening research, technological development and innovation (project POCI-01-0145-FEDER-007491). JLM acknowledges funding from the Symbit project (reg. no. CZ.02.1.01/0.0/0.0/15_003/0000477) financed by the ERDF.

References

- [1] Bates, P. J., Laber, D. a., Miller, D. M., Thomas, S. D., and Trent, J. O. (2009) Discovery and development of the G-rich oligonucleotide AS1411 as a novel treatment for cancer. *Experimental and Molecular Pathology* 86, 151–164. <https://doi.org/10.1016/j.yexmp.2009.01.004>.
- [2] Bates, P. J., Reyes-Reyes, E. M., Malik, M. T., Murphy, E. M., O’Toole, M. G., and Trent, J. O. (2017) G-quadruplex oligonucleotide AS1411 as a cancer-targeting agent: Uses and mechanisms. *Biochimica et Biophysica Acta - General Subjects* 1861, 1414–1428. <https://doi.org/10.1016/j.bbagen.2016.12.015>.

- [3] Mongelard, F., and Bouvet, P. (2007) Nucleolin: a multiFACeTed protein. *Trends in Cell Biology* 17, 80–86. <https://doi.org/10.1016/j.tcb.2006.11.010>.
- [4] Reyes-Reyes, E. M., Šalipur, F. R., Shams, M., Forsthoefel, M. K., and Bates, P. J. (2015) Mechanistic studies of anticancer aptamer AS1411 reveal a novel role for nucleolin in regulating Rac1 activation. *Molecular Oncology* 9, 1392–1405. <https://doi.org/10.1016/j.molonc.2015.03.012>.
- [5] Shieh, Y. A., Yang, S. J., Wei, M. F., and Shieh, M. J. (2010) Aptamer-based tumor-targeted drug delivery for photodynamic therapy. *ACS Nano* 4, 1433–1442. <https://doi.org/10.1021/nn901374b>.
- [6] Carvalho, J., Pereira, E., Marquevielle, J., Campello, M. P. C., Mergny, J.-L., Paulo, A., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2018) Fluorescent light-up acridine orange derivatives bind and stabilize KRAS-22RT G-quadruplex. *Biochimie* 144, 144–152. <https://doi.org/10.1016/j.biochi.2017.11.004>.
- [7] Pereira, E., Do Quental, L., Palma, E., Oliveira, M. C., Mendes, F., Raposinho, P., Correia, I., Lavrado, J., Di Maria, S., Belchior, A., Vaz, P., Santos, I., and Paulo, A. (2017) Evaluation of Acridine Orange Derivatives as DNA-Targeted Radiopharmaceuticals for Auger Therapy: Influence of the Radionuclide and Distance to DNA. *Scientific Reports* 7, 42544. <https://doi.org/10.1038/srep42544>.
- [8] Mergny, J.-L., Li, J., Lacroix, L., Amrane, S., and Chaires, J. B. (2005) Thermal difference spectra: A specific signature for nucleic acid structures. *Nucleic Acids Research* 33, 1–6. <https://doi.org/10.1093/nar/gni134>.
- [9] Kumar, P., Nagarajan, A., and Uchil, P. D. (2018) Analysis of Cell Viability by the Lactate Dehydrogenase Assay. *Cold Spring Harbor Protocols* 465–469. <https://doi.org/10.1101/pdb.proto95497>.
- [10] De Rache, A., and Mergny, J.-L. (2015) Assessment of selectivity of G-quadruplex ligands via an optimised FRET melting assay. *Biochimie* 115, 194–202. <https://doi.org/10.1016/j.biochi.2015.06.002>.
- [11] Loureiro, A., Noro, J., Abreu, A. S., Soares, D., Silva, C., and Cavaco-Paulo, A. (2018) Absence of Albumin Improves in Vitro Cellular Uptake and Disruption of Poloxamer 407-Based Nanoparticles inside Cancer Cells. *Molecular Pharmaceutics* 15, 527–535 Article. <https://doi.org/10.1021/acs.molpharmaceut.7b00893>.

- [12] Gao, S., Zheng, X., Jiao, B., and Wang, L. (2016) Post-SELEX optimization of aptamers. *Analytical and Bioanalytical Chemistry* 4567–4573. <https://doi.org/10.1007/s00216-016-9556-2>.
- [13] Olsen, C. M., and Marky, L. A. (2009) Energetic and hydration contributions of the removal of methyl groups from thymine to form uracil in G-quadruplexes. *Journal of Physical Chemistry B*. <https://doi.org/10.1021/jp808526d>.
- [14] Dailey, M. M., Clarke Miller, M., Bates, P. J., Lane, A. N., and Trent, J. O. (2010) Resolution and characterization of the structural polymorphism of a single quadruplex-forming sequence. *Nucleic Acids Research* 38, 4877–4888. <https://doi.org/10.1093/nar/gkq166>.
- [15] Karsisiotis, A. I., Hessari, N. M. A., Novellino, E., Spada, G. P., Randazzo, A., and Webba Da Silva, M. (2011) Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. *Angewandte Chemie - International Edition* 50, 10645–10648. <https://doi.org/10.1002/anie.201105193>.
- [16] Pedersen, E. B., Nielsen, J. T., Nielsen, C., and Filichev, V. V. (2011) Enhanced anti-HIV-1 activity of G-quadruplexes comprising locked nucleic acids and intercalating nucleic acids. *Nucleic Acids Research* 39, 2470–2481. <https://doi.org/10.1093/nar/gkq1133>.
- [17] Bing, T., Zheng, W., Zhang, X., Shen, L., Liu, X., Wang, F., Cui, J., Cao, Z., and Shanguan, D. (2017) Triplex-quadruplex structural scaffold: A new binding structure of aptamer. *Scientific Reports*. <https://doi.org/10.1038/s41598-017-15797-5>.
- [18] Lecarme, L., Prado, E., Rache, A. De, Bonnet, R., Heyden, A. Van Der, Philouze, C., Gomez, D., Mergny, J., Defrancq, E., Jarjays, O., and Thomas, F. (2014) Interaction of Polycationic Ni (II)-Salophen Complexes with G-Quadruplex DNA.
- [19] Hasegawa, H., Savory, N., Abe, K., and Ikebukuro, K. (2016) Methods for improving aptamer binding affinity. *Molecules* 21. <https://doi.org/10.3390/molecules21040421>.
- [20] D'Onofrio, J., Petraccone, L., Erra, E., Martino, L., Di Fabio, G., De Napoli, L., Giancola, C., and Montesarchio, D. (2007) 5'-Modified G-quadruplex forming oligonucleotides endowed with anti-HIV activity: Synthesis and biophysical properties. *Bioconjugate Chemistry* 18, 1194–1204. <https://doi.org/10.1021/bc070062f>.
- [21] Bates, P. J., Kahlon, J. B., Thomas, S. D., Trent, J. O., and Miller, D. M. (1999) Antiproliferative activity of G-rich oligonucleotides correlates with protein binding.

[22] Girvan, A. C., Teng, Y., Casson, L. K., Thomas, S. D., Jv^oliger, S., Ball, M. W., Klein, J. B., Pierce, W. M., Barve, S. S., and Bates, P. J. (2006) AGRO100 inhibits activation of nuclear factor- κ B (NF- κ B) by forming a complex with NF- κ B essential modulator (NEMO) and nucleolin. *Molecular Cancer Therapeutics* 5, 1790–1799. <https://doi.org/10.1158/1535-7163.mct-05-0361>.

[23] Arzumanov, A., Walsh, A. P., Rajwanshi, V. K., Kumar, R., Wengel, J., and Gait, M. J. (2001) Inhibition of HIV-1 Tat-dependent trans activation by steric block chimeric 2'-O-methyl/LNA oligoribonucleotides. *Biochemistry*. <https://doi.org/10.1021/bi011279e>.

[24] Rai, Y., Pathak, R., Kumari, N., Sah, D. K., Pandey, S., Kalra, N., Soni, R., Dwarakanath, B. S., and Bhatt, A. N. (2018) Mitochondrial biogenesis and metabolic hyperactivation limits the application of MTT assay in the estimation of radiation induced growth inhibition. *Scientific Reports* 8, 1–15. <https://doi.org/10.1038/s41598-018-19930-w>.

[25] Yoshimasu, T., Ohashi, T., Oura, S., Kokawa, Y., Kawago, M., Hirai, Y., Miyasaka, M., Nishiguchi, H., Kawashima, S., Yata, Y., Honda, M., Fujimoto, T., and Okamura, Y. (2015) A theoretical model for the hormetic dose-response curve for anticancer agents. *Anticancer Research* 35, 5851–5855.

[26] Ursini, C. L., Cavallo, D., Freseigna, A. M., Ciervo, A., Maiello, R., Casciardi, S., Tombolini, F., Buresti, G., and Iavicoli, S. (2012) Study of Cytotoxic and Genotoxic Effects of Hydroxyl-Functionalized Multiwalled Carbon Nanotubes on Human Pulmonary Cells. *Journal of Nanomaterials* 2012, 1–9. <https://doi.org/10.1155/2012/815979>.

[27] Kornienko, A., Rastogi, S. K., Lefranc, F., and Kiss, R. (2013) Therapeutic Agents Triggering Nonapoptotic Cancer Cell Death. *Journal of Medicinal Chemistry* 56, 4823–4839.

[28] Fan, X., Sun, L., Wu, Y., Zhang, L., and Yang, Z. (2016) Bioactivity of 2'-deoxyinosine-incorporated aptamer AS1411. *Scientific Reports* 6, 25799. <https://doi.org/10.1038/srep25799>.

[29] Fan, X., Sun, L., Li, K., Yang, X., Cai, B., Zhang, Y., Zhu, Y., Ma, Y., Guan, Z., Wu, Y., Zhang, L., and Yang, Z. (2017) The Bioactivity of D-/L-Isonucleoside- and 2'-Deoxyinosine-Incorporated Aptamer AS1411s Including DNA Replication/MicroRNA

Expression. *Molecular Therapy - Nucleic Acids* 9, 218–229. <https://doi.org/10.1016/j.omtn.2017.09.010>.

[30] Dapić, V., Abdomerović, V., Marrington, R., Peberdy, J., Rodger, A., Trent, J. O., and Bates, P. J. (2003) Biophysical and biological properties of quadruplex oligodeoxyribonucleotides. *Nucleic Acids Research* 31, 2097–2107. <https://doi.org/10.1093/nar/gkg316>.

[31] Hurley, L. H., Von Hoff, D. D., Siddiqui-Jain, A., and Yang, D. (2006) Drug Targeting of the c-MYC Promoter to Repress Gene Expression via a G-Quadruplex Silencer Element. *Seminars in Oncology* 33, 498–512. <https://doi.org/10.1053/j.seminoncol.2006.04.012>.

[32] Chauhan, A., Paul, R., Debnath, M., Bessi, I., Mandal, S., Schwalbe, H., and Dash, J. (2016) Synthesis of Fluorescent Binaphthyl Amines That Bind c-MYC G-Quadruplex DNA and Repress c-MYC Expression. *Journal of Medicinal Chemistry* 59, 7275–7281. <https://doi.org/10.1021/acs.jmedchem.6b00328>.

[33] Sharma, V. R., Thomas, S. D., Miller, D. M., Rezzoug, F., Sharma, V. R., Thomas, S. D., Miller, D. M., and Rezzoug, F. (2018) Nucleolin Overexpression Confers Increased Sensitivity to the Anti-Nucleolin Aptamer, AS1411. *Cancer Investigation* 36, 475–491. <https://doi.org/10.1080/07357907.2018.1527930>.

[34] Reyes-Reyes, E. M., Teng, Y., and Bates, P. J. (2010) A new paradigm for aptamer therapeutic AS1411 action: Uptake by macropinocytosis and its stimulation by a nucleolin-dependent mechanism. *Cancer Research* 70, 8617–8629. <https://doi.org/10.1158/0008-5472.CAN-10-0920>.

Chapter 9

Aptamer-guided Acridine Derivatives for Cervical Cancer

This chapter was published in:

Carvalho, J. [†], Lopes-Nunes, J. [†], Lopes, A. C. [†], Campello, M. P. C., Paulo, A., Queiroz, J. A., and Cruz, C. (2019) Aptamer-guided acridine derivatives for cervical cancer. *Organic and Biomolecular Chemistry* 17, 2992–3002. <https://doi.org/10.1039/c9ob00318e>.

[†] Equal contribution

Aptamer-guided Acridine Derivatives for Cervical Cancer

Abstract

DNA aptamers represent a way to target cancer cells at a molecular level and continue to be developed with a view to improve treatment and imaging in cancer medicine. AT11-LO, derived from the DNA sequence AT11, forms a single major parallel G4 conformation and exhibits an anti-proliferative activity similar to that of AT11 and AS1411 aptamers. By the other side, acridine orange derivatives represent a valuable class of G4 ligands. Herein, we evaluate AT11-LO G4 as a supramolecular carrier for the delivery of acridine ligands C₃, C₅ and C₈ to HeLa cancer cells. The CD titrations suggest no changes on chiroptical signal upon addition of an excess of ligands maintaining the parallel G4 topology and C₈ stabilizes the structure in more than 20 °C. All the ligands exhibit high affinity (micromolar range) towards AT11-LO G4, as well as the respective complexes against nucleolin (nanomolar range) suggesting that the ligands do not negatively affect the recognition of the nucleolin by AT11-LO G4. NMR studies showed that AT11-LO forms a G4 containing four G-tetrad layers. Ligand C₈ binds AT11-LO G4 through π - π stacking of the acridine moiety onto the top-tetrad with involvement of additional interactions with ligand's side chain and iodobenzene ring. In vitro, the complexes lowered the ligand's cytotoxicity towards non-malignant cells but have a weak inhibitory effect in HeLa cancer cells, except for AT11-LO-C₅ complex. All complexes are efficiently internalized into nucleolin-positive HeLa cells. Overall, these results suggest that AT11-LO can act as an aptamer by targeting nucleolin and a delivery system of cytotoxic ligands for cervical cancer.

Introduction

DNA aptamers can fold into specific secondary structures capable of specific high-affinity binding. This facilitates molecular recognition events in diagnostics and therapeutics [1]. Compared to antibodies, aptamers have low molecular weight, fast tumor penetration and are more rapidly cleared from the blood, showing rapid biodistribution, high stability, low toxicity and immunogenesis [2]. They can be synthesized reproducibly with low cost and modified easily with fluorophores for molecular recognition [3]. These advantages make

aptamers promising molecules for medical and pharmaceutical research, such as drug development, disease diagnosis, and targeted therapy. Several aptamers have been generated against cancer-related proteins, such as PDGF, VEGF, HER3, NFkB, tenascin-C, or PSMA [4,5].

The most advanced anti-cancer aptamer is AS1411, formerly developed as AGRO100 by Aptamera (Louisville, KY). It was reported to target nucleolin, a protein overexpressed in many tumor types that seems to be crucial for the anticancer activity of this oligonucleotide [6]. AS1411 has inhibitory activity against a large number of malignant cell lines, in the low micromolar range, and underwent a Phase II clinical trial [6]; however, the trial was terminated due to low response rates and suboptimal pharmacokinetics. The development of AS1411 derivatives capable of improving its *in vitro* and *in vivo* efficacy can be the strategy to reduce the gap between laboratory research and clinical applications. Previous studies identified a DNA sequence named AT11 and derived from AS1411, which formed a single G4 conformation; however, AT11 exhibited a similar anti-proliferative activity as AS1411 [7]. Modifications on AT11 by removing one thymine in the linker gives sequence AT11-Lo, forming a G4 conformation with higher thermal stability than that of AT11 (melting temperature increase of 8 °C).

In this study, we evaluated the biophysical properties and biological activities of the sequence AT11-Lo, in terms of its nucleolin targeting ability and as a potential drug delivery system to carry acridine orange ligands to cervical cancer cells.

Material and Methods

Oligonucleotides and ligands

AT11-Lo (5'-TGGTGGTGGTTGTTGGGTGGTGGTGGT-3') and Cy5-AT11-Lo were purchased from Eurogentec (Belgium) with HPLC-grade purification and purity of 98%. Stock solutions of about 1 mM were prepared using Milli-Q water and stored at -20 °C until used. For all experiments, the oligo was annealed by heating to 95 °C for 10 min and cooling down. The nucleolin (ref. RPC242Hu03) was purchased from Arium - Sistemas de Diagnóstico e Informática, Lda with HPLC-grade purification and purity > 95%. Hoechst 33342 was purchased from Thermo Scientific (USA). Ligands 10-(3-(4-iodobenzamide)propyl))-3,6-bis(dimethylamino) acridinium iodide (C₃), 10-(5-(4-iodobenzamide)pentyl))-3,6-bis(dimethylamino) acridinium iodide (C₅) and 10-(8-(4-iodobenzamide)octyl))-3,6-bis(dimethylamine) acridinium iodide (C₈) were synthesized as

previously described [8,9]. The ligands were dissolved in DMSO to obtain a 10 mM stock solution and further diluted in Milli-Q water.

Circular dichroism spectroscopy

CD spectra were acquired on a Jasco J-815 spectropolarizer equipped with a Peltier temperature controller (model CDF-426S/15). Before the CD spectrum measurement, the AT11-Lo was annealed in 20 mM potassium phosphate buffer containing 65 mM KCl by heating to 95°C for 10 min, then rapidly cooled on ice for 30 min. Spectra acquisition was performed in 1 mm quartz cuvette at AT11-Lo concentration of 10 µM. Spectral width was set to 220-340 nm, with a scan speed of 100 nm/min, 1 nm bandwidth, 1 s integration time over 4 averaged accumulations. During titrations, the required volume of ligand solution was added to the quartz cell.

The CD melting studies were performed by monitoring the maximum ellipticity wavelength through temperatures ranging from 20 to 110°C with a heating rate of 2 °C/min in 20 mM potassium phosphate buffer containing 65 mM KCl. Data were converted into fraction folded (θ) plots, fitted to a Boltzmann distribution using OriginPro 2016 and the melting temperatures (T_m) determined from the two-state transition model.

Fluorescence Titrations

Fluorescence titrations were conducted on a Horiba FluoroMax4 fluorometer (Japan) using a high-precision quartz suprasil cuvette (light path 10 mm × 4 mm).

All spectra were scanned with an integration time of 0.5 s, an emission and excitation slit width of 1.3 nm and step size of 1 nm. The fluorescence spectra were obtained between 530-700 nm and averaged over three scans. The excitation wavelength of ligands C₃, C₅ and C₈ was 498 nm. Titrations were performed with successive additions of concentrated and pre-annealed AT11-Lo at 500 µM in 20 mM potassium phosphate buffer containing 65 mM KCl to a ligand solution at 5 µM in the same buffer. Measurements were performed after 10 min equilibration time. Fluorescence data was converted into fraction of bound ligand (α) plots using the following equation:

$$\alpha = \frac{I - I_{\lambda}^{free}}{I_{\lambda}^{bound} - I_{\lambda}^{free}}$$

where I , is the fluorescence intensity at 530 nm at each ligand:DNA ratio and I_{free} and I_{bound} are the fluorescence intensity of the free and fully bound ligand, respectively. Data

points were then fitted to a hyperbolic function (OriginPro 8) and K_D values were determined from the following saturation binding model [10]:

$$\alpha = \frac{[DNA]}{K_D + [DNA]}$$

wh

ere α is the fraction of ligand bound and $[DNA]$ is the concentration of the DNA. When required, the saturation binding Hill slope model was used:

$$\alpha = \frac{[DNA]^n}{K_D + [DNA]^n}$$

wh

ere n is the Hill constant which describes cooperativity of ligand binding.

Cy5-labelled AT11-Lo at 1 μ M was titrated with nucleolin in the absence and presence of C_3 , C_5 and C_8 to evaluate the effect of ligand binding in the affinity of AT11-Lo towards the protein. The fluorescence emission spectra were obtained between 660 and 800 nm and averaged over three scans.

NMR spectroscopy

Standard zgesgp ^1H NMR spectrum ($P_1=10.44 \mu$ s, $SW=20.03$ ppm, $TD=64$ K, $AQ=2.73$ s, $D_1=2$ s, $NS=64$) was acquired on a 600 MHz Bruker Avance III spectrometer with QCI cryoprobe operating at proton frequency of 600 MHz at 310 K. AT11-Lo concentration was 200 μ M in 3 mm tube with a total volume of 150 μ L. The titrations were done with compound C_8 and pre-annealed AT11-Lo G4 at 250 μ M in 10% D_2O solution containing 20 mM potassium phosphate and 65 mM KCl.

2D ^1H - ^1H NOESY and ^1H - ^1H TOCSY NMR spectra were also recorded. Water suppression was achieved with double pulse-field gradient spin echo (dpgse) methods. NOESY spectra was acquired with a mixing time (d8) of 250 ms ($SW=18.03$ ppm, $TD=4096$, $NS=256$, $AQ=0.19$ s). Processed to 4096×2048 points using a qsine function ($SSB=2$). DIPSI-2 sequence was used for TOCSY experiments with a mixing time (d1) of 1.5 s ($SW=20.03$ ppm, $TD=2048$, $NS=128$, $AQ=0.09$ s). Processed to 2048×1024 points using a qsine function ($SSB=2$). All spectra were acquired and processed with the software Topspin 3.1, and further analysed with Sparky. The G4 protons assignment was performed based on the deposited NMR data of AT11 on BioMagResBank (accession code: 2N3M). Ligand assignment was made based on the intramolecular NOE signals, TOCSY and ^1H NMR spectra. Chemical shifts (δ) were measured in ppm.

Serum stability assay

Pre-annealed (95 °C for 10 min, following by a slowly cooling in ice for 30 min) AT11-Lo G4 and AT11-Lo G4-C₈ in 20 mM potassium phosphate containing 65 mM KCl were incubated in cell culture medium RPMI supplemented with 10 % (m/v) FBS and 1 % (m/v) penicillin-Streptomycin at final concentration of 30 µM. For each sample, 1 µL of AT11-Lo G4 or AT11-Lo G4-C₈ were mixed with 9 µL of RPMI medium and incubated for 0, 7, 24 and 48 h at 37 °C. 10 µL of each mixture was used for agarose gel electrophoresis, which was carried out using 1 % agarose gel. Additional experiments were conducted by incubating 30 µM oligonucleotide in 50% human serum (Sigma-Aldrich, EUA) for 3, 7, 12 and 48 h at 37 °C. Gels were stained with 0.01% Green Safe in 1×TAE buffer (Tris-acetate-EDTA) and visualized through UV light.

Cytotoxicity Assay

HeLa cancer cells were grown in DMEM medium, supplemented with 10% FBS and 1% streptomycin-penicillin antibiotic. NHDF cells were grown in RPMI medium, supplemented with 10% FBS, 1% streptomycin-penicillin antibiotic, 0.01 M HEPES, 0.02 M L-glutamine and 0.001 M sodium pyruvate. Upon confluence, cells were seeded on 48-well culture plates (HeLa at 0.5×10^4 /mL and NHDF 1×10^4 /mL) in 300 µL of medium. Ligands C₃, C₅ were added at 2 µM, C₈ at 1 µM and AT11-Lo aptamer at a concentration of 15 µM, in triplicate. The complexes aptamer-ligand were prepared by mixing pre-annealed AT11-Lo G4 (1 µM) with the corresponding amount of C₃, C₅ and C₈ and incubated for 10 min before being added to the wells. After 7 days of incubation at controlled humidified atmosphere at 37 °C and 5% CO₂, 50 µL of 5 mg/mL MTT solution was added to each well and further incubated for 1 h in HeLa or 4 h in NHDF cells. The resulting formazan crystals of HeLa cells were dissolved in 250 µL of DMSO and of NHDF in 100 µL of DMSO and the optical density (OD) was recorded at 570 nm. Cell viability was determined using wells without compounds as control.

Confocal fluorescence Microscopy

For the fluorescence confocal microscopy assays, human cancer HeLa cells were seeded at a density of 0.3×10^4 cells/mL and NHDF at 5×10^4 cells/mL in a treated µ-slide 8 well (IBIDI, Germany) and grown at 37 °C in a 95% air/5% CO₂ humidified atmosphere until 80% confluence. Cells were incubated with ligands C₃, C₅ and C₈ at 0.5 µM and AT11-Lo at 1 µM for 5 days. Cells were then washed three times with phosphate-buffered saline and nucleus was then stained with 1 µM nuclear probe Hoechst 33342 for 15 min. After that probe was washed off by rinsing with PBS three times and the cells were imaged using a

Zeiss AxioObserver LSM 710 microscope with 405, 488 and 564 nm laser excitation for Hoechst 33342, C₈ and Cy5-AT11-LO, respectively.

For nucleolin immunocytochemistry, HeLa and NHDF cells were incubated with the aptamer- ligand complex for 2 h. Nucleolin antibody (1:100) was added during 2 h and secondary antibody (Alexa Fluor 546®, 1:1000) for 1 h. Then, nucleus was stained with 1 μ M nuclear probe Hoechst 33342 for 15 min. The probe was washed off by rinsing with PBS three times and the cells were imaged using a Zeiss AxioObserver LSM 710 microscope. Images were processed with ImageJ software.

Results and Discussion

Aptamers can be effectively internalized into cancer cells via binding to target molecules with high affinity and specificity. For this reason, aptamers started to be envisaged as potential carriers of anticancer drugs, searching for more selective therapeutic effects without compromising the drug efficacy. In the current study, we have evaluated the sequence AT11-LO as a delivery system of acridine ligands C₃, C₅ and C₈ (**Figure 9.1**) and assessed its therapeutic effects on cervical cancer cells HeLa which have high expression levels of nucleolin [12].

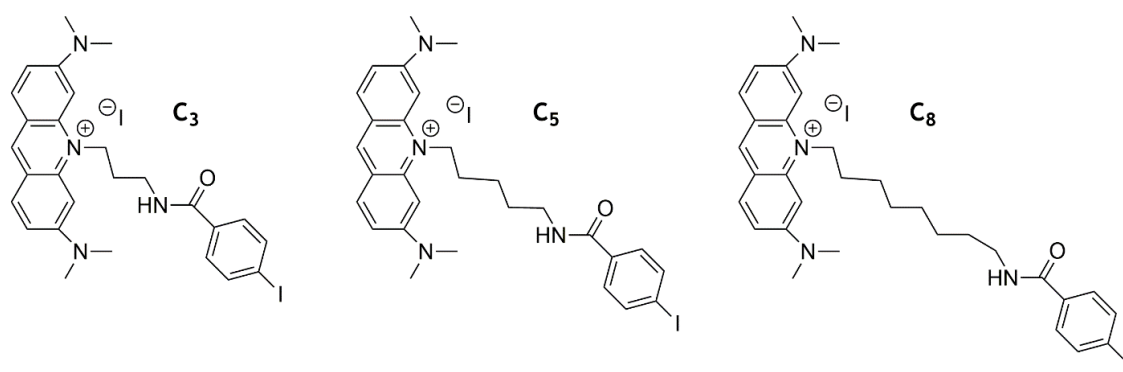


Figure 9.1 - Chemical structures of the acridine ligands C₃, C₅ and C₈.

Firstly, the G4 formation was evaluated by CD in 20 mM potassium phosphate buffer containing 65 mM KCl. The CD spectrum is presented in Figure S1 (**Appendix F**) showing the typical CD signature of parallel G4 topology with a positive peak at 260 nm and a negative peak at 240 nm.

Thereafter, the complexes formed by AT11-LO and acridine ligands were prepared using a supramolecular strategy and their binding and stability were determined by CD and fluorescence titrations. CD titrations were first performed to assess complex formation and/or structural modifications of the chiral G4 structure by ligand binding (**Figure 9.2**).

No relevant variations of AT11-Lo chiroptical signal were observed upon addition of an excess of ligands (4 equivalents relative to AT11-Lo), thus suggesting that the investigated ligands do not alter the parallel G4 topology. The results are presented in **Figure 9.2**. Increasing concentrations of ligand C₈ produced an enhancement of the band at 260 nm that might reflect the occurrence of intermolecular association and stabilization of AT11-Lo.

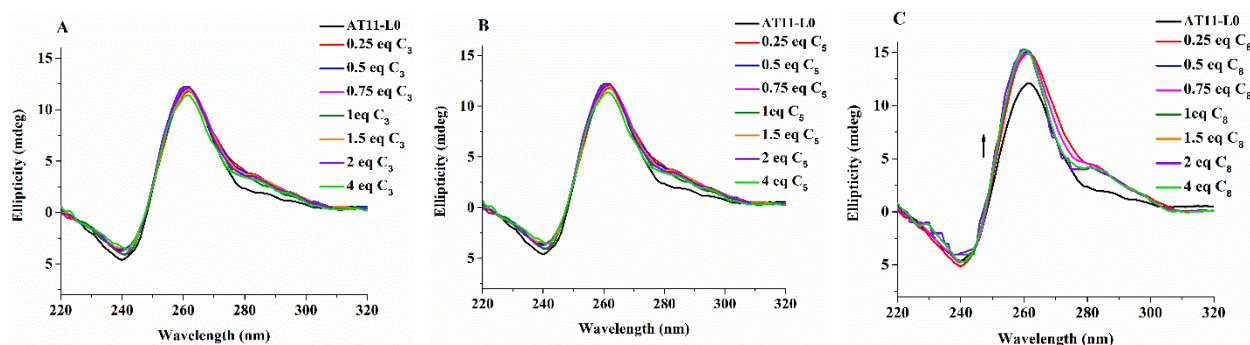


Figure 9.2 - CD spectra of AT11-Lo in the presence of increasing amounts of ligands (A) C₃ (B) C₅ and (C) C₈. Spectra acquisition was performed in 20 mM potassium phosphate buffer containing 65 mM KCl.

Then, the AT11-Lo stabilizing properties of the ligands were evaluated by measuring the ligand-induced change in the melting temperature (ΔT_m) of the pre-folded G4 (Figure S2). Firstly, the salt conditions of AT11-Lo G4 was adjusted for the melting temperature (T_m) which was 51 °C in 20 mM potassium phosphate buffer containing 65 mM KCl significantly higher than AT11 G4 (T_m = 42 °C, under the same conditions) due to the modification introduced in the sequence (removal of one thymine at position 16). Upon addition of 4 molar equivalents of acridine orange ligands, C₈ has a much higher stabilization ability with a ΔT_m of 23.9 °C; the ΔT_m values induced by C₃ and C₅ at the same ligand:G4 ratio were 4.2 °C and 7.7 °C, respectively. The effect of the side chain length of the ligands clearly affect the thermal stability of the G4 structure.

Fluorescence spectroscopy was used to further determine the binding strength between the ligands C₃, C₅ and C₈ and pre-folded AT11-Lo G4. The emission spectra of ligands C₃, C₅ and C₈ were recorded in 20 mM potassium phosphate buffer containing 65 mM KCl with excitation at 498 nm. The fluorescence titration curves are presented in **Figure 9.3**. Free ligands exhibited very low fluorescence. With every addition of the pre-folded AT11-Lo G4, the emission of the ligands increased significantly, namely for ligand C₈. This behaviour may be related to the protection of the ligand from quenching by polar water molecules, as the hydrophobic environment inside the AT11-Lo G4 favors the recovery of ligands fluorescence [13].

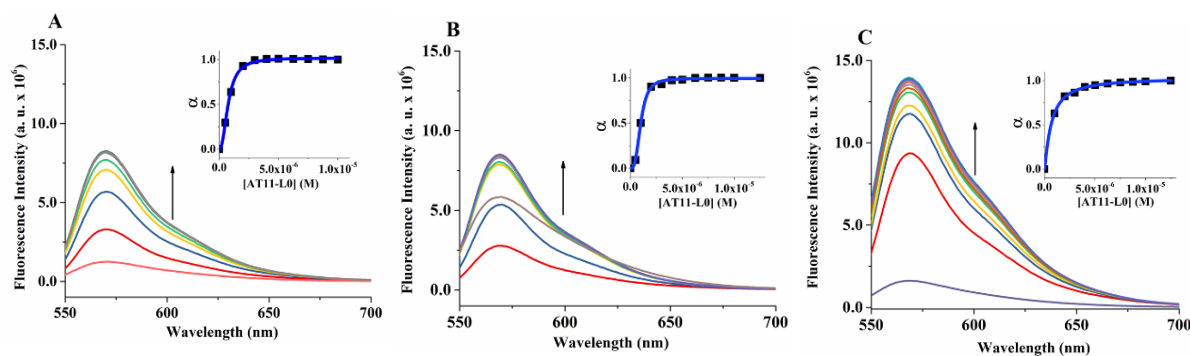


Figure 9.3 - Fluorescence titration spectra of ligands (A) C_3 , (B) C_5 and (C) C_8 with AT11-Lo in 20 mM potassium phosphate buffer containing 65 mM KCl, and relative enhancement of the fluorescence intensities of ligands vs. the increasing concentration of AT11-Lo G4; λ_{ex} : 498 nm.

The results from the fluorescence titrations were then fitted to a saturation binding model, which allowed the determination of the dissociation constants K_D for C_3 , C_5 and C_8 . The calculated values obtained were 7.7×10^{-7} ($R^2 = 0.99$), 1.0×10^{-6} ($R^2 = 0.99$) and 6.9×10^{-7} M ($R^2 = 0.99$) ($SD \leq 2.6 \times 10^{-8}$ M), respectively, indicating that the ligands have a high affinity for AT11-Lo G4. The Hill coefficient (n) which is a measure of cooperativity in a binding process was 2.27 ± 0.11 , 3.13 ± 0.16 and 1.22 ± 0.09 for C_3 , C_5 and C_8 , respectively. This reveals that the binding of C_3 and C_5 to AT11-Lo G4 exhibits positive cooperativity ($n > 1$); the Hill coefficient calculated for C_8 is closer to 1 and indicates a more probable independent binding. Additionally, the n coefficient can be interpreted as the average number of bound ligands per G4, thus suggesting a 2:1, 3:1 and 1:1 ligand:DNA ratio for C_3 , C_5 and C_8 , respectively.

To study the interaction between the acridine orange ligands and AT11-Lo G4, NMR experiments were performed using ligand C_8 (**Figure 9.4**) because it promoted the higher thermal stabilization ($\Delta T_m = 23.9$ °C), showed a high binding affinity ($K_D = 6.9 \times 10^{-7}$ M) and an independent 1:1 binding stoichiometry as suggested by the Hill coefficient. AT11-Lo derives from AT11 by eliminating a thymine residue in the linker between the two propeller-type parallel-stranded subunits (**Figures 9.4A** and **9.4B**) [7]. This modification in the linker is not expected to affect the overall conformation of the G-tetrads, as indicated by the similarity of the ^1H NMR imino region and the one reported by Phan and colleagues [7], showing 16 peaks in the range between 11.0–11.7 ppm (**Figure 9.4C**) which is consistent with two subunits of two G-tetrads each [7]. The imino signals were assigned based on the reported assignment for AT11 G4 structure [7]. Indeed, similar behavior was already described for other G4 structures upon a single thymine deletion [14,15]. Upon titration with the ligand C_8 , there is a significant upfield shift of most imino signals pointing out for the ligand binding through π - π stacking interactions, as expected

due to the presence of the acridine moiety [8,16]. Particularly, the top-tetrad G5-H8 and G8-H8 atoms undergo a slow exchange regime between the free and bound state which suggests a strong binding [17], in agreement with the fluorescence spectroscopy experiments. T1-H6 which caps the top-tetrad is shifted downfield, supporting that the G2-G5-G8-G12 5'-tetrad is the principal binding site. Besides, according to the nucleotide resonances, a 1:1 ligand:DNA binding ratio is indicated in agreement with the determined Hill coefficient. Several broad signals belonging to the ligand acridine moiety and NH group arose in the aromatic region (7-8.5 ppm) (**Figure 9.4C and 9.4D**). The signals shown on the top spectra (**Figure 9.4D**) were assigned based on 2D NOESY and TOCSY spectra (see below for further discussion).

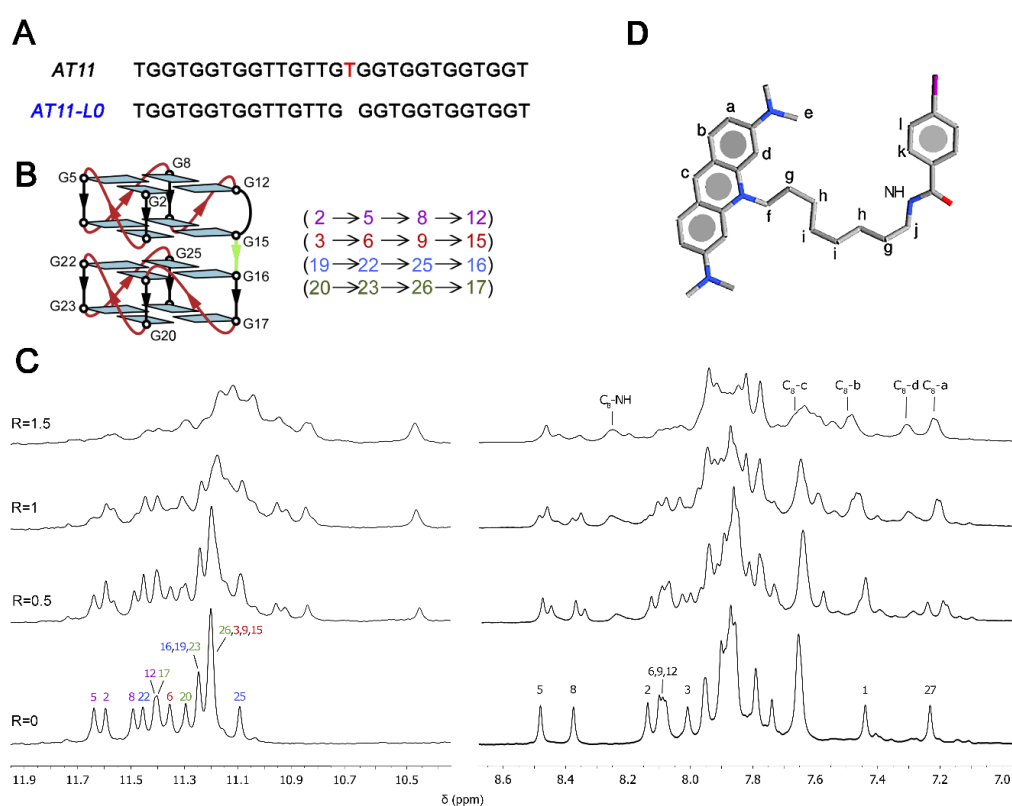


Figure 9.4 - A) Sequences of the AT11 and AT11-Lo oligonucleotides. (B) Schematic representation of AT11-Lo G4 showing the G-tetrad alignments and folding topology. Red arrows denote thymine loops while the green arrow highlights the linker between the two subunits where a thymine residue was deleted to derive AT11-Lo from AT11. Adapted with permission from reference 7. C) Imino H1 and aromatic H8/H6 region of the ¹H NMR spectra of AT11-Lo, including the spectra obtained upon titration with ligand C₈ at increasing ligand:DNA ratios (R). Labelled residues were indicated with their corresponding residue number. Top tetrad residues are shown in purple, middle tetrads in red and blue, and bottom tetrad in green. C₈ protons from the acridine moiety and NH group are shown on the top spectrum as broad signals. These were assigned based on NOESY and TOCSY spectra. (D) Chemical structure of ligand C₈ with respective protons indicated by letters.

Additional 2D NOESY NMR spectra were acquired to further elucidate the binding mode. Despite the overlapping signals, several broad cross-peaks indicating exchange were observed between the C₈ protons and those from the AT11-Lo G4 structure (**Figure 9.5**).

The ligand protons were assigned based on the previously reported data of the complex formed by C₈ and KRAS-22RT G4 [8], the compound intramolecular NOEs (Figure S3, S4 and S6) and TOCSY experiments (Figure S5).

The intramolecular NOEs between the aliphatic protons of the ligand side-chain and the iodobenzene moiety (protons i, h and protons k, l) suggest that the ligand is locked in a crescent-shape conformation (**Figure 9.4C**) which is highly favourable for G-tetrad overlap as observed for other ligands in the literature [18-20]. The observed intermolecular cross-peaks showed that the ligand's side chain interacts with the imino groups from the G4 structure (**Figure 9.5**), namely through the aliphatic protons i, h, f and also the iodobenzene moiety (protons k). Similar binding mode was already observed for C₈ with other G4 structures [8] and emphasizes the importance of the side chain functionality. These protons appear well shifted upfield due to shielding of protons upon stacking interactions.

NOEs between the iodobenzene protons are also observed with thymine residues CH₃ groups (Figure S4), presumably with T1 capping residue, further suggesting end stacking. The large shifting and overlapping of the imino region hindered the unambiguous assignment of the NOE signals but there is a clear evidence of the ligand interaction with the G-tetrads. Additional NOE cross-peaks were observed between G8-H8, T1-H6 and C₈ NH and acridine proton b, respectively, (Figure S6) in agreement with the observed chemical shifts thus confirming the π - π stacking of the acridine moiety onto the top-tetrad.

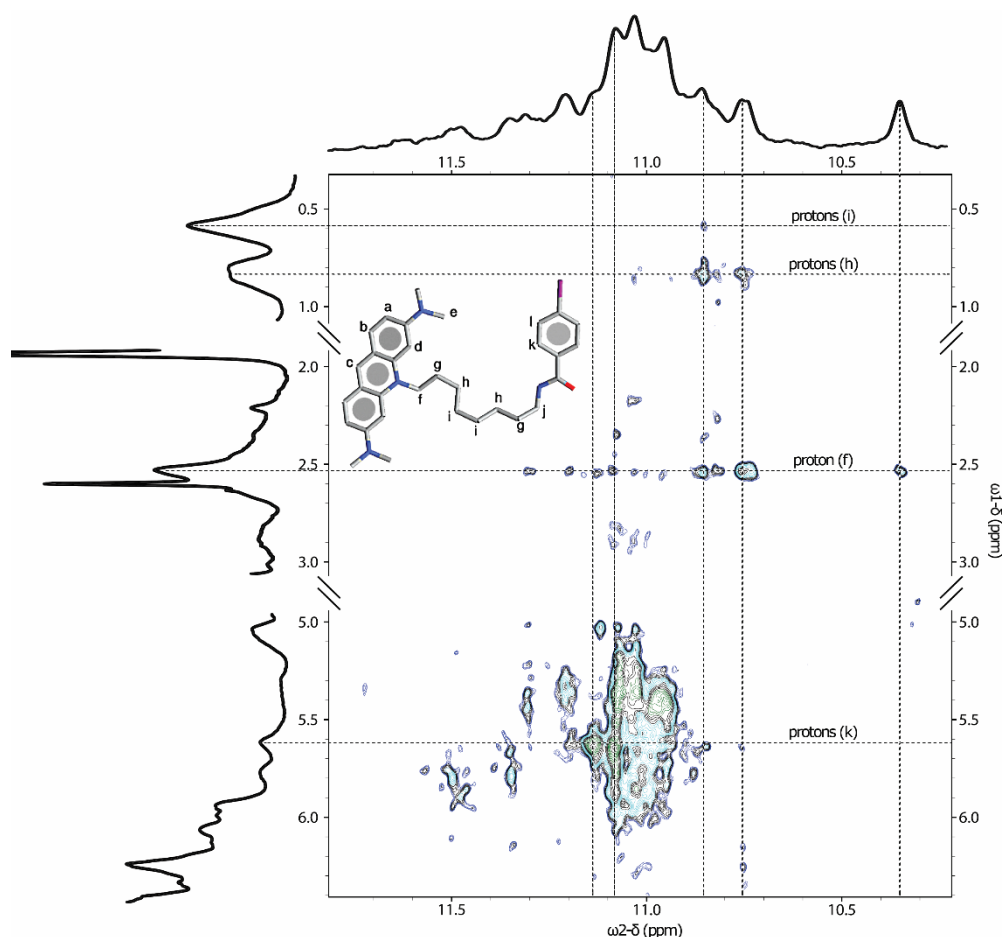


Figure 9.5 - Expanded region of the 2D NOESY spectrum showing the intermolecular NOEs (highlighted with red arrows) between the AT11-Lo imino protons (not assigned due to spectral overlap) and ligand C₈, namely the alkylamide chain aliphatic protons i, h and f of the ligand, and the iodobenzene k protons (inset).

Degradation by nucleases in serum is a possible drawback when aptamers are administered intravenously. To determine the stability of AT11-Lo G4, serum stability assay was carried out by incubating the AT11-Lo G4 in culture medium supplemented with 10 % FBS (**Figure 9.6A**). Additionally, in order to obtain a more representative data for in vivo applications, serum stability experiments were also conducted in 50% (**Figure 9.6B**) and 100% human serum. In full human serum, due to molecular crowding of the highly concentrated solutions, the samples were not resolved properly on the agarose gel (data not shown), as previously observed in a similar study [21]. Since AT11-Lo G4-C₈ complex exhibited a noteworthy antiproliferative activity, its resistance in biological environments was evaluated between 3 h and 48 h of incubation. In culture medium supplemented with 10 % FBS, the electrophoretic mobility of the DNA bands remained unchanged for up to 48 h of incubation with an increase in band intensities when C₈ is complexed with AT11-Lo G4 (**Figure 9.6A**). These results indicate that the free AT11-Bo G4 is stable in culture medium supplemented with 10 % FBS which may be explained by its higher thermal stability when compared to AT11 G4. In the presence of C₈ the stability was not affected. The results shown in **Figure 9.6B** suggest that after 3 h incubation in

50% serum, free AT11-LO G4 is readily degraded as seen by the absence of any band below 500 pb. In the presence of ligand C₈, the band corresponding to the AT11-LO G4-C₈ complex was observed (size confirmed by the control experiment without serum, C⁺) suggesting that the ligand significantly increases the serum stability of the aptamer up to 48 h.

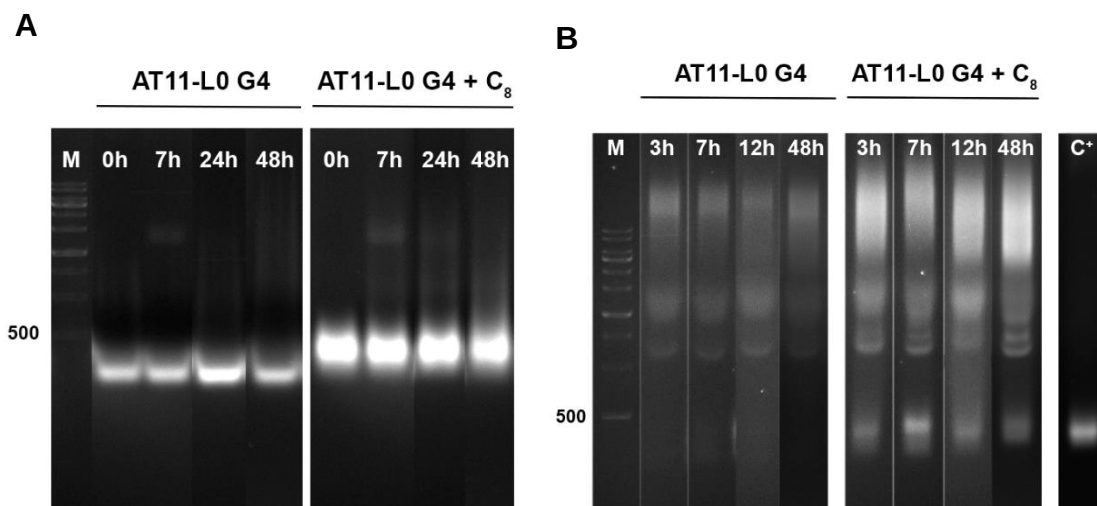


Figure 9.6 - Serum stability of AT11-LO G4 and AT11-LO G4-C₈ in RPMI medium containing (i) 10 % FBS and (ii) 50% human serum and resolved in a 1% agarose gel electrophoresis. In ii) the large smearing seen on the top of each lane corresponds to the serum proteins aggregation. On the right, control lane (C⁺) of AT11-LO G4-C₈ complex in the absence of human serum.

Previous reports showed that AS1411 aptamers have antiproliferative activity against various cancer cell lines [6,21]. The acridine ligands C₃, C₅ and C₈ also have antitumor activity, including against cervical cancer cell lines [8]. The *in vitro* cytotoxic effects of the complexes formed by AT11-LO with the acridine ligands, using a supramolecular strategy, were tested on human cervical cancer cells (HeLa) vs. NHDF cells, using the MTT assay. Overall, the AT11-LO-C₃, AT11-LO-C₅ and AT11-LO-C₈ complexes showed lower cellular toxicity in non-malignant cells (**Figure 9.7**) than the free acridine ligands C₃, C₅ and C₈. AT11-LO has a weak inhibitory effect in HeLa cancer cells, needing more than 5 days to induce a slight effect on cell viability, even at 15 μ M. This result is similar to that reported for the AS1411 aptamer [22]. The complexes AT11-LO-C₃ (68% mean viability). and AT11-LO-C₈ (53 % mean viability). showed reduced and similar cytotoxicity towards HeLa and NHDF cells. In the case of AT11-LO-C₅ it presented a similar cytotoxicity for NHDF cells (64 % mean viability), however in the case of HeLa cells its cytotoxicity was significantly higher (48% mean viability).

The lower cytotoxicity of C₃, C₅ and C₈ towards NHDF cells when complexed with AT11-LO may be due to the mechanism involved in the selective accumulation of oligonucleotide in non-malignant cells because they are being constantly eliminated by efflux or exocytosis

and the ligands cannot perform their cytotoxic effect [6,23]. Additionally, as proposed by Bates and colleagues, due to the low levels of nucleolin in non-malignant cells, the complexes may be entrapped in endosomes or suffer from lysosomal degradation [23]. Relatively to the cancer cells, an attenuation of the cell toxicity caused by the acridine ligands was also observed after AT11-Lo complex formation. The more pronounced cytotoxic effect on HeLa than on NHDF cells for AT11-Lo-C₅ may be due to the lower affinity of C₅ to AT11-Lo compared with the other ligands, being released more easily, thus exerting more efficiently its effect. Overall, the MTT results showed that the complex formed by AT11-Lo and C₅ reduced the most the ligand cytotoxicity towards non-malignant cells (about 20% reduction) while still offering a lower IC₅₀ value towards cancer cells, improving the free aptamer toxicity.

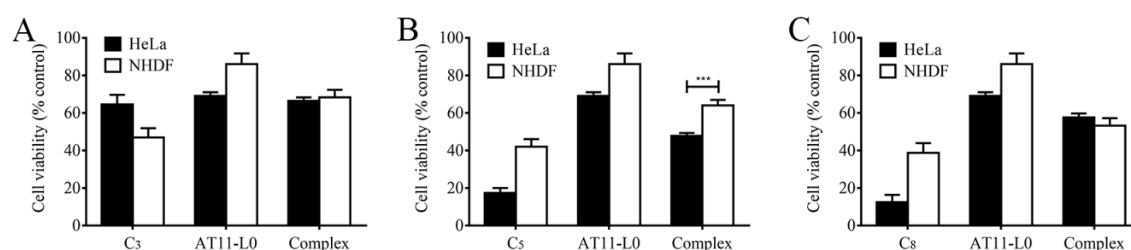


Figure 9.7 - Percentage of viable HeLa and NHDF cells after incubation with free AT11-Lo aptamer (15 μM) or conjugated with (A) C₃ at 2 μM, (B) C₅ at 2 μM and (C) C₈ at 1 μM for 7 days. ***p<0.001.

Since our objective is to deliver the acridine orange ligands, as potential anticancer drugs, by using AT11-Lo as a supramolecular carrier that binds to nucleolin, we assessed if the incorporation of the ligand alters the nucleolin-binding properties of the aptamer. We determined the equilibrium dissociation constants (K_D) for the binding of Cy5-labelled AT11-Lo to nucleolin, in the presence or absence of C₃, C₅ and C₈. These determinations were done using fluorescence spectroscopy and by adding increasing concentrations of nucleolin to solutions of the aptamer, complexed or not with the acridine ligands. The saturation binding plots are presented in **Figure 9.8**. The K_D values obtained are 3.73 ± 1.68 nM, 8.84 ± 2.66 nM, 8.03 ± 2.89 nM and 25.89 ± 8.43 nM for free AT11-Lo G4 and bound to C₃, C₅ and C₈, respectively. The AT11-Lo G4 displayed K_D values in the nanomolar range indicating high affinity to nucleolin, being lower in the case of AT11-Lo complexed with the acridine ligands, particularly in the case of the C₈ complex. The binding strength to nucleolin found in this work for AT11-Lo is higher than that reported in the literature for AS1411 ($K_D = 16.36 \pm 10.30$ nM) and APTA-12 ($K_D = 14.37 \pm 8.93$ nM) aptamers [24]. Altogether, these data suggest that the ligands do not significantly affect the recognition of nucleolin by AT11-Lo G4. Although the AT11-Lo-ligand complexes presented a slightly lower affinity to nucleolin than the free AT11-Lo G4, the measured K_D

values are still in the high affinity range, meaning that the complexes can recognize and bind nucleolin.

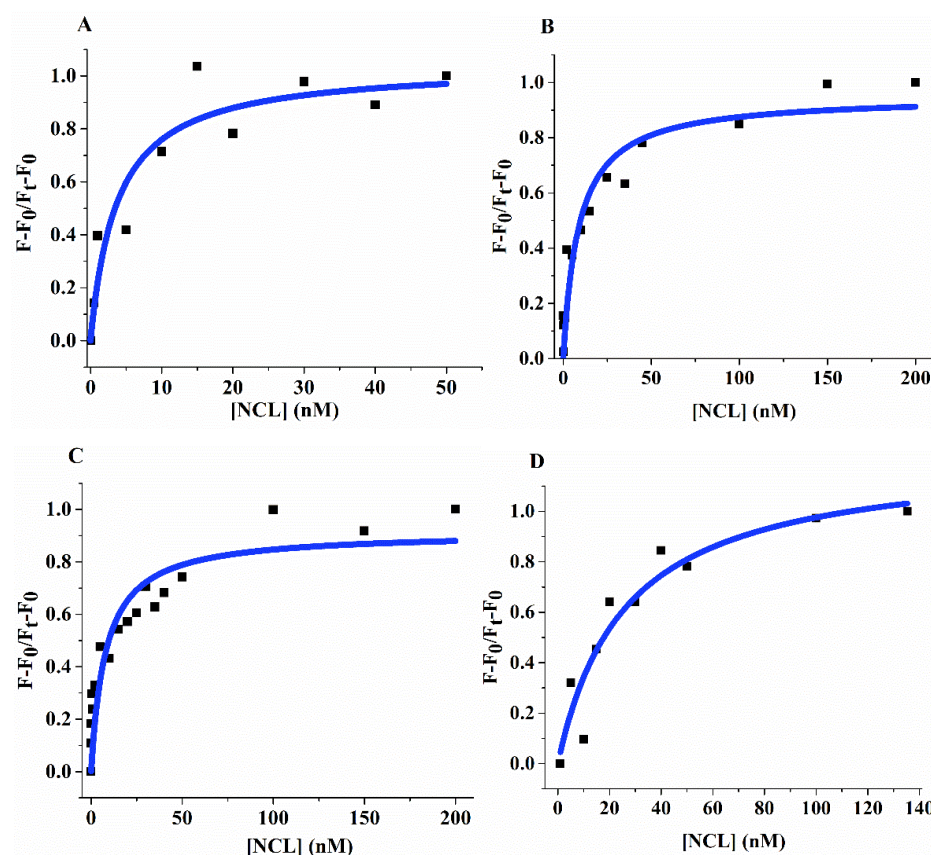


Figure 9.8 - Saturation binding plots obtained by fluorescence spectroscopy and fitted to Michaelis-Menten model. Fluorescence intensity of (A) free AT11-Lo G4 labeled with Cy5 or complexed with 2 eq of (B) C₃, (C) C₅ and (D) C₈ obtained after addition of increasing amounts of nucleolin.

To investigate the binding specificity of AT11-Lo, HeLa cancer cells were imaged by confocal microscopy after incubation with Cy5-labeled AT11-Lo. Non-malignant NHDF cells were used as a negative control to determine nonspecific binding. Considering the biophysical results, we selected complex AT11-Lo-C₈ for the confocal microscopy studies and evaluated the possibility of selective delivery of C₈ to cancer cells through AT11-Lo nucleolin targeting. The primary anti-nucleolin antibody conjugated with the secondary antibody AlexaFluor 546® were used to localize cell surface nucleolin. The images presented in **Figure 9.9A** clearly showed that Cy5-AT11-Lo and Cy5-AT11-Lo-C₈ were efficiently internalized into nucleolin-positive HeLa cells after an incubation period of 2 h, showing a non-random distribution and localization of nucleolin (blue) and AT11-Lo (red); however, the staining of NHDF-surface nucleolin is not observed as expected (**Figure 9.9B**).

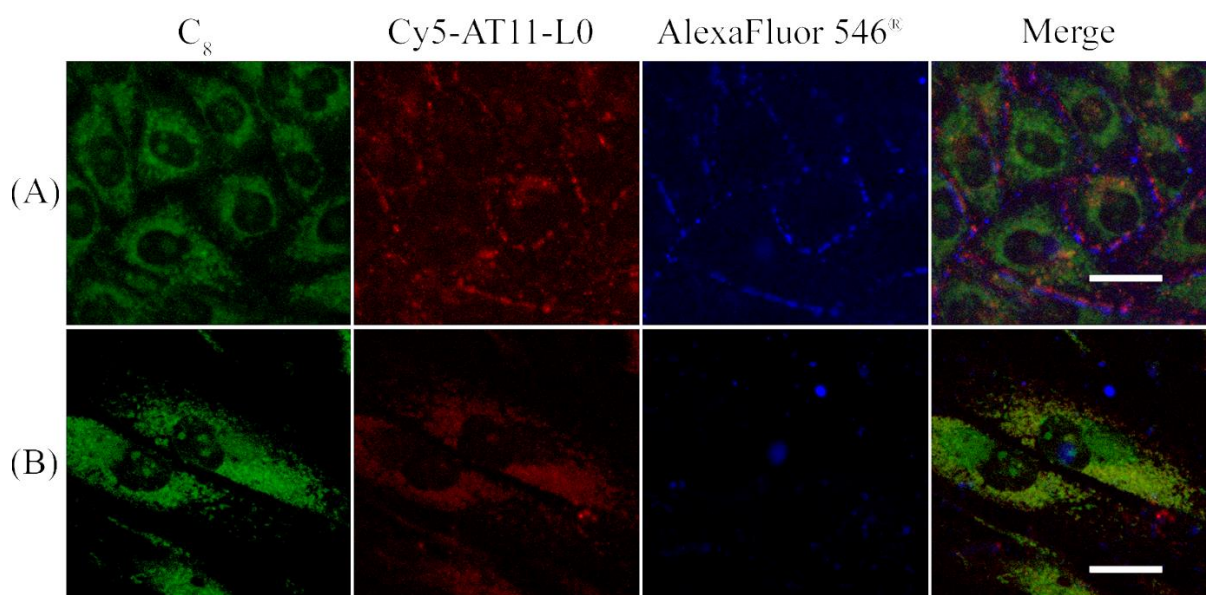


Figure 9.9 - Confocal fluorescence images of (A) HeLa and (B) NHDF cells incubated with Cy5-AT11-LO- C_8 after 2h at 37°C. Nucleolin is stained with AlexaFluor 546®, C_8 emits green fluorescence and Cy5- AT11-LO is shown in red. Scale bar: 25 μ m. Brightness was adjusted for all images equally.

The localization of Cy5-AT11-LO and Cy5-AT11-LO- C_8 complex was evaluated in the mentioned cell lines. As presented in **Figure 9.10A**, the yellow/orange staining of Cy5-AT11-LO- C_8 in the cytoplasm of HeLa cells suggests that the complex remains intact in intracellular environment. Additionally, compared to free AT11-LO (Figure S7), the complex has a slightly higher fluorescence in the cytoplasm that may be due to a preferential endocytosis of the AT11-LO when complexed with C_8 possibly due to the stabilization of its G4 structure. Moreover, C_8 can be seen in the nucleoli in a free state which may suggest that after complex internalization, C_8 can be released to the same compartment (nucleoli) as previously reported [8].

After 5 days of AT11-LO- C_8 incubation, the changes observed in the cancer cell morphology were consistent with irregular nucleus and vacuolization, which suggests that methuosis is occurring, similarly to the effects produced by the AS1411 aptamer in several cancer cells [23]. Similar findings were observed for NHDF cells (**Figure 9.10B**). Cy5-AT11-LO and Cy5-AT11-LO- C_8 complex were also internalized, the overlapping of the emitted C_8 fluorescence (green) and Cy5-AT11-LO (red) fluorescence suggests that the complex is still maintained intracellularly (yellow) during cell internalization and trafficking. Emission of C_8 fluorescence can be seen in the nucleoli but not Cy5-AT11-LO fluorescence, which indicates that C_8 is accumulating in the free form in this organelle. A similar effect was previously observed in other non-malignant cell line (Hs27 cells) with AS1411 aptamer [23]. The authors suggested that this might happen because normal cells present high levels of aptamer uptake due to endosomal entrapment or lysosomal degradation [23].

However, the aptamer can be gradually eliminated from normal cells by efflux or exocytosis [6]. In our case, cells are constantly in contact with Cy5-AT11-Lo and Cy5-AT11-Lo-C₈, which may mask the above-mentioned effects.

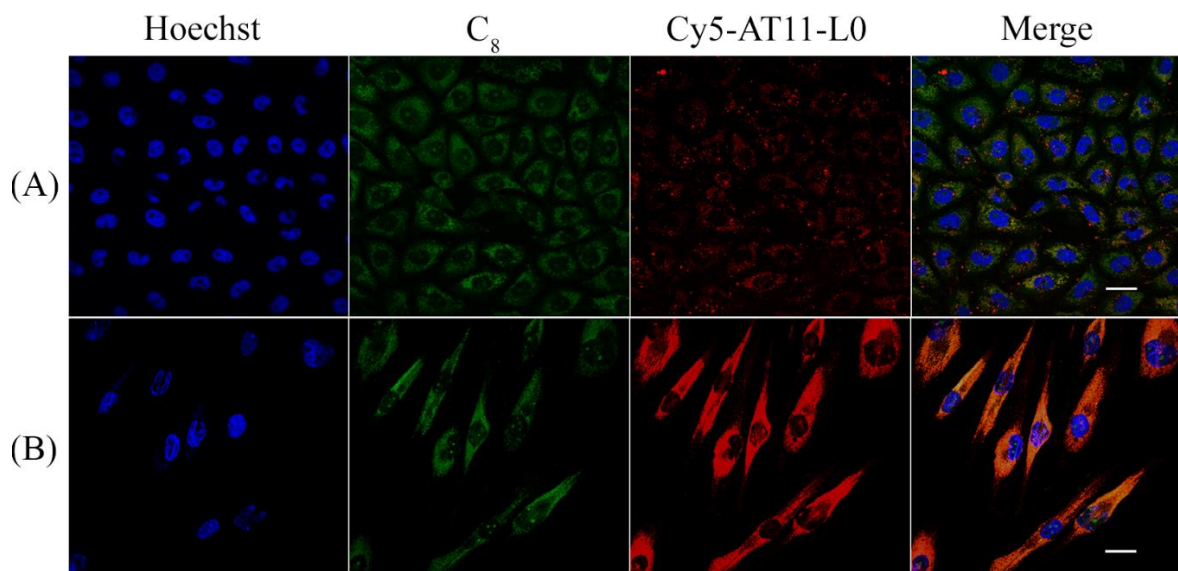


Figure 9.10 - Confocal microscopy images of (A) HeLa cells and (B) NHDF cells incubated with Cy5-AT11-Lo-C₈ complex for 5 days. Cell nuclei are stained with Hoechst 33342 (blue), C₈ emits green fluorescence and Cy5-AT11-Lo is shown in red. Overlapping of the two stains, observed as yellow/orange regions, can be seen in the merged images of the cells. Scale bar: 25 μ m.

Conclusion

The AT11-Lo sequence, derived from AT11 by elimination of the thymine residue between the two G4 subunits, retains a G4 parallel topology. We have shown that the binding of AT11-Lo G4 by the acridine derivatives C₃, C₅ and C₈ involves stack on the external top G-quartet and interactions with the long alkylamide chain and iodobenzene ring, which sustain and ameliorate the G4 binding. C₃, C₅ and C₈ strongly associate with AT11-Lo G4 with K_D values in the micromolar range. Additionally, the acridine derivatives are able to thermally stabilize the AT11-Lo G4, possible enhancing its properties as a G4-forming oligonucleotide and its cellular internalization which is an issue in the development of aptamers as therapeutics components. The binding affinity between AT11-Lo-C₃, AT11-Lo-C₅ and AT11-Lo-C₈ and nucleolin was determined by fluorescence spectroscopy being obtained K_D values in the nanomolar range, suggesting that ligands complexation with AT11-Lo G4 did not affect the recognition of nucleolin. In line with these results, we have found that the AT11-Lo-C₈ complex is efficiently internalized into nucleolin-positive HeLa cells. The MTT assays have shown that complexation of C₃, C₅ and C₈ with AT11-Lo G4

decreases its cytotoxic activity in NHDF cells and also in human cervical cancer HeLa cells, although in a less extent in the case of the AT11-Lo-C₅ complex.

Overall, our results suggest that AT11-Lo is a G4-forming oligonucleotide, being a promising anticancer delivery system via nucleolin targeting.

Acknowledgments

J. Carvalho acknowledges a doctoral fellowship grant from the FCT – Foundation for Science and Technology ref. SFRH/BD/122953/2016. Jéssica Lopes-Nunes acknowledges the fellowship reference UTAP-EXPL/NTec/0015/2017. This work was supported by project “Ações Integradas Luso-Francesas” ref. TC-15/17, FCT project ref. IF/00959/2015 financed by Fundo Social Europeu and Programa Operacional Potencial Humano, FCT project ref. UID/Multi/04349/2013, project Fundação Luso-Americana (FLAD) Healthcare 2020 ref. 45/2018, MIT Portugal project BIODEVICE ref. MIT-EXPL/BIO/0008/2017 and UTAustin FCT project DREAM ref. UTAP-EXPL/NTec/0015/2017. This work was also supported by POCI – COMPETE 2020 – Operational Programme Competitiveness and Internationalisation in Axis I – Strengthening research, technological development and innovation (project POCI-01-0145-FEDER-007491). C2TN/IST authors gratefully acknowledge the FCT support through the UID/Multi/04349/2013 project.

References

- [1] Ni, X., Castanares, M., Mukherjee, A., and Lupold, S. E. (2011) Nucleic acid aptamers: clinical applications and promising new horizons. *Current Medicinal Chemistry* 18, 4206–4214. <https://doi.org/10.2174/092986711797189600>.
- [2] Blind, M., and Blank, M. (2015) Aptamer Selection Technology and Recent Advances. *Molecular Therapy - Nucleic Acids* 4, e223. <https://doi.org/10.1038/mtna.2014.74>.
- [3] Shangguan, D., Cao, Z. C., Li, Y., and Tan, W. (2007) Aptamers evolved from cultured cancer cells reveal molecular differences of cancer cells in patient samples. *Clinical Chemistry* 53, 1153–1155. <https://doi.org/10.1373/clinchem.2006.083246>.
- [4] Lupold, S. E., Hicke, B. J., Lin, Y., and Coffey, D. S. (2002) Identification and characterization of nuclease-stabilized RNA molecules that bind human prostate cancer cells via the prostate-specific membrane antigen. *Cancer Research* 62, 4029–4033. <https://doi.org/10.1158/0008-5472.CAN-12-2152>.

- [5] Ireson, C. R., and Kelland, L. R. (2006) Discovery and development of anticancer aptamers. *Molecular Cancer Therapeutics* 5, 2957–2962. <https://doi.org/10.1158/1535-7163.MCT-06-0172>.
- [6] Bates, P. J., Reyes-Reyes, E. M., Malik, M. T., Murphy, E. M., O'Toole, M. G., and Trent, J. O. (2017) G-quadruplex oligonucleotide AS1411 as a cancer-targeting agent: Uses and mechanisms. *Biochimica et Biophysica Acta - General Subjects* 1861, 1414–1428. <https://doi.org/10.1016/j.bbagen.2016.12.015>.
- [7] Do, N. Q., Chung, W. J., Truong, T. H. A., Heddi, B., and Phan, A. T. (2017) G-quadruplex structure of an anti-proliferative DNA sequence. *Nucleic Acids Research* 45, 7487–7493. <https://doi.org/10.1093/nar/gkx274>.
- [8] Carvalho, J., Pereira, E., Marquevielle, J., Campello, M. P. C., Mergny, J.-L., Paulo, A., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2018) Fluorescent light-up acridine orange derivatives bind and stabilize KRAS-22RT G-quadruplex. *Biochimie* 144, 144–152. <https://doi.org/10.1016/j.biochi.2017.11.004>.
- [9] Pereira, E., Do Quental, L., Palma, E., Oliveira, M. C., Mendes, F., Raposinho, P., Correia, I., Lavrado, J., Di Maria, S., Belchior, A., Vaz, P., Santos, I., and Paulo, A. (2017) Evaluation of Acridine Orange Derivatives as DNA-Targeted Radiopharmaceuticals for Auger Therapy: Influence of the Radionuclide and Distance to DNA. *Scientific Reports* 7, 42544. <https://doi.org/10.1038/srep42544>.
- [10] Pollard, T. D. (2010) A guide to simple and informative binding assays. *Molecular biology of the cell* 21, 4061–4067. <https://doi.org/10.1091/mbc.E10-08-0683>.
- [11] Lavrado, J., Brito, H., Borralho, P. M., Ohnmacht, S. a., Kim, N.-S., Leitão, C., Pisco, S., Gunaratnam, M., Rodrigues, C. M. P., Moreira, R., Neidle, S., and Paulo, A. (2015) KRAS oncogene repression in colon cancer cell lines by G-quadruplex binding indolo[3,2-c]quinolines. *Scientific Reports* 5, 9696. <https://doi.org/10.1038/srep09696>.
- [12] Sharma, V. R., Thomas, S. D., Miller, D. M., Rezzoug, F., Sharma, V. R., Thomas, S. D., Miller, D. M., and Rezzoug, F. (2018) Nucleolin Overexpression Confers Increased Sensitivity to the Anti-Nucleolin Aptamer , AS1411 Nucleolin Overexpression Confers Increased Sensitivity to the Anti-Nucleolin. *Cancer Investigation* 0, 1–17. <https://doi.org/10.1080/07357907.2018.1527930>.
- [13] Maji, B., Kumar, K., Kaulage, M., Muniyappa, K., and Bhattacharya, S. (2014) Design and synthesis of new benzimidazole-carbazole conjugates for the stabilization of human

telomeric DNA, telomerase inhibition, and their selective action on cancer cells. *Journal of Medicinal Chemistry* 57, 6973–6988. <https://doi.org/10.1021/jm500427n>.

[14] Greco, M. L., Kotar, A., Rigo, R., Cristofari, C., Plavec, J., and Sissi, C. (2017) Coexistence of two main folded G-quadruplexes within a single G-rich domain in the EGFRpromoter. *Nucleic Acids Research* 45, 10132–10142. <https://doi.org/10.1093/nar/gkx678>.

[15] Liy, Z., Lechy, C. J., and Phan, A. T. (2014) Sugar-modified G-quadruplexes: Effects of LNA-, 2oF-RNA-and 2oF-ANA-guanosine chemistries on G-quadruplex structure and stability. *Nucleic Acids Research* 42, 4068–4079. <https://doi.org/10.1093/nar/gkt1312>.

[16] Carvalho, J., Nottelet, P., Mergny, J.-L., Queiroz, J. A., Salgado, G. F., and Cruz, C. (2017) Study of the interaction between indole-based compounds and biologically relevant G-quadruplexes. *Biochimie* 135, 186–195. <https://doi.org/10.1016/j.biochi.2017.02.005>.

[17] Becker, W., Bhattiprolu, K. C., Gubensäk, N., and Zangger, K. (2018) Investigating Protein–Ligand Interactions by Solution Nuclear Magnetic Resonance Spectroscopy. *ChemPhysChem* 19, 895–906. <https://doi.org/10.1002/cphc.201701253>.

[18] Lin, C., Wu, G., Wang, K., Onel, B., Sakai, S., Shao, Y., and Yang, D. (2018) Molecular Recognition of the Hybrid-2 Human Telomeric G-Quadruplex by Epiberberine: Insights into Conversion of Telomeric G-Quadruplex Structures. *Angewandte Chemie - International Edition* 57, 10888–10893. <https://doi.org/10.1002/anie.201804667>.

[19] Dai, J., Carver, M., Hurley, L. H., and Yang, D. (2011) Solution structure of a 2:1 quindoline-c-MYC G-quadruplex: Insights into G-quadruplex-interactive small molecule drug design. *Journal of the American Chemical Society* 133, 17673–17680. <https://doi.org/10.1021/ja205646q>.

[20] De Cian, A., DeLemos, E., Mergny, J.-L., Teulade-Fichou, M.-P. P., and Monchaud, D. (2007) Highly efficient G-quadruplex recognition by bisquinolinium compounds. *Journal of the American Chemical Society* 129, 1856–1857. <https://doi.org/10.1021/ja067352b>.

[21] Bates, P. J., Laber, D. a., Miller, D. M., Thomas, S. D., and Trent, J. O. (2009) Discovery and development of the G-rich oligonucleotide AS1411 as a novel treatment for cancer. *Experimental and Molecular Pathology* 86, 151–164. <https://doi.org/10.1016/j.yexmp.2009.01.004>.

[22] Girvan, A. C., Teng, Y., Casson, L. K., Thomas, S. D., Julinger, S., Ball, M. W., Klein, J.

B., Pierce, W. M., Barve, S. S., and Bates, P. J. (2006) AGRO100 inhibits activation of nuclear factor- κ B (NF- κ B) by forming a complex with NF- κ B essential modulator (NEMO) and nucleolin. *Molecular Cancer Therapeutics* 5, 1790–1799. <https://doi.org/10.1158/1535-7163.MCT-05-0361>.

[23] Reyes-Reyes, E. M., Teng, Y., and Bates, P. J. (2010) A new paradigm for aptamer therapeutic AS1411 action: Uptake by macropinocytosis and its stimulation by a nucleolin-dependent mechanism. *Cancer Research* 70, 8617–8629. <https://doi.org/10.1158/0008-5472.CAN-10-0920>.

[24] Park, J. Y., Cho, Y. L., Chae, J. R., Moon, S. H., Cho, W. G., Choi, Y. J., Lee, S. J., and Kang, W. J. (2018) Gemcitabine-Incorporated G-Quadruplex Aptamer for Targeted Drug Delivery into Pancreas Cancer. *Molecular Therapy - Nucleic Acids* 12, 543–553. <https://doi.org/10.1016/j.omtn.2018.06.003>.

Chapter 10

AS1411 Derivatives as Carriers of G-Quadruplex Ligands for Cervical Cancer Cells

This chapter was published in:

Figueiredo, J. [†], Lopes-Nunes, J. [†], **Carvalho, J.** [†], Antunes, F., Ribeiro, M., Campello, M. P. C., Paulo, A., Paiva, A., Salgado, G. F., Queiroz, J. A., Mergny, J.-L., and Cruz, C. (2019) AS1411 derivatives as carriers of G-quadruplex ligands for cervical cancer cells. *International Journal of Pharmaceutics* 568, 118511. <https://doi.org/10.1016/j.ijpharm.2019.118511>.

[†] Equal contribution

AS1411 Derivatives as Carriers of G-Quadruplex Ligands for Cervical Cancer Cells

Abstract

Nucleic acid aptamers can specifically bind to target molecules on the cell membrane that mediate their entrance into the cells. Their small size, high binding affinity, specificity, good biocompatibility, stability and low immunogenicity make them ideal drug delivery systems for cancer therapy. These biopharmaceuticals have potential for the delivery of anticancer compounds to diseased tissues, increasing their effectiveness while mitigating the off-target toxicity towards healthy cells. Herein, we have studied two quadruplex-forming DNA sequences derived from the nucleolin-targeted aptamer AS1411 as supramolecular carriers for the cancer-selective delivery of acridine orange derivatives (C_3 , C_5 and C_8) in cervical cancer cells. The devised delivery strategy relied on the non-covalent association of the acridine derivatives and the G4 structures. This association is done with a high binding strength, as suggested by the obtained K_D values in the $10^{-6} - 10^{-7}$ M range, leading to the thermal stabilization of the G4 structures, particularly for C_8 . The stability of the resulting supramolecular conjugates was evaluated in fetal bovine serum, which proved their resistance against serum nucleases up to 48 h. Previous studies showed that the tested acridine orange derivatives were cytotoxic towards cervical cancer cells (HeLa) and non-malignant cells. However, when conjugated to AS1411 derivatives, the cytotoxicity of the free ligands towards non-malignant cells was restrained. Furthermore, conjugated C_3 showed an enhanced cytotoxicity against HeLa cancer cells. Confocal microscopy indicated that both G4 sequences appear to colocalize with nucleolin, suggesting their ability to recognize and bind nucleolin on the cell surface. Additionally, the results confirmed the internalization of these delivery systems into HeLa cancer cells and their sustained cell trafficking, although being able to dissociate intracellularly to deliver C_8 to the nucleoli. Overall, we showed that AS1411-derived G4s can be used as a potential cancer drug delivery system for cervical cancer.

Introduction

Conventional chemotherapeutics suffer from a narrow therapeutic window due to their poor specificity towards tumor/healthy cells, leading to the so called “off-target” side effects [1]. There is a need of developing targeted delivery systems to improve drug

selective accumulation in cancer cells while increasing its effectiveness. One approach to tackle such need is the use of cell-specific carriers that guide the delivery of the drug to the cancer cell. Nucleic acid aptamers have emerged as an attractive class of drug delivery systems due to their ability to bind with high affinity to specific ligands, their high chemical flexibility as well as tissue penetration capability [1]. They have several advantages such as a relatively small size, flexible structure, inexpensive and reproducible chemical production when compared to antibodies, ease of chemical modification, high stability and lack of immunogenicity [1]. Taking advantage of such properties, nucleic acid aptamers have been used to selectively convey ligands with therapeutic or imaging purposes into cancer cells [2,3].

Several aptamers have already been developed against a wide array of therapeutically relevant targets [1–3]. Among these, DNA aptamer AS1411 that targets the nucleolin protein has been extensively used as a cancer-targeting agent [4,5]. Nucleolin is considered a tumor biomarker as it is overexpressed in the surface of cancer cells, while in non-malignant cells it is found primarily in the nuclei and cytoplasm [6]. In vivo studies showed that AS1411-linked ligands selectively accumulate in cancer cells following systemic administration, emphasizing the potential of using this aptamer as a drug delivery system [4]. Due to its G-rich nature, AS1411 is able to fold into a G4 structure [4]. The aptamer structure is however highly polymorphic with at least eight different folding patterns detected by size exclusion chromatography [7], complicating the identification of the structure(s) relevant for its biological behavior. More recently, Phan and colleagues reported the NMR structures of a series of AS1411 aptamer derivatives that adopted a single G4 structure in solution with identical antiproliferative activity [8,9]. These AS1411 derivatives, namely AT11 and AT11-Bo (a more thermally stable version of AT11) [9], are potentially promising for the development of cancer-targeted systems for the delivery of anticancer ligands.

Herein, using AT11 and AT11-Bo derivatives, we studied aptamer-guided approaches for the delivery of ligands with anticancer potential into HeLa cervical cancer cells. The ligands used were acridine orange derivatives (C_3 , C_5 and C_8) that showed potent inhibitory effect towards HeLa cells (low micromolar IC_{50} values at 24 h incubation) [10,11]. However, the ligands had a similar antiproliferative effect on NHDF and their application in further *in vitro* or *in vivo* assays would benefit from a cancer-selective drug delivery system. The ligands were also shown to bind and stabilize a variety of G4 structures such as telomeric and oncogene promoter G4s [10]. Using a supramolecular strategy relying on the G4-binding ability of the acridine orange derivatives, lacking any

chemical modifications of both sequences and ligands, the ability of AT11 and AT11-Bo to selectively convey C₃, C₅ and C₈ to HeLa human cervical cancer cells was assessed using an array of spectroscopic and biochemical techniques.

Materials and Methods

Acridine Orange Derivatives and Oligonucleotides

Acridine orange ligands 10-(3-(4-iodobenzamide)propyl))-3,6-bis(dimethylamino)acridinium iodide (C₃), 10-(5-(4-iodobenzamide)pentyl))-3,6-bis(dimethylamino)acridinium iodide (C₅) and 10-(8-(4-iodobenzamide)octyl))-3,6-bis(dimethylamine)acridinium iodide (C₈) were synthesized and purified as previously described [11]. The ligands were dissolved in DMSO to obtain a 10 mM stock solution and the subsequent dilutions were made with Milli-Q water. AT11 (5'-TGGTGGTGGTTGTTGTGGTGGTGGTGGT-3'), Cy5-AT11, AT11-Bo (5'-TGGTGGTGGTTGTTGTGGTGGTGGTGGT-3') and Cy5-AT11-Bo were obtained from Eurogentec (Belgium) and StabVida (Portugal) with HPLC-grade purification. Stock solutions were prepared using Milli-Q water and stored at -20 °C until used. For all the experiments, the oligonucleotides were annealed by heating to 95°C during 10 min and then cooled down in ice until used.

Thermal Difference Spectra

TDS experiments were performed using a Thermo Scientific™ Evolution™ 201 UV-Visible Spectrophotometer (Thermo Fisher Scientific, USA). UV spectra were recorded in the 220-335 nm range with a scan rate of 200 nm/min and 1 nm data intervals, above and below the melting temperature, at 90 °C and 20 °C, respectively. Pre-folded AT11 or AT11-Bo were used at 3 μM concentration in 20 mM phosphate buffer containing 65 mM KCl. The TDS spectrum was calculated by subtracting the 20 °C spectrum from the spectrum obtained at 90 °C. The data was normalized relatively to the maximum absorbance.

Circular Dichroism Spectroscopy

CD spectra were acquired in a Jasco J-815 spectropolarimeter equipped with a Peltier temperature control system. AT11 or AT11-Bo oligonucleotides were dissolved in 20 mM potassium phosphate buffer containing 65 mM KCl. For titrations experiments, the required amount of acridine orange ligands solution was added to the 1 mm quartz cuvette containing 10 μM of pre-folded AT11 or AT11-Bo solutions. Spectra were recorded through

wavelengths ranging from 220 to 340 nm, with a scan speed of 100 nm/min, 1 nm bandwidth, 1 s integration time over 4 averaged accumulations.

For CD melting procedures, AT11-Bo solutions were prepared in 10 mM KCl and 90 mM LiCl in order to obtain a melting temperature around 50 °C. The denaturation process was analyzed by monitoring the wavelength of maximum ellipticity (260 nm) through temperatures ranging from 10-20 to 110 °C (temperature set by the Peltier system) with a heating rate of 2 °C/min, in the presence or absence of acridine orange derivatives. Data were converted into fraction folded (θ) plots according to Equation 1, fitted to a Boltzmann distribution using OriginPro 2016:

$$\theta = \frac{CD - CD_{\lambda}^{min}}{CD_{\lambda}^{max} - CD_{\lambda}^{min}} \quad (\text{Equation 1})$$

The melting temperatures (T_m) were determined from a two-state transition model where CD is the ellipticity at 260 nm at each temperature and CD^{min} and CD^{max} are the lowest and highest ellipticities, respectively.

Fluorescence Binding Studies

Fluorescence titrations were performed on a Horiba FluoroMax4 fluorometer (Japan) using a high-precision quartz suprasil cuvette (light path 10 mm × 4 mm). For these experiments, previously annealed AT11 or AT11-Bo at concentration of 500 μM in 20 mM potassium phosphate buffer containing 65 mM KCl were successively added to a solution of the same buffer containing 5 μM of one of the acridine orange ligands. After each addition the mixture was left for 10 min for equilibration. All spectra were acquired with an integration time of 0.5 s, an emission and excitation slit fixed at 2 nm and step size of 1 nm, averaged over three scans. The acridine orange ligands were excited at 498 nm and the emission spectra were collected from 530 to 700 nm. Additionally, Cy5-labelled AT11 and AT11-Bo at 1 μM were titrated with nucleolin in the absence and presence of C₈ to evaluate the effect of ligand binding in the affinity of the G4 structures towards the protein. The fluorescence emission spectra were obtained between 660 and 800 nm and averaged over three scans.

The obtained fluorescence data was converted into fraction of ligand bound (α) plots according to Equation 2:

$$\alpha = \frac{I - I_{\lambda}^{free}}{I_{\lambda}^{bound} - I_{\lambda}^{free}} \quad (\text{Equation 2})$$

where I is the fluorescence intensity at 568 nm at each ligand:DNA ratio, and I_{free} and I_{bound} are the fluorescence intensity of the free and fully bound ligand, respectively. Data points were then fitted to Hill saturation binding model (OriginPro 8):

$$\alpha = \frac{[DNA]^n}{K_D + [DNA]^n} \quad (\text{Equation 3})$$

where K_D is the apparent equilibrium dissociation constant, $[DNA]$ is the concentration of the DNA and n is the Hill constant which describes cooperativity of ligand binding. For the protein binding experiments, data points were fitted to Michaelis-Menten model (OriginPro 8):

$$\alpha = \frac{[DNA]}{K_D + [DNA]} \quad (\text{Equation 4})$$

Serum Stability Assay

Following annealing, 30 μM of free AT11 or AT11-Bo or complexed with 30 μM C₈ (in 20 mM potassium phosphate buffer containing 65 mM KCl) were incubated in RPMI cell culture medium supplemented with 10 % FBS. For each sample, 1 μL of free AT11 or AT11-Bo or complexed with C₈ were mixed with 9 μL of RPMI medium with FBS and incubated for 0, 7, 24 and 48 h at 37 °C. Then, 10 μL of each sample were loaded in a 1 % agarose gel. Additional experiments were performed in 50% FBS for 0, 7, 24, 48 and 72 h at 37 °C. The electrophoresis was run for 1 h at 153 V. Gels were stained with 0.01% Xpert Green DNA stain (GRISP, Portugal) in 1 $\times$ Tris-acetate-EDTA buffer and visualized through UV light. GRS ladder 1kb (GRISP, Portugal) was used to confirm DNA size.

Cell Viability Assay

NHDF cells were grown in RPMI medium, supplemented with 10% FBS, 1% streptomycin-penicillin antibiotic, 0.01 M HEPES, 0.02 M L-glutamine and 0.001 M sodium pyruvate. Additionally, cervical cancer cell line (HeLa) was grown in DMEM medium, supplemented with 10% FBS and 1% streptomycin-penicillin antibiotic. Cell cultures were maintained in a controlled humidified atmosphere at 37 °C and 5% CO₂. Cells were plated in 48-well culture plates (HeLa or NHDF were seeded at a density of 0.5 $\times$ 10⁴ cells/mL or 1 $\times$ 10⁴cells/mL, respectively) in 300 μL of medium. In the next day, both cell lines were treated with the pre-formed AT11 or AT11-Bo -ligand complexes (2 μM of C₃/C₅ or 1 μM of C₈ with 15 μM of AT11 or AT11-Bo for 10 min) and incubated for 7 days. At the end of incubation, 50 μL of MTT solution was added to each well and further incubated for 1 h in HeLa or 4 h in NHDF cells. The resulting formazan crystals were solubilized with 250 μL

of DMSO and the optical density (OD) was measured at 570 nm. Cell viability was normalized to control condition (cells without ligands). Statistical analysis was performed by using Student's unpaired t test. The values of $p < 0.05$ were considered statistically significant. Data analysis was performed with GraphPad Prism 6 software (San Diego, CA, USA).

Confocal Fluorescence Microscopy

To better understand the internalization ability of free AT11 or AT11-Bo and complexes, HeLa and NHDF cells were seeded at a density of 0.3×10^4 and 5×10^4 cells/mL, respectively, in treated 8 well μ -slides (IBIDI, Germany). Cells were grown at 37 °C in 5% CO₂ humidified atmosphere and treated the day after with the pre-formed complex (0.5 μ M of acridine orange ligands (C₃, C₅ or C₈) and 1 μ M of AT11 and AT11-Bo), ligand-only control (0.5 μ M) and DNA-only (AT11 and AT11-Bo) controls (1 μ M). After 3 days of incubation, cells were washed with phosphate-buffered saline and nuclei were then stained with 1 μ M Hoechst 33342 nuclear probe for 15 min. Cells were then observed using a Zeiss AxioObserver LSM 710 microscope.

For nucleolin immunocytochemistry, HeLa cells were seeded at a density of 1.5×10^4 cells/mL in 8 well μ -slides (IBIDI, Germany) and incubated with free AT11 and AT11-Bo, or the ligand complexes for 2 h. Nucleolin antibody (1:100, ref. PA3-16875, Invitrogen) was then added during 2 h and secondary antibody (Alexa Fluor 546®, 1:1000) for 1 h, both at room temperature. The excess fluorophores were washed off by rinsing with PBS three times and the cells were imaged using a Zeiss AxioObserver LSM 710 microscope. The degree of complexes colocalization is expressed in Pearson correlation coefficients (r_p) values [12], which were calculated using the colocalization analysis tool in the ImageJ software.

Flow cytometry

HeLa and NHDF cells were seeded in 12-well plates at a density of 50×10^4 cells/well and incubated overnight for cell adhesion. Then, cells were treated for 24 h with the preformed C₈-aptamer complexes and free C₈ as control. After the incubation period, the wells were washed by rinsing with PBS three times and cells were then trypsinized, resuspended in PBS and analyzed in a BD FACSCanto™ II flow cytometry system (BD Life Sciences, US) to evaluate the uptake of the aptamer-ligand complexes. C₈ fluorescence was monitored

using the FITC channel. Non-specific coloring and debris were excluded by analyzing FSC vs SSC density plots.

G4 model structures

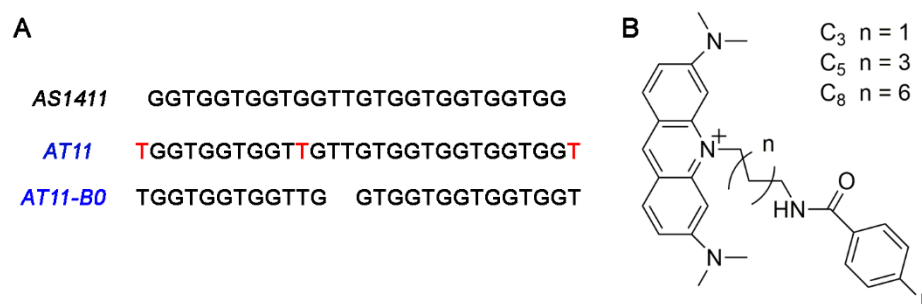
Since 3D structure of AT11-Bo G4 is not available, a predicted model was built based on the AT11 G4 structure (PDB: 2N3M) as previously reported [13]. Briefly, residues T13 and T14 were deleted from 2N3M using Swiss-PDB Viewer mutation tool and converted the nucleotide sequence of AT11 to AT11-Bo (please see **Figure 10.1** for further details). The built structure was optimized for further experiments by running fully solvated molecular dynamics (MD) simulations using GROMACS 2016.3 with the following parameters. The AT11-Bo G4 structure was initially centered in an octahedral solute box and filled with TIP3P water molecules and K⁺ atoms. The structure was then submitted to 5000-step energy minimization with no restraints. A 100 ps equilibration was then performed under a modified Berendsen thermostat, with a temperature gradient from 0 to 300 K, followed by another 100 ps of equilibration done under aforementioned thermostat and a Parrinello-Raman barometer. Finally, 200 ns of MD were carried out. The final snapshot of the 200 ns simulation was used as optimized model for further experiments.

Molecular Dynamics

Both the AT11 (PDB: 2N3M, Figure S6) and AT11-Bo (built as described above, Figure S7) G4 structures, and nucleolin RNA-binding domains (RBD1,2 – PDB: 2KRR) were optimized for docking using Dock Prep tool of Chimera 1.11.2. After assigning polar hydrogens and Gasteiger charges, rigid body docking calculations were carried with AutoDock 4.2 program. The G4 structure was centered in a box sized $125 \times 125 \times 125$ Å along the x, y, and z axes, respectively, with a grid spacing of 0.6 Å. For each G4 structure, 25 runs under Genetic Algorithm were performed setting an initial population of 150 random individuals, a maximum number of evaluations of 2.5×10^6 , rate of mutation and crossover of 0.02 and 0.8, respectively, and elitism value of 1. The most representative structures were selected based on available literature [14] and further processed with MD simulations the all-atom force field AMBER99SB of GROMACS 2016.3. For MD simulations, the G4-nucleolin complexes were centered in an octahedral box and solvated with TIP3P water molecules and K⁺ atoms to neutralize the system. After an energy minimization of 1000 steps using the steepest descent algorithm, system was gradually heated to 300 K in 100 ps under the control of Berendsen thermostat followed by 100 ps isobaric simulation under the control of the Parrinello-Rahman barostat. Finally, explicit solvent MD runs of 50 ns were performed for each G4-nucleolin complex at constant

temperature (300 K) and pressure (1 atm). The Particle mesh Ewald (PME) method was used for calculating long range electrostatic interactions. A 1 nm cut-off was applied to short range Lennard-Jones interactions. Coordinates were collected in trajectory files every 10 ps. All molecular images were rendered using UCSF Chimera 1.11.2.

Nucleic acid aptamers have been employed to deliver an extensive array of ligands with therapeutic interest such as doxorubicin or 5-fluorouracil to cancer cells [15]. Our idea is based on the fact that sequences with the potential to form G4s are attractive nucleolin binders, since nucleolin has high affinity to G4s and is overexpressed in HeLa cervical cancer cells [16]. AT11 and AT11-Bo G4s can gain intracellular access by endocytosis after binding to nucleolin present at the cell surface, and benefit of nucleolin shuttling to the nucleus. Based on the presumed mechanism, we have evaluated if AT11 and AT11-Bo G4s can still recognize nucleolin while carrying ligands for an intracellular targeted delivery with anticancer effects.



Firstly, we studied the formation and topology of the AT11 and AT11-Bo G4s by TDS absorbance spectroscopy. The difference between the folded and unfolded spectra of a oligonucleotide may provide valuable information about its conformation [17]. The TDS spectra of AT11 and AT11-Bo are depicted in **Figure 10.2**. A common feature to both spectra is the negative signal at 295 nm which is indicative of a G4 structure [17,18]. The

TDS factor $\Delta A_{240\text{nm}}/\Delta A_{295\text{nm}}$ for AT11 and AT11-Bo was 3.5 and 2.3, respectively. The values above 2 suggest parallel-stranded G4 topologies for both AT11 and AT11-Bo sequences [18].

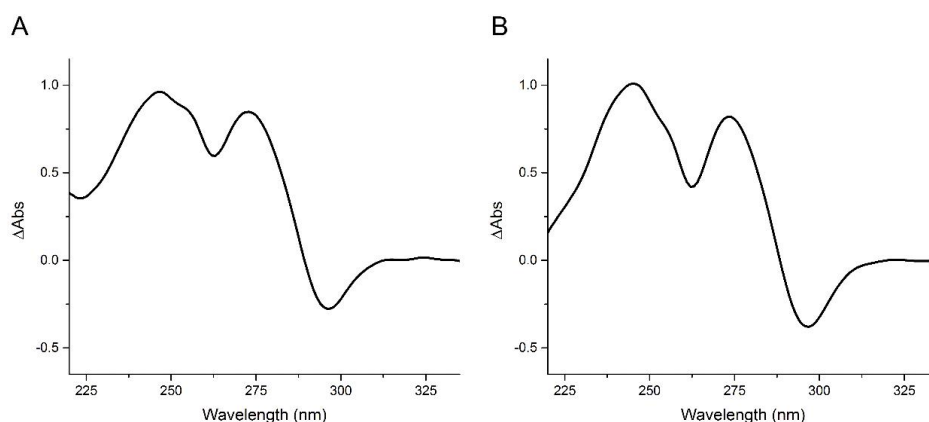


Figure 10.2 - TDS spectra obtained by subtracting the spectrum at 20 °C from the spectrum at 90 °C for (A) AT11 and (B) AT11-Bo. Spectra acquired in 20 mM potassium phosphate buffer containing 65 mM KCl, pH 7.

Afterwards, the formation of complexes AT11- and AT11-Bo G4s-ligand was studied using CD and steady state fluorescence spectroscopies. The CD spectra of AT11 and AT11-Bo G4s are shown in **Figure 10.3**. CD spectroscopy is a useful technique to infer on the G4 topology of a given G-rich oligonucleotide and to evaluate the effect of ligand binding on the pre-formed topology [19]. Both AT11 and AT11-Bo present the characteristic spectral features of parallel-stranded G4s as seen by a positive peak around 260 nm and a negative peak around 240 nm [9], in agreement with the conclusions reached by TDS analysis. Upon titrating the ligands with the AT11 G4 solution, an overall increase in ellipticity was observed together with the retention of the characteristic bands (**Figure 10.3A**). This indicates that the ligands are able to bind the G4 structure without disrupting/changing its pre-folded topology. In the case of AT11-Bo G4, ligands C₃ and C₅ had little or no effect on the ellipticity while C₈ promoted an increase (**Figure 10.3B**).

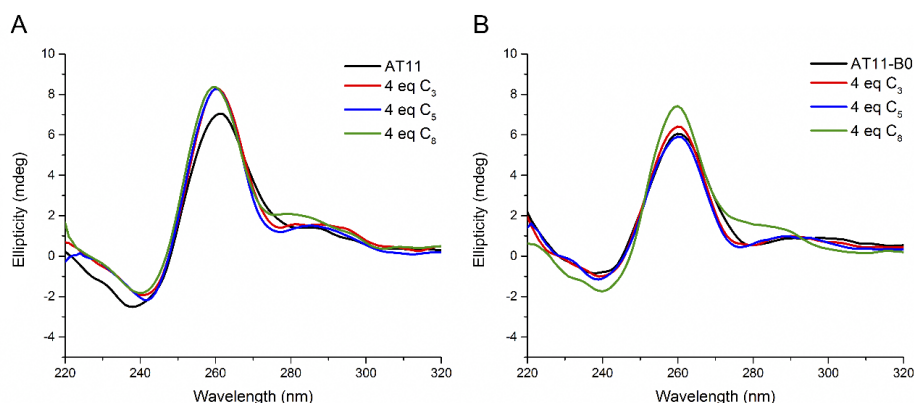


Figure 10.3 - CD spectra of (A) AT11 G4 and (B) AT11-Bo G4 in the absence and presence of 4 molar eq. of each of the acridine orange ligand. Spectra acquired in 20 mM potassium phosphate buffer containing 65 mM KCl, pH 7.

Then, the effect of ligand binding on the thermal stability of the AT11 and AT11-Bo G4 structures was assessed by CD melting experiments. The potential stabilization of AT11 and AT11-Bo G4 structures may be relevant in two different ways: first, it may be used as a measure of how stable the complex is in solution; and second, it might indicate how the interaction with the ligand may enhance the targeting properties of the aptamers as G4 conformations [16]. The CD melting data of each ligand is shown in **Figure 10.4**. The melting temperatures (T_m) of AT11 and AT11-Bo G4s were 42.0 and 53.7, respectively. Despite the adjustment of the salt conditions in the melting experiments for each aptamer (65 mM KCl vs 10 mM KCl for AT11 and AT11-Bo, respectively), the T_m value of AT11-Bo G4 is significantly higher than that of AT11 G4 due to the modification introduced in the sequence. AT11-Bo derives from AT11 by the elimination of the bulge residues which leads to an increased thermal stability [9]. Upon addition of 4 molar equivalents of acridine orange ligands, a clear increase in the T_m values was observed which indicates ligand-induced stabilization of the G4 structures (**Table 10.1**).

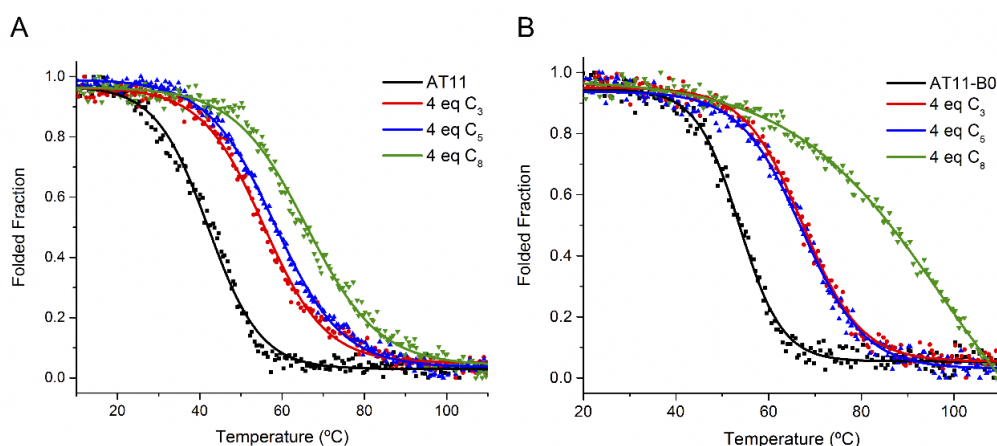


Figure 10.4 - CD melting curves of (A) AT11 G4 and (B) AT11-Bo G4 in the absence and presence of 4 molar eq. of each of the acridine orange ligands. Data points were recorded at 260 nm. The Boltzmann curve fitting of the data points is shown.

In the case of AT11 G4 (**Figure 10.4A**), C_8 was the most stabilizing ligand ($\Delta T_m = 24.0$ °C) while C_3 stabilized the least ($\Delta T_m = 13.5$ °C). As reported previously for this set of acridine orange derivatives, the increase in the alkylamide chain enhances the thermal stabilization induced by the ligands [10]. Similarly, C_8 promoted a stabilization of over 30 °C for AT11-Bo G4 (absolute value not determined due to lack of convergence in data fitting) while C_3 and C_5 equally stabilized the G4 structure by ca. 13.7 °C (**Figure 10.4B**).

Table 10.1 - Ligand-induced thermal stabilization of the AT11 or AT11-Bo measured by CD melting experiments.

Ligand	ΔT_m (°C) ^a	
	AT11	AT11-Bo
C₃	13.5 ± 0.2	13.7 ± 0.2
C₅	16.1 ± 0.1	13.7 ± 0.2
C₈	24.0 ± 0.2	> 30

^a ΔT_m represents difference in melting temperature [$\Delta T_m = T_m$ (DNA + 4 eq ligand) – T_m (DNA)]. The T_m values for DNA alone are 42.0 ± 0.2 °C [AT11 in 65 mM KCl] and 53.7 ± 0.1 °C [AT11-Bo in 10 mM KCl + 90 mM LiCl].

Following CD studies, the apparent equilibrium dissociation constants (K_D) for the complexes were determined by fluorimetric titrations (**Figure 10.5**). The ligands produce a broad band centered at 568 nm after excitation at 498 nm. Upon adding pre-folded AT11 G4 to the ligands solution, an enhancement in the ligands fluorescence was observed (**Figure 10.5, top**). Up to 10-fold fluorescence enhancement was observed which is indicative of strong association between the ligands and the AT11 G4 or AT11-Bo G4 [20]. Fitting the fluorimetric data to Hill saturation binding model gave K_D values of 1.4×10^{-6} , 1.6×10^{-6} and 9.9×10^{-7} M for C₃, C₅ and C₈, respectively (**Table 10.2**). These K_D values in the 10^{-6} - 10^{-7} M range suggest medium-strong complex formation. Ligand C₈ presented the highest affinity and ΔT_m value. Hill coefficients (n) above 1 suggest a cooperative binding mode.

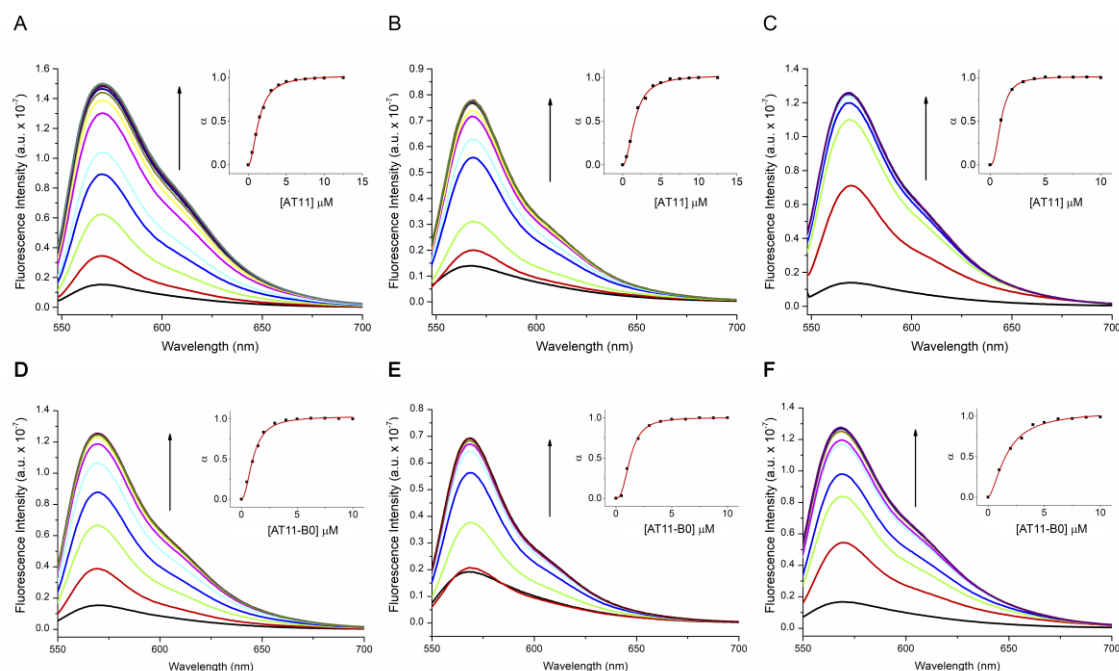


Figure 10.5 - Fluorescence emission spectra of (A, D) C₃, (B, E) C₅ and (C, F) C₈ with increasing concentration of AT11 G4 (top) or AT11-Bo G4 (bottom). Insets: fraction of ligand bound plots fitted to the Hill saturation binding model. Spectra acquired in 20 mM potassium phosphate buffer containing 65 mM KCl, pH 7.

Regarding AT11-Bo G4, similar fluorescence enhancement was observed for all ligands (**Figure 10.5, bottom**), with K_D values of 1.1×10^{-6} , 1.3×10^{-6} and 1.7×10^{-6} M for C₃, C₅ and C₈, respectively (**Table 10.2**). Likewise, the complex formation is in the moderate-high affinity range which supports the ligand-induced stabilization observed in CD melting experiments. The Hill coefficients (n) suggest a cooperative binding mode. Overall, the ligands bind AT11 G4 and AT11-Bo G4 with similar affinities.

Table 10.2 - Apparent dissociation constants (K_D) obtained from fluorescence titration experiments.

Ligand	AT11		AT11-Bo	
	K_D (M)	n	K_D (M)	n
C ₃	1.4×10^{-6}	1.9	1.1×10^{-6}	2.0
C ₅	1.6×10^{-6}	2.1	1.3×10^{-6}	2.6
C ₈	9.9×10^{-7}	2.5	1.7×10^{-6}	1.6

One of the issues of using oligonucleotides as aptamers for drug delivery is serum degradation and susceptibility to nuclease [21]. In order to evaluate whether the AT11 G4 and AT11-Bo G4 were stable in serum or if the ligand conjugation was able to protect from serum degradation due to G4 stabilization, free AT11 G4 and AT11-Bo G4 and their complexes were incubated in culture media supplemented with 10% or 50% FBS. Taking into account that aptamers may take several days to internalize [22], this experiment was also performed at 72 h of incubation. As C₈ promoted the highest thermal stabilization, the assay was only carried in the presence of this ligand.

The results revealed by agarose gel electrophoresis are shown in **Figure 10.6**. Free AT11 G4 seems to undergo some degradation after a 48 h incubation period as indicated by the progressive disappearance/decrease of the corresponding band (**Figure 10.6A**). In the presence of C₈ no apparent degradation can be observed suggesting that the complex is stable in 10% FBS (**Figure 10.6B**). On the other hand, free AT11-Bo (**Figure 10.6C**) is resistant to serum nucleases, probably because of its higher thermal stability when compared to AT11 G4 and the lack of bulge in the G4 structure [9]. In the presence of C₈ the stability was not affected (**Figure 10.6D**). The results shown in **Figure 10.6E** suggest that in 50% FBS free AT11 G4 is readily degraded as seen from the absence of any band below 500 pb. In contrast, in the presence of ligand C₈, the band corresponding to the AT11 G4-C₈ complex was observed (**Figure 10.6F**). These results suggest that the ligand significantly increases the serum stability of the aptamer up to 72 h. Free AT11-Bo G4 showed to be resistant to serum nucleases in 50% FBS (**Figure 10.6G**) and the presence of ligand C₈ did not affect this stability (**Figure 10.6H**).

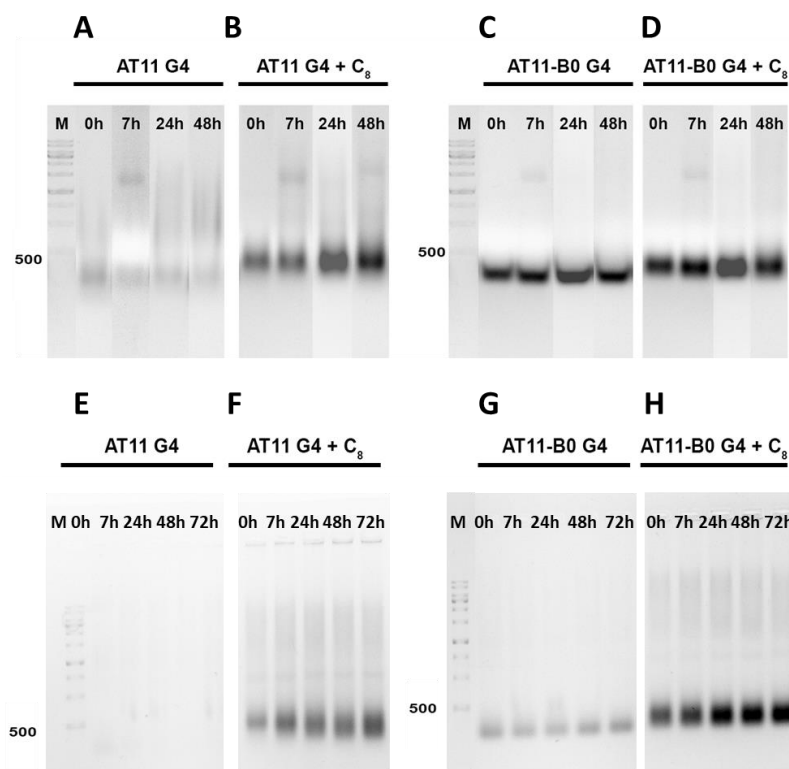


Figure 10.6 - Serum stability of AT11 G4, AT11 G4-C₈, AT11-Bo G4 and AT11-Bo G4-C₈ in 10% FBS at 37 °C for 0, 7, 24 and 48 h (A, B, C and D) and 50% FBS at 37 °C for 0, 7, 24, 48 and 72 h (E, F, G and H). The samples were resolved in a 1% agarose gel, stained with 0.01% Xpert Green DNA and visualized under UV light.

As both AT11 G4 and AT11-Bo G4 sequences are AS1411 derivatives, the affinity of these two G4 structures towards nucleolin was assessed by fluorimetric titrations using Cy5-labelled sequences. The saturation binding curves shown in **Figures 10.7A and B** suggest fast saturation at low protein concentration which is indicative of high affinity. Indeed, the fitting of the data yielded K_D values of 9.1×10^{-12} and 9.5×10^{-12} M for Cy5-AT11 G4 and Cy5-AT11-Bo G4, respectively, which are in the same range of values reported for other aptamers [1]. To ascertain whether ligand binding would decrease the affinity of the G4 sequences towards nucleolin, a similar experiment was conducted with AT11-C₈ and AT11-Bo-C₈ complexes. As seen in **Figures 10.7C and D**, fast saturation was also attained in the presence of the ligand, yielding K_D values of 5.2×10^{-12} and 3.3×10^{-12} M for Cy5-AT11 G4 C₈ and Cy5-AT11-Bo G4 C₈, respectively. The K_D values for the complexes are in the same picomolar range as the free G4 structures, suggesting that the presence of the ligand does not negatively affect the recognition of nucleolin by AT11 G4 and AT11-Bo G4.

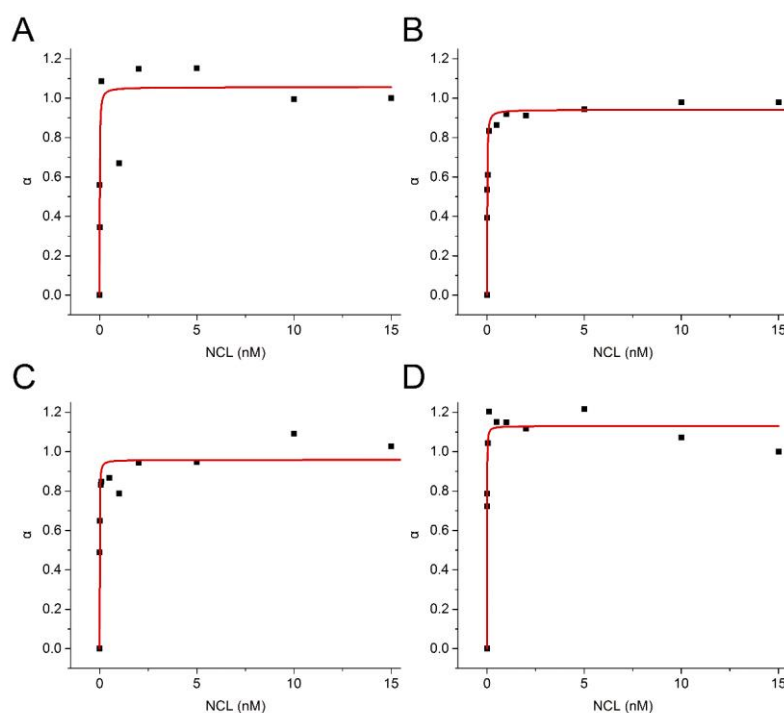


Figure 10.7 - Saturation binding plots obtained by fluorescence titrations with nucleolin and fitted to Michaelis-Menten model for free Cy5-AT11 G4 and Cy5-AT11-Bo G4 (A and B, respectively), or for Cy5-AT11 and Cy5-AT11-Bo in the presence of ligand Cs (C and D, respectively).

The interaction of AT11 and AT11-Bo G4s with nucleolin was assessed via docking experiments followed by MD simulations. AT11 G4 3D structure (Figure S6) has already been determined by NMR [9] and is available through PDB 2N3M. As no 3D model for AT11-Bo G4 is available, we manually edited the sequence of PDB 2N3M and re-optimized the structure for MD studies to remove unfavorable van der Waals interactions and probe for incorrect atom names and connectivity. The final snapshot of the 200 ns MD run for optimization of the AT11-Bo G4 model is shown in Figure S7. AT11 PDB structure and AT11-Bo constructed model were used as starting scaffolds for docking experiments. Previously, it has been shown that AS1411 aptamer binds preferentially the RNA-binding domains 1 and 2 (RBD1,2) of nucleolin [14]. As AT11 and AT11-Bo are both AS1411 derivatives, we have used RBD1,2 for the calculations. The complexes formed by AT11-RBD1,2 and AT11-Bo-RBD1,2 are depicted in **Figure 10.8**. Stable complexes were attained during the 50 ns simulations as seen by the RMSD graphs (Figure S7 and S8). After an initial increase in the RMSD values, both complexes converged at around 3 ns remaining stable relatively to this conformer throughout the remaining of the simulation. RBD1,2 was found interacting with the loop region of both G4 structures, establishing multiple H-bonds (green lines) with the G4 residues, particularly the loops thymines (depicted in grey). The 12-residue linker between both RBD seems to create a binding pocket for the G4 structures, with additional interactions with RBD units themselves. In the case of AT11 G4 (**Figure 10.8A**), thymines T4, T14, T16 and T19 interact with RBD1,2

thus demonstrating the preference of protein nucleolin for the loop residues. AT11-Bo on its hand (**Figure 10.8B**), was found similarly bound to RBD1,2 through T4, T17 and T20. These results agree with previous MD studies using a predicted model of AS1411 aptamer [14] and experimental reports using loop-modified AS1411 sequences [23].

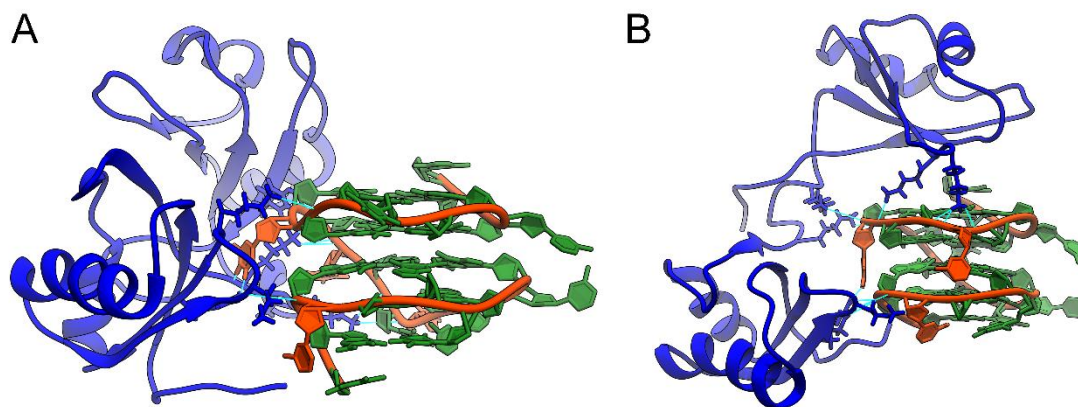


Figure 10.8 - Final snapshots of 50 ns MD simulations of the complexes formed by (A) AT11 G4- RBD1,2 and (B) AT11-Bo G4- RBD1,2. G4 structures are depicted in green with loop thymines responsible for binding highlighted in orange while RBD1,2 is represented in blue ribbon (except for atoms participating in H-bonding). H-bonds are shown as cyan lines.

To determine if AT11 G4 and AT11-Bo G4 are able to bind specifically to nucleolin, HeLa cells were used and analyzed by confocal microscopy after incubation with Cy5-AT11 or Cy5-AT11-Bo. For comparative purposes, complexes AT11-C₈ and AT11-Bo-C₈ were selected for confocal microscopy studies to analyze if the conjugates can recognize nucleolin *in vitro*. The primary anti-nucleolin antibody conjugated with the secondary antibody AlexaFluor 546® were used to localize cell surface nucleolin. Free Cy5-AT11 and Cy5-AT11-Bo, as well as their conjugates with C₈, were able to colocalize with nucleolin-positive regions after a short incubation (2 h) (**Figures 10.9** and **S1 Appendix G**). This assay demonstrated that AT11 and AT11-Bo are anti-nucleolin aptamers, because they were able to co-stain this protein thus suggesting cell surface binding. Moreover, the conjugation of C₈ with the G4 sequences does not seem to negatively affect their nucleolin binding as similar behavior was observed for free and conjugated aptamers.

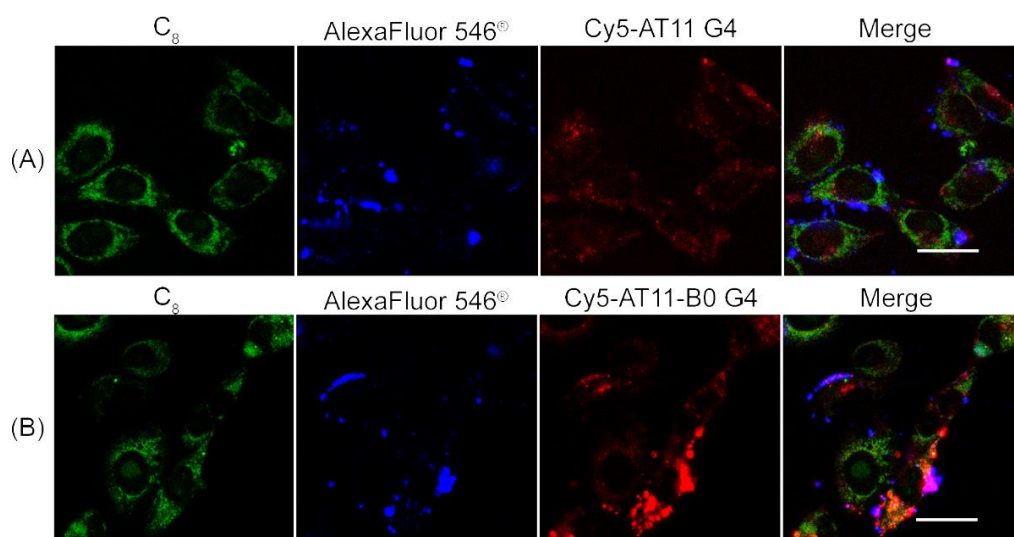


Figure 10.9 - Confocal fluorescence images of HeLa cells incubated with (A) Cy5-AT11-C₈ and (B) Cy5-AT11-Bo-C₈ after 2 h at 37°C. Nucleolin is stained with AlexaFluor 546®, C₈ emits green fluorescence, Cy5-AT11 and Cy5-AT11-Bo are shown in red. Scale bar: 25 µm. Brightness was adjusted for all images equally.

To evaluate the internalization ability of AT11 G4 and AT11-Bo G4 in cervical cancer and non-malignant cell lines, the cellular uptake of Cy5-labelled AT11 and AT11-Bo or ligand-conjugated was monitored by confocal microscopy. Previously, it was demonstrated that cell death starts to occur after prolonged exposure to AT11 (as to AS1411), more specifically between the second and fourth day, depending on the cells type used [9,16]. Taking this into account, cells were visualized after three days of incubation with the complexes AT11 G4-C₈ and AT11-Bo G4-C₈ in order to avoid excessive cell death rate and to enabling the imaging of their internalization in the used cell lines. After 3 days of Cy5-AT11 and -AT11-Bo incubation, cells were able to internalize them in their free (red coloration, Figure S2) and C₈-conjugated forms. The Pearson coefficients were determined and are presented in Table S1. For HeLa cell line, the Pearson coefficients were 0.67 for AT11 G4-C₈ and 0.6 for AT11-Bo-C₈, respectively (**Figure 10.10**). Interestingly, it was possible to observe that the complex seems to be more internalized by HeLa cells when compared to the free AT11 G4 or AT11-Bo G4. As indicated by the spectroscopic studies, C₈ increases the stability of the AT11 or AT11-Bo G4 structures and this effect may lead to a higher uptake as the G4 conformation is involved in their biological activity [16]. Additionally, the yellow/orange staining in HeLa cytoplasm (**Figure 10.10**) suggests that the complexes remain stable during cell trafficking and the green coloration in the nucleoli demonstrates that C₈ can be released and reach the same compartment that it attained when administrated alone [10]. Moreover, the complexes promoted changes in the malignant cells' morphology, presenting vacuolization and irregular nuclei, which are an evidence of methuosis phenotype (Figure S3) [22].

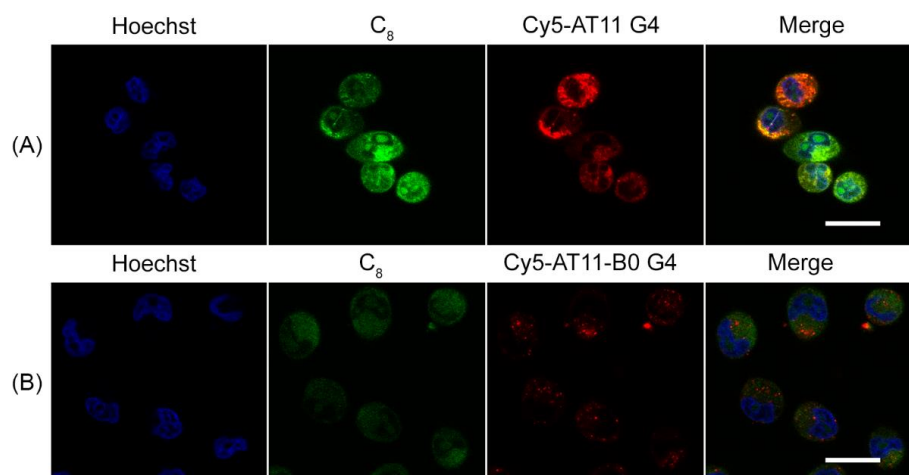


Figure 10.10 - Confocal microscopy images of HeLa cells incubated with (A) Cy5-AT11 G4-C₈ and (B) Cy5-AT11-Bo G4-C₈ for 3 days. Cell nuclei are stained with Hoechst 33342 (blue), C₈ emits green fluorescence and Cy5- AT11 G4 or Cy5-AT11-Bo G4 are shown in red. Overlapping of the two stains, observed as yellow/orange regions, can be seen in the merged images of the cells. Scale bar: 25 μ m.

The internalization of the free AT11 G4 and AT11-Bo G4 (Figure S4) or AT11 G4-C₈ and AT11-Bo G4-C₈ (**Figure 10.11**) also occurs in NHDF cells, despite the absence of overexpressed nucleolin at their surface. This effect was already observed in normal cells after AS1411 incubation [22]. The AT11 G4 and AT11-Bo G4 are taken up by endocytosis being latter expelled by efflux or trapped in lysosomes and degraded [22]. As the cells were constantly kept in medium with the free AT11 G4 and AT11-Bo G4 or the complexes throughout the incubation period, the efflux was most likely not observed in the images obtained by confocal microscopy. The Pearson coefficients for NHDF cell line were 0.87 and 0.83 for AT11 G4-C₈ and AT11-Bo G4-C₈, respectively, suggesting internalization even without overexpression of nucleolin at their surface (**Figure 10.11**).

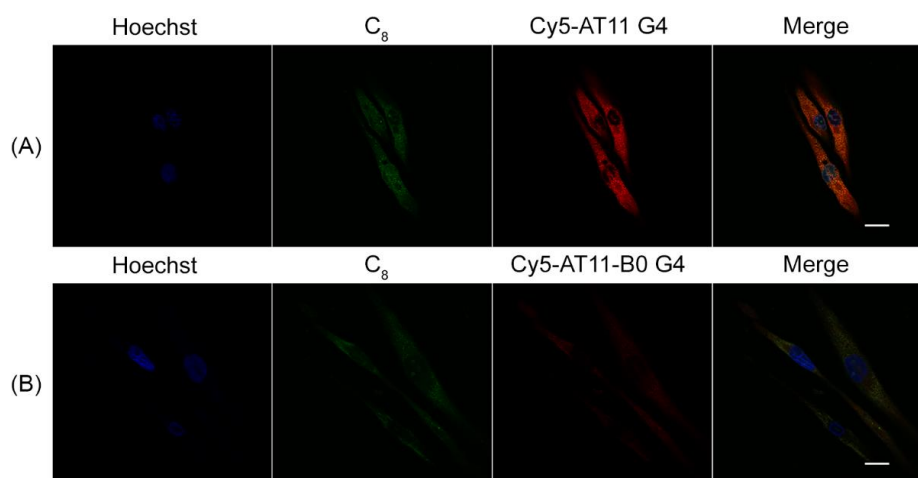


Figure 10.11 - Confocal microscopy images of NHDF cells incubated for 3 days with (A) Cy5-AT11 G4-C₈ and (B) Cy5-AT11-Bo G4-C₈. Cell nuclei are stained with Hoechst 33342 (blue), C₈ emits green fluorescence and Cy5- AT11 G4 or -AT11-Bo G4 are shown in red. The combination of the two stains, observed as yellow/orange regions, can be seen in the merged images of the cells. Scale bar: 25 μ m.

To support the confocal microscopy data for the internalization of the complexes, flow cytometry experiments were conducted. The results for HeLa and NHDF cells are shown in **Figures 10.12** and **10.13**, respectively. When compared to free C₈ (**Figures 10.12A**, **10.13A**), AT11-C₈ complex seems to be internalized to a similar extent in HeLa cells (**Figure 10.12B**), whereas in NHDF cells the complex seems to be less internalized (**Figure 10.13B**) which may be attributed to the differential expression of surface nucleolin between cancerous and non-malignant cells, which may lead to increased accumulation of the complex in cancer cells. Regarding AT11-Bo-C₈ complex, the results suggest less internalization in both cell lines (**Figures 10.12C**, **10.13C**) when compared to free C₈. Altogether, these data indicate that the uptake of conjugated-C₈ is mediated by the aptamer, most probably involving the interaction with nucleolin at the cell surface.

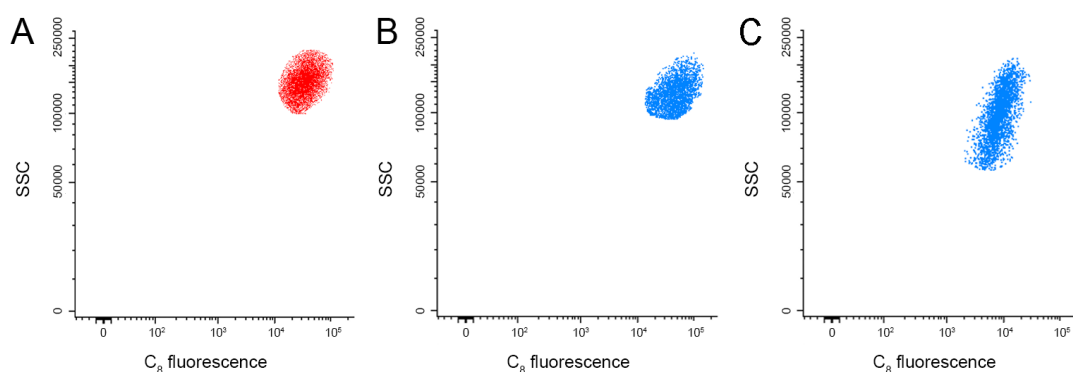


Figure 10.12 - Flow cytometry analysis of HeLa cells incubated with (A) free C₈, (B) AT11 G4-C₈ complex and (C) AT11-Bo G4-C₈ complex. Cells were incubated for 24 h with the preformed aptamer-ligand complex at a C₈ concentration of 1 μ M.

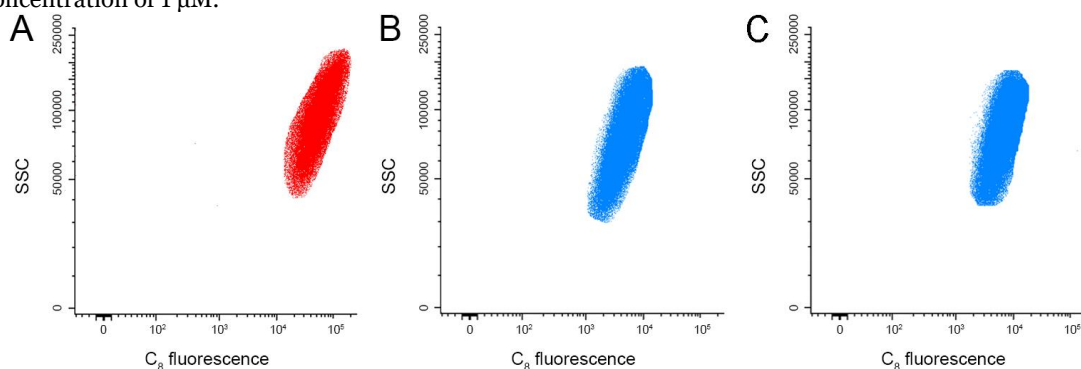


Figure 10.13 - Flow cytometry analysis of NHDF cells incubated with (A) free C₈, (B) AT11 G4-C₈ complex and (C) AT11-Bo G4-C₈ complex. Cells were incubated for 24 h with the preformed aptamer-ligand complex at a C₈ concentration of 1 μ M.

Following the determination of the subcellular localization of the complexes, MTT studies were performed to determine whether the AT11 G4-ligands (**Figure 10.14**) and AT11-Bo G4-ligands (**Figure 10.15**) could mitigate the cytotoxicity of the free acridine orange derivatives against non-malignant cells. The free AT11 G4 and AT11-Bo G4 were also evaluated in both HeLa and NHDF cell lines (Figure S5). The free sequences presented

similar toxicity and showed a more pronounced effect on the viability of cancer cells (ca. 60% *vs* 80% for HeLa and NHDF, respectively). This result agrees with previous studies that report AT11 G4 as being selective for cancer cells [9]. In the NHDF cell line and at the same dosage, free C₃ and AT11 G4-C₃ complex presented similar degrees of cytotoxicity, however in the case of HeLa cells the cytotoxicity of the latter was significantly higher. AT11-Bo G4-C₃ also presented a higher cytotoxicity against cancer cells but a reduced cytotoxicity against non-malignant cells. AT11 G4-C₅ showed similar cytotoxicity in HeLa cells when compared with free C₅. However, complex formation greatly reduced the toxicity of the C₅ ligand towards NHDF cells. Similar effect occurred in the case of AT11-Bo G4-C₅, although its cytotoxicity towards the cancer cell line was also reduced (47% of cell viability). The results obtained for C₈ were not so encouraging as those above discussed for the congener C₃ and C₅. In fact, the C₈-complexes induce lower cytotoxicity towards malignant cells than free C₈. In the case of AT11-Bo G4 this may be due to decreased internalization of the complex when compared to free C₈ as suggested by the flow cytometry results. Indeed, when compared to AT11-C₈, the AT11-Bo-C₈ complex showed a highest cytotoxicity reduction in HeLa cells relatively to free C₈. By contrast, AT11 G4-C₈ and AT11-Bo G4-C₈ showed a rather similar cytotoxicity against non-malignant NHDF cells, which is quite comparable to the cytotoxicity exhibited by free C₈ in the same cell line.

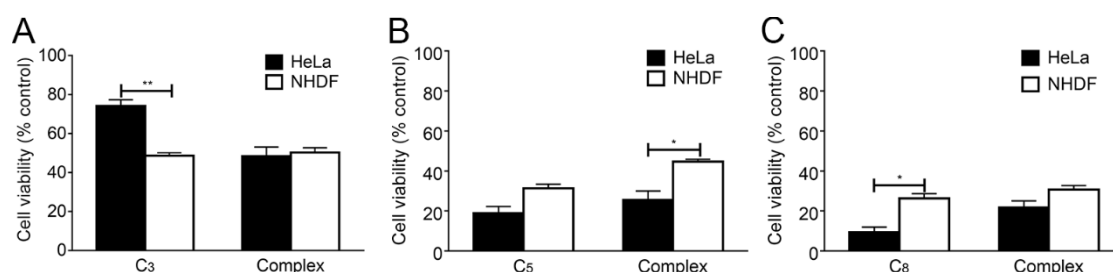


Figure 10.14 - Percentage of viable HeLa and NHDF cells after incubation with free C₃ (2 μ M), C₅ (2 μ M) and C₈ (1 μ M), or the AT11 G4 -ligands (pre-formed AT11-ligand conjugates by 10 min incubation) for 7 days. Results are presented as mean $\pm$ SEM.

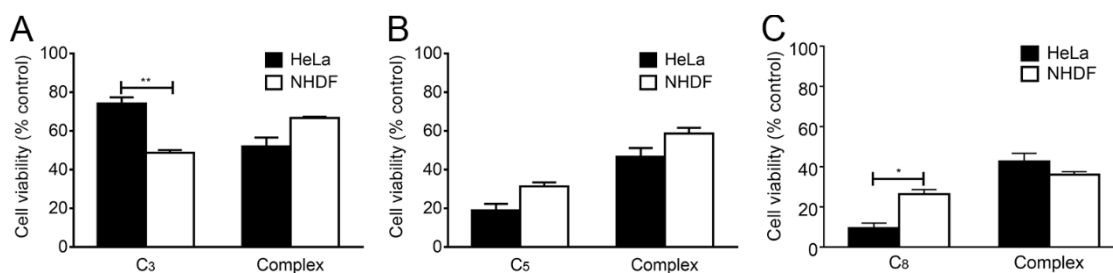


Figure 10.15 - Percentage of viable HeLa and NHDF cells after incubation with free C₃ (2 μ M), C₅ (2 μ M) and C₈ (1 μ M), or the AT11-Bo G4-ligands (pre-formed AT11-Bo-ligand conjugates by 10 min incubation) for 7 days. Results are presented as mean $\pm$ SEM.

Conclusions

Herein we reported a non-covalent conjugation approach between AT11 and AT11-Bo G4 and acridine orange derivatives in order to promote their selective delivery into HeLa cervical cancer cells. The resulting complexes were studied by CD, UV and fluorescence spectroscopies. The serum stability of the complexes versus free DNA was also determined. Additionally, the aptamers and their conjugates were tested in a cellular environment to assess their nucleolin binding potential, cell penetration ability and cytotoxic effect against HeLa cancer and healthy cells. The results showed that the interaction of the acridine derivatives with the aptamers is characterized by a medium-high affinity (K_D values in the $10^{-6} - 10^{-7}$ M range), leading to the stabilization of their G4 structure (> 10 °C for all ligands). The formation of supramolecular complexes reduced the cytotoxicity of the free ligands towards non-malignant cells. Additionally, both AT11 and AT11-Bo seemed to bind nucleolin on the cell surface of HeLa cells, as suggested by the colocalization experiments using anti-nucleolin antibody. In summary, the complexes AT11 G4-C₃, AT11 G4-C₅, AT11-Bo G4-C₃ and AT11-Bo G4-C₅ seem more promising candidates as therapeutic tools for cervical cancer than the C₈ congeners. Eventually, this trend might reflect the highest stability of the complexes AT11 G4-C₈ and AT11-Bo G4-C₈ that can hinder the intracellular release of C₈ to exert its therapeutic action. Our results emphasize the potential of using G4-aptamers as drug delivery systems for cervical cancer.

Acknowledgments

J. Carvalho acknowledges a doctoral fellowship grant from the FCT – Foundation for Science and Technology ref. SFRH/BD/122953/2016. J. Figueiredo acknowledges a fellowship from the project PINFRA/22161/2016 PTNMR. J. Nunes acknowledges a fellowship from the project UTAustin FCT project DREAM ref. UTAP-EXPL/NTec/0015/2017. This work was supported by project “Ações Integradas Luso-Francesas” ref. TC-15/17, FCT project ref. IF/00959/2015 financed by Fundo Social Europeu and Programa Operacional Potencial Humano, Project Pessoa 5079/2019, FCT project ref. UID/Multi/04349/2019, project Fundação Luso-Americana (FLAD) Healthcare 2020 ref. 45/2018, MIT Portugal project BIODEVICE ref. MIT-EXPL/BIO/0008/2017 and UTAustin FCT project DREAM ref. UTAP-EXPL/NTec/0015/2017. This work was also supported by POCI – COMPETE 2020 – Operational Programme Competitiveness and Internationalisation in Axis I – Strengthening research, technological development and innovation (project POCI-01-0145-FEDER-007491) and the SYMBIT project (reg. no.

CZ.02.1.01/0.0/0.0/15_003/0000477) financed by the ERDF (to J.L.M.). C2TN/IST authors gratefully acknowledge the FCT support through the UID/Multi/04349/2013 project.

References

- [1] Zhou, J., and Rossi, J. (2017) Aptamers as targeted therapeutics: Current potential and challenges. *Nature Reviews Drug Discovery* 16, 181–202. <https://doi.org/10.1038/nrd.2016.199>.
- [2] Alshaer, W., Hillaireau, H., and Fattal, E. (2018) Aptamer-guided nanomedicines for anticancer drug delivery. *Advanced Drug Delivery Reviews* 134, 122–137. <https://doi.org/10.1016/j.addr.2018.09.011>.
- [3] Vandghanooni, S., Eskandani, M., Barar, J., and Omid, Y. (2018) Recent advances in aptamer-armed multimodal theranostic nanosystems for imaging and targeted therapy of cancer. *European Journal of Pharmaceutical Sciences*. <https://doi.org/10.1016/j.ejps.2018.02.027>.
- [4] Bates, P. J., Reyes-Reyes, E. M., Malik, M. T., Murphy, E. M., O'Toole, M. G., and Trent, J. O. (2017) G-quadruplex oligonucleotide AS1411 as a cancer-targeting agent: Uses and mechanisms. *Biochimica et Biophysica Acta - General Subjects* 1861, 1414–1428. <https://doi.org/10.1016/j.bbagen.2016.12.015>.
- [5] Vandghanooni, S., Eskandani, M., Barar, J., and Omid, Y. (2018) AS1411 aptamer-decorated cisplatin-loaded poly(lactic-co-glycolic acid) nanoparticles for targeted therapy of miR-21-inhibited ovarian cancer cells. *Nanomedicine*. <https://doi.org/10.2217/nnm-2018-0205>.
- [6] Mongelard, F., and Bouvet, P. (2007) Nucleolin: a multiFACeTed protein. *Trends in Cell Biology* 17, 80–86. <https://doi.org/10.1016/j.tcb.2006.11.010>.
- [7] Dailey, M. M., Clarke Miller, M., Bates, P. J., Lane, A. N., and Trent, J. O. (2010) Resolution and characterization of the structural polymorphism of a single quadruplex-forming sequence. *Nucleic Acids Research* 38, 4877–4888. <https://doi.org/10.1093/nar/gkq166>.
- [8] Chung, W. J., Heddi, B., Schmitt, E., Lim, K. W., Mechulam, Y., and Phan, A. T. (2015) Structure of a left-handed DNA G-quadruplex. *Proceedings of the National Academy of Sciences* 112, 2729–2733. <https://doi.org/10.1073/pnas.1418718112>.

- [9] Do, N. Q., Chung, W. J., Truong, T. H. A., Heddi, B., and Phan, A. T. (2017) G-quadruplex structure of an anti-proliferative DNA sequence. *Nucleic Acids Research* 45, 7487–7493. <https://doi.org/10.1093/nar/gkx274>.
- [10] Carvalho, J., Pereira, E., Marquevielle, J., Campello, M. P. C., Mergny, J.-L., Paulo, A., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2018) Fluorescent light-up acridine orange derivatives bind and stabilize KRAS-22RT G-quadruplex. *Biochimie* 144, 144–152. <https://doi.org/10.1016/j.biochi.2017.11.004>.
- [11] Pereira, E., Do Quental, L., Palma, E., Oliveira, M. C., Mendes, F., Raposinho, P., Correia, I., Lavrado, J., Di Maria, S., Belchior, A., Vaz, P., Santos, I., and Paulo, A. (2017) Evaluation of Acridine Orange Derivatives as DNA-Targeted Radiopharmaceuticals for Auger Therapy: Influence of the Radionuclide and Distance to DNA. *Scientific Reports* 7, 42544. <https://doi.org/10.1038/srep42544>.
- [12] Manders, E. M. M., Verbeek, F. J., and Aten, J. A. (1993) Measurement of co-localization of objects in dual-colour confocal images. *Journal of Microscopy*. <https://doi.org/10.1111/j.1365-2818.1993.tb03313.x>.
- [13] Carvalho, J., Santos, T., Carrilho, R., Sousa, F., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2020) Ligand screening to pre-miRNA 149 G-quadruplex investigated by molecular dynamics. *Journal of Biomolecular Structure and Dynamics* 38, 2276–2286. <https://doi.org/10.1080/07391102.2019.1632743>.
- [14] Fan, X., Sun, L., Wu, Y., Zhang, L., and Yang, Z. (2016) Bioactivity of 2'-deoxyinosine-incorporated aptamer AS1411. *Scientific Reports* 6, 25799. <https://doi.org/10.1038/srep25799>.
- [15] Zhu, G., Niu, G., and Chen, X. (2015) Aptamer–Drug Conjugates. *Bioconjugate Chemistry* 26, 2186–2197. <https://doi.org/10.1021/acs.bioconjchem.5b00291>.
- [16] Bates, P. J., Laber, D. a., Miller, D. M., Thomas, S. D., and Trent, J. O. (2009) Discovery and development of the G-rich oligonucleotide AS1411 as a novel treatment for cancer. *Experimental and Molecular Pathology* 86, 151–164. <https://doi.org/10.1016/j.yexmp.2009.01.004>.
- [17] Mergny, J.-L., Li, J., Lacroix, L., Amrane, S., and Chaires, J. B. (2005) Thermal difference spectra: A specific signature for nucleic acid structures. *Nucleic Acids Research* 33, 1–6. <https://doi.org/10.1093/nar/gni134>.

- [18] Karsisiotis, A. I., Hessari, N. M. A., Novellino, E., Spada, G. P., Randazzo, A., and Webba Da Silva, M. (2011) Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. *Angewandte Chemie - International Edition* 50, 10645–10648. <https://doi.org/10.1002/anie.201105193>.
- [19] Carvalho, J., Queiroz, J. A., and Cruz, C. (2017) Circular dichroism of G-Quadruplex: A laboratory experiment for the study of topology and ligand binding. *Journal of Chemical Education* 94, 1547–1551. <https://doi.org/10.1021/acs.jchemed.7b00160>.
- [20] Zhang, X., Wei, Y., Bing, T., Liu, X., Zhang, N., Wang, J., He, J., Jin, B., and Shangguan, D. (2017) Development of squaraine based G-quadruplex ligands using click chemistry. *Scientific Reports* 7, 4766. <https://doi.org/10.1038/s41598-017-04344-x>.
- [21] Carvalho, J., Lopes-Nunes, J., Lopes, A. C., Cabral Campello, M. P., Paulo, A., Queiroz, J. A., and Cruz, C. (2019) Aptamer-guided acridine derivatives for cervical cancer. *Organic and Biomolecular Chemistry* 17, 2992–3002. <https://doi.org/10.1039/c9ob00318e>.
- [22] Reyes-Reyes, E. M., Teng, Y., and Bates, P. J. (2010) A new paradigm for aptamer therapeutic AS1411 action: Uptake by macropinocytosis and its stimulation by a nucleolin-dependent mechanism. *Cancer Research* 70, 8617–8629. <https://doi.org/10.1158/0008-5472.CAN-10-0920>.
- [23] Kim, S., Lee, K. Y., Kang, H., Ryu, S. H., Lee, D. S., and Lee, J. H. (2010) Bioimaging of nucleolin aptamer-containing 5-(N -benzylcarboxamide)- 2'-deoxyuridine more capable of specific binding to targets in cancer cells. *Journal of Biomedicine and Biotechnology* 2010. <https://doi.org/10.1155/2010/168306>.

Chapter 11

Nanoaggregate-Forming Lipid-Conjugated AS1411 Aptamer as a Tumour-targeted Delivery System of Anticancer Agents for Cervical Cancer Therapy

Nanoaggregate-Forming Lipid-Conjugated AS1411 Aptamer as a Tumor-targeted Delivery System of Anticancer Agents for Cervical Cancer Therapy

Abstract

Recent investigations suggest that acridine orange derivative C₈ exhibits anticancer and antiviral activities against cervical cancer cell lines and Human Papillomavirus, respectively. Nanoparticles offer targeted delivery of drugs with minimal toxicity to surrounding healthy tissue and have great potential in the management of HPV-related diseases. Pursuing a novel nanotherapy approach, we synthesized lipid-modified AS1411 aptamers capable of spontaneously forming nanoaggregates in solution containing Mg²⁺. The nanoaggregates presented suitable properties for pharmaceutical applications such as small size (around 100 nm), negative charge, and drug release. The nanoaggregates were loaded with C₈ for its specific delivery into cervical cancer cell lines and HPV-positive tissue biopsies. This improved inhibition of HeLa proliferation and cell uptake without significantly affecting healthy cells. Finally, the nanoaggregates were incorporated in a semisolid gel formulation with promising tissue retention properties aiming at developing a local delivery strategy of the nanoaggregates in the female genital tract. Collectively, these findings suggest that our novel nanoformulation protocol has great potential for the delivery of both anticancer and antiviral agents, becoming a novel modality for cervical cancer management.

Introduction

Cervical cancer is the 3rd most common cancer among women worldwide, with estimated 569,847 new cases and 311,365 deaths in 2018 alone (GLOBOCAN) [1]. The main cause for the development of cervical cancer is the human papillomavirus (HPV). High-risk HPV types 16 and 18 are responsible for the development of 70% of all cancers. The oncogenic potential of HPV is associated with the expression of oncoproteins E6 and E7 that inactivate the tumor suppressor system p53 and members of the pRb family. Despite recent efforts in implementing a combination of screening and vaccination programs for

HPV, which reduced cervical cancer incidence and mortality, the need of a treatment for patients already infected and in precancerous/cancerous states is still a priority. Treatment options for precancerous states are usually excision or ablation of the lesion tissue [2]. These procedures are usually effective; however, a high rate of recurrence is observed due to failure in eradicating the virus. As for cancerous lesions, the therapeutic options available are surgery, radio and chemotherapy [2]. These methods are invasive, with limited effectiveness and undesirable side-effects, and regrettably, about 50% of the cancer patients will die.

AS1411 (recently renamed ACT-GRO-777) is a 26-nucleotide oligonucleotide that acts as an aptamer of nucleolin, a protein found in the nucleus of proliferating cells (both normal and malignant) but also in the cytoplasm and particularly on the cell surface of cancer cells favouring the transport and cellular internalization of AS1411 [3,4]. The AS1411 aptamer induces proliferation arrest and cell death in a variety of cancer cells but has little effect on normal cells, which was recently suggested to be due to a overstimulated nucleolin-dependent micropinocytosis uptake mechanism in cancer cells, which is different from that of normal cells (mainly by endocytosis) [5]. AS1411 reached Phase II clinical trials as monotherapy for renal cell carcinoma (trial ID: NCT00740441) and in combination with cytarabine for acute myeloid leukaemia (trials ID: NCT00512083 and NCT01034410) [6–8]. Although these clinical trials did not lead to Phase III studies, the results showed that AS1411 was well tolerated for up to 7 days, with steady-state plasma concentrations, and induced long-lasting clinical responses in at least 7 patients [6–8]. Regardless of being safe for clinical application, its suboptimal pharmacology and low potency still remain a major obstacle for its progression as an approved drug.

More recently, the interesting AS1411's cancer-targeting properties have been extensively exploited in the therapy and diagnostic fields, particularly by using it as a cancer-selective delivery agent of drugs and imaging agents [3,9]. The aptamer has been used as a targeting agent of a variety of nanosystems such as gold nanoparticles [10], liposomes [11], micelles [12] and other polymeric nanoaggregates [13,14] by attaching it to the surface of the systems. Additionally, AS1411 has been employed to non-covalently convey molecules with anticancer potential such as the cationic porphyrin TMPyP4 [15] and the acridine orange derivative C₈ [16]. These works suggest that AS1411-linked nanosystems/drugs have higher cancer cell internalization and/or tumour accumulation compared to untargeted systems or free drugs. Ultimately, this leads to the reduction of toxicity to normal tissues which is a common problem for most chemotherapeutic drugs.

Due to its guanine-rich nature, AS1411 is able to fold into multiple mono- and bimolecular G-quadruplex (G4) conformations which are important for its biological activities [3]. Trent *et al.* reported the existence of at least eight monomeric interconvertible G4 conformations for AS1411 [17]. To minimize the conformational polymorphism and attempt to find the bioactive conformation, several research groups have developed related oligonucleotides harbouring nucleotide substitutions and/or additions, as well as sequences with modified nucleosides or conjugation to different chemical groups [18,19]. Recently, Barthélémy *et al.* showed that the lipid modification of G4 prone sequences has the potential of controlling their folding and resulting conformation [20]. A similar strategy was used by Montesarchio *et al.* when they conjugated a saturated 18-carbon alkyl chain to AS1411 5'-end which favoured the formation of parallel G4 structures exclusively [21]. However, these studies have not addressed the potential of using these lipid-modified sequences in nanotechnology applications such as drug delivery.

In this work, we synthesized AS1411 sequences with a 5' lipidic tail (mono octadecyl chain, C₁₈H₃₇) and explored their ability to form nanoaggregates. Nanoaggregates have recently emerged as promising supramolecular colloidal carriers in molecular targeted therapies [22]. Nanoaggregates possess some fundamental properties such as nanosize, high loading efficiency, and *in vivo* stability. These are suitable for encapsulation of poorly water-soluble drugs by physical encapsulation [22]. Due to physical encapsulation, a high amount of drug can be loaded into the nanoaggregates, which helps to achieve a maximum drug concentration at the target site [22]. Once the formation of nanoaggregates was attained, we employed these lipophilic nanosystems to selectively deliver compounds with therapeutic potential both *in vitro* and *ex vivo*, taking advantage of AS1411 as a targeting ligand. The compound used in this study, the acridine orange derivative C₈, has already been shown to have promising anticancer and antiviral effects against both cervical cancer cell lines [23] and replicative HPV18 in organotypic raft cultures [24]. Altogether, our findings demonstrate the potential of AS1411-functionalized nanosystems as effective drug delivery systems for the development of alternative therapeutic strategies for HPV-induced cervical cancer.

Materials and Methods

DNA synthesis and purification

5'-octadecyl-AS1411 (C₁₈H₃₇-5'-GGTGGTGGTGGTTGTGGTGGTGGTGG-3') was synthesized using the phosphoramidite methodology on an automated H8 (K&A

Laborgeraete, Germany) DNA synthesizer (K&A Laborgeraete, Germany) at the μmol scale on 1000 Å primer support following the previously published method [20]. The lipid-oligonucleotides were purified on a preparative C4-reverse phase HPLC column (Waters, Xbridge Protein BEH, 300 Å, 5 μm , 30mm x 50mm) followed by dialysis against a 10 mM LiCl solution (to disfavour G4 formation, as Li⁺ ions do not support stable G4 formation) followed by Milli-Q water. The identity and purity of the oligomers were proved by MALDI-TOF mass spectrometry. Sample concentration was determined from the absorbance value at 260 nm and the oligonucleotide molar extinction coefficient following Beer-Lambert relationship. The epsilon was calculated using the Integrated DNA Technology online oligo analyser tool which uses the standard nearest neighbour method.

Nanoaggregates formation

The nanoaggregates formed spontaneously following the resuspension of the lipid-oligonucleotide conjugate in 20 mM potassium phosphate buffer pH 6.9 containing 50 mM NaCl, 5 mM KCl and 5 mM MgCl₂ (hereafter referred to as 1× buffer). Nanoaggregates were formed in both the absence and presence of C₈. In this case, 0.5 μM of C₈ was added to 1× buffer prior to the addition of the lipid-oligonucleotide conjugate. Nanoaggregates formation was detected by agarose gel electrophoresis.

Agarose gel electrophoresis

Gel electrophoresis was carried out using 1% agarose gels in 1× Tris-acetate-EDTA (TAE) buffer at 120 V for 50 min. After staining with Xpert Green DNA Stain (GRiSP, Lda., Portugal) gels were imaged with a UVitec FireReader system (UVitec, UK). HyperLadder™ 25bp (Bioline, UK) was used as a DNA molecular weight marker.

Characterization of nanoaggregates

The nanoaggregates size distribution and the zeta potential were determined using dynamic light scattering (DLS) in a Zetasizer Nano ZS instrument (Malvern Instruments, UK). The samples were analysed using a disposable cuvette with a dip cell (ZEN1002), at 25 °C. The morphology and size of the nanoaggregates was further evaluated by transmission electron microscopy (TEM) using a Hitachi HT7700 microscope. In brief, the nanoaggregate samples were placed on formvar-coated copper grids and allowed to dry at room temperature, then the images were acquired at a voltage of 80 kV with 0.204 nm resolution.

Circular dichroism

All CD spectra were recorded using a Jasco J-815 spectropolarimeter (Jasco, Japan) equipped with a Peltier-type temperature control system. Lipid-oligonucleotides were used at 10 μM concentration in 1 $\times$ buffer. The measurements were made in a 10 mm pathlength quartz cell at 25 $^{\circ}\text{C}$ using an instrument scanning speed of 100 nm/min through wavelengths ranging from 200 to 340 nm. The spectra were signal-averaged over four scans and baseline corrected by subtracting a buffer spectrum. CD melting experiments were performed in the 15-95 $^{\circ}\text{C}$ temperature range, with a heating rate of 1 $^{\circ}\text{C}/\text{minute}$ by monitoring the ellipticity at 265 nm. Data was converted into fraction folded plots and fitted to a Boltzmann distribution (OriginPro 8) to determine the melting temperatures (T_m) from the two-state transition model as previously reported [25].

Nuclear Magnetic Resonance

^1H NMR spectra were recorded on a Bruker Advance III 600 MHz spectrometer (Bruker, EUA) equipped with a QCI cryoprobe at 298 K using standard zgpg30 experiments. Oligonucleotide concentration was 500 μM in 3 mm tubes in a total volume of 180 μL . Samples were prepared by diluting the lipid-oligonucleotide in 1 $\times$ buffer, supplemented with 10% D_2O . All spectra were acquired and processed with Topspin 3.1 software. Chemical shifts (δ) were measured in ppm.

Critical micellar concentration (CMC) determination

The CMC value of the AS1411-decorated nanoaggregates was determined using the standard pyrene encapsulation method [26]. Briefly, a series of pyrene solutions at a fixed concentration of 2 μM were prepared containing increasing concentrations of 5'-octadecyl-AS1411 (0-20 μM). Samples were then excited at 338 nm and the emission was recorded between 350-550 nm using a FluoroMax 4 spectrofluorimeter (Horiba Scientific, Japan). When pyrene enters nanoaggregates from the aqueous phase, it alters the intensity ratio of the first and third vibrational peaks, *i.e.* the $I_{\text{I}}/I_{\text{III}}$ ratio. The CMC value was taken from the intersection of the linear fits at the point of inflection of the data points at low and high lipid concentrations.

Drug release

The characterization of the nanoaggregates drug release profile was performed according to a method previously described in the literature [27]. Briefly, 100 μL of nanoaggregate sample were inserted in a Slide-A-LyzerTM MINI Dialysis Device (Thermo Fisher Scientific,

USA) with a 3.5 kDa molecular cut-off. The dialysis was performed in 1.5 mL tubes under constant stirring using an HulaMixer Sample Mixer (Thermo Fisher Scientific, USA) in 1 ml of 1× buffer solution. At different time points up to 72h, 100 µL samples were collected and the same volume of fresh 1× buffer was added to the dialysis medium, in order to maintain constant volume during the experiment. Released C₈ content was quantified by measuring the fluorescence of the samples using a Spectramax Gemini XS spectrofluorometer (Molecular Devices LLC, USA), using excitation and emission wavelengths of 498 and 530 nm, respectively.

Nanoaggregate cellular uptake

Normal human dermal fibroblasts (NHDF) were grown in RPMI-1640 medium supplemented with 0.01 M HEPES, 0.02 M L-glutamine, 0.001 M sodium pyruvate, 10% fetal bovine serum, and 1% penicillium/streptomycin antibiotic. Human cervical cancer cells (HeLa) were grown in DMEM medium supplemented with 10% fetal bovine serum, and 1% penicillium/streptomycin antibiotic. HeLa and NHDF cells were cultured in Ibidi 8-well µ-slides (5×10^4 cells/mL) overnight at 37 °C in a humidified atmosphere containing 5% CO₂. After cell adhesion, cells were incubated with 5 µM of nanoaggregates loaded with 0.5 µM of C₈ for up to 72 h. After the incubation period, the wells were washed by rinsing with PBS three times and cell nucleoli were then stained with 1 µM nuclear probe Hoechst 33342 for 15 min. Prior to visualization, excess probe was washed off by rinsing with PBS three times. Images were acquired using a Zeiss AxioObserver LSM 710 microscope (Carl Zeiss, Germany) with 405 and 488 nm laser excitation for Hoechst 33342 and C₈, respectively. For colocalization studies with anti-NCL antibody (PA3-16875, Thermo Scientific), cells were incubated for 2 h at room temperature with primary antibody at a 1:100 dilution followed by 1 h incubation with anti-rabbit secondary antibody conjugated to Alexa Fluor™ 647 fluorophore diluted at 1:1000. Appropriate emission bands were selected for the three fluorescent channels, considering the spectral overlap and potential bleed-through between channels. Images were processed with Zeiss ZEN Black software.

Nanoaggregate tissue penetration

Cervix tissue biopsies of women with cervical cancer were obtained from the Institute of Molecular Pathology and Immunology of the University of Porto (Ipatimup) according with the protocol approved by the Ethical Committee of University Hospital Center Cova da Beira (study number 92/2018). The tissue fragments were fixed in 10% formaldehyde and embedded in paraffin. Serial sections of 5 µm were cut from each block, dewaxed and

the antigen retrieval done using Trilogy™ reagent (Sigma Aldrich, USA) following the manufacturer protocol. Sections were then blocked with 2% FBS and incubated with the C₈-loaded nanoaggregates and anti-NCL antibody as described above. Images were acquired using a Zeiss AxioObserver LSM 710 microscope (Carl Zeiss, Germany).

Flow cytometry

HeLa and NHDF cells were seeded in 12-well plates (5×10^5 cells/mL) and incubated overnight for cell adhesion at 37 °C in a humidified atmosphere containing 5% CO₂. Then, cells were incubated with 5 µM of nanoaggregate loaded with 0.5 µM of C₈ for 24 h or free C₈ as control. After the incubation period, the wells were washed by rinsing with PBS three times. Cells were then trypsinized, resuspended in PBS and analysed in a BD FACSCanto™ II flow cytometry system (BD Life Sciences, US) to evaluate the uptake of the nanoaggregates. C₈ fluorescence was used as an indicator of cellular uptake of nanoaggregates. Non-specific colouring and debris were excluded by analysing FSC vs SSC density plot as dead cells and debris have lower forward scatter levels.

In vitro antitumoral effect

For MTT assays, HeLa and NHDF cells were seeded in 48-well plates (2×10^4 cells/mL) and incubated overnight for cell adhesion. Then, cells were incubated for 72h with 5 µM of blank nanoaggregate or C₈-loaded nanoaggregate at a drug concentration of 0.5 µM. Wells containing untreated cells were used as control. At the end of incubation, the media was replaced with fresh media containing 3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide salt (MTT) and further incubated at 37 °C for 1 h. The resultant formazan crystals were then dissolved in DMSO and the absorbance recorded in a BioRad xMark™ microplate reader (BioRad, USA) at 570 nm. Cell viability was expressed as mean ± SEM from at least three different plates with each condition tested in triplicate wells. All data treatment was performed using GraphPad Prism 6.

Gel formulation preparation

Gel formulations were prepared according to Tien *et al.* [28] by dissolving 1.50 g of hydroxypropylmethylcellulose (Methocel K100M) in 90.36 g of water containing 0.14 g of sorbic acid. Two different formulation with 8 and 10% propylene glycol were prepared by adding 8 and 10 g, respectively. The final pH was adjusted to 5.5 by adding HCl, and the gel was centrifuged at 3000 rpm for 5 minutes and stored at 4 °C. Prior to each experiment, the formulations were mixed with C₈-loaded nanoaggregates with drug concentration at 0.5 µM.

Ex vivo vaginal permeation assay

Porcine female reproductive genital tract was kindly provided by a local slaughterhouse from 3 female pigs up to 6 months (non-reproductive animals) and processed within 3 h after the collection. Vaginal tissue was isolated from the other genital organs using surgical scissors. Then, the tissues were washed in 0.9% NaCl. Before the permeation study, tissues were manually dermatomed to a thickness of a maximum of 300 μm (BBraun equipped with Aesculap blades, Germany). Tissue thickness was confirmed using a micrometer (Vogel, Germany). Tissue diameter was adjusted using a surgical scalp in order to fit the Franz cell diameter. A Static Franz Diffusion Vertical Cell equipment (PermeGear, GmbH Analysesysteme, Germany) was used to assembly 6 cells in the same experimental set. Each experiment was performed at least three times. An infinite dose (300 mg), *i.e.* an excess of formulation able to achieve and maintain a maximum rate of absorption, was applied uniformly to the tissue surface (epithelial side). The experiment was performed for 24 h, and 200 μL aliquots were collected at 0, 6, 12 and 24 h, being immediately replaced by fresh receptor media (PBS) pre-heated at 37 °C. Tissues were also collected and extracted by a 20 min sonication with 2 $\mu\text{L}/\text{mg}$ tissue of PBS.

Chromatographic method development and validation

To determine C_8 concentration in the vaginal tissue, donor and receptor chambers, a simple high performance-liquid chromatography coupled to a fluorescence detector (HPLC-FLD) method was developed and validated according to the guiding principles of the Food and Drug Administration (FDA) [29]. A 1290 HPLC with a binary pump coupled to a 1260 FLD from Agilent Technologies (Soquimica, Lisboa, Portugal) was used. C_8 was separated with an YMC-Triart PFP (5 μm , 4.6 $\times$ 150 mm) analytical column coupled to a Guard-c holder (4 $\times$ 10 mm) containing Triart PFP (5 μm , 3 $\times$ 10 mm) pre-column, all from YMC Europe GMBH (Solítica, Portugal). The HPLC-FLD worked on isocratic mode with a mobile phase composed of methanol/0.1% trifluoroacetic acid (9:1, v/v), the flow rate was 0.5 mL/min, and the injection volume was 50 μL . The sampler and column temperatures were set at 4 and 25 °C, respectively, being the analyte detected using 498 and 530 nm as excitation and emission wavelengths, respectively. The chromatogram presented in Fig. S1 shows the compound's peak at 9.57 minutes.

Analysis of tissue permeation/penetration

The amount of C_8 released from the nanoaggregate-loaded formulation and permeated/penetrated the vaginal tissue was determined by HPLC-FLD method developed and validated as described above. Samples (50 μL) collected from the Franz

cells were injected onto the chromatographic system using the method previously described. In order to quantify C₈ the method was validated according to FDA guidelines in terms of linearity, limit of quantitation and detection, precision and accuracy [29].

Statistical analysis

The statistical analysis was performed by using Student's unpaired t-test or ANOVA followed by Dunnet's test for multiple comparisons. A *p* value < 0.05 was considered statistically significant. Data analysis was performed in GraphPad Prism 6 (San Diego, USA).

Results and Discussion

There is strong evidence that suggests AS1411 is an effective targeting ligand for nanoparticle-based drug delivery systems [3]. The majority of nanosystems found in the literature are composed of two or more copolymers, and require several reactions and synthetic steps for their preparation [12,30,31]. In this work, we show that a simple lipid modification of AS1411 allows the formation of nanoaggregates that have an improved speed of assembly and tuneable stability (**Figure 11.1**). These nanoaggregates were used to selectively deliver an acridine orange derivative, compound C₈, with antitumoral [23] and antiviral activity [24]. The simplicity of these AS1411-linked nanoaggregates could facilitate their manufacturing, scale-up, and ultimately regulatory approval.

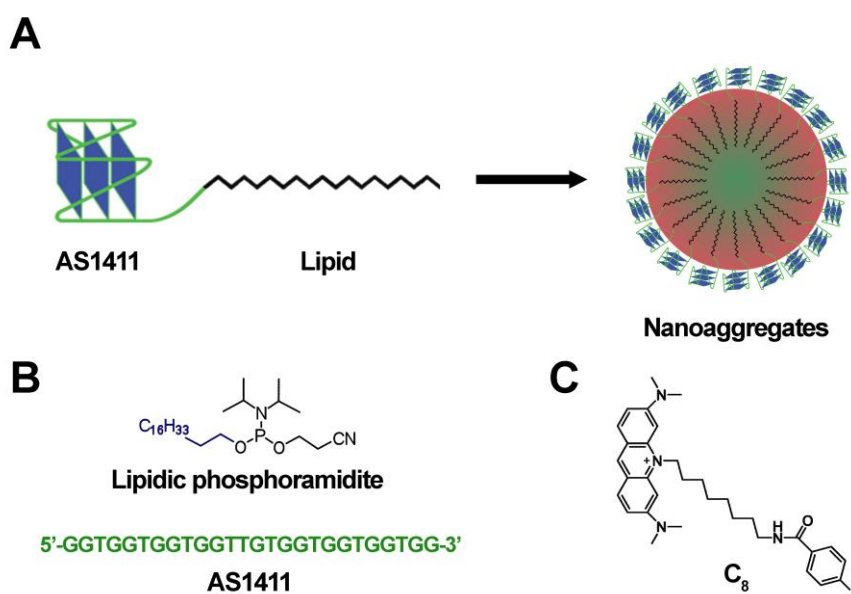


Figure 11.1 - (A) Lipid-modified AS1411 oligonucleotides self-assemble in nanoaggregates decorated with G4 structures. The nanoaggregates form spontaneously by resuspending the lipid-oligonucleotide conjugate in 1× buffer. Similarly, nanoaggregates were also formed in the presence of C₈. In this case, C₈ was added to 1× buffer prior to the addition of the lipid-oligonucleotide conjugate. (B) Chemical structure of the octadecyl

lipidic phosphoramidite used and the sequence of AS1411; (C) Chemical structure of the acridine orange derivative C₈ used in this study.

Initially, the ability of the synthetic lipid-modified AS1411 oligonucleotides to form nanoaggregates was evaluated by agarose gel electrophoresis. The importance of buffer composition, particularly in terms of cations, was evaluated by using increasing concentrations of potassium (K⁺), sodium (Na⁺) and magnesium (Mg²⁺) (**Figure 11.2**). Nanoaggregate assemblies were observed in the presence of all types of cations used as suggested by the retardation of the oligonucleotide bands, particularly in the presence of Mg²⁺. This may be due to a Mg²⁺-mediated neutralization of the nanoaggregate anionic charges with concomitant entropic gain [20]. At concentrations greater than 50 mM Mg²⁺ the bands suggest that larger aggregates are formed. Nanoaggregate formation was also observed in water in the absence of salts, but to a lesser extent and with monomeric species clearly visible further demonstrating the strong cation-dependence. Based on these results, the buffer conditions for the remaining experiments were defined as 50 mM NaCl, 5 mM KCl and 5 mM MgCl₂ (hereafter referred to as 1× buffer), which also took into account previously published results [20]), which promoted the formation of stable nanoaggregates (Figure S2). The fact that the nanoaggregates survived the sieving conditions of electrophoresis further indicated that these were rather stable even in the absence of salts. An 18 carbon atoms lipid tail (such as the octadecyl used herein) was already shown to be very effective at stabilizing nucleolipids within liposomes or other lipid nanosystems by hydrophobic interactions [21].

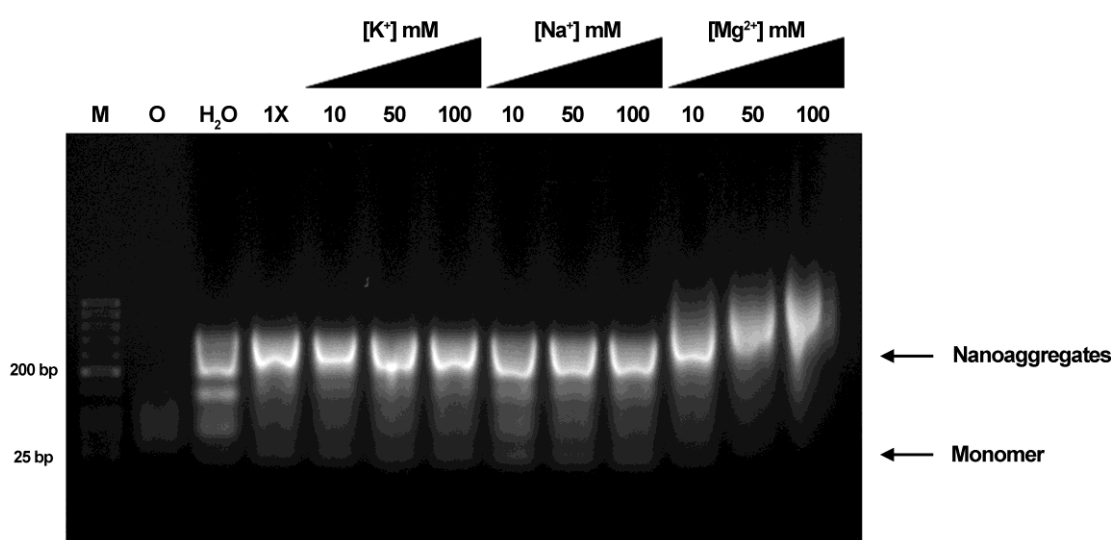


Figure 11.2 - Native agarose gel electrophoresis of 5'-lipid-AS1411 in different ionic conditions or in the absence of any added salt.

The effect of the lipid conjugation on the AS1411 G4 structure was then evaluated by circular dichroism and nuclear magnetic resonance (NMR) (**Figure 11.3** and S2). The CD spectra of 5'-lipid-AS1411 in 1× buffer showed a positive band at around 260 nm and a negative band at 240 nm (**Figure 11.3A**) both of which are characteristic of predominantly parallel stranded G4 structures [21,25]. Thus, attaching the lipid to the 5'-end of the sequence does not affect the ability of AS1411 to fold into a stable G4 structure. Furthermore, the overall folding topology seems to be maintained when compared to the unmodified AS1411 CD profile [12]. This was further confirmed by ¹H NMR experiments whose spectra demonstrated the appearance of the typical H1 imino signals in the 10-12 ppm range (Figure S3), corresponding to the guanine residues involved in Hoogsteen hydrogen bonding in G-quartets [32]. The broad unresolved signals in the H1 imino, H8 aromatic and H1' sugar regions are probably due to the formation of nanoaggregate assemblies [33]. The formation of slow tumbling species on a NMR time scale (μs to ns) leads to less spectral resolution and extensive line broadening and may also be indicative of the formation of a nanoaggregate that presents a larger molecular weight than the lipid-oligo monomers [34]. The thermal stability of the G4 structure formed by 5'-lipid-AS1411 was monitored by CD melting experiments (**Figure 11.3B**). The *T_m* value determined for 5'-lipid-AS1411 was 69.0 ± 0.3 °C, again indicative of stable G4 structures, even at lower K⁺ concentrations (5 mM) when compared to unmodified AS1411 sequences found in the literature and previously reported by our group (56.3 ± 0.1 °C at 10 mM K⁺) [16,21].

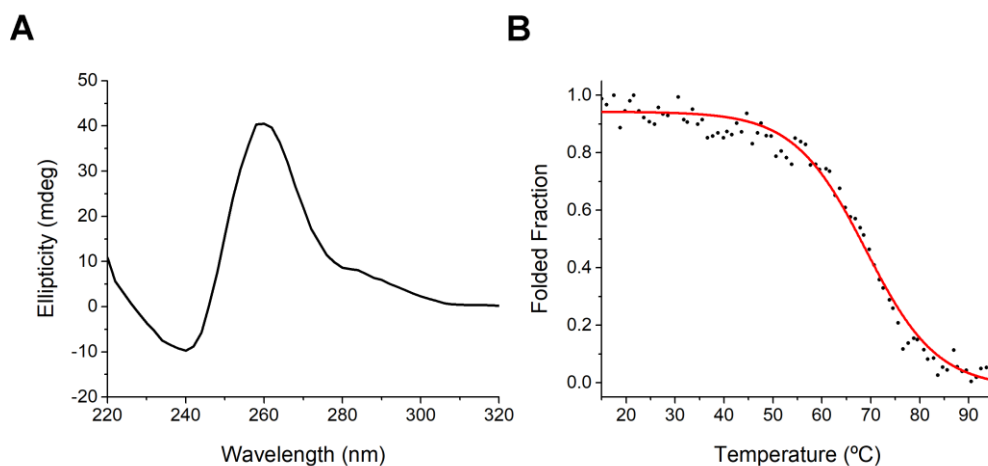


Figure 11.3 – (A) CD spectrum of 5'-lipid-AS1411 G4 structure showing the typical bands of a parallel stranded G4; (B) CD melting profile of 5'-lipid-AS1411 in 1× buffer obtained by monitoring the ellipticity at 260 nm at increasing temperatures.

Once the ability of the 5'-lipid-AS1411 to fold into a stable G4 structure was demonstrated, the nanoaggregates formed by the sequence were characterized in terms of its CMC, size, surface charge and shape. The CMC was determined using the pyrene method [26]. The

fluorescence of pyrene changes significantly when it enters the hydrophobic core (or the vicinity) of the nanoaggregates. The CMC value can thus be determined from the intersection of the linear fits through the points at low and high 5'-lipid-AS1411 concentrations. The CMC was determined in both water and 1× buffer (Figure S4). The CMC value of 5'-lipid-AS1411 in water was 6.03 μM (Figure S4A) while in 1× buffer the value was 2.84 μM (Figure S4B). The lower CMC value in 1× buffer emphasizes the importance of the salts for nanoaggregate formation and stability, as previously suggested by the agarose gel electrophoresis experiments. Nonetheless, both CMC values are low suggesting that 5'-lipid-AS1411 should self-assemble with considerable stability in both water and 1× buffer, as already showed by surviving the sieving conditions of gel electrophoresis.

Following the determination of the CMC values, the nanoaggregates were then characterized in terms of size, shape and surface charge. The particle size, polydispersity index (PDI) and zeta potential of the nanoaggregates are listed in **Table 11.1**.

Table 11.1 - Characterization of AS1411 nanoaggregates in terms of size and surface charge.

Sample	Particle size (nm)	PDI	Zeta potential (mV)
Blank nanoaggregates	86.9 $\pm$ 7.3	0.263	-23.3 $\pm$ 0.7
C ₈ -loaded nanoaggregates	104.0 $\pm$ 0.6	0.218	-24.8 $\pm$ 1.9

The nanoaggregates formed by 5'-lipid-AS1411 exhibited a mean particle size of 86.9 $\pm$ 7.3 nm (**Figure 11.4A**) with a spherical structure as demonstrated by TEM micrographs (**Figure 11.4B**). The zeta potential of the blank nanoaggregates was -23.3 $\pm$ 0.7 mV. Similarly, nanoaggregates encapsulating the acridine orange derivative C₈ were also prepared and characterized. The mean particle size of C₈-loaded nanoaggregates was 104.0 $\pm$ 0.6 nm (**Figure 11.4C**) while a zeta potential of -24.8 $\pm$ 1.9 mV was measured. The increase in size after C₈ incorporation was attributed to the physical encapsulation of the drug in the hydrophobic core of the nanoaggregate, which maintained an overall spherical structure (**Figure 11.4D**). There were however no significant changes in the zeta potential of the nanoaggregates. Overall, both blank and C₈-loaded nanoaggregates presented small particle sizes (<200 nm) and negative surface charges (<-20 mV), which are important factors for passive tumour accumulation via the enhanced permeability and retention (EPR) effect, longer circulation times and less accumulation in the liver, spleen and kidneys [22,35]. The stability of the nanoaggregates was monitored for 7 days at room temperature. As indicated by DLS and TEM analysis (Figure S5), there was a significant particle size increase (189.6 $\pm$ 0.2 nm) but the nanoaggregates remained smaller than 200

nm and maintained their overall structure, as suggested by the micrographs. These results indicate great stability for these nanosystems which is an important feature for clinical applicability.

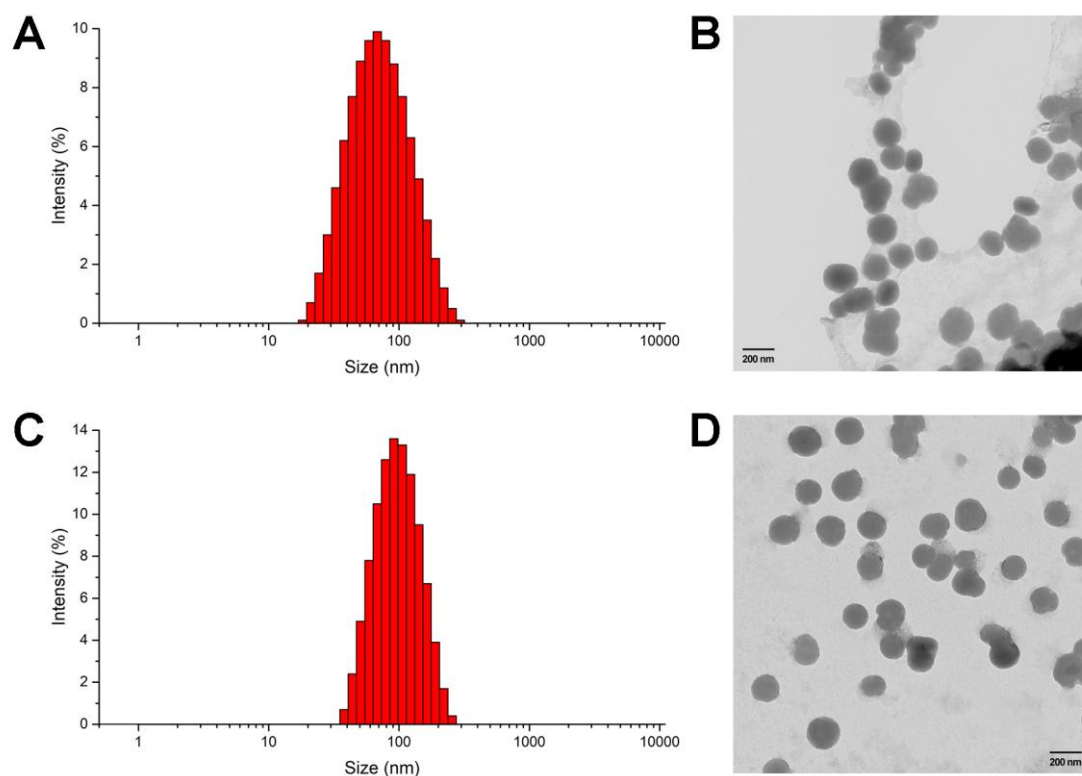


Figure 11.4 - (A) DLS and (B) TEM analysis (12,000 $\times$ magnification) of nanoaggregates formed by 5'-lipid-AS1411 in 1 $\times$ buffer; (C) DLS and (D) TEM (12,000 $\times$ magnification) analysis of 5'-lipid-AS1411 nanoaggregates encapsulating the acridine orange derivative C₈. Scale bar is 200 nm.

The release of C₈ from the nanoaggregates was monitored over time and is shown in **Figure 11.5**. C₈-loaded nanoaggregates had a sustained drug release over 72 h, with a mean cumulative release of about 80% at the end of the incubation period. A burst of release was observed in the first 6 h with an ensuing steady release in the following hours.

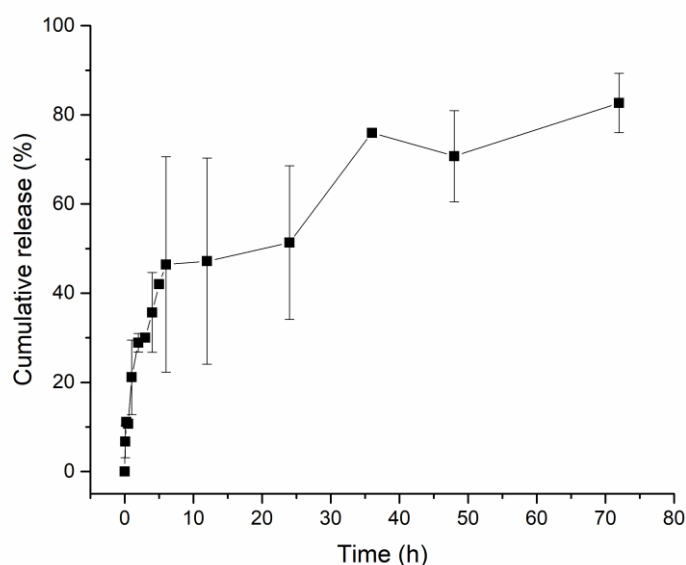


Figure 11.5 - Cumulative release profile of C₈ from the AS1411-decorated nanoaggregates in 1× buffer for 72h. Values are presented as mean ± SD.

The *in vitro* antitumoral effect of the AS1411-decorated nanoaggregates was then evaluated in human cervical cancer cells (HeLa) and normal human fibroblasts (NHDF). The results of a MTT assay using both blank and C₈-loaded nanoaggregates are presented in **Figure 11.6**. Blank nanoaggregates elicited only a minor effect on the viability of both HeLa and NHDF cells (81.6% and 81.9% mean viability, respectively), demonstrating the desirable safety of the vehicle. For the C₈-loaded nanoaggregates, there was a significantly higher effect on the viability of HeLa cancer cells (41.6% mean viability) when compared to NHDF normal cells (76.4% mean viability). The cancer selectivity is presumably conferred by the presence of AS1411 on the surface of the nanoaggregates, as its cancer-targeting properties and selective accumulation due to the overexpression of nucleolin on the surface of cancer cells [36] have already been described [3,5]. Furthermore, and since blank nanoaggregate promoted a similar decrease in the viability of both cancer and healthy cells, other possible explanation for the cancer-selectivity of C₈-loaded nanoaggregates includes the fact that the tumour cells microenvironment is different from that of non-malignant cells, possibly promoting C₈ release by nanoaggregates destabilization [37].

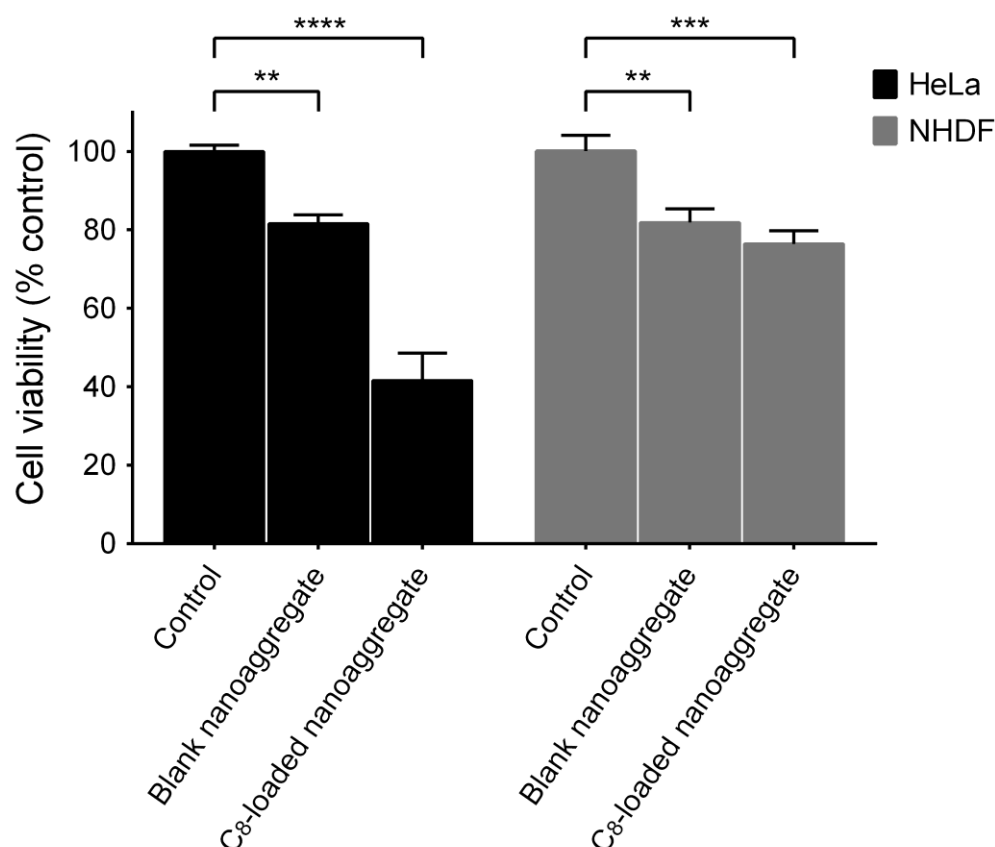


Figure 11.6 - Relative cell viability of HeLa and NHDF cells incubated for 72 h with either blank nanoaggregate or C₈-loaded nanoaggregate at a drug concentration of 0.5 μ M. Values are presented as mean $\pm$ SEM.

In order to further characterize the cancer-selective targeting of the nanoaggregates, *in vitro* cellular uptake was characterized by flow cytometry using both HeLa and NHDF cells (**Figure 11.7A-B**). For HeLa cancer cells the cellular uptake of C₈-loaded nanoaggregates was considerably higher than that of free C₈, suggesting an enhanced internalization of the nanoaggregates (**Figure 11.7A**). In contrast, NHDF cells did not show any difference in the internalization of the C₈-loaded nanoaggregates when compared to the free drug (**Figure 11.7B**). These results may provide an explanation for the cancer-specific antiproliferative effect observed in the MTT assays (**Figure 11.6**), since the overexpression of nucleolin on the surface of HeLa cancer cells may facilitate the cellular uptake of the C₈-loaded nanoaggregates, as described before for other AS1411-decorated systems [12,30]. The cellular internalization and trafficking of C₈-loaded nanoaggregates was also evaluated by fluorescence confocal microscopy in HeLa cancer cells (**Figure 11.7C**). Microscopy images in which C₈ fluorescence was monitored showed that the nanoaggregates were able to penetrate the cancer cells and localized mainly the cytoplasm, particularly around the nuclei, suggesting effective drug delivery.

In order to assess whether the nanoaggregates were in fact able to target cell surface nucleolin, additional experiments were performed in tandem with a labelled nucleolin-specific antibody (Figure S6). The cells were incubated with the C₈-loaded nanoaggregates for a shorter period (15 min) so that the systems could be observed during cell membrane binding and trafficking. The yellow dots on the surface of HeLa cells indicated likely co-localization of the fluorophores (C₈ in green and nucleolin in red), and thus nucleolin-specific trafficking.

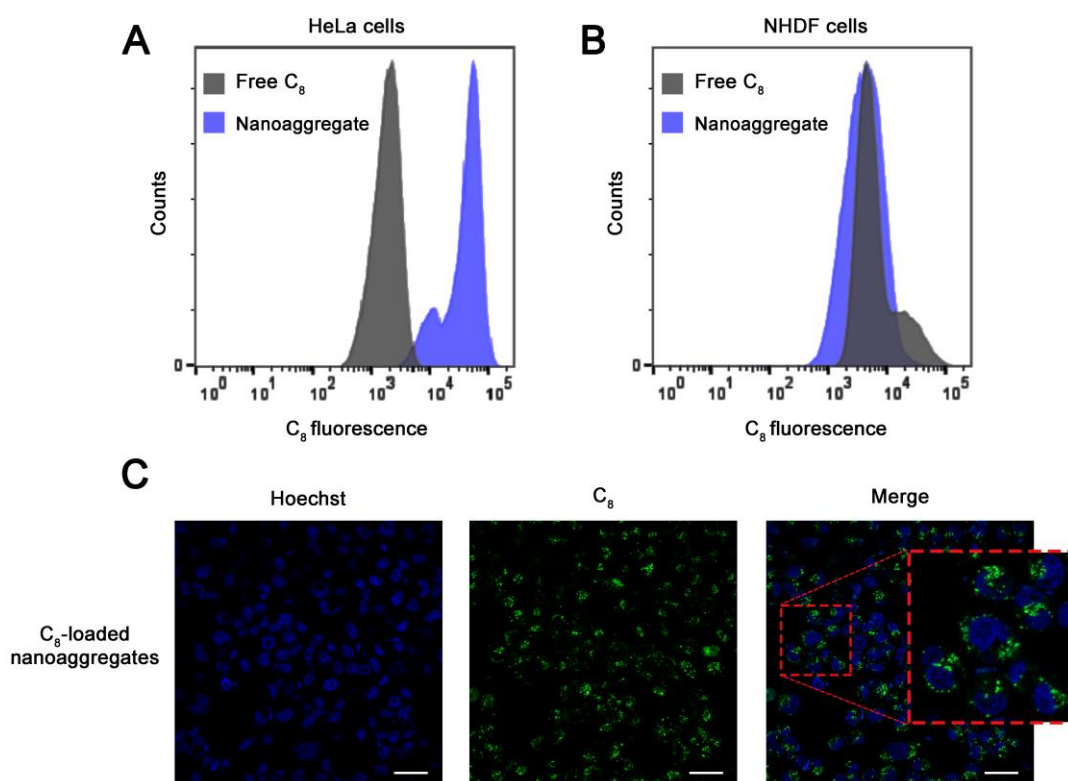


Figure 11.7 - Cellular internalization of the C₈-loaded nanoaggregates determined (A-B) by flow cytometry for HeLa and NHDF cells, respectively, and (C) fluorescence confocal microscopy. The experiments were performed by incubating the cells for 24 h with C₈-loaded nanoaggregates prior to data acquisition. Scale bar is 20 μ m.

One of the particular features of cervical cancer and HPV infection is the development of precancerous lesions commonly referred to as cervical intraepithelial neoplasia (CIN). CINs are divided into three stages: CIN1, CIN2 or CIN3 [38]. If left untreated, CIN2 and CIN3 can progress to cervical cancer [38]. In the management of CINs, the ideal is to use a ‘screen-and-treat’ approach in which the treatment decision is based on a screening test and treatment is provided immediately after a positive is identified. To determine whether early stage CIN lesions could be identified, tissue biopsies of patients with CIN1 lesions were obtained from Ipatimup and used in fluorescence confocal microscopy experiments. Staining of the tissues with a labeled nucleolin antibody indicated that nucleolin is already

overexpressed on the membrane of cells in precancerous state (**Figure 11.8**, third panel). Incubation with C₈-loaded nanoaggregates demonstrated ready internalization into the tissue (**Figure 11.8**, second and fourth panels). The engineered nanosystems could therefore potentially be used for the detection and treatment of precancerous lesions, for instance by topical application in the cervix, especially as these systems are able to penetrate the tissue and deliver drug intracellularly.

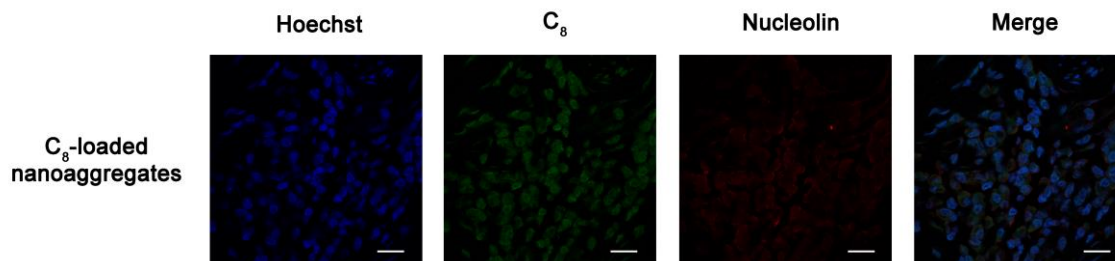


Figure 11.8 - Fluorescence confocal microscopy images of tissue biopsies of CIN1 patients incubated with C₈-loaded nanoaggregates for 3 h. Cell nuclei are stained with Hoechst 33342 in the blue channel, C₈ emits green fluorescence and nucleolin was stained with Alexa Fluor™ 647 antibody.

In order to explore options for the transdermal delivery of C₈-loaded nanosystems to an appropriate target tissue (cervix), the nanoaggregates were incorporated into a semisolid gel formulation suitable for vaginal self-administration. The base formulation was based on a Universal Placebo containing hydroxyethylcellulose, which has previously been used in clinical trials for HIV treatment with an adequate safety profile [28]. To achieve desirable viscosity and mucoadhesion profiles for prolonged tissue retention and penetration, two different formulations at 8 and 10% propyleneglycol (PG) were prepared. Tissue penetration was assessed *ex vivo* in Franz cells using porcine vaginal tissue, which is considered the gold standard animal model for vaginal permeation studies as both porcine and human vagina share similarities such as pH, vaginal fluid, permeability barrier and histology [39]. The amount of C₈-loaded nanoaggregate that penetrated/permeated the vaginal tissue was determined by the HPLC analysis of C₈ in the receptor chamber at different time points (0, 6, 12 and 24h) and in the tissue at the end of the experiment.

A simple and effective method for the analysis of C₈ using HPLC was first developed. The investigated chromatographic conditions (methanol:0.1% TFA (9:1, v/v), column temperature at 25 °C, sample injection volume of 50 µL, and isocratic flow rate of 0.5 mL/min) provided a selective and sensitive method to determine C₈. In order to quantify the samples, daily calibration curves were constructed by plotting the peak area of the target compound against concentration. As acceptance criteria, a determination coefficient (R²) of at least 0.99 and the calibrators' accuracy within a ± 15% interval from the nominal

value ($\pm 20\%$, for the limit of quantitation-LOQ), were adopted. Method linearity was obtained in the range of 0.98 to 500 ng/mL (ten calibrators evenly distributed). However, weighted least squares regression ($1/\times$) was used to compensate for heteroscedasticity. The LOQ was considered the lowest concentration that could be measured with a relative error within $\pm 20\%$ of the nominal concentration and a relative standard deviation equal or lower than 20%. This value was 0.98 ng/mL. Concerning the remaining parameters, the method proved to be precise and accurate, which is suitable for routine analyses. Figure S7 presents the representative chromatograms of a sample.

The results of the HPLC analysis are presented in **Table 11.2**. Both formulations were able to promote the retention of the C₈-loaded nanoaggregates in the vaginal tissue as demonstrated by the presence of C₈ in the tissue after 24h, but not on the receptor chamber throughout the assay, suggesting that the drug-loaded nanoaggregates did not readily permeate the tissue. This could be important for the controlled and prolonged retention of the drug loaded nanoaggregates on the vaginal tract to maximize drug accumulation in the target tissue before absorption into the bloodstream. The formulation prepared with 8% PG was slightly more effective at retaining the drug-loaded nanoaggregates in the tissue (2.58 ng/mL) when compared to the 10% PG formulation (1.58 ng/mL). These results are corroborated by the quantitation of C₈ in the donor chambers at the end of the experiment, being that C₈ was more concentrated in the 10% PG cell (2.76 ng/mL) than the 8% PG cell (1.90 ng/mL). These results are in line with what has been previously observed in other studies which demonstrate that under infinite dose conditions, a thicker, more viscous formulation offers less retention of the drug due to the diffusional resistance in the formulation [36].

Table 11.2 - HPLC quantitation of C₈ in the samples recovered from Franz cells and porcine vaginal tissue.

Sample	Donor Chamber	C ₈ concentration (ng/mL) ^a				
		Receptor Chamber				Tissue
		0h	6h	12h	24h	
8% PG	1.90 $\pm$ 0.44	-	-	-	-	2.58 $\pm$ 0 ^b
10% PG	2.76 $\pm$ 0.88	-	-	-	-	1.59 $\pm$ 0.99

^a The results are presented as mean $\pm$ SEM of three independent experiments. Symbol “-” means C₈ was not detected by HPLC in the samples.

^b C₈ was only detected above the LOQ in one experiment

Conclusions

In the current work, C₈-loaded nanoaggregates were successfully produced with AS1411-lipid conjugates that could fold into an active G4 structure. The C₈-loaded nanoaggregates were evaluated for particle size, zeta potential, in vitro drug release and antiproliferative effect against HeLa cervical cancer cells. The developed nanoaggregates presented small sizes (about 100 nm), negative surface charge, suitable PDI for pharmaceutical applications, high stability, and an adequate in vitro drug release profile. The C₈-loaded nanoaggregates showed cancer-specific accumulation and antiproliferative effects, probably due to the functionalization of the lipids with the cancer-selective guiding agent AS1411, which was especially effective given the large amounts of nucleolin presented in CIN1 lesions. Since transdermal drug delivery is a promising route for the delivery of drugs (especially as it offers several advantages, such as avoidance of the first-pass effect, easy access, potential for mucoadhesion-prolonged residence time, and ease of self-administration [40] into vaginal tissue [41]), we further studied C₈-loaded nanoaggregates formulated as a semisolid gel for local application. *Ex vivo* permeation studies using porcine vaginal tissue demonstrated the efficacy of the gel formulation in retaining and accumulating the nanosystems in the vaginal epithelia. In conclusion, these studies provided evidence that the C₈-loaded nanoaggregates are very promising therapeutic agents against HPV-induced disease and are suitable to be used in transdermal drug delivery via semisolid formulations. In the near future, these formulations will be characterized in terms of safety and physico-chemical properties to confirm its potential for *in vivo* applications.

Acknowledgments

J. Carvalho acknowledges a doctoral fellowship grant from the FCT – Foundation for Science and Technology ref. SFRH/BD/122953/2016. We thank B. Vialet and P. Barthélémy for synthesizing and providing the lipid-oligos, C. L. Ferreira for the assistance in immunohistochemistry experiments, A. Paiva and Centro Hospitalar Universitário de Coimbra for the access to the flow cytometry facilities. This work was supported by UTAustin Portugal Program Exploratory project DREAM ref. UTAP-EXPL/NTec/0015/2017, project “Ações Integradas Luso-Francesas” ref. TC- 15/17, FCT project “Projeto de Investigação Exploratória” ref. IF/00959/2015 entitled “NCL targeting by G-quadruplex aptamers for cervical cancer therapy” financed by Fundo Social Europeu

e Programa Operacional Potencial Humano, and FEDER funds through the POCI - COMPETE 2020 - Operational Programme Competitiveness and Internationalisation in Axis I - Strengthening research, and to CICS-UBI (UIDB/00709/2020). The NMR spectrometers are part of Rede Nacional de RMN (PTNMR), supported by FCT-MCTES (ROTEIRO/0031/2013 - PINFRA/22161/2016) co-funded by FEDER through COMPETE 2020, POCI, and PORL and FCT through PIDDAC.

References

- [1] Bruni, L., Albero, G., Serrano, B., Mena, M., Gómez, D., Muñoz, J., Bosch, F. X., and de Sanjosé, S. (2019) Human Papillomavirus and Related Diseases in the World. *ICO/IARC Information Centre on HPV and Cancer (HPV Information Centre)*.
- [2] Vora, C., and Gupta, S. (2019) Targeted therapy in cervical cancer. *ESMO Open* 3, e000462. <https://doi.org/10.1136/esmoopen-2018-000462>.
- [3] Bates, P. J., Reyes-Reyes, E. M., Malik, M. T., Murphy, E. M., O'Toole, M. G., and Trent, J. O. (2017) G-quadruplex oligonucleotide AS1411 as a cancer-targeting agent: Uses and mechanisms. *Biochimica et Biophysica Acta - General Subjects* 1861, 1414–1428. <https://doi.org/10.1016/j.bbagen.2016.12.015>.
- [4] Sharma, V. R., Thomas, S. D., Miller, D. M., Rezzoug, F., Sharma, V. R., Thomas, S. D., Miller, D. M., and Rezzoug, F. (2018) Nucleolin Overexpression Confers Increased Sensitivity to the Anti-Nucleolin Aptamer, AS1411. *Cancer Investigation* 36, 475–491. <https://doi.org/10.1080/07357907.2018.1527930>.
- [5] Reyes-Reyes, E. M., and Bates, P. J. (2019) Characterizing Oligonucleotide Uptake in Cultured Cells: A Case Study Using AS1411 Aptamer, in *Oligonucleotide-Based Therapies: Methods and Protocols*, pp 173–186. https://doi.org/10.1007/978-1-4939-9670-4_10.
- [6] Rosenberg, J., Figlin, R. A., Smith, G. W., Choueiri, T., and Erlandsson, F. (2015) A phase II trial of the nucleolin-targeted DNA aptamer AS1411 in metastatic refractory renal cell carcinoma. *Invest New Drugs* 32, 178–187. <https://doi.org/10.1007/s10637-013-0045-6.A>.
- [7] Rosenberg, J. E., Bambury, R. M., Van Allen, E. M., Drabkin, H. a., Lara, P. N., Harzstark, A. L., Wagle, N., Figlin, R. a., Smith, G. W., Garraway, L. a., Choueiri, T., Erlandsson, F., and Laber, D. a. (2014) A phase II trial of AS1411 (a novel nucleolin-targeted DNA aptamer) in metastatic renal cell carcinoma. *Investigational New Drugs* 32, 178–187. <https://doi.org/10.1007/s10637-013-0045-6>.

- [8] Stuart, R. K., Stockerl-Goldstein, K., Cooper, M., Devetten, M., Herzig, R., Medeiros, B., Schiller, G., Wei, A., Acton, G., and Rizzieri, D. (2009) Randomized phase II trial of the nucleolin targeting aptamer AS1411 combined with high-dose cytarabine in relapsed/refractory acute myeloid leukemia (AML). *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 27, 7019.
- [9] Giudice, V., Mensitieri, F., Izzo, V., Filippelli, A., and Selleri, C. (2020) Aptamers and antisense oligonucleotides for diagnosis and treatment of hematological diseases. *International Journal of Molecular Sciences* 21, 3252. <https://doi.org/10.3390/ijms21093252>.
- [10] Malik, M. T., O 'toole, M. G., Casson, L. K., Thomas, S. D., Bardi, G. T., Reyes-Reyes, E. M., Ng, C. K., Kang, K. A., and Bates, P. J. (2015) AS1411-conjugated gold nanospheres and their potential for breast cancer therapy. *Oncotarget* 6, 22270–22281. <https://doi.org/10.18632/oncotarget.4207>.
- [11] Li, L., Hou, J., Liu, X., Guo, Y., Wu, Y., Zhang, L., and Yang, Z. (2014) Nucleolin-targeting liposomes guided by aptamer AS1411 for the delivery of siRNA for the treatment of malignant melanomas. *Biomaterials* 35, 3840–3850. <https://doi.org/10.1016/j.biomaterials.2014.01.019>.
- [12] Yang, S., Ren, Z., Chen, M., Wang, Y., You, B., Chen, W., Qu, C., Liu, Y., and Zhang, X. (2018) Nucleolin-Targeting AS1411-Aptamer-Modified Graft Polymeric Micelle with Dual pH/Redox Sensitivity Designed to Enhance Tumor Therapy through the Codelivery of Doxorubicin/TLR4 siRNA and Suppression of Invasion. *Molecular Pharmaceutics* 15, 314–325. <https://doi.org/10.1021/acs.molpharmaceut.7b01093>.
- [13] Alibolandi, M., Ramezani, M., Abnous, K., and Hadizadeh, F. (2016) AS1411 Aptamer-Decorated Biodegradable Polyethylene Glycol-Poly(lactic-co-glycolic acid) Nanopolymersomes for the Targeted Delivery of Gemcitabine to Non-Small Cell Lung Cancer in Vitro. *Journal of Pharmaceutical Sciences* 105, 1741–1750. <https://doi.org/10.1016/j.xphs.2016.02.021>.
- [14] Vandghanooni, S., Eskandani, M., Barar, J., and Omid, Y. (2018) AS1411 aptamer-decorated cisplatin-loaded poly(lactic- co -glycolic acid) nanoparticles for targeted therapy of miR-21-inhibited ovarian cancer cells. *Nanomedicine* 13, 2729–2758. <https://doi.org/10.2217/nnm-2018-0205>.

- [15] Shieh, Y. A., Yang, S. J., Wei, M. F., and Shieh, M. J. (2010) Aptamer-based tumor-targeted drug delivery for photodynamic therapy. *ACS Nano* 4, 1433–1442. <https://doi.org/10.1021/nn901374b>.
- [16] Carvalho, J., Paiva, A., Cabral Campello, M. P., Paulo, A., Mergny, J.-L., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2019) Aptamer-based Targeted Delivery of a G-quadruplex Ligand in Cervical Cancer Cells. *Scientific Reports* 9, 7945. <https://doi.org/10.1038/s41598-019-44388-9>.
- [17] Dailey, M. M., Clarke Miller, M., Bates, P. J., Lane, A. N., and Trent, J. O. (2010) Resolution and characterization of the structural polymorphism of a single quadruplex-forming sequence. *Nucleic Acids Research* 38, 4877–4888. <https://doi.org/10.1093/nar/gkq166>.
- [18] Do, N. Q., Chung, W. J., Truong, T. H. A., Heddi, B., and Phan, A. T. (2017) G-quadruplex structure of an anti-proliferative DNA sequence. *Nucleic Acids Research* 45, 7487–7493. <https://doi.org/10.1093/nar/gkx274>.
- [19] Fan, X., Sun, L., Li, K., Yang, X., Cai, B., Zhang, Y., Zhu, Y., Ma, Y., Guan, Z., Wu, Y., Zhang, L., and Yang, Z. (2017) The Bioactivity of D-/L-Isonucleoside- and 2'-Deoxyinosine-Incorporated Aptamer AS1411s Including DNA Replication/MicroRNA Expression. *Molecular Therapy - Nucleic Acids* 9, 218–229. <https://doi.org/10.1016/j.omtn.2017.09.010>.
- [20] Violet, B., Gissot, A., Delzor, R., and Barthélémy, P. (2017) Controlling G-quadruplex formation: Via lipid modification of oligonucleotide sequences. *Chemical Communications* 53, 11560–11563. <https://doi.org/10.1039/c7cc05693a>.
- [21] Riccardi, C., Musumeci, D., Russo Krauss, I., Piccolo, M., Irace, C., Paduano, L., and Montesarchio, D. (2018) Exploring the conformational behaviour and aggregation properties of lipid-conjugated AS1411 aptamers. *International Journal of Biological Macromolecules* 118, 1384–1399. <https://doi.org/10.1016/j.ijbiomac.2018.06.137>.
- [22] Sharma, V. K., Jain, A., and Soni, V. (2013) Nano-aggregates: Emerging delivery tools for tumor therapy. *Critical Reviews in Therapeutic Drug Carrier Systems* 30, 535–563. <https://doi.org/10.1615/CritRevTherDrugCarrierSyst.2013007706>.
- [23] Carvalho, J., Pereira, E., Marquevielle, J., Campello, M. P. C., Mergny, J.-L., Paulo, A., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2018) Fluorescent light-up acridine orange

derivatives bind and stabilize KRAS-22RT G-quadruplex. *Biochimie* 144, 144–152. <https://doi.org/10.1016/j.biochi.2017.11.004>.

[24] Carvalho, J., Lopes-Nunes, J., Paula Cabral Campello, M., Paulo, A., Milici, J., Meyers, C., Mergny, J.-L., Salgado, G. F., Queiroz, J. A., and Cruz, C. (2020) Human Papillomavirus G-Rich Regions as Potential Antiviral Drug Targets. *Nucleic Acid Therapeutics*. <https://doi.org/10.1089/nat.2020.0869>.

[25] Carvalho, J., Queiroz, J. A., and Cruz, C. (2017) Circular dichroism of G-Quadruplex: A laboratory experiment for the study of topology and ligand binding. *Journal of Chemical Education* 94, 1547–1551. <https://doi.org/10.1021/acs.jchemed.7b00160>.

[26] Dominguez, Ana; Fernandez, Aurora; Gonzalez, Noemi; Iglesias, Emilia; Montenegro, L. (1997) Determination of Critical Micelle Concentration of Some Surfactants by Three Techniques. *Journal of Chemical Education* 74, 1227–1231.

[27] Moreira, A. F., Rodrigues, C. F., Reis, C. A., Costa, E. C., Ferreira, P., and Correia, I. J. (2018) Development of poly-2-ethyl-2-oxazoline coated gold-core silica shell nanorods for cancer chemo-photothermal therapy. *Nanomedicine* 13, 2611–2627. <https://doi.org/10.2217/nnm-2018-0179>.

[28] Tien, D., Schnaare, R. L., Kang, F., Cohl, G., McCormick, T. J., Moench, T. R., Doncel, G., Watson, K., Buckheit, R. W., Lewis, M. G., Schwartz, J., Douville, K., and Romano, J. W. (2005) In vitro and in vivo characterization of a potential universal placebo designed for use in vaginal microbicide clinical trials. *AIDS Research and Human Retroviruses* 21, 845–853. <https://doi.org/10.1089/aid.2005.21.845>.

[29] U.S. Food & Drug Administration. (2015) Analytical Procedures and Methods Validation for Drugs and Biologics.

[30] Zhang, J., Chen, R., Fang, X., Chen, F., Wang, Y., and Chen, M. (2015) Nucleolin targeting AS1411 aptamer modified pH-sensitive micelles for enhanced delivery and antitumor efficacy of paclitaxel. *Nano Research* 8, 201–218. <https://doi.org/10.1007/s12274-014-0619-4>.

[31] Li, X., Yu, Y., Ji, Q., and Qiu, L. (2015) Targeted delivery of anticancer drugs by aptamer AS1411 mediated Pluronic F127 / cyclodextrin-linked polymer composite micelles. *Nanomedicine: Nanotechnology, Biology, and Medicine* 11, 175–184. <https://doi.org/10.1016/j.nano.2014.08.013>.

- [32] Webba da Silva, M. (2007) NMR methods for studying quadruplex nucleic acids. *Methods* 43, 264–277. <https://doi.org/10.1016/j.ymeth.2007.05.007>.
- [33] Huc, I., and Oda, R. (1999) Gemini surfactants: studying micellisation by ^1H and ^{19}F NMR spectroscopy. *Chemical Communications* 2025–2026. <https://doi.org/10.1039/a906141j>.
- [34] Kay, L. E., Torchia, D. A., and Bax, A. (1989) Backbone dynamics of proteins as studied by nitrogen-15 inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease. *Biochemistry* 28, 8972–8979. <https://doi.org/10.1021/bio0449a003>.
- [35] Blanco, E., Shen, H., and Ferrari, M. (2015) Principles of nanoparticle design for overcoming biological barriers to drug delivery. *Nature Biotechnology* 33, 941–951. <https://doi.org/10.1038/nbt.3330>.
- [36] Reyes-Reyes, E. M., Teng, Y., and Bates, P. J. (2010) A new paradigm for aptamer therapeutic AS1411 action: Uptake by macropinocytosis and its stimulation by a nucleolin-dependent mechanism. *Cancer Research* 70, 8617–8629. <https://doi.org/10.1158/0008-5472.CAN-10-0920>.
- [37] Mo, R., and Gu, Z. (2016) Tumor microenvironment and intracellular signal-activated nanomaterials for anticancer drug delivery. *Materials Today*. <https://doi.org/10.1016/j.mattod.2015.11.025>.
- [38] Schiffman, M., Doorbar, J., Wentzensen, N., De Sanjosé, S., Fakhry, C., Monk, B. J., Stanley, M. A., and Franceschi, S. (2016) Carcinogenic human papillomavirus infection. *Nature Reviews Disease Primers* 2, 16086. <https://doi.org/10.1038/nrdp.2016.86>.
- [39] Machado, R. M., Palmeira-de-Oliveira, A., Gaspar, C., Martinez-de-Oliveira, J., and Palmeira-de-Oliveira, R. (2015) Studies and methodologies on vaginal drug permeation. *Advanced Drug Delivery Reviews* 92, 14–26. <https://doi.org/10.1016/j.addr.2015.02.003>.
- [40] Machado, R. M., Palmeira-de-Oliveira, A., Breitenfeld, L., Martinez-de-Oliveira, J., and Palmeira-de-Oliveira, R. (2019) Optimization and Application of In Vitro and Ex Vivo Models for Vaginal Semisolids Safety Evaluation. *Journal of Pharmaceutical Sciences* 108, 3289–3301. <https://doi.org/10.1016/j.xphs.2019.05.026>.

[41] Caramella, C. M., Rossi, S., Ferrari, F., Bonferoni, M. C., and Sandri, G. (2015) Mucoadhesive and thermogelling systems for vaginal drug delivery. *Advanced Drug Delivery Reviews* 92, 39–52. <https://doi.org/10.1016/j.addr.2015.02.001>.

Chapter 12

General Conclusions and Future Perspectives

General Conclusions and Future Perspectives

General Conclusions

The World Health Organization (WHO) has set the ambitious goal of eliminating cervical cancer as a public health problem by 2030. Cervical cancer is still one of the most common cancers in women, particularly in lower income developing countries, due to inefficiency in establishing HPV vaccination programmes and widespread screening routines. Some high-income countries have also struggled to attain HPV vaccination and cervical cancer screening rates that are necessary to eliminate the disease due to fragmented health-care systems that limit the access to these programmes. Other current issue includes vaccine hesitancy due to myths and misconceptions (commonly referred to as anti-vaxxers), which is increasingly recognized as a key factor in suboptimal vaccine uptake. Taking into account the aforementioned issues, together with the fact that to this date there is still no effective cure for HPV infection and persistence, the development of novel therapies against HPV diseases are on great demand.

To this end, G4 nucleic acids provide exciting opportunities for the discovery of novel therapeutic targets in the HPV life cycle, and the development of specific antiviral strategies to control HPV virulence and cervical cancer progression. The therapeutic strategy proposed in this Thesis holds a dual antiviral and anticancer aspect. On the antiviral side, targeting the G4 regions of the HPV genome with small compounds capable of stabilizing these structures has the potential of interfering with viral gene transcription and DNA replication which can clear the infection in its latent non-productive stages, stopping it from progressing into high-grade lesions that ultimately cause cervical cancer. Furthermore, G4 aptamers such as AS1411 that target nucleolin may affect the protein's functions in the HPV life cycle, particularly in the case of HPV16 and 18, where it was found to regulate viral transcription and genome maintenance. Regarding the anticancer aspect, oncogene promoters and telomeres are known for also being enriched with G4 structures which targeting with ligands was already shown to produce antitumoral effects. nucleolin, whose oncogenic function are known, was also found to be a promising anticancer target as AS1411 reached Phase 2 clinical trials as an antitumoral drug. The main objective of this Thesis was to combine all this knowledge by developing G4-targeted ligands and an G4 aptamer-based strategy to specifically deliver these ligands into HPV-positive cells in either precancerous or cancerous stages.

In **Chapters 4-6**, a variety of compounds belonging to different aromatic scaffolds (indole, phenanthroline and acridine) were studied in terms of their ability to bind and stabilize biologically relevant G4 structures demonstrated by several biophysical and biochemical assays. The compounds were able to inhibit the activity of helicase and polymerase enzymes *in vitro*, which demonstrated their potential to inhibit key cellular processes of cancer cells such as DNA replication and gene transcription. Furthermore, some ligands were shown to inhibit the growth of cervical cancer cells with low IC₅₀ values in the micromolar range when compared to known clinically approved anticancer drugs such as 5-FU (IC₅₀>100 μ M against HeLa cells). These results emphasized the importance of the larger aromatic core, positive charge, and side-chain functionality in ligands' structure which were invariably the ones with greater affinity towards G4 sequences, particularly phenanthroline derivatives [16]phenN₂ and [32]phen₂N₄, and acridine derivative C₈. The analysis of the interaction of these ligands with the G4 structures at atomic level resolution supplied useful insights and data for the lead optimization of other G4 ligands with drug-like properties. Altogether these chapters pointed out new G4 ligands to be tested against putative G4s found in the genome of HPV.

In **Chapter 7**, the most promising G4 ligands developed in the previous chapters, together with commercially available ligands, were tested for their ability to bind and stabilize HPV G4s. All compounds were found to successfully recognize HPV G4s and these were able to induce structural transitions between different topologies in some of the G4s. The observed structural polymorphism upon ligand binding may have great impact in the regulation of the life cycle as the G4 structures were found in important regulatory regions of the HPV genome. Notably, high risk HPV16, which folds into multiple non-canonical DNA structures, was converted into a major G4 conformation upon interaction with PhenDC₃. Surprisingly, PhenDC₃ had no effect on the growth of tissue infected with HPV16, or on the viral replication, which may be due to the ligand having more preference for the G4 structures formed in the genome of the cancer cells rather than the viral G4s. In contrast, C₈ was found to inhibit the growth of HPV18 infected tissue. The effect was shown to be HPV-dependent as the ligand reduced both viral replication and encapsidation. These effects may correlate with G4-binding as the studied G4 sequence is found in the L2 capsid protein coding sequence. Luciferase reporter assays to experimentally prove this G4-dependence are currently undergoing. This work presented the first evidence of the potential use of G4 ligands as anti-HPV agents and the regulatory roles of G4 structures in the viral life cycle.

In **Chapters 8-10**, I present how we developed an aptamer-based delivery strategy to selectively convey the promising G4 ligands into cervical cancer cells. The strategy was based on the non-covalent conjugation of AS1411 and its derivatives AT11, AT11-Lo, and AT11-Bo, with acridine-based G4 ligands previously shown to be promising antitumoral and antiviral agents. This simple non-covalent strategy provided a cancer-specific accumulation of the G4 ligands in cervical cancer cell lines with concomitant antiproliferative effects. The cancer-selective uptake of the aptamer-ligand complexes was suggested to be dependent of the overexpression of cell-surface nucleolin, and that aptamer-ligand interaction did not affect the affinity of the aptamers towards nucleolin. Indeed, as AS1411 and its derivatives are also guanine-rich sequences capable of forming G4 structures, the presence of the ligands enhanced the aptamers' properties such as cellular penetration and serum nuclease resistance. This was the first indication that G4 ligands may enhance the pharmacological properties of G-rich aptamers. The efficiency of the aptamer-based delivery was demonstrated by the downregulation of *c-MYC* at transcriptional level which suggests that the acridine ligands were able to detach from the aptamer once inside the cells, reaching the nucleus of the cervical cancer cells. This is further corroborated by the fact that free C₈ was observed within the cell nucleus. These results demonstrated the potential of using G4 aptamers as drug delivery systems for cervical cancer.

Finally, in **Chapter 11**, bringing together all the knowledge and expertise acquired in the previous chapters, the development of more complex and improved drug carriers was achieved. Particularly, by using AS1411-lipid conjugates, we developed micellar carriers to selectively deliver ligand C₈ into HPV-positive cervical cancer cells. The resulting functionalized and targeted nanocarriers showed high selectivity towards cancer cells, and negligible uptake and antiproliferative effect in normal cells. The C₈-loaded nanoaggregates were also found to penetrate tissue biopsies of women infected with HPV, which demonstrated high levels of surface nucleolin expression even in CIN1 stages, pointing to the potential of using this approach to treat precancerous lesions. These promising results led us to formulate the C₈-loaded nanoaggregates as a gel for topical application in the cervix, in collaboration with LABFIT-HPRD, Lda, a company dedicated to the development of pharmaceutical products for vaginal application. The formulation proved to be a suitable platform for the local administration of the micellar systems in the female reproductive tract. These preliminary results demonstrate the potential of this type of formulations for the safe administration of nanoparticles with therapeutic potential against HPV-diseases.

The proposed objectives of this Thesis were successfully attained. Overall, by using a hypothesis driven approach, my doctoral work led to state-of-the-art advances in the various stages of the development of targeted therapeutics for HPV persistence and cervical cancer treatment. In addition, my work impacted significantly in the development of novel drug administration vectors which have a potential to positively impact HPV diseases in the near future.

Future Perspectives

G4s are remarkable tools for the development of therapeutics against HPV persistence and cervical cancer. However, there are still questions to be answered regarding the different development stages of the aptamer-based therapeutic strategies proposed herein:

- i. Regarding the ligand's biological effects, these should be further studied, particularly the G4-dependence of the antiviral effects demonstrated in organotypic raft cultures. There are probes capable of visualizing G4s in living cells, which may aid in the identification of the ligand's targets. Furthermore, the use of reporter assays and the analysis of gene expression, both at transcriptional and translational levels, are highly suggested to help elucidate the antiviral mechanisms of C₈ and other G4 ligands that may follow. These effects should also be studied *in vivo* using HPV infection models such as the K14-HPV16 transgenic mice model.
- ii. The role of the G4 structures in the HPV life cycle is still unknown. Several guanine-rich sequences were found and proved to fold into stable G4s, however their impact on the regulation of viral processes is yet to be elucidated. The C₈ results in organotypic cultures with replicating HPV18 prompt us to speculate that the G4 structures found in the L2 gene may regulate the assembly of the virus capsid. However, further studies should be pursued to confirm this hypothesis, for instance using ChIP-seq methods may help unravel the mechanistic and molecular details of how G4s influence viral replication and transcription. Also, knowing that nucleolin is also a remarkable G4 binder, its importance for the HPV life cycle, should be further studied in the future.
- iii. Regarding the aptamer-based nanocarriers, as the follow-up of this research it would be relevant to study the pharmacokinetic/pharmacodynamic profile of the C₈-loaded nanoaggregates (or even the aptamer-ligand complexes)

following intravenous injection in cervical cancer animal models such as the K14-HPV16 transgenic mice model. Additionally, the vaginal formulation of the nanoparticles should be further developed and optimized by adding for instance bioadhesive polymers such as Carbopol to increase drug retention, in compliance with the regulatory requirements and the clinical needs. In the future, the formulations should be tested for their safety and physico-chemical properties (pH, viscosity, bioadhesion, osmolality and cytotoxicity), and ideally tested *in vivo*.

In a broader perspective for the field of G4 nucleic acids, the following outstanding questions should be addressed in the future:

- Are G4s a simple nuisance or indispensable regulatory elements? Under what conditions do they form? Are they regulated by neighbouring G4 structures?
- What is the interplay between G4 and other unusual structures such as R-loops or i-motifs? Do they form differently at distinct cellular developmental stages? How does the supramolecular assembly of these non-canonical structures affect protein binding and hold back G4 formation?
- What level of selectivity (vs duplexes, or vs other G4) is required? Can genomic G4 structures be studied in individual terms or should they be perceived in context of many G4-triggered mechanisms? Is it possible/necessary to achieve perfect selectivity for a ligand?
- How G4 aptamers possess such high target specificity and binding affinity? Is sequence rather than the folding (or the opposite) important for target recognition? Can we use that information to design novel and potent ligands/aptamers?
- Which cancer cells would be more vulnerable to G4-based treatments? What is the phenotype that makes certain patients more prone for G4-targeted anticancer ligands and G4 aptamers?

Appendices

Appendix A

Supporting Information for Chapter 4

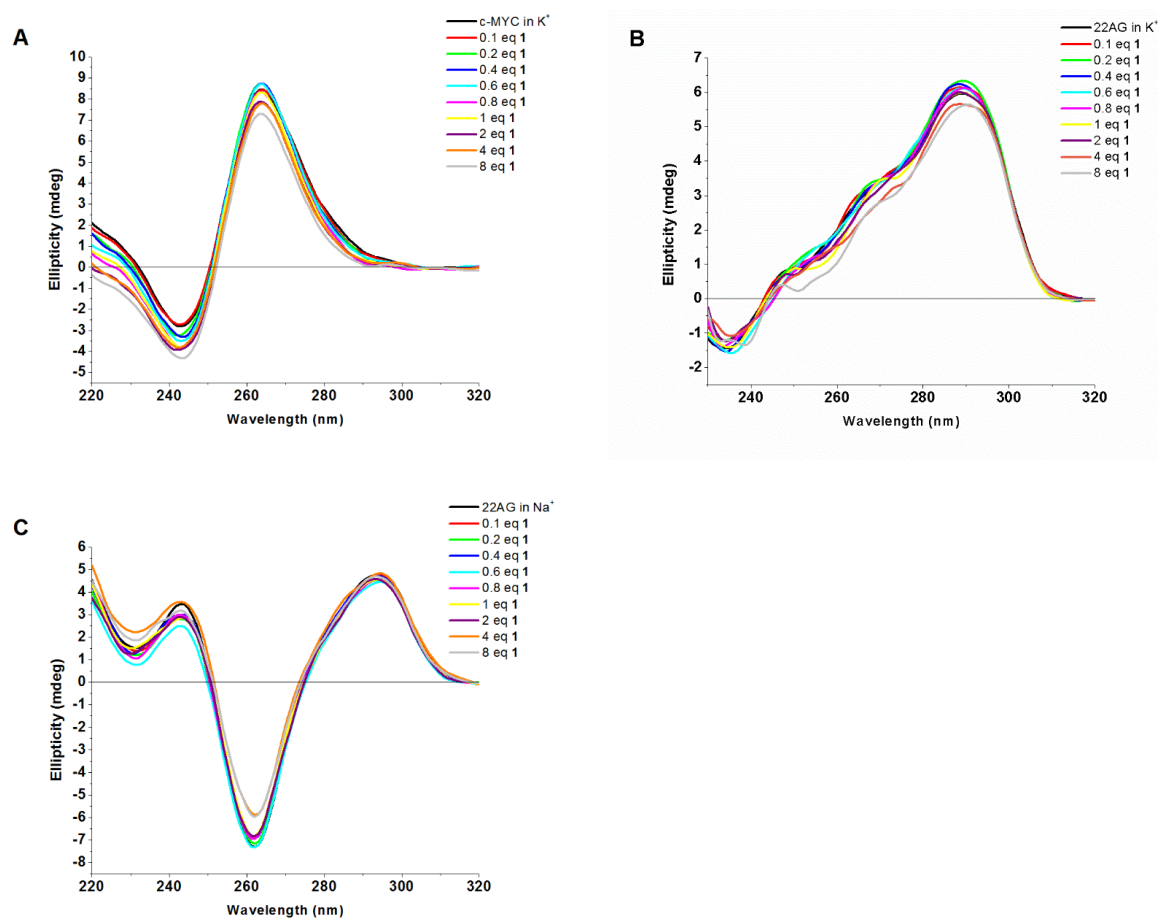


Figure S1 - CD titration spectra of (A) c-MYC, (B) 22AG in K⁺ and (C) 22AG in Na⁺ with compound **1**.

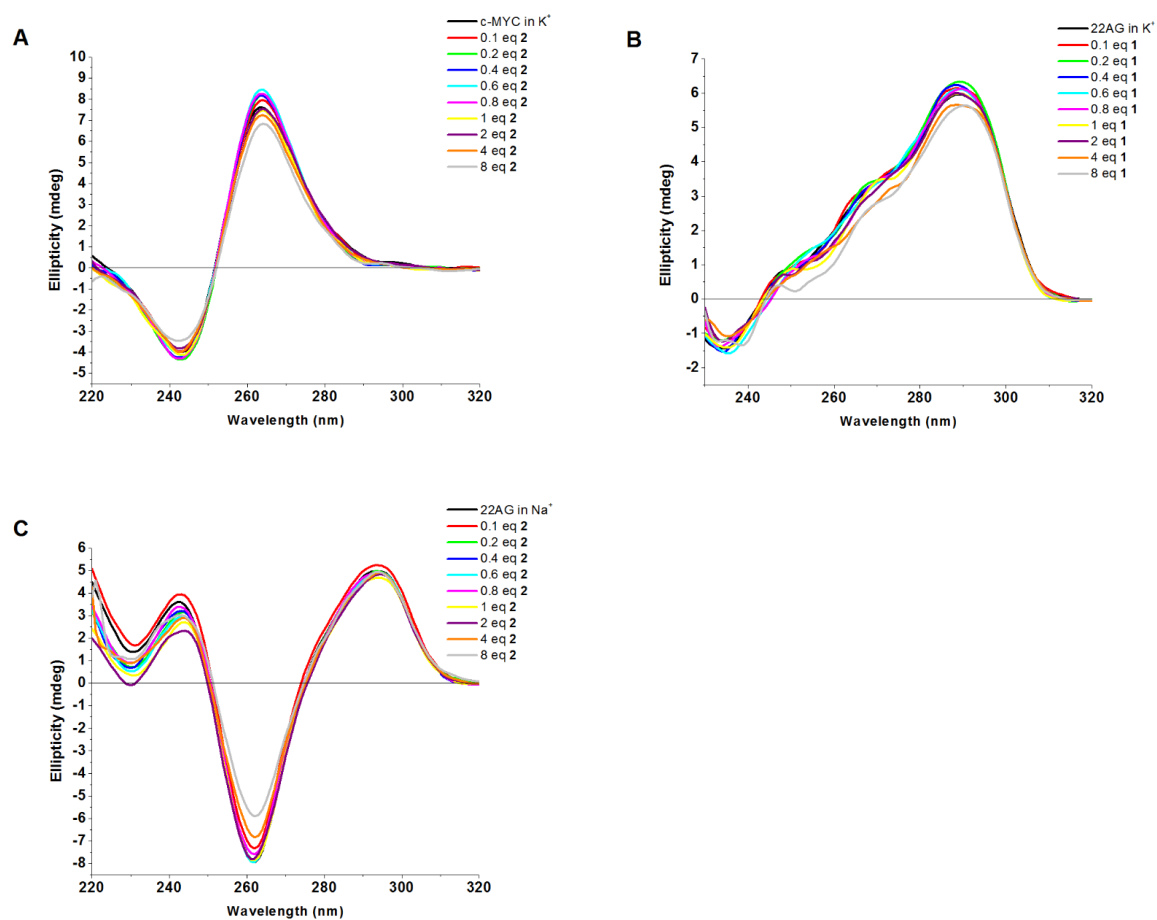


Figure S2 - CD titration spectra of (A) c-MYC, (B) 22AG in K^+ and (C) 22AG in Na^+ with compound **2**.

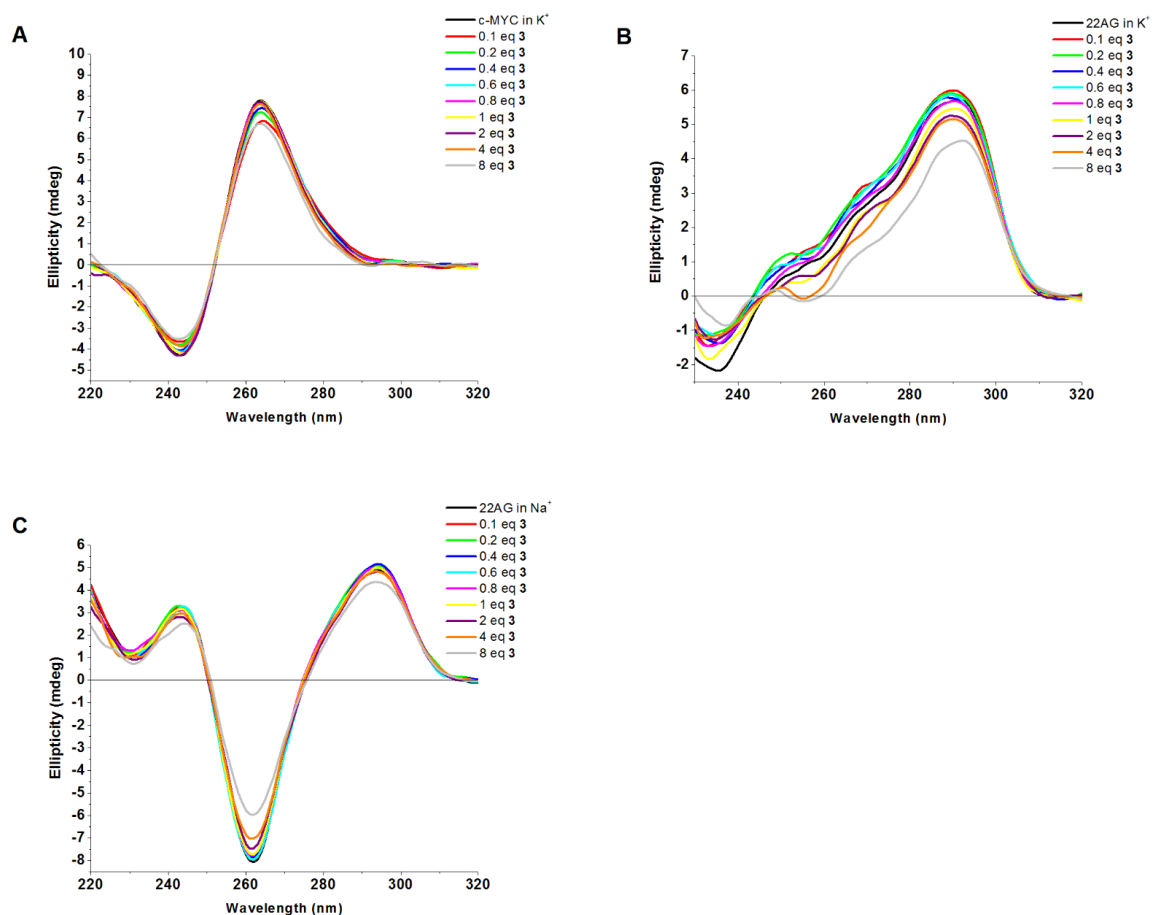


Figure S3 - CD titration spectra of (A) c-MYC, (B) 22AG in K^+ and (C) 22AG in Na^+ with compound **3**.

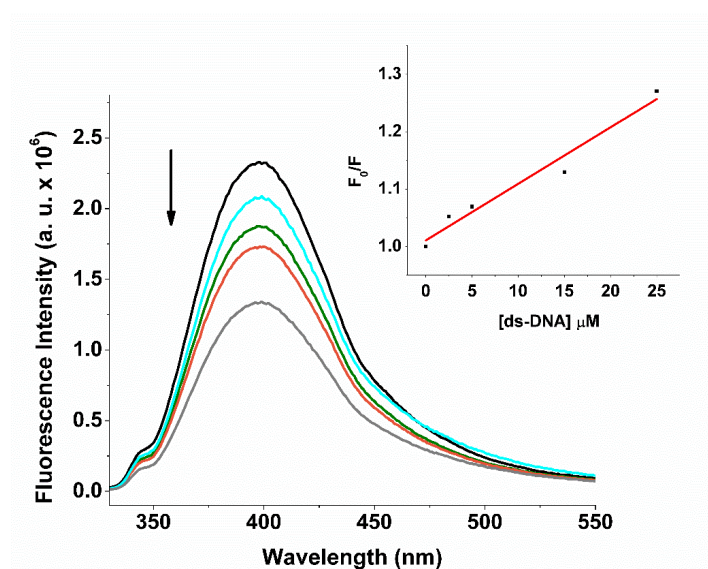


Figure S4. Fluorescence emission spectra of compound **3** with successive additions of ds-DNA. Inset: Stern-Volmer plots obtained by plotting F_0/F as a function of DNA concentration and fitting to the adequate equation.

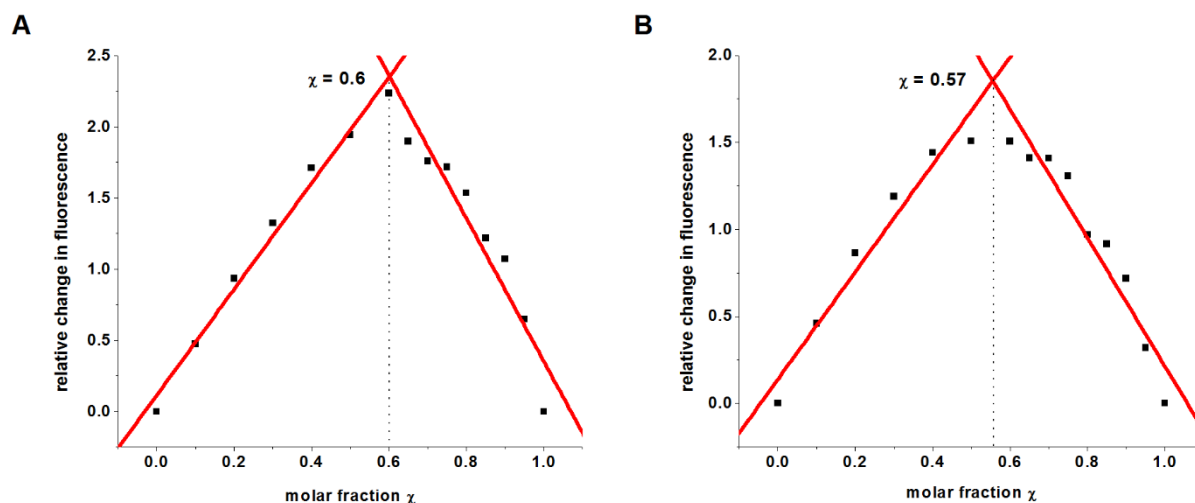


Figure S5 - Job plots resulting from the method of continuous variation analysis for compound **3** with (A) c-MYC and (B) hybrid-1 22AG.

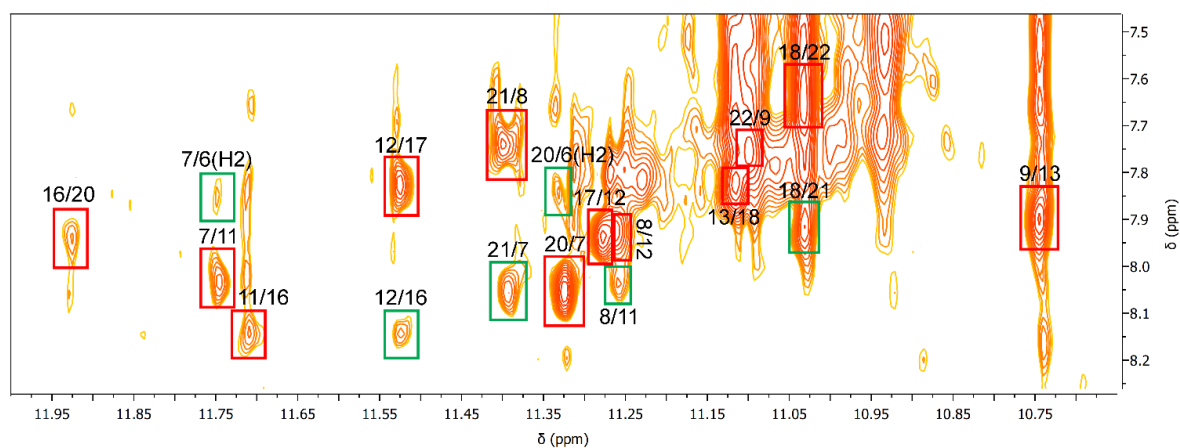


Figure S6 - The expanded H1-H8/H2 region of the 2D-NOESY spectrum of the ligand:DNA 2:1 complex at 310 K. Cross-peaks are labeled using the pairing system [guanine residue number(H1)/residue number (H8/H2)]. The interresidue NOEs within a tetrad plane are highlighted with red boxes while green boxes highlight additional NOEs.

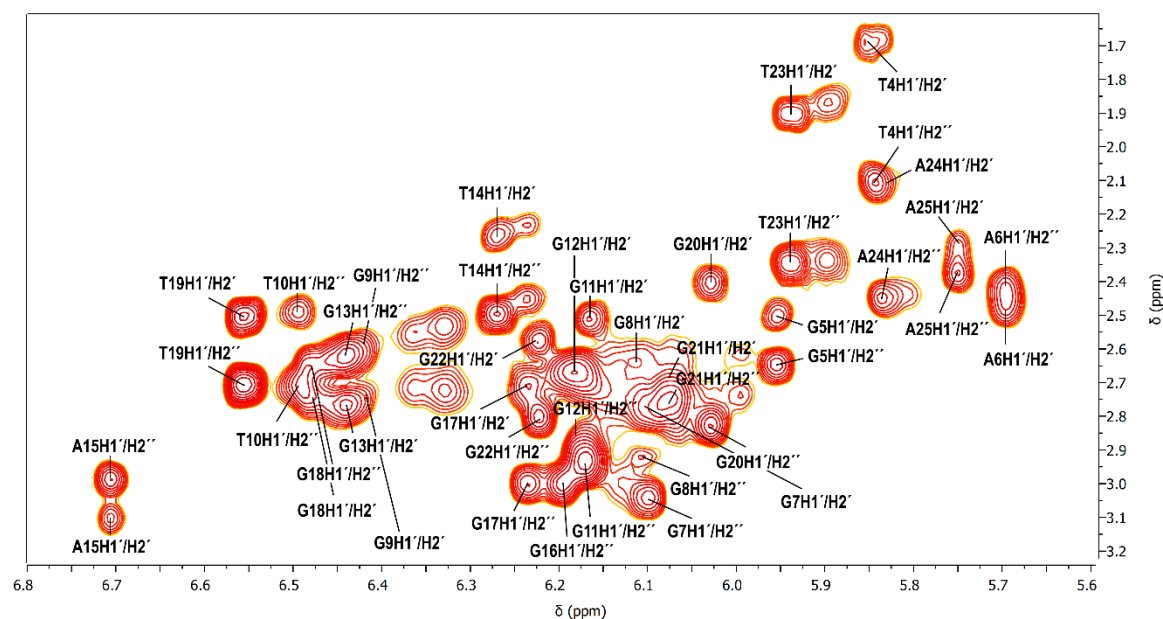


Figure S7 - The expanded H1'-H2'/H2'' region of the 2D-NOESY spectrum of the ligand:DNA 2:1 complex at 310 K.

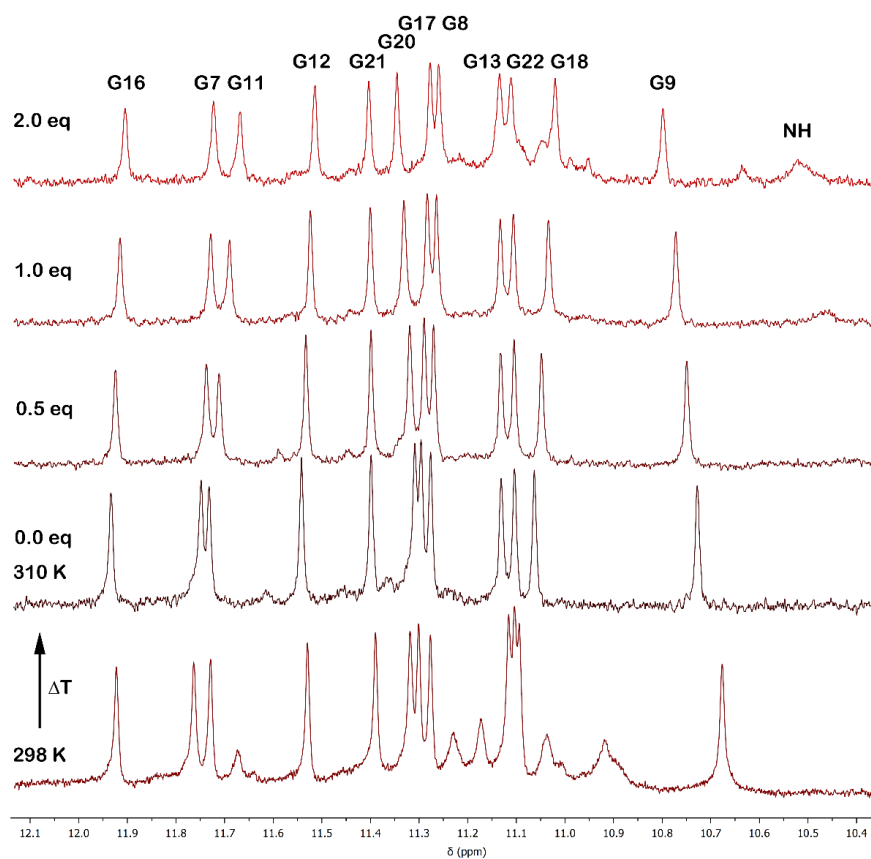


Figure S8 - Imino region of c-MYC ^1H spectra at 298 K and 310 K showing titration of compound 3 at 310 K. The increase of temperature favors the formation of a single G4 conformation as the set of peaks correspondent to minor conformations are almost inexistent. The proton signal corresponding to the NH group of the indole moiety of compound 3 appears as a broad peak during titration.

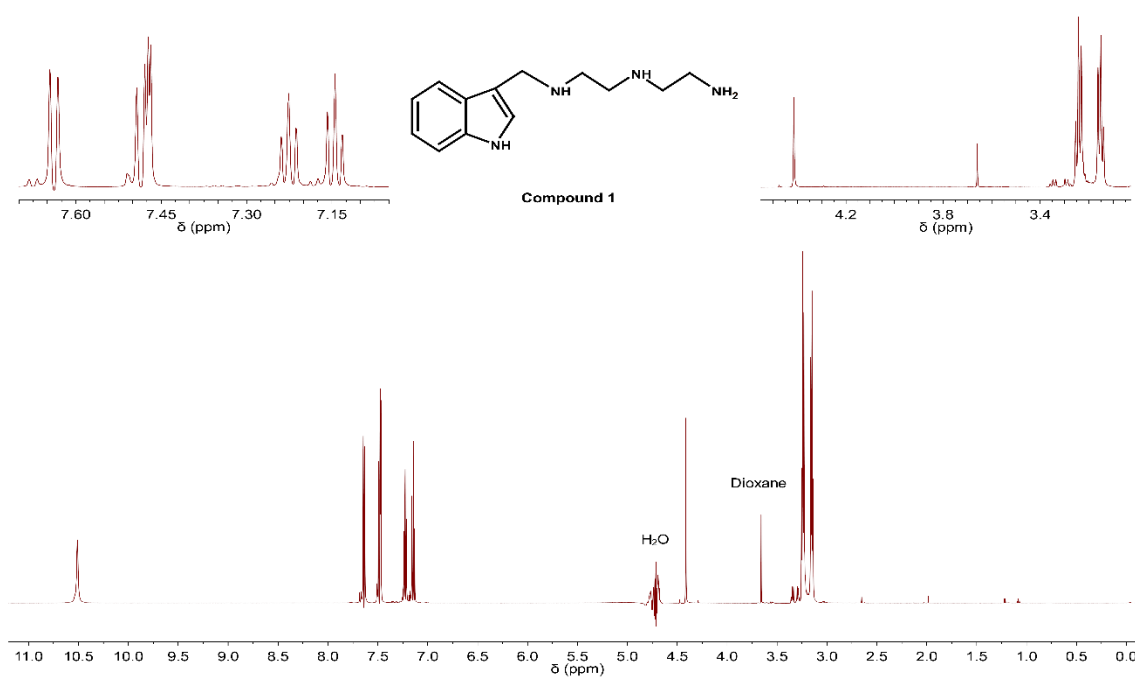


Figure S9 - ¹H NMR spectra of compound **1**.

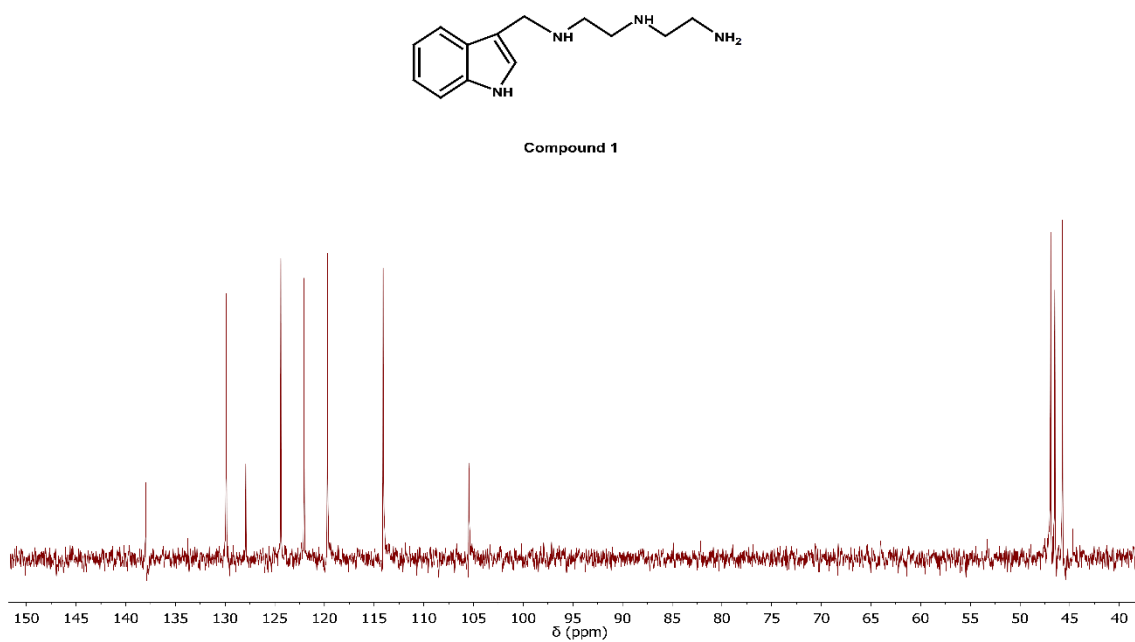


Figure S10 - ¹³C NMR spectra of compound **1**. Signals were confirmed by dept-135 spectra.

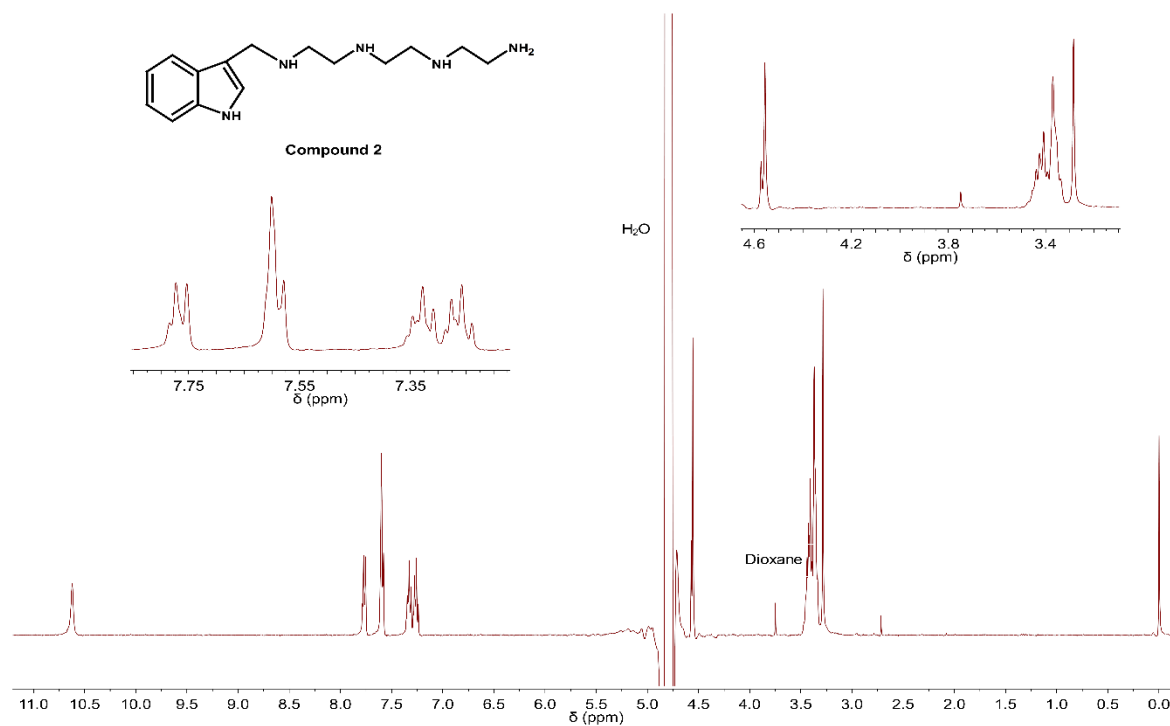


Figure S11 - ¹H NMR spectra of compound **2**.

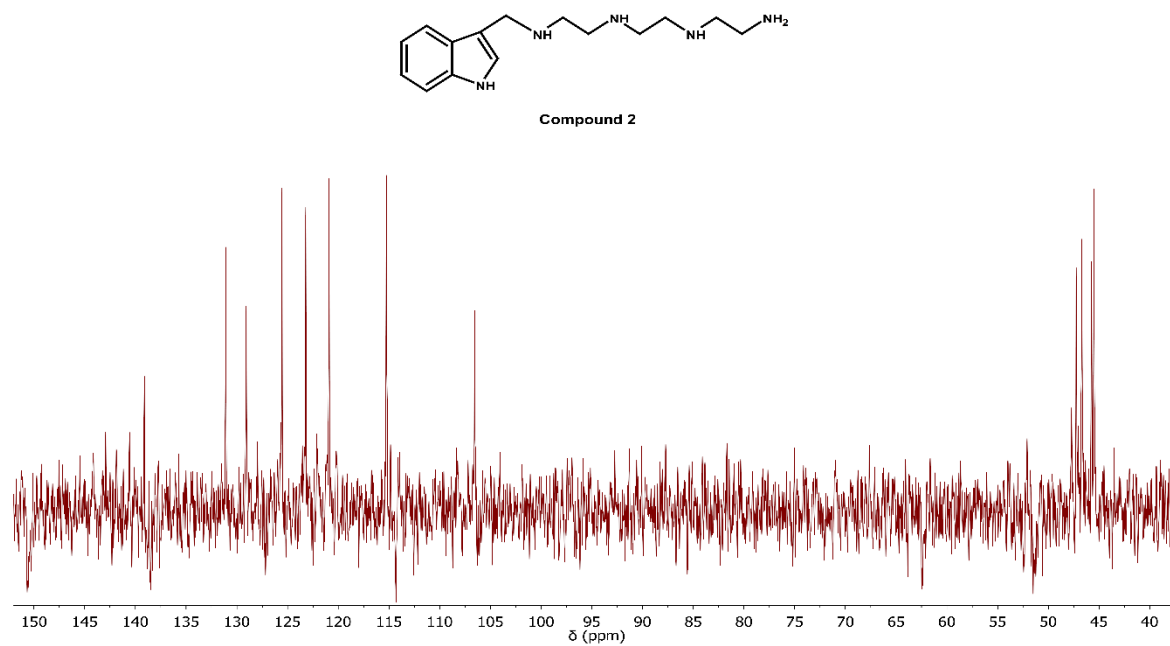


Figure S12 - ¹³C NMR spectra of compound **2**. Signals were confirmed by dept-135 spectra.

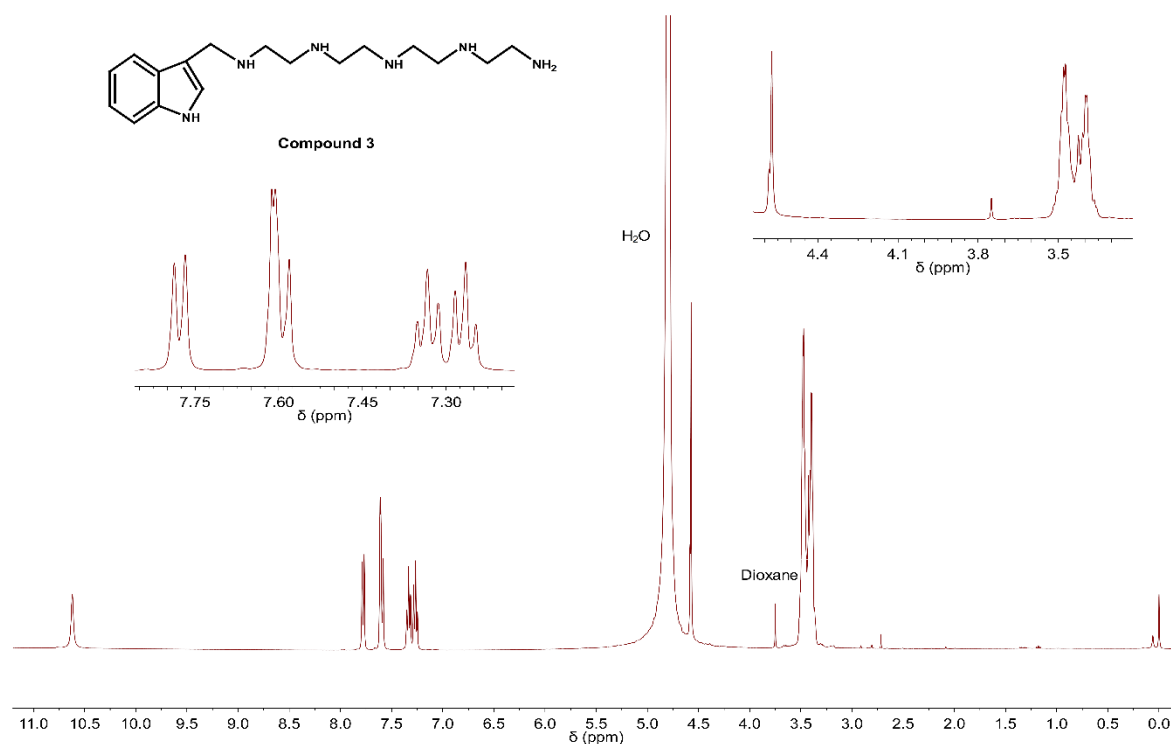


Figure S13 - ¹H NMR spectra of compound **3**.

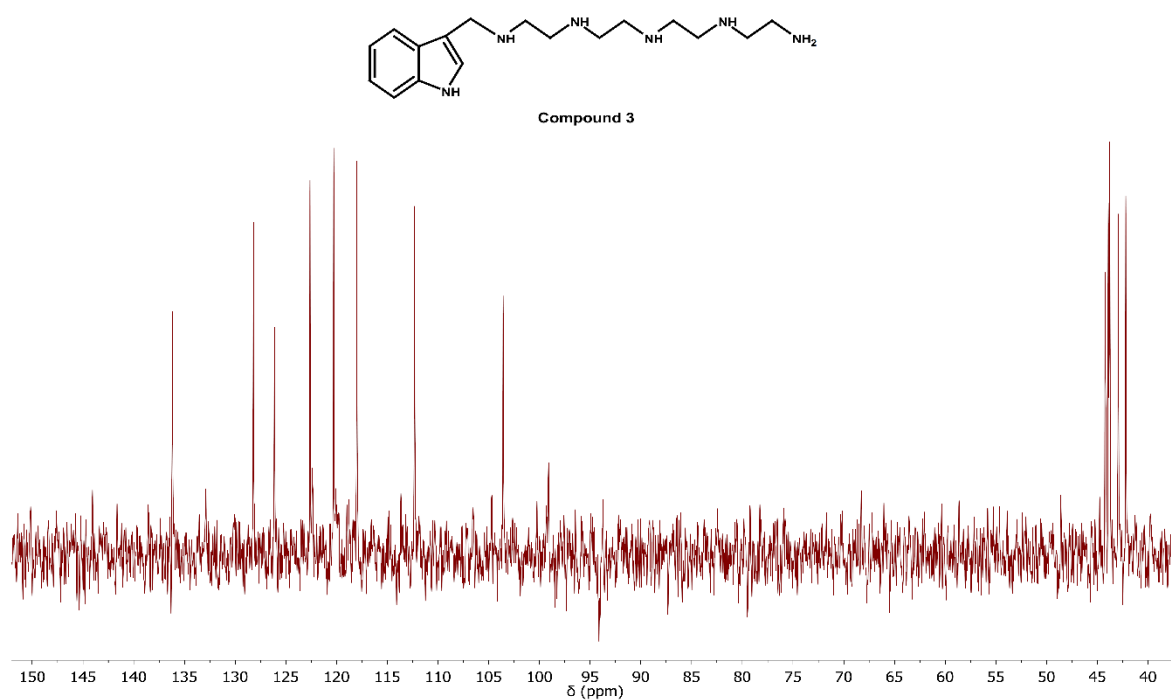


Figure S14 - ¹³C NMR spectra of compound **3**. Signals were confirmed by dept-135 spectra.

Appendix B

Supporting Information for Chapter 5

Synthesis

[16]phenN₂: Yield 92%. m.p. 208–209 °C. ¹H NMR (600 MHz, D₂O): δ (ppm) 1.45 (m, 4H, NCH₂CH₂CH₂), 1.81 (m, 4H, NCH₂CH₂CH₂), 3.23 (m, 4H, NCH₂CH₂CH₂), 4.66 (s, 4H, phenCH₂N), 7.78 (d, 2H, phen), 7.99 (s, 2H, phen) and 8.53 (d, 2H, phen); ¹³C NMR (150 MHz, D₂O): δ (ppm) 25.46 (NCH₂CH₂CH₂), 25.65 (NCH₂CH₂CH₂), 47.59 (NCH₂CH₂CH₂), 51.64 (phenCH₂N), 123.02, 127.07, 129.10, 138.88, 144.12 and 151.37 (phen). HRMS-ESI *m/z* calcd for C₂₀H₂₅N₄ [L+H]⁺ 321.2074, found 321.2128.

[32]phen₂N₄: Yield 70%. m.p. 201–202 °C. ¹H NMR (600 MHz, D₂O): δ (ppm) 1.37 (m, 8H, NCH₂CH₂CH₂), 1.73 (m, 8H, NCH₂CH₂CH₂), 3.16 (m, 8H, NCH₂CH₂CH₂), 4.61 (s, 8H, phenCH₂N), 7.73 (d, 4H, phen), 7.88 (s, 4H, phen) and 8.45 (d, 4H, phen); ¹³C NMR (150 MHz, D₂O): δ (ppm) 25.39 (NCH₂CH₂CH₂), 25.58 (NCH₂CH₂CH₂), 47.68 (NCH₂CH₂CH₂), 51.22 (phenCH₂N), 123.31, 127.21, 129.01, 139.49, 144.12 and 151.12 (phen). HRMS-ESI *m/z* calcd for C₄₀H₄₉N₈ [L+H]⁺ 641.4075, found 641.4088.

[30]phen₂N₆: Yield 40%. m.p. 224–225 °C. ¹H NMR (600 MHz, D₂O): δ (ppm) 3.34 (t, 8H, NCH₂CH₂N), 3.53 (t, 8H, NCH₂CH₂N), 4.65 (s, 8H, phenCH₂N), 7.46 (d, 4H, phen), 7.81 (s, 4H, phen), 8.18 (d, 4H, phen); ¹³C NMR (150 MHz, D₂O): δ (ppm) 43.81 (NCH₂CH₂N), 45.45 (NCH₂CH₂N), 51.18 (phenCH₂N), 67.83 (dioxane), 122.15, 127.33, 128.84, 139.07, 143.44, 151.35 (phen). HRMS-ESI *m/z* calcd for C₃₆H₄₃N₁₀ [L+H]⁺ 615.3667, found 615.3708.

[34]phen₂N₆: Yield: 50%; m.p. 227–228 °C; ¹H NMR (600 MHz, D₂O, 25 °C): δ (ppm) 2.31 (m, 8H; NCH₂CH₂CH₂N), 3.27 (t, 16H; NCH₂CH₂CH₂N), 3.43 (t, 16H; NCH₂CH₂CH₂N), 4.71 (s, 8H; phenCH₂N), 7.82 (d, 4H; phen), 8.01 (s, 4H; phen), 8.56 ppm (d, 4H; phen); ¹³C NMR (150 MHz, D₂O): δ (ppm) 44.26 (NCH₂CH₂CH₂N), 44.52 (NCH₂CH₂CH₂N), 44.72 (NCH₂CH₂CH₂N), 48.99 (methanol), 52.08 (phenCH₂N), 67.97 (dioxane), 123.43, 127.26, 129.32, 139.32, 144.32, 151.09 ppm (phen). HRMS-ESI *m/z* calcd for C₄₀H₅₁N₁₀ [L+H]⁺ 671.4293, found 671.4341.

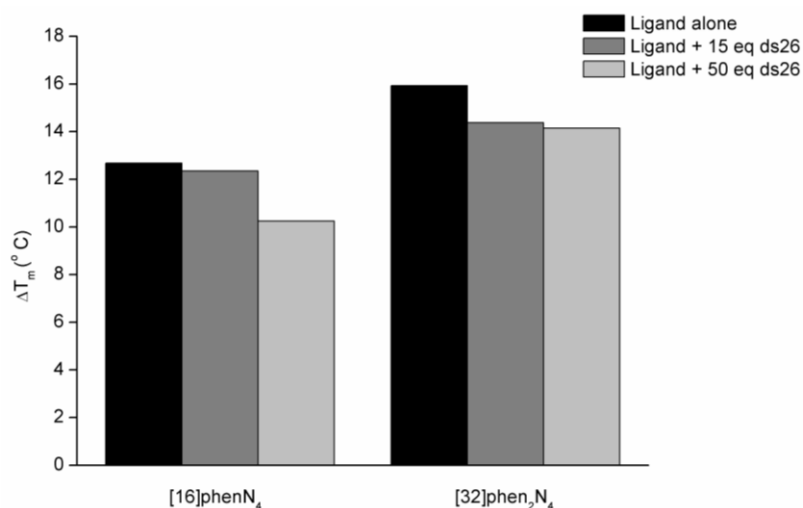


Figure S1 - FRET-melting stabilization by [16]phenN₂ and [32]phen₂N₄ (1 μ M) of the labelled F21T sequence (0.2 μ M) in buffer 10 mM KCl + 90 mM LiCl and in the presence of increasing concentrations of competitor ds26 (0, 15 and 50 equivalents).

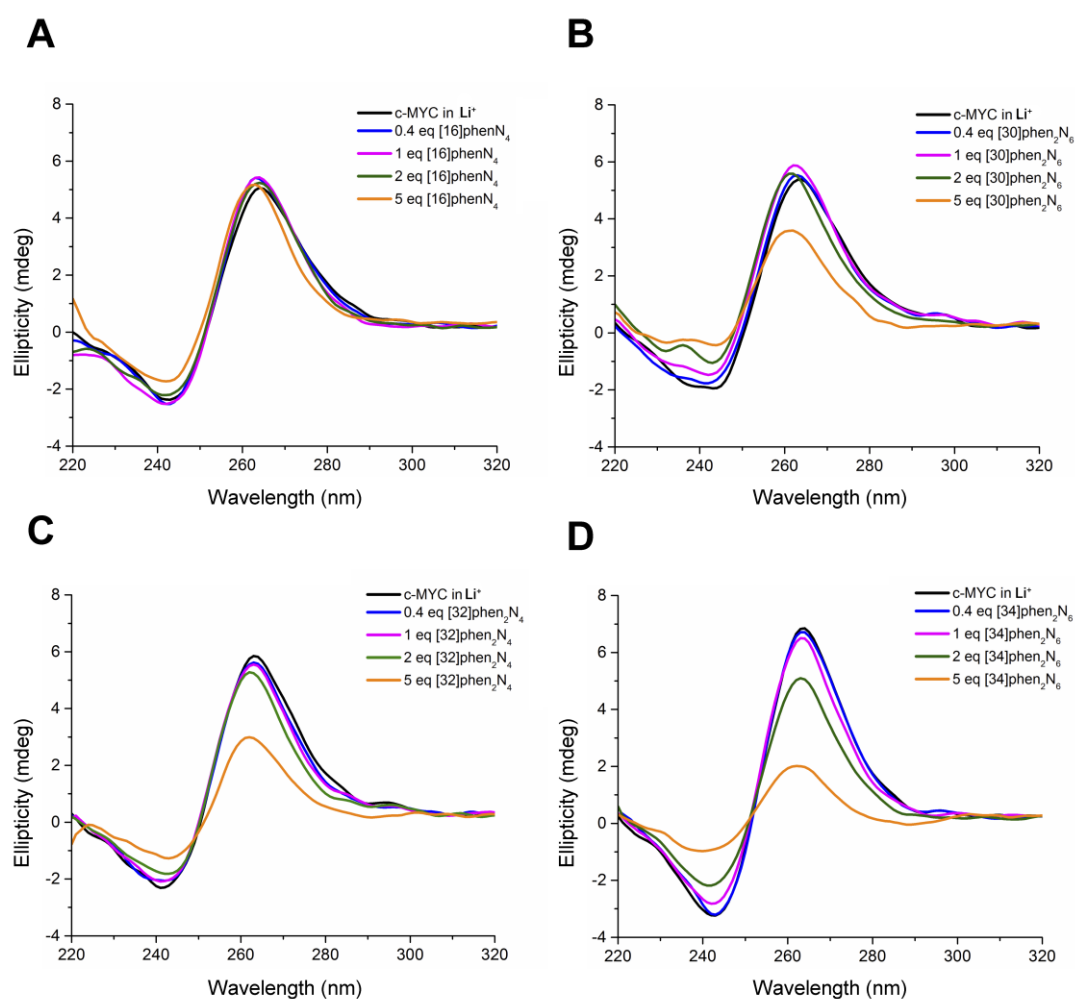


Figure S2 - CD titration spectra of (A) [16]phenN₂, (B) [30]phen₂N₆, (C) [32]phen₂N₄ and (D) [34]phen₂N₆ with c-MYC G4 in 100 mM LiCl. Some of the intermediate spectra were removed for better clarity.

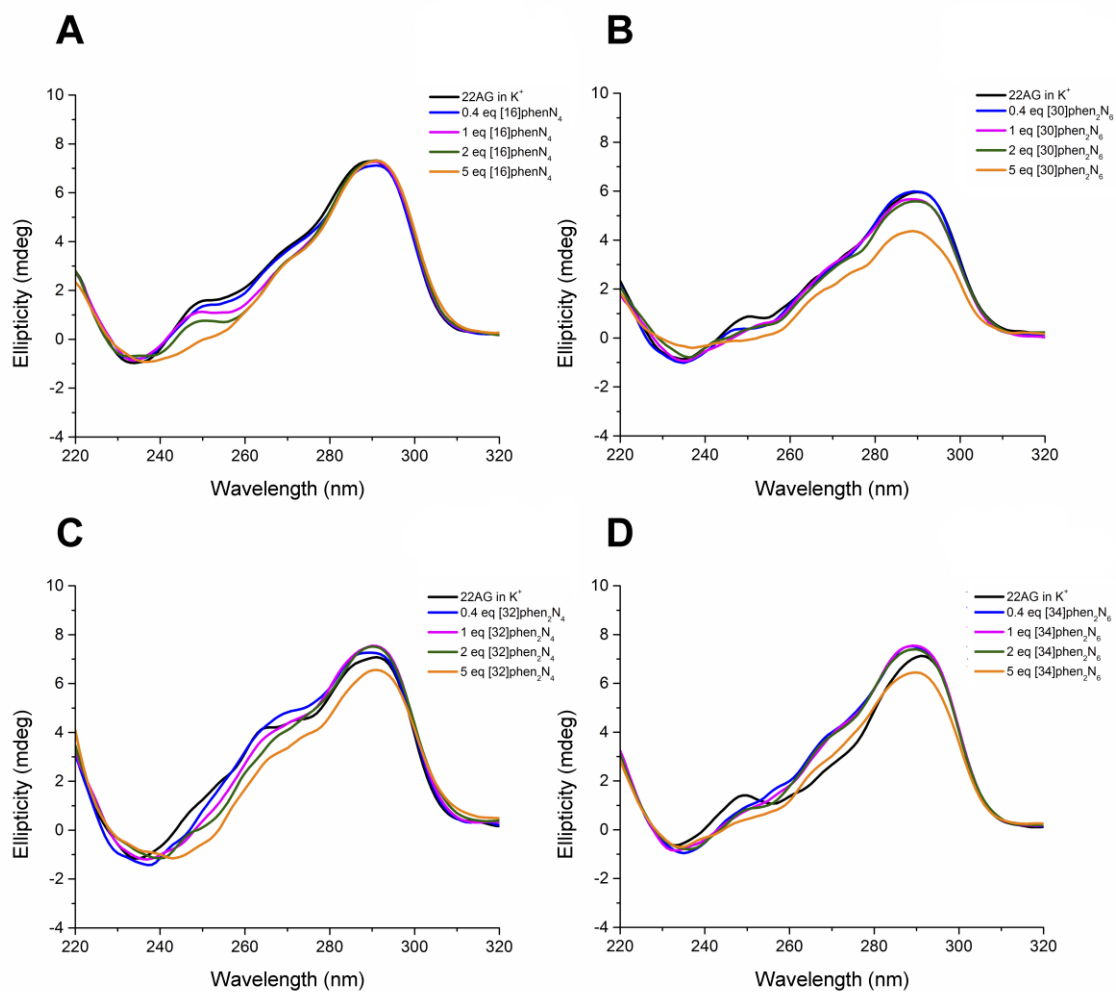


Figure S3 - CD titration spectra of (A) [16]phenN₂, (B) [30]phen₂N₆, (C) [32]phen₂N₄ and (D) [34]phen₂N₆ with 22AG G4 in 10 mM KCl + 90 mM LiCl. Some of the intermediate spectra were removed for better clarity.

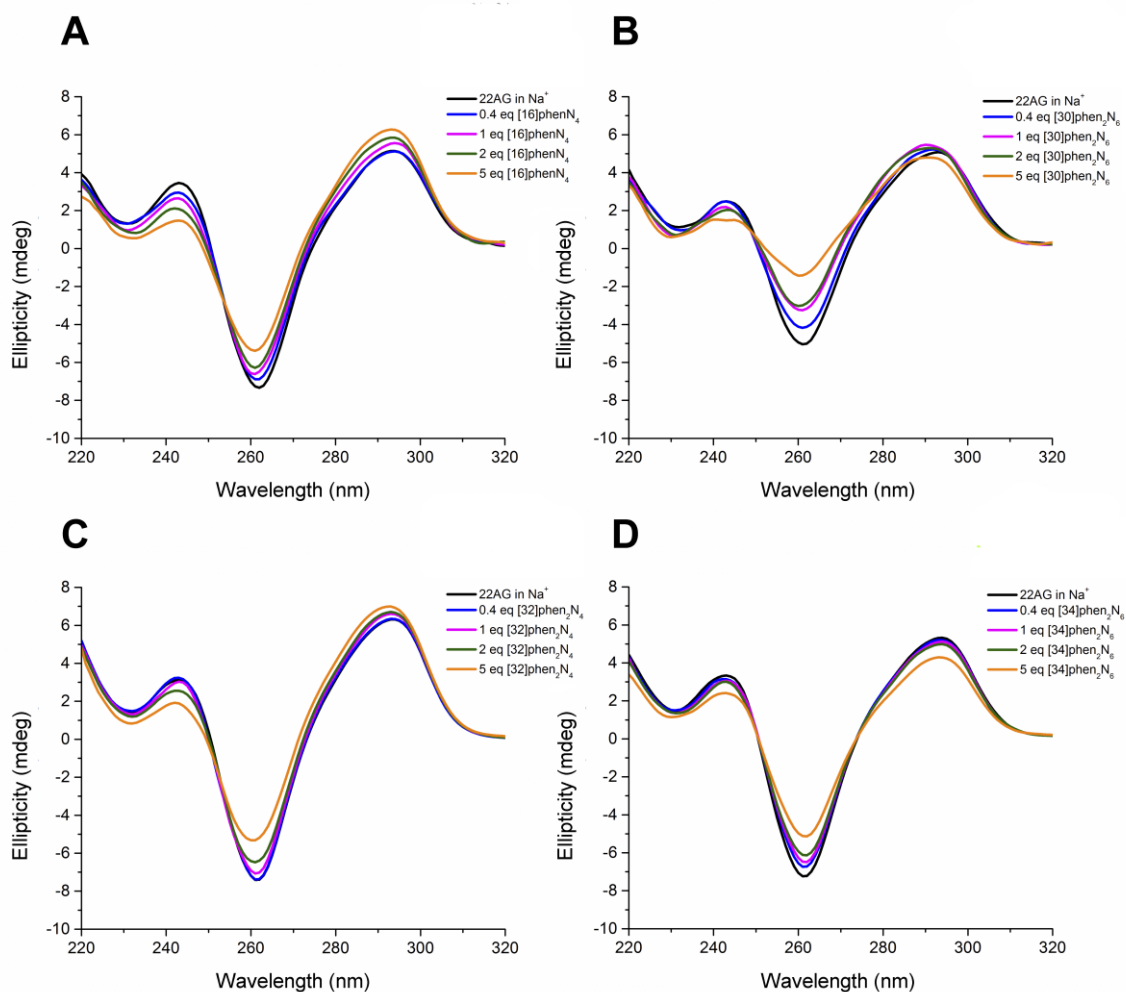


Figure S4 - CD titration spectra of (A) [16]phenN₂, (B) [30]phen₂N₆, (C) [32]phen₂N₄ and (D) [34]phen₂N₆ with 22AG G4 in 100 mM NaCl. Some of the intermediate spectra were removed for better clarity.

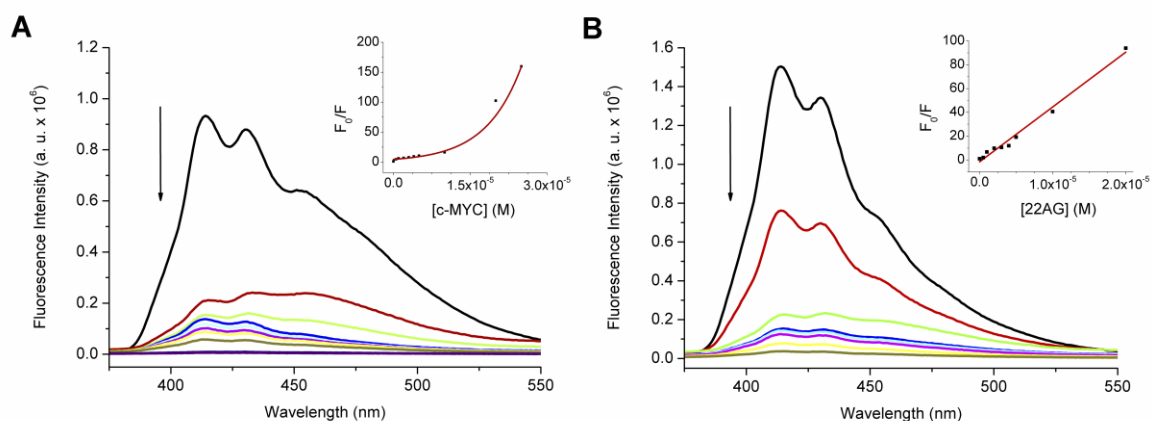


Figure S5 - Fluorescence titrations of the ligand [16]phenN₂ with increasing concentration of (A) c-MYC and (B) 22AG G4s [5 μ M in 100 mM KCl, pH 7.2, concentrations 0–25 μ M]. Insets: Stern-Volmer plots obtained by plotting F_0/F as a function of DNA concentration.

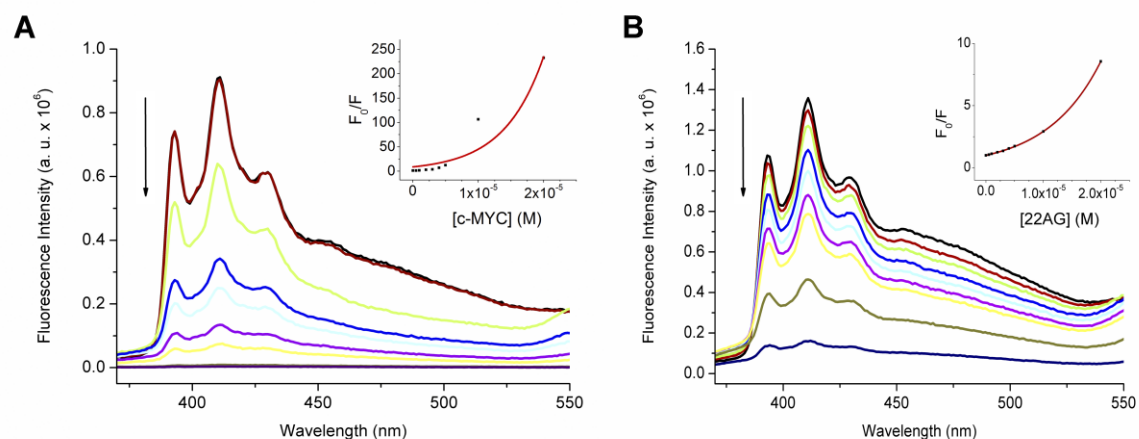


Fig S6 Fluorescence titrations of the ligand [30]phen₂N₆ with increasing concentration of (A) c-MYC and (B) 22AG G4s [5 μ M in 100 mM KCl, pH 7.2, concentrations 0–25 μ M]. Insets: Stern-Volmer plots obtained by plotting F_0/F as a function of DNA concentration.

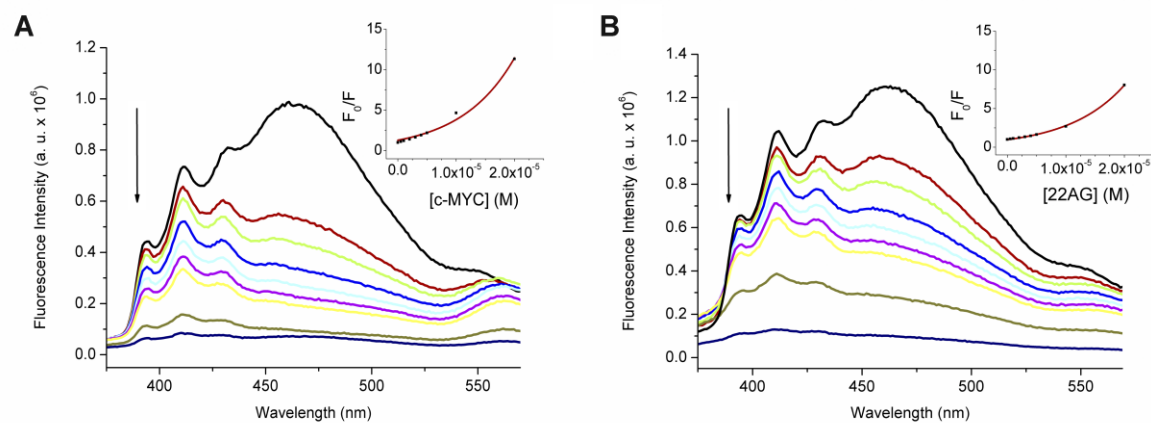


Figure S7 - Fluorescence titrations of the ligand [34]phen₂N₆ with increasing concentration of (A) c-MYC and (B) 22AG G4s [5 μ M in 100 mM KCl, pH 7.2, concentrations 0–25 μ M]. Insets: Stern-Volmer plots obtained by plotting F_0/F as a function of DNA concentration.

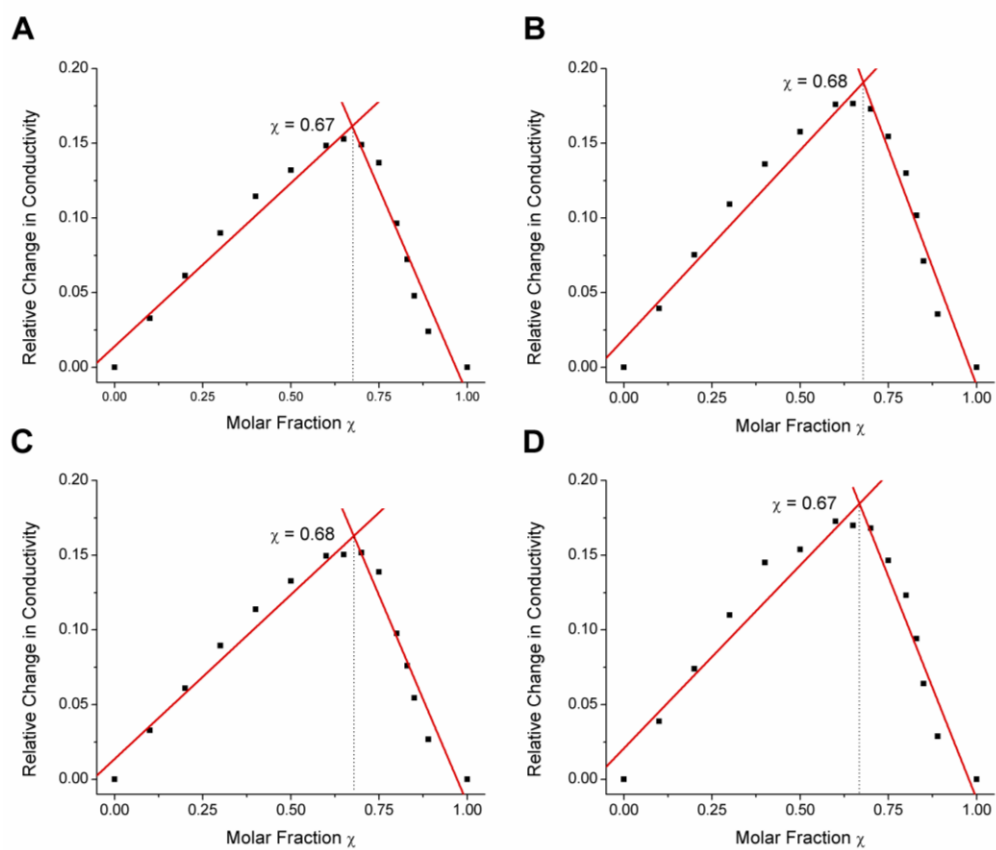
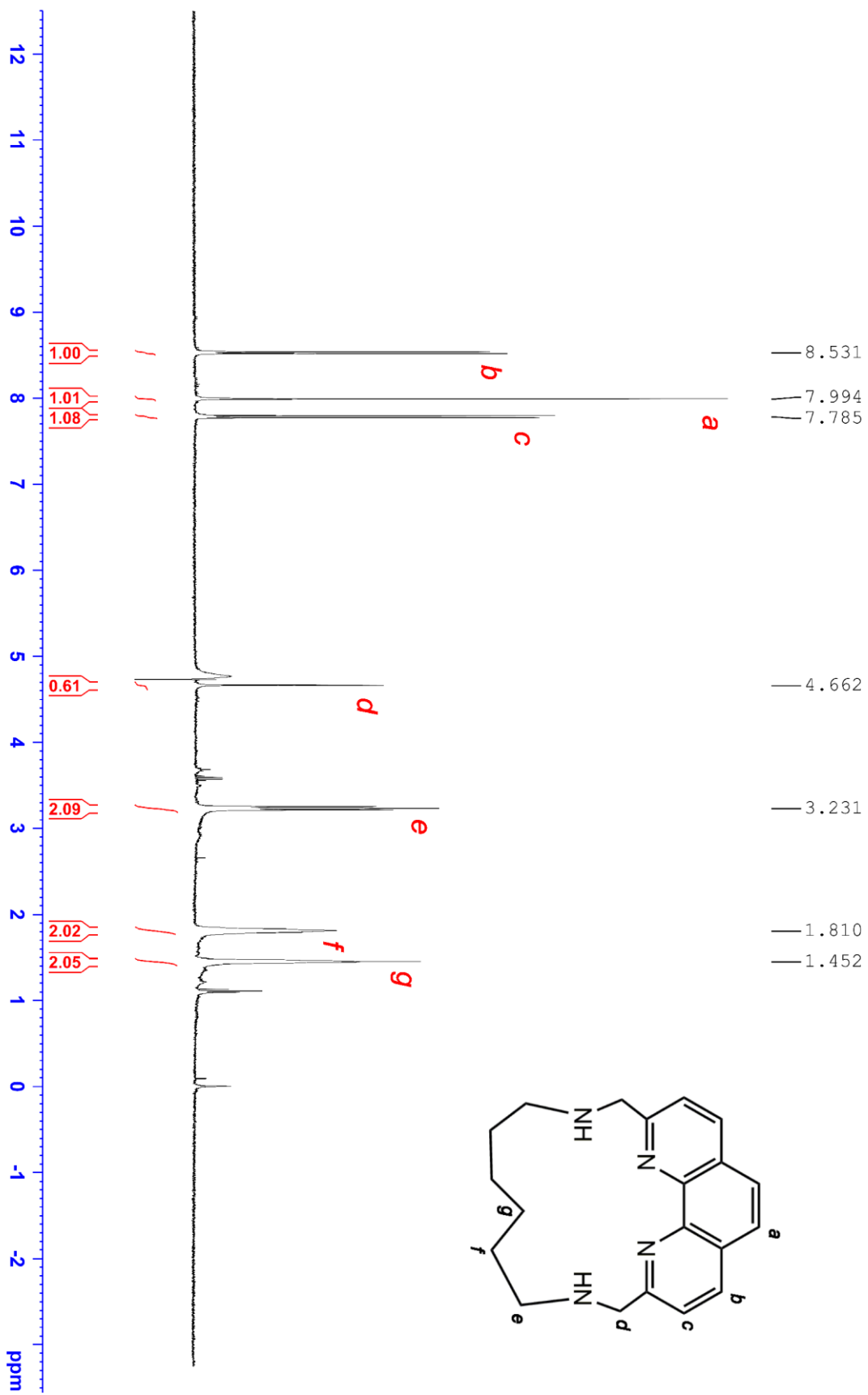
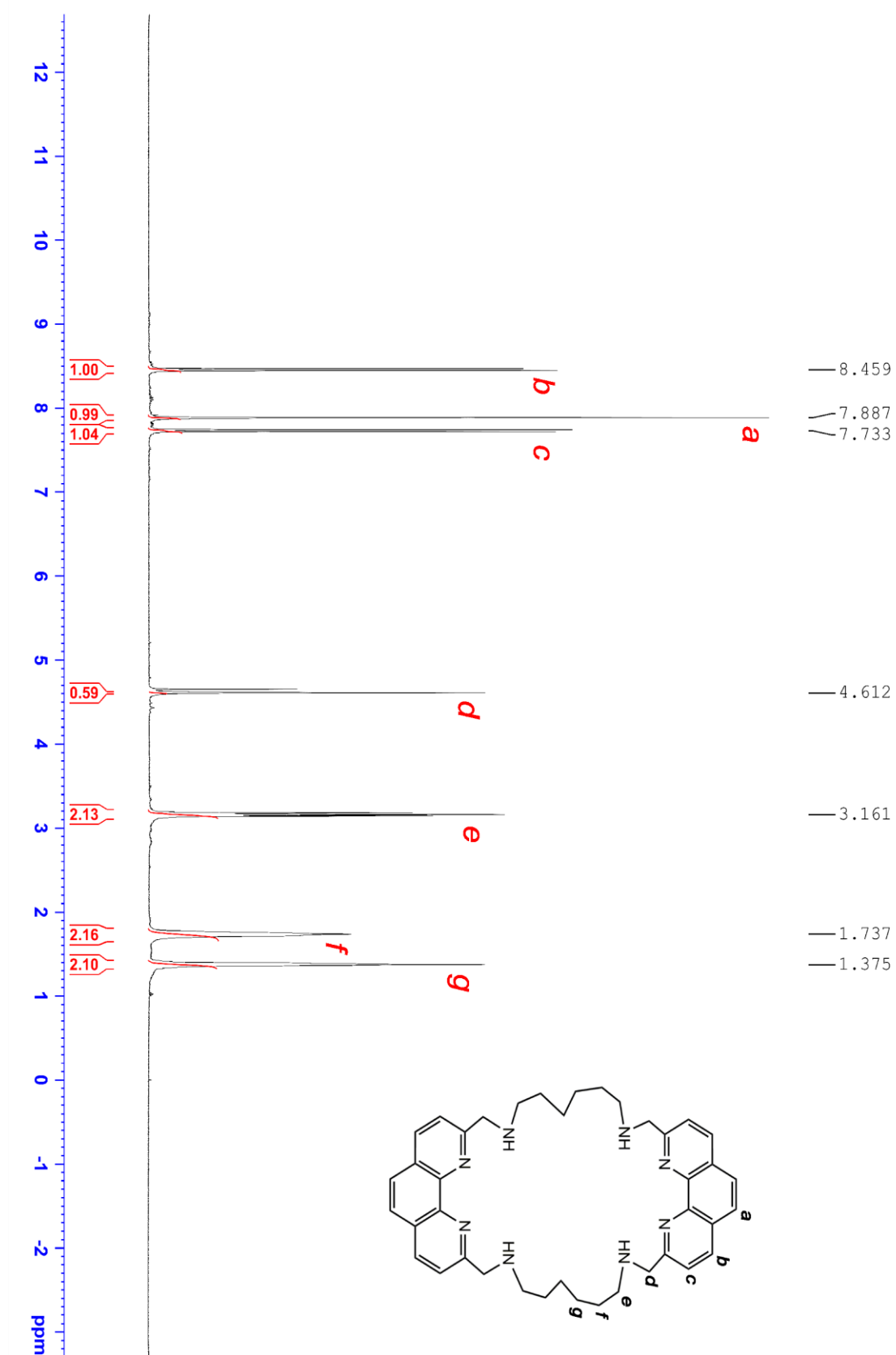


Figure S8 - Job plots resulting from the method of continuous variation analysis for (A) ligand [16]phenN₂ with c-MYC G4, (B) ligand [16]phenN₂ with 22AG G4, (C) [32]phen₂N₄ with c-MYC G4 and (D) [32]phen₂N₄ with 22AG G4.

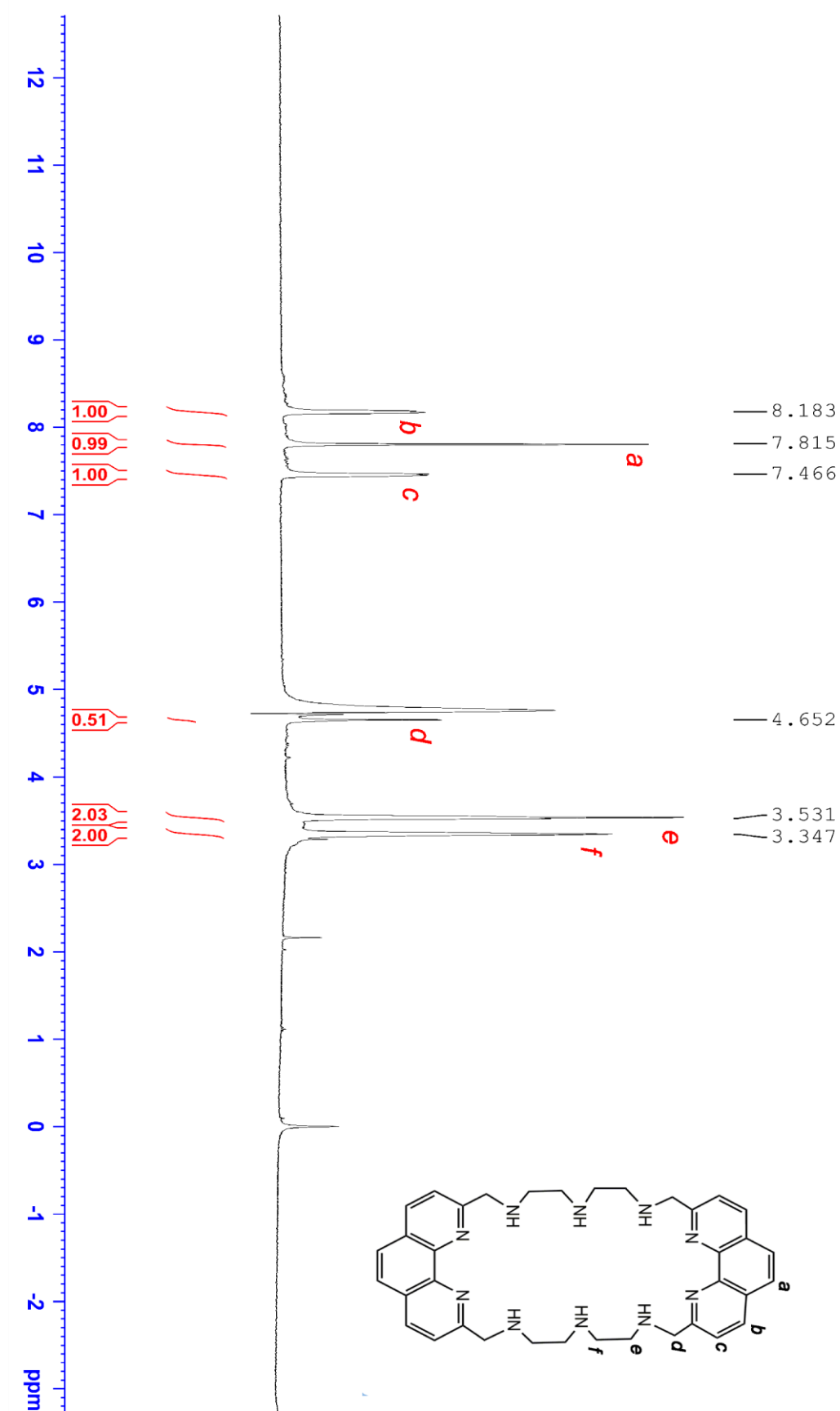
(A)



(B)



(C)



(D)

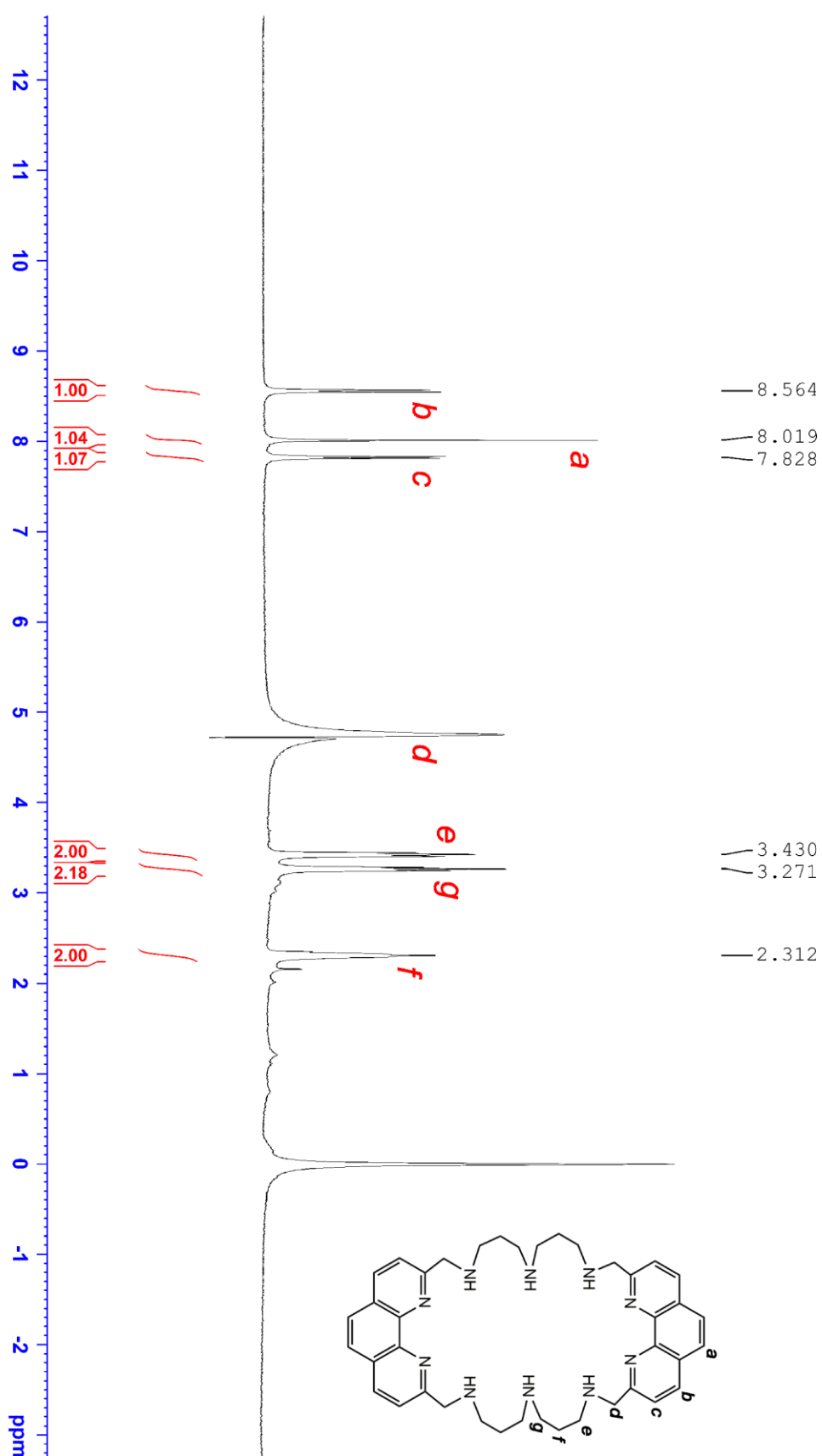
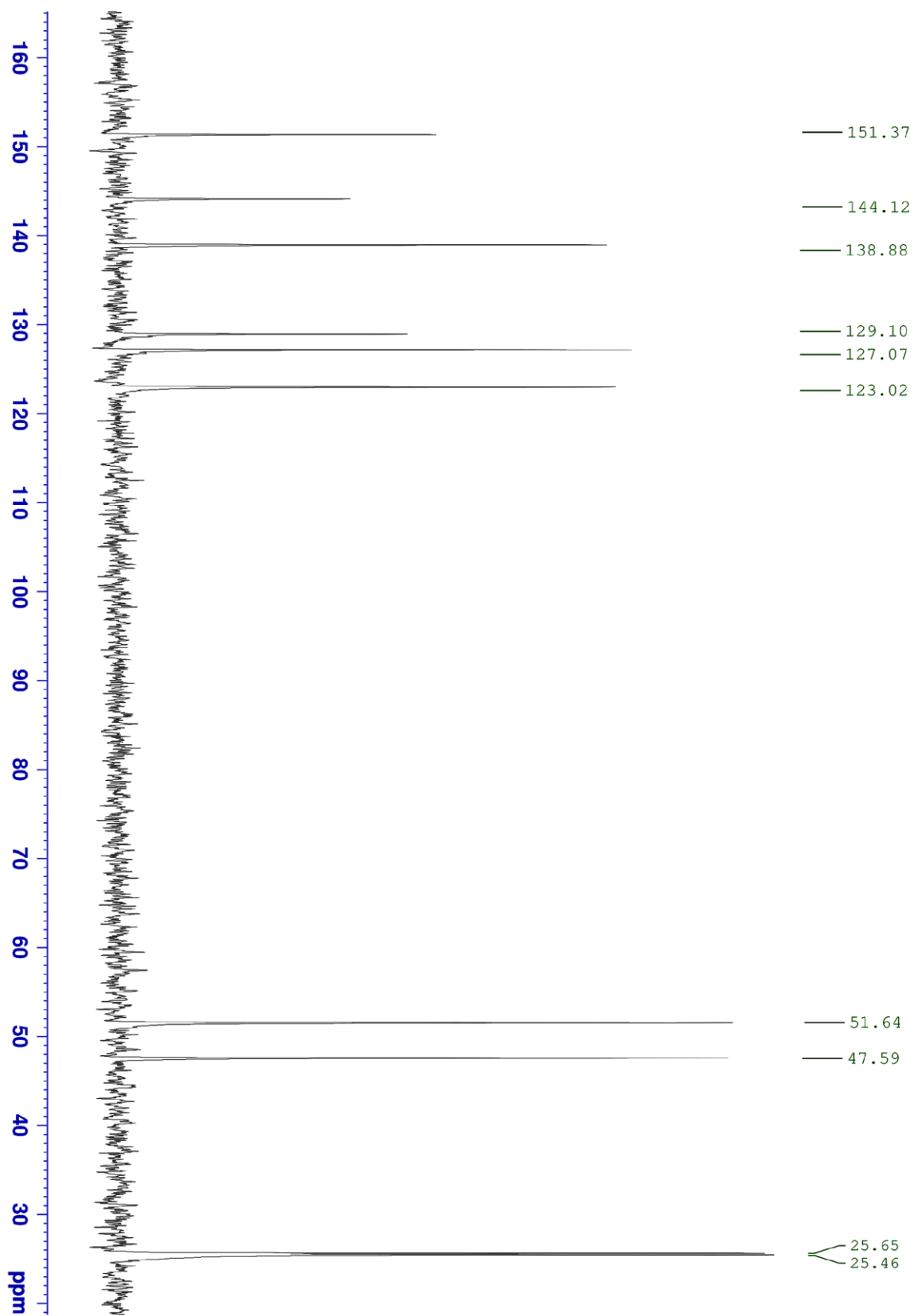
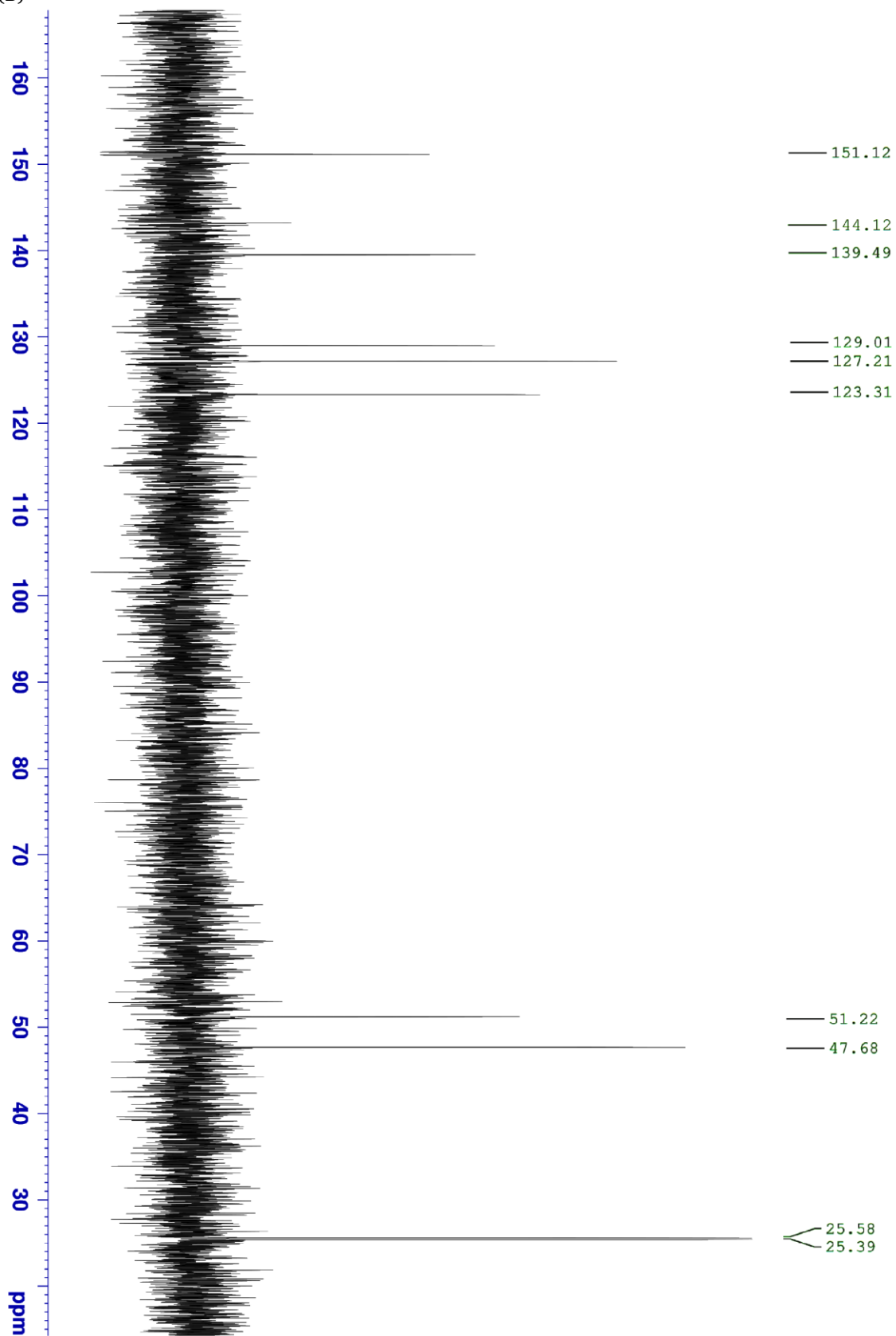


Figure S9 - ¹H NMR spectra of the ligands (A) [16]phenN₂, (B) [32]phen₂N₄ (C) [30]phen₂N₆ and (D) [34]phen₂N₆ (protons d are underneath the water signal and could not be assigned).

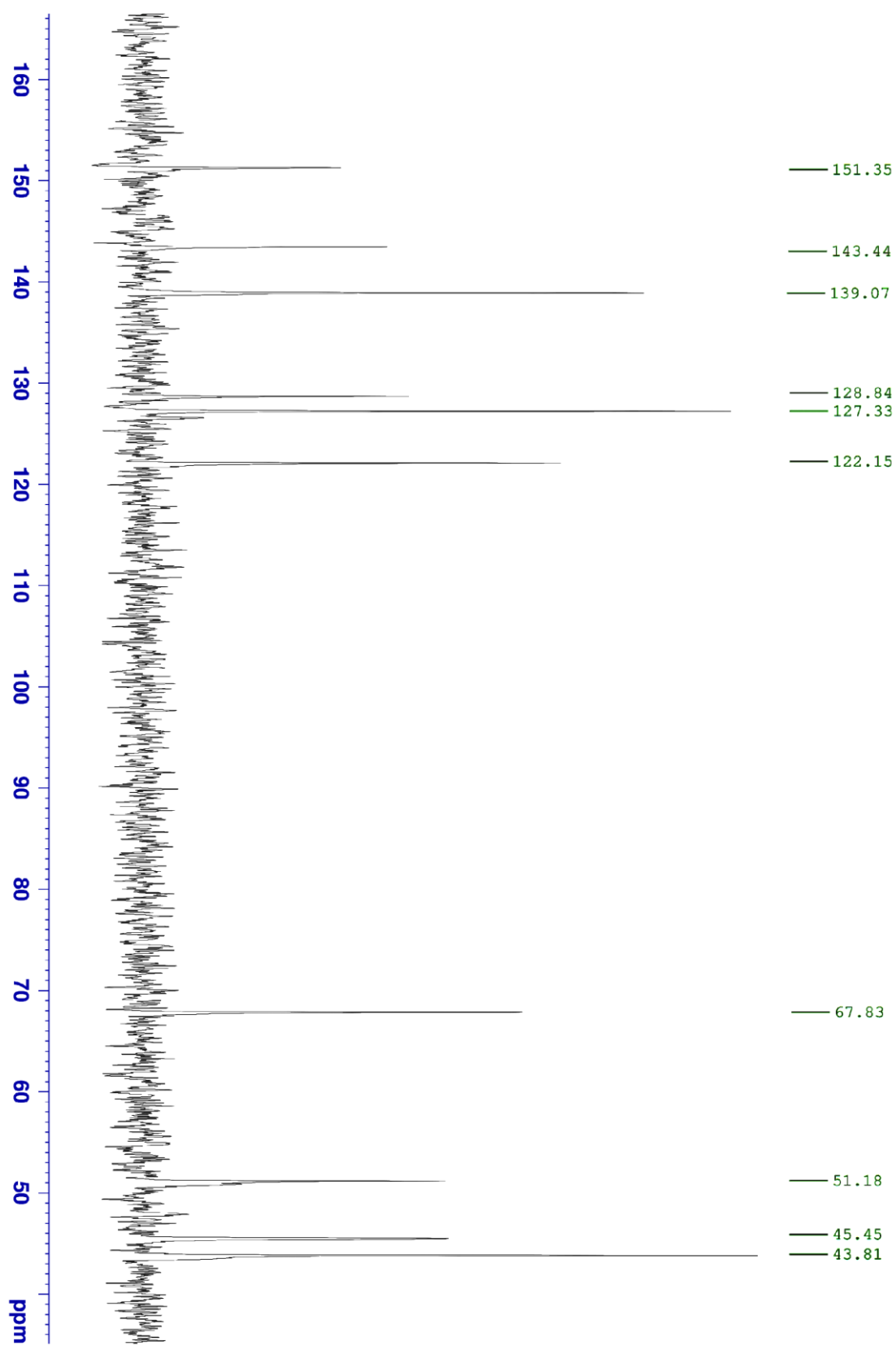
(A)



(B)



(C)



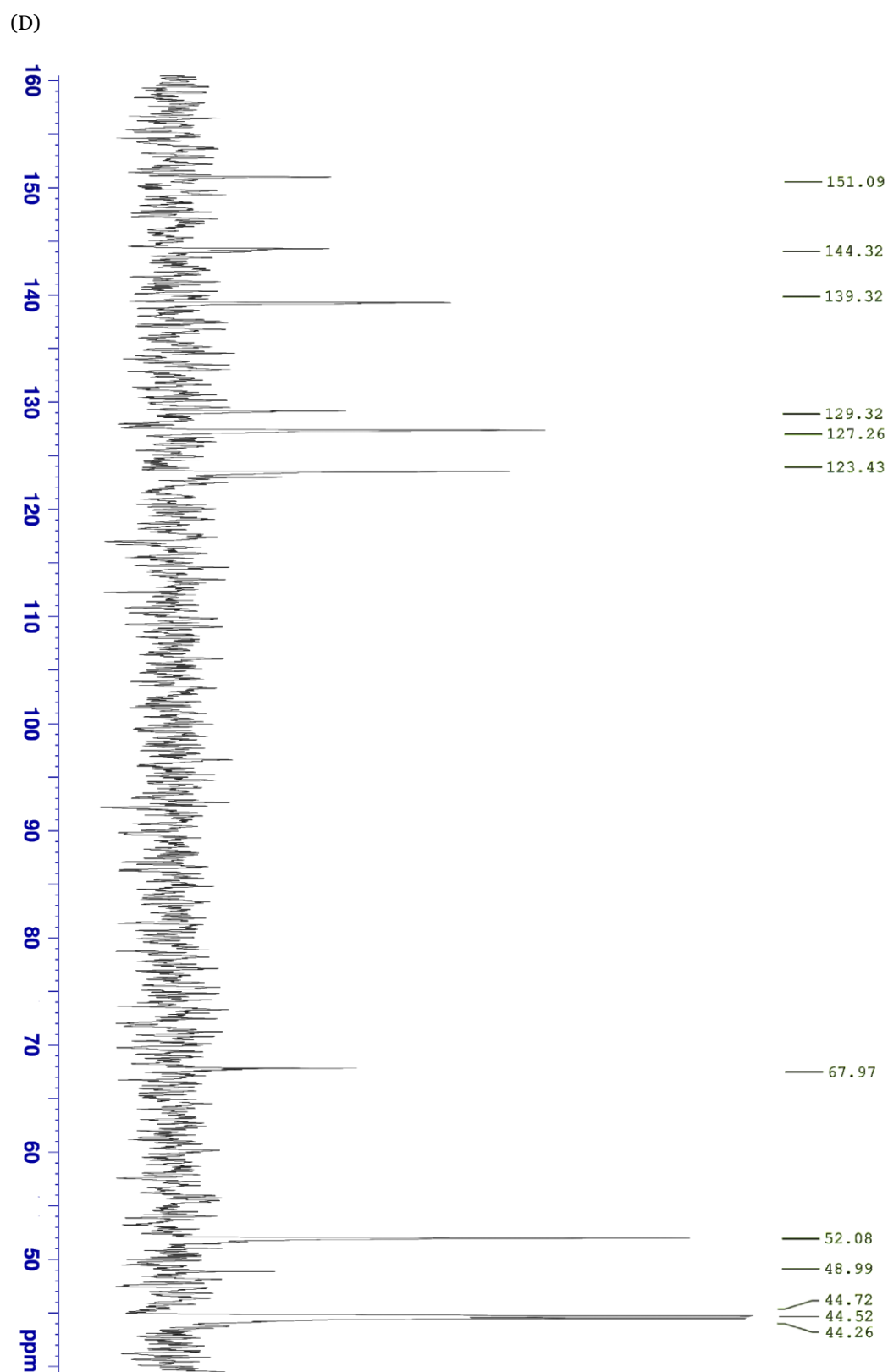


Figure S10 - ^{13}C NMR spectra of the ligands (A) [16]phenN₂, (B) [32]phen₂N₄ (C) [30]phen₂N₆ and (D) [34]phen₂N₆.

Appendix C

Supporting Information for Chapter 6

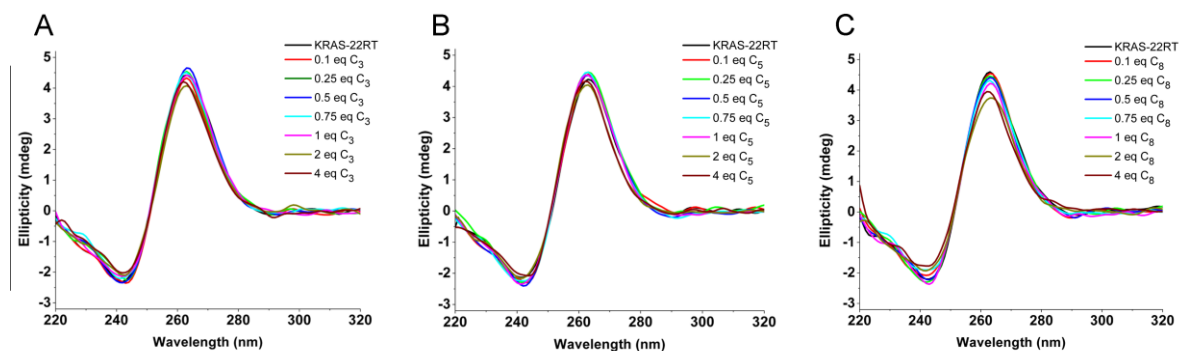


Figure S1 - CD titration spectra of KRAS-22RT G4 at 25 °C in 10 mM K-phosphate buffer and 50 mM KCl with increasing concentrations of (A) C₃, (B) C₅ and (C) C₈.

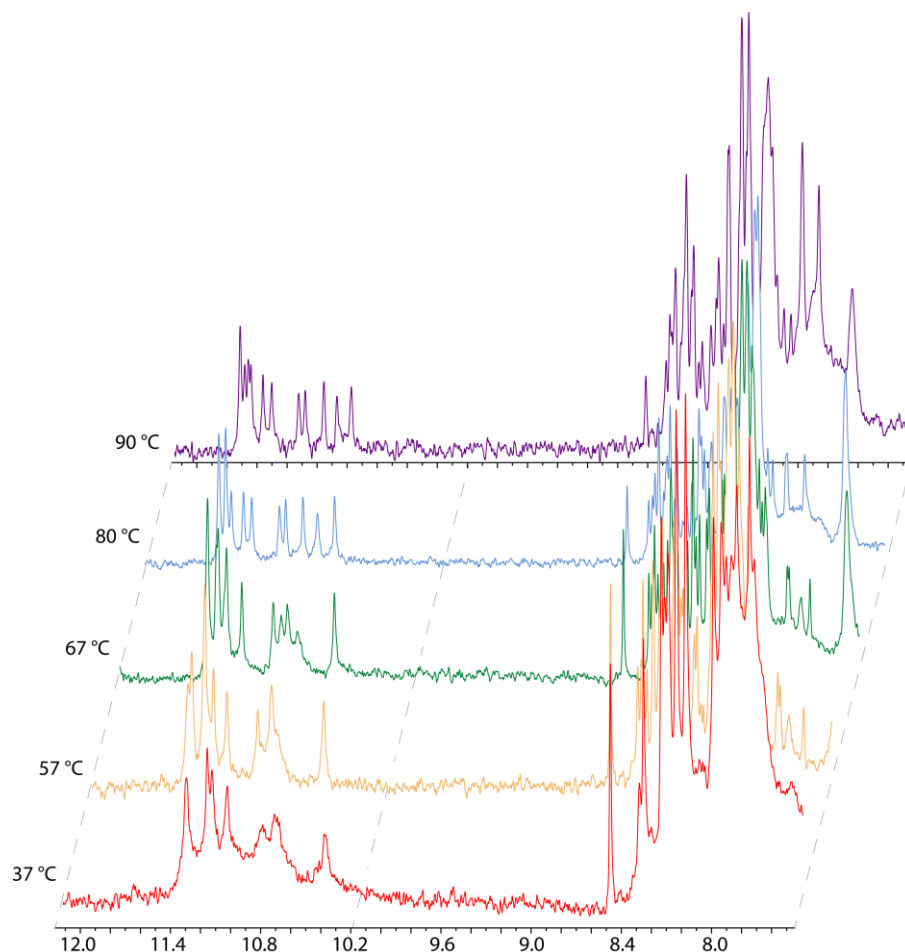


Figure S2 - NMR spectra at different temperatures of KRAS-22RT G4 at 350 μM in presence of 4 molar equivalents of C₈. From the figure, we can appreciate a substantial amount of imino peaks ($\approx 10.4 - 12$ ppm) formed at 90 °C in presence of ligand.

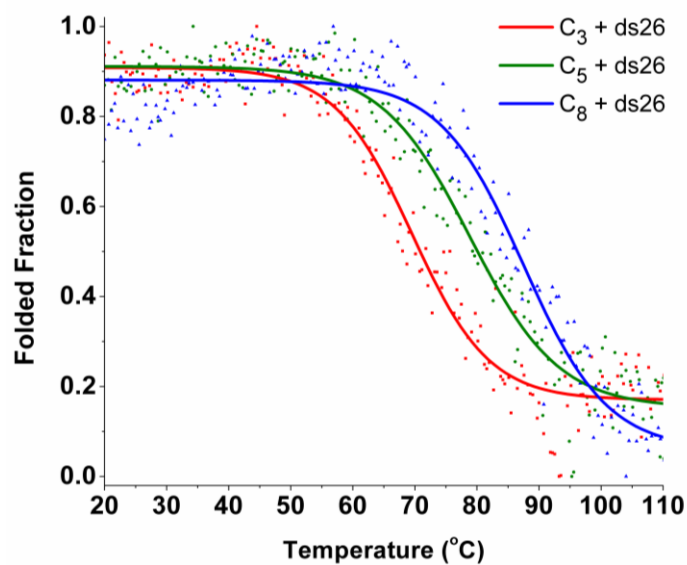


Figure S3. CD melting curves of KRAS-22RT G4 at 5 μ M in the presence of 2 molar equivalents of ligands and 20 μ M of ds26 competitor, in 10 mM K-phosphate buffer and 50 mM KCl. The Boltzmann curve fitting of the data points is shown. The end baseline is around 0.2 folded fraction due to the CD contribution coming from ds26 DNA competitor.

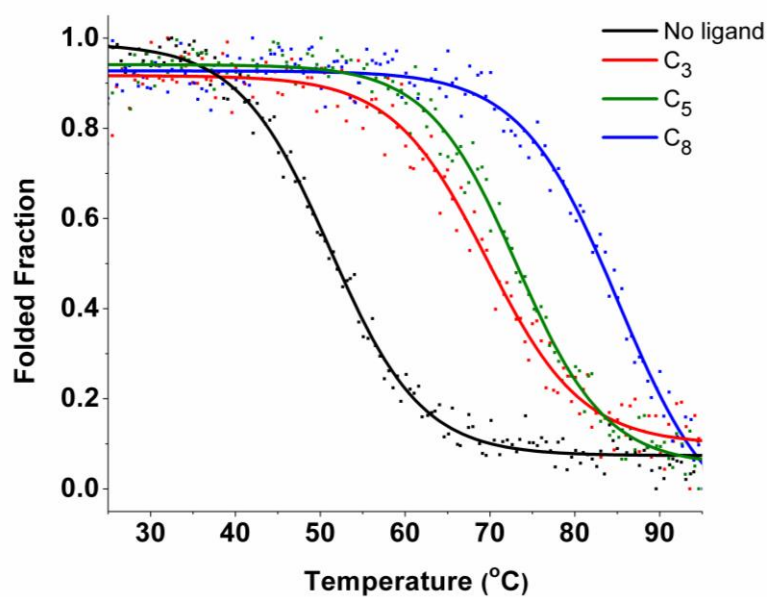


Figure S4 - CD melting curves of c-MYC G4 at 5 μ M in the absence and presence of 2 molar equivalents of ligands, in 10 mM K-phosphate buffer and 100 mM LiCl. The Boltzmann curve fitting of the data points is shown.

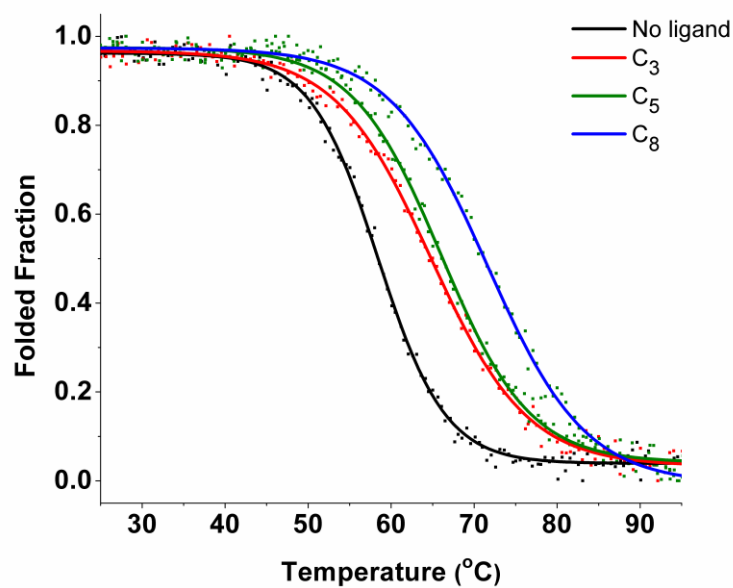


Figure S5 - CD melting curves of hybrid-1 22AG G4 at 5 μM in the absence and presence of 2 molar equivalents of ligands, in 10 mM K-phosphate buffer and 10 mM KCl + 90 mM LiCl. The Boltzmann curve fitting of the data points is shown.

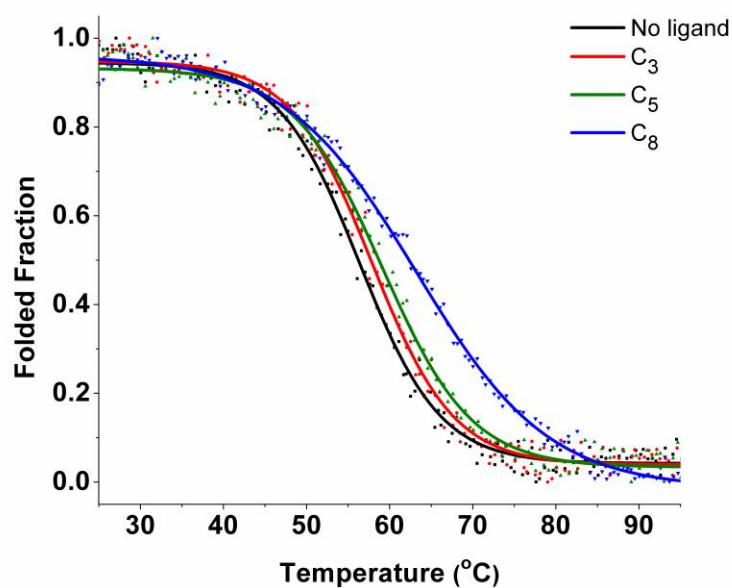


Figure S6 - CD melting curves of antiparallel 22AG G4 at 5 μM in the absence and presence of 2 molar equivalents of ligands, in 10 mM K-phosphate buffer and 100 mM NaCl. The Boltzmann curve fitting of the data points is shown.

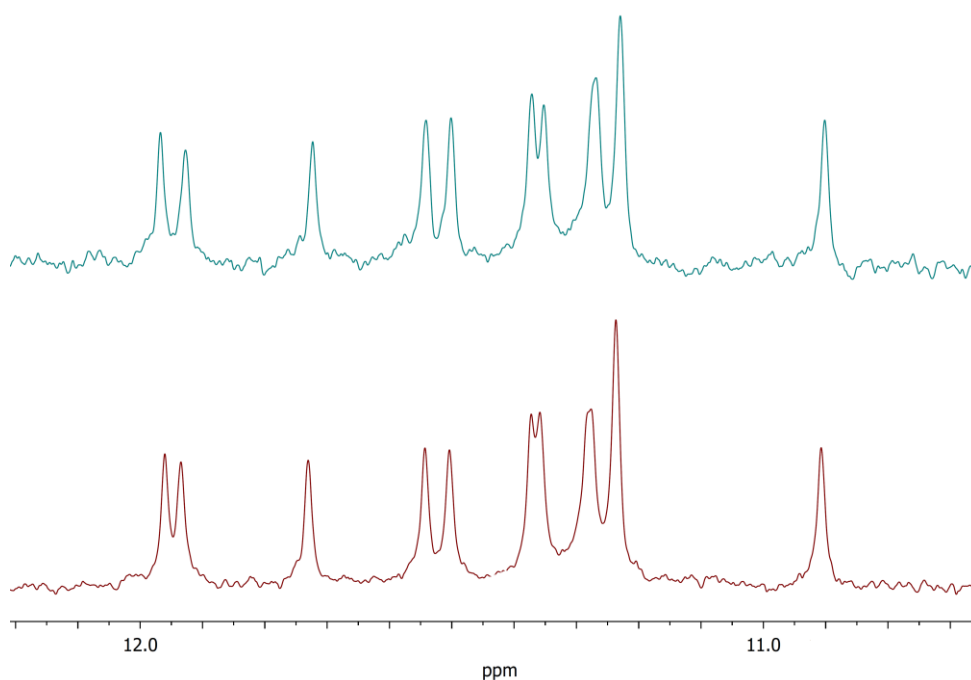


Figure S7 - Imino region of KRAS22-RT G4 ¹H spectra at different KCl concentrations, namely 10 mM (green) and 120 mM (brown). No perceptible differences in chemical shift are observed thus suggesting the same structure of the G4.

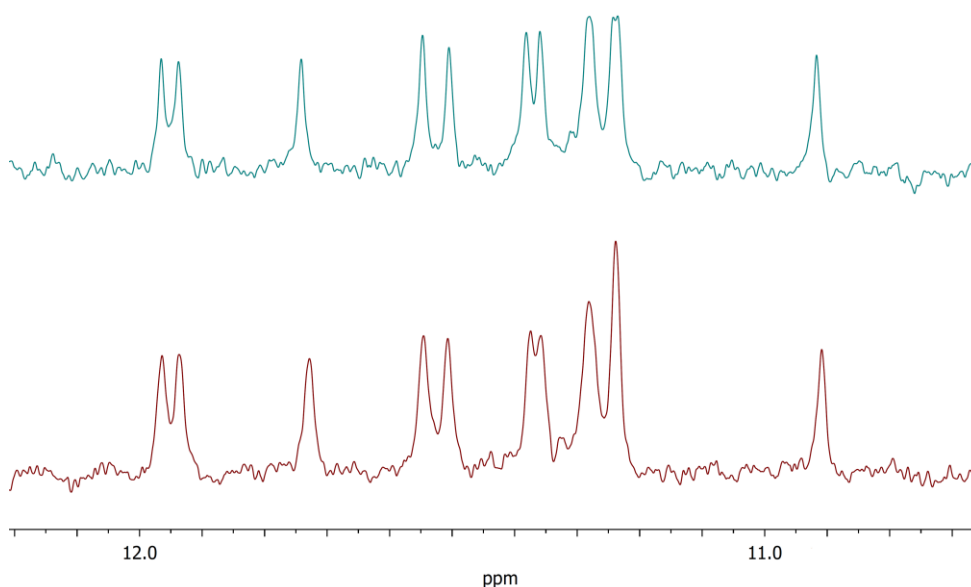


Figure S8 - Imino region of KRAS22-RT G4 ¹H spectra at different pH, namely pH 6.0 (green) and pH 7.0 (brown). No perceptible differences in chemical shift are observed thus suggesting the same structure of the G4.

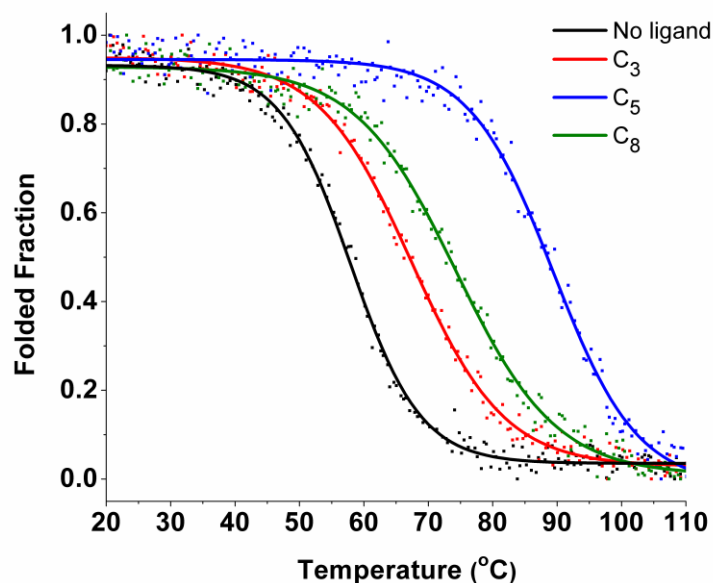


Figure S9 - CD melting curves of KRAS-22RT G4 at 5 μM in the absence and presence of 2 molar equivalents of ligands, in physiologically relevant conditions buffer mimicking the intracellular salt environment (25 mM HEPES (pH = 7.2), 10.5 mM NaCl, 140 mM KCl, 130 nM CaCl_2 , 2 mM MgCl_2). Data points were recorded at 263 nm. The Boltzmann curve fitting of the data points is shown. Temperature corresponds to the temperature set by the Peltier system.

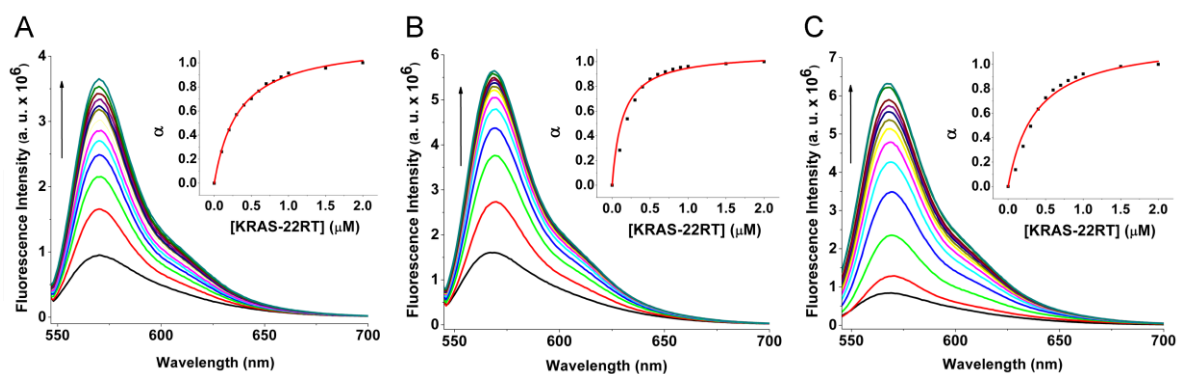


Figure S10 - Fluorescence emission spectra of (A) C_3 , (B) C_5 and (C) C_8 with increasing concentration of KRAS-22RT G4 (0.1 – 2.0 μM) in physiological buffer conditions (25 mM HEPES (pH = 7.2), 10.5 mM NaCl, 140 mM KCl, 130 nM CaCl_2 , 2 mM MgCl_2). Insets: fraction of ligand bound plots fitted to the saturation binding model (Equation 3).

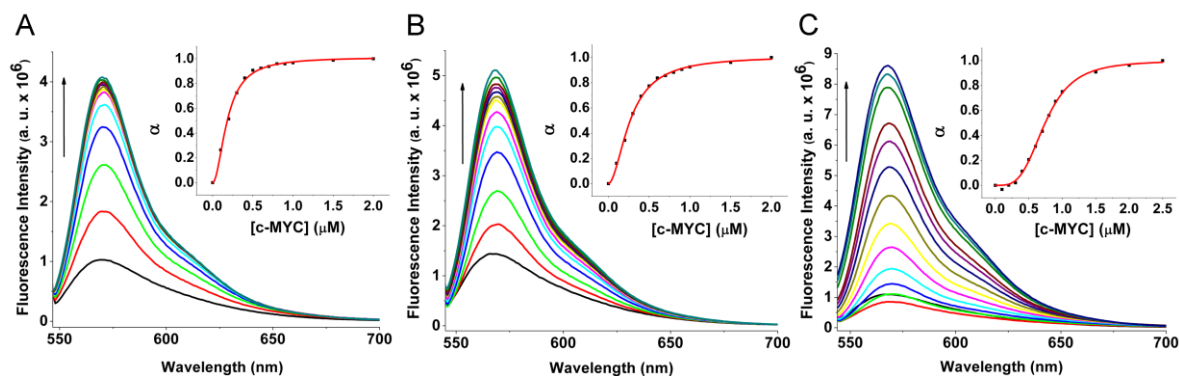


Figure S11 - Fluorescence emission spectra of (A) C_3 , (B) C_5 and (C) C_8 with increasing concentration of c-MYC G4 (0.1 – 2.5 μM) in 10 mM K-phosphate buffer and 100 mM KCl. Insets: fraction of ligand bound plots fitted to the Hill saturation binding model (Equation 4).

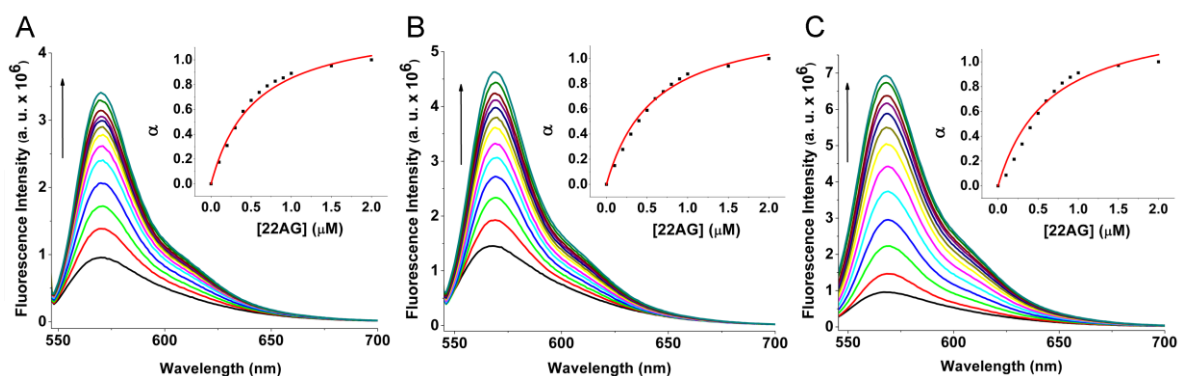


Figure S12 - Fluorescence emission spectra of (A) C_3 , (B) C_5 and (C) C_8 with increasing concentration of 22AG G4 (0.1 – 2.0 μM) in 10 mM K-phosphate buffer and 100 mM KCl. Insets: fraction of ligand bound plots fitted to the saturation binding model (Equation 3).

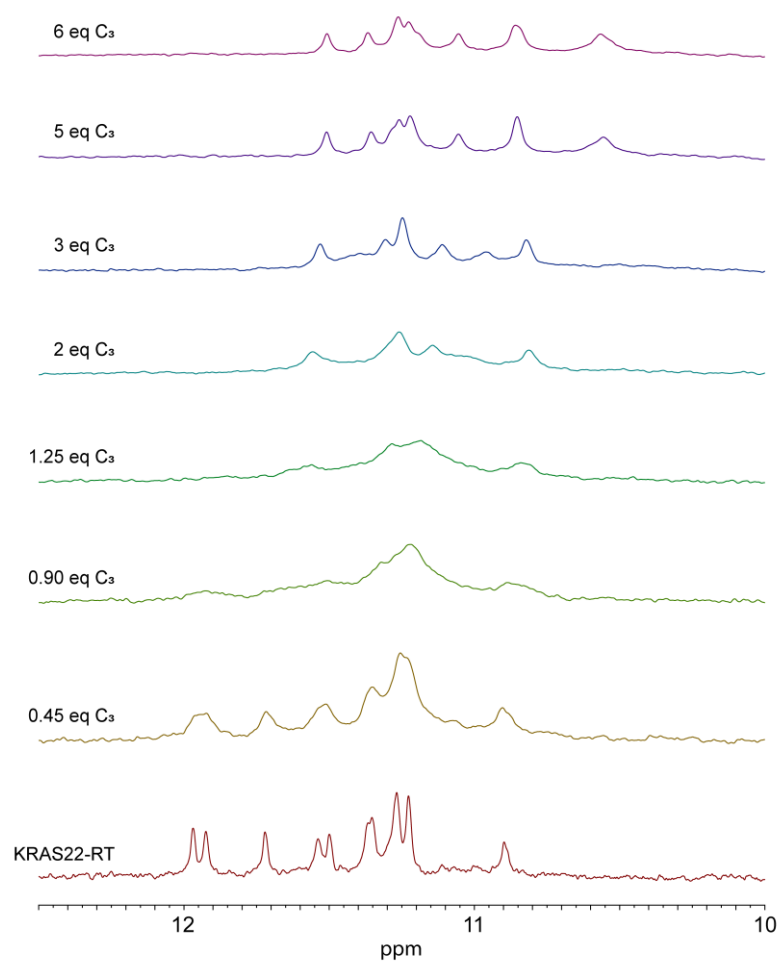


Figure S13 - Imino regions of KRAS22-RT G4 ^1H spectra upon titration of compound C3 at 310 K.

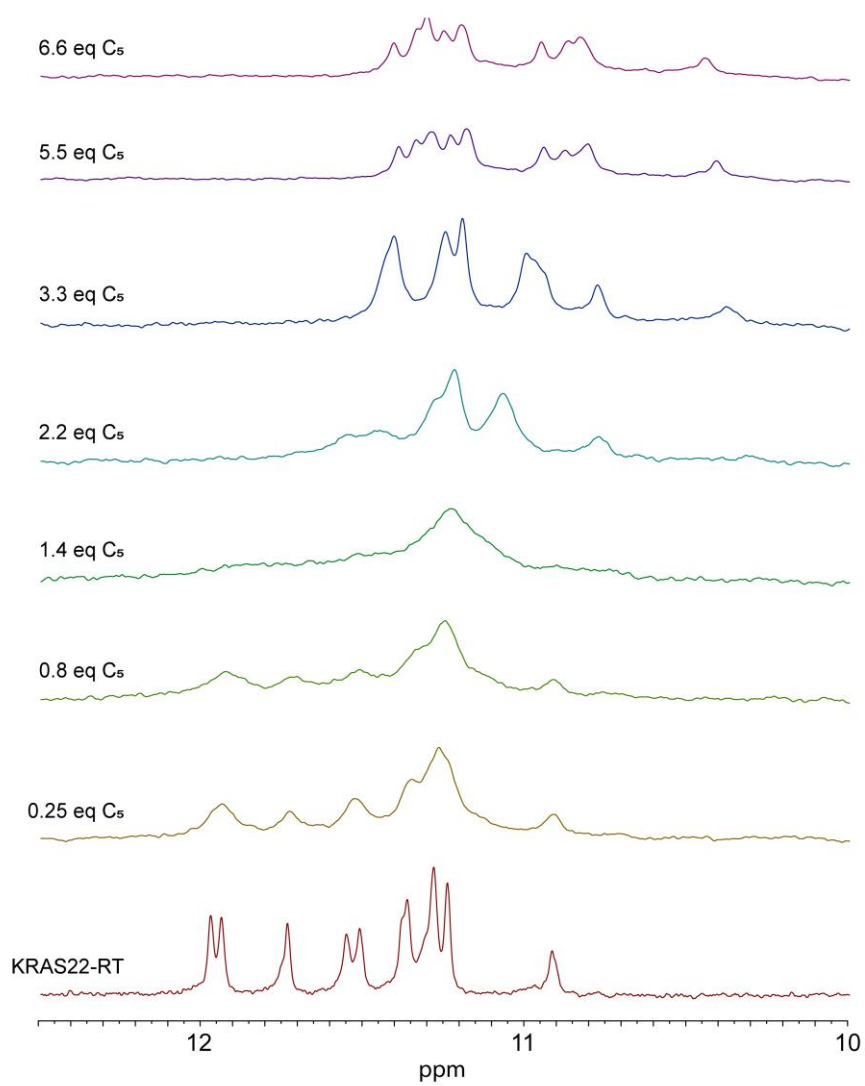


Figure S14 - Imino regions of KRAS22-RT G4 ¹H spectra upon titration of compound C₅ at 310 K.

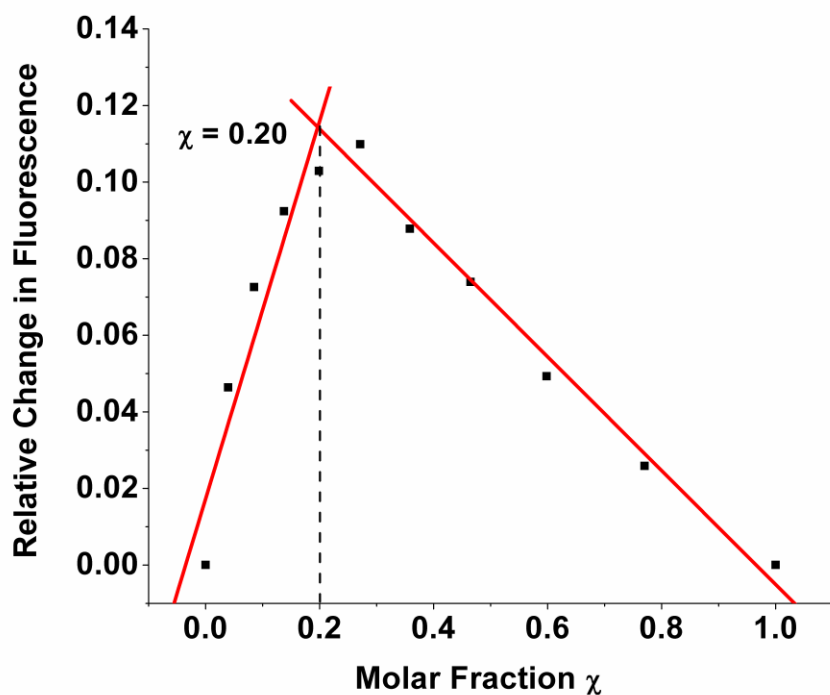


Figure S15 - Job plot analysis of C₈ in 10 mM K-phosphate buffer and 50 mM KCl. The molar fraction (χ) of 0.2 suggests a 4:1 ligand:DNA binding ratio.

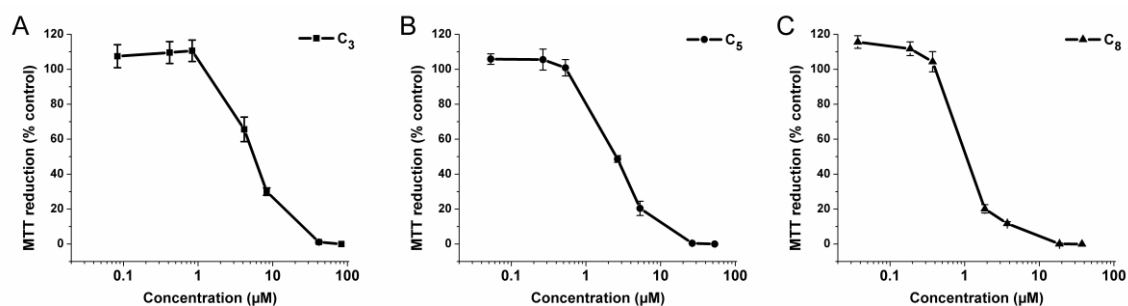


Figure S16 - Cell viability determined by MTT assay for compounds (A) C₃, (B) C₅ and (C) C₈. Figure shows results as mean $\pm$ SEM from at least three independent experiments using HeLa cells.

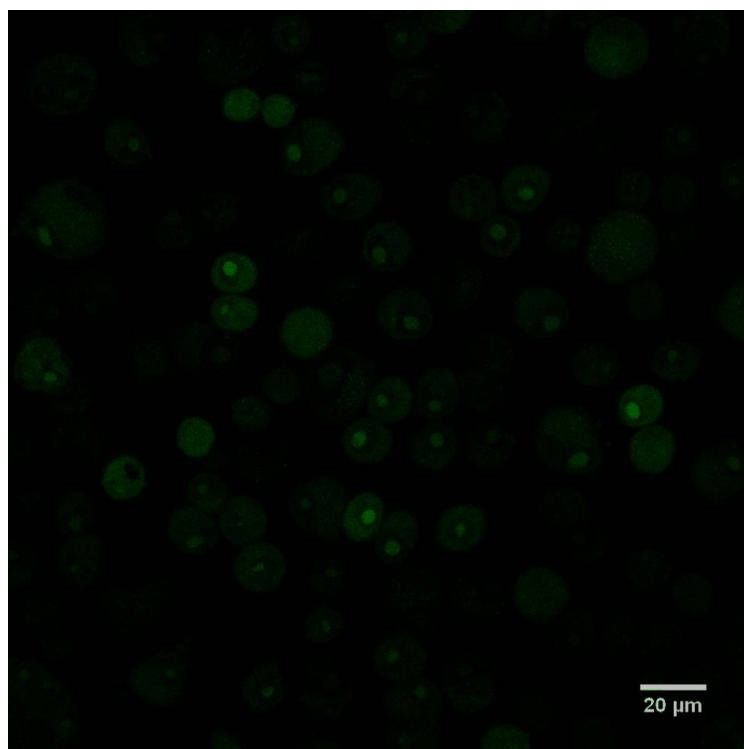


Figure S17 - Confocal microscopy images of HeLa cells treated for 5 min with 10 μM C₈. A considerable accumulation of this compound is seen in a short time. Scale bar is 20 μm .

Molecular Modelling

Experimental

Molecular docking studies were performed with KRAS-22RT NMR structure (PDB code: 5I2V) and most promising ligand C₈ using AutoDockTools package [1]. Both the G4 structure and ligands were first optimized for docking using DockPrep tool of UCSF Chimera software. After assigning empirical atomic partial charges (Gasteiger) and polar hydrogens, docking simulations were carried with AutoDock 4.2 program using Lamarckian genetic algorithm. Adenine residues A1 and A17 were allowed flexibility in order to facilitate the access of the ligand to 5' tetrad. The remaining nucleotide residues were kept rigid. The grid was placed on the upper tetrad of KRAS-22RT G4 based on the NOE contacts and constrained to $115 \times 40 \times 55$ Å in the x, y and z axes, respectively, with a grid spacing of 0.375 Å. Each run consisted of an initial population of 150 random individuals, a maximum number of evaluations set to 2.5×10^7 , the rate of mutation and crossover set to 0.02 and 0.8, respectively, and elitism value of 1. All torsions of the ligands were set as flexible while the structure of the G4 was kept rigid. Each run provided 10 different ligand structures.

The most energetically favorable binding poses and in agreement with the NOESY experiments were energy minimized using GROMACS 2016.2 package [2]. Ligand topology was generated using ACPYPE server which employs ANTECHAMBER module of AMBER [3]. Energy minimization was carried using AMBER99SB force field and TIP3P water model. DNA-ligand complexes were solvated in a 10 Å octahedral box. K⁺ counterions were added to neutralize the system total charge. Energy minimization (steepest descent algorithm) was run for a maximum of 50000 steps until the maximum force < 10.0 kJ/mol. Resulting complexes were analyzed with UCSF Chimera software.

Results and Discussion

Restraint-assisted docking procedures used a single flexible ligand molecule during the conformational search and energy calculation. The grid was placed on the top tetrad and the 5'-cap adenine residues A1 and A17 were set as flexible in order to represent the different conformers found at those adenines in the structure at 37°C (results not shown). All the other residues were maintained rigid. From NMR data, it was observed that both adenine and thymine residues were the most flexible, which is not reflected by using the rigid PDB structure. Therefore, we introduced some mobility to the bases that have less restraints observed in the NOESY spectra. The flexibility of adenines also facilitates access of the ligand to the tetrad surface. The best four conformers obtained by docking, in terms of binding free energy and also in agreement with the NOESY restraints experiments, are shown in Figure S18. Note that the models reflect only the binding of one ligand, despite the binding ratio being indicative of a 4:1 ligand:DNA ratio; we have used only the restraints coming from strong peaks whose assignment was done unambiguously evidencing the principal binding site. The binding free energies of complexes A-D are -9.32, -7.24, -9.17 and -8.40 kcal/mol, respectively, which were among the five lowest energy complexes obtained in the docking experiments with binding free energies ranging between -9.32 and -6.04 kcal/mol. The acridine moiety seems to interact with the 5' tetrad formed by G2-G6-G11-G18 mainly by stacking onto the planar guanine bases. These results are in agreement with the upfield chemical shift of the imino protons observed in ^1H NMR titrations. The position of the acridine moiety may be in part due to the protonated nitrogen atom being placed near the central carbonyl groups of the G-tetrads which have negative electrostatic potential. The iodinated benzene ring is found in two main positions among the different conformers found by docking. In the most abundant family, this ring is inserted inside the four-base propeller loop, occupying a small opening created in this region (Figures S18A, B, D). The other ligand-family of conformers have the iodinated benzene group placed in the groove between G7 and G11 (Figure S18C). Moreover, the

benzene ring is also capable of establishing stacking interactions with the loop bases. The position of the alkylamide side chain is in agreement with the NOE interactions previously discussed and concerning the ligand C8 seems to be of optimal size to place the acridine moiety in direct contact with the tetrad and the alkylamide chain as much as possible excluded from water molecules (Figure S19).

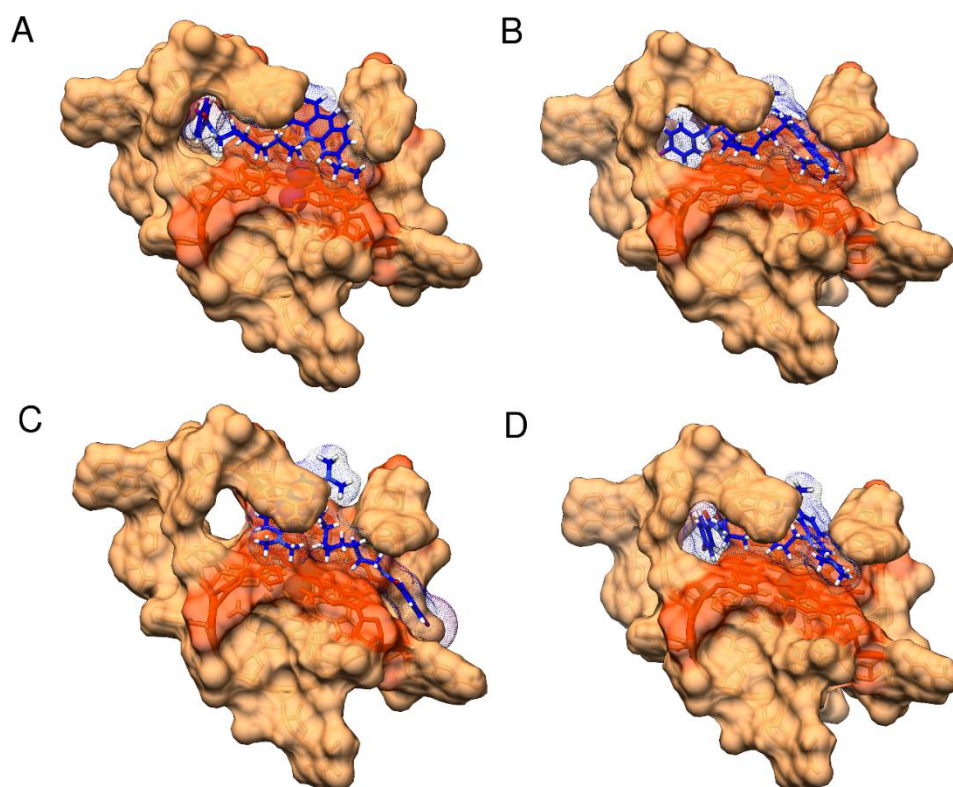


Figure S18 - Models based on lowest energy of complex C₈/KRAS22-RT produced by docking after energy minimization. All the binding poses show the ligand interacting with top 5' G-tetrad via stacking of the acridine moiety with the guanine residues within the tetrad. Panels A-D illustrate four best conformers obtained with different binding free energies, -9.32, -7.24, -9.17 and -8.40 kcal/mol, respectively. G4 structure is depicted as a light orange surface with the 5' tetrad highlighted in darker orange while the ligand is represented in blue.

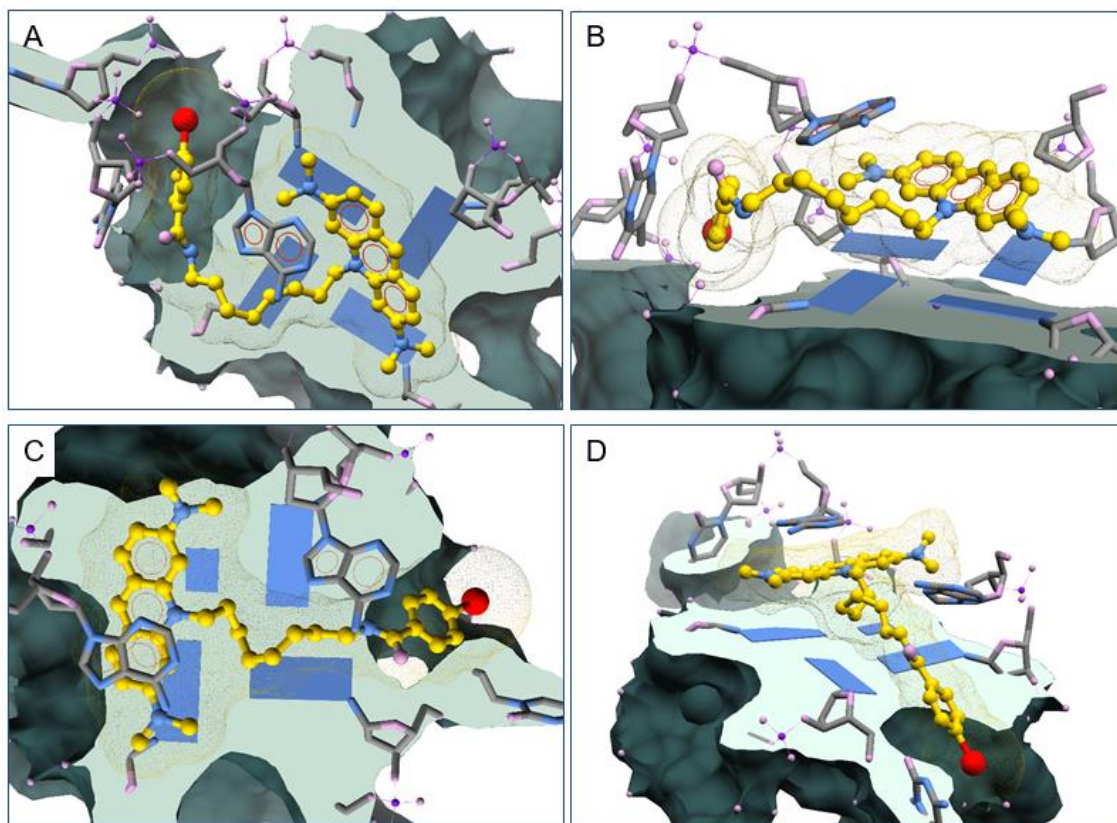


Figure S19 - Expansion of ligand C₈ bound to KRAS-22RT. Typical conformer found from restrain-assisted docking experiments highlighting iodobenzene moiety inserted in loop cavity top and side views, panels A and B respectively. Panels C and D depict top and side view of ligand iodobenzene moiety inserted in groove. Iodine is represented by small red sphere.

References for Supporting Information

- [1] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility., *J. Comput. Chem.* 30 (2009) 2785–91. doi:10.1002/jcc.21256.
- [2] S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M.R. Shirts, J.C. Smith, P.M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit, *Bioinformatics*. 29 (2013) 845–854. doi:10.1093/bioinformatics/btt055.
- [3] A.W. Sousa da Silva, W.F. Vranken, ACPYPE - AnteChamber PYthon Parser interface, *BMC Res. Notes*. 5 (2012) 367. doi:10.1186/1756-0500-5-367.

Appendix D

Supporting Information for Chapter 7

Table S1 - Ligand-induced thermal stabilization (ΔT_m) measured by FRET melting experiments at 0.2 μ M ligand concentration

	ΔT_m (°C) ^a at 0.2 μ M ligand concentration						
Ligand	FHPV9T	FHPV16T	FHPV18T	FHPV32T	FHPV52T	FHPV57T	FHPV58T
PhenDC ₃	14.2 ± 0.6	10.3 ± 0.1	13.7 ± 2.0	9.1 ± 0.2	21.2 ± 1.1	12.5 ± 0.9	> 30
[16]phenN ₂	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
BRACO-19	2.0 ± 0.7	2.0 ± 0.5	1.7 ± 0.9	10.5 ± 0.5	28.4 ± 0.4	1.9 ± 0.6	17.5 ± 0.9
C ₈	0.9 ± 0.4	1.0 ± 1.2	1.5 ± 0.2	12.5 ± 2.6	6.5 ± 1.2	0.5 ± 0.7	13.1 ± 5.6
PDS	1.5 ± 0.3	2.5 ± 0.6	1.4 ± 0.1	0.0 ± 0.0	6.0 ± 0.1	0.0 ± 0.0	11.8 ± 0.7
360A	10.0 ± 1.5	1.9 ± 0.3	1.1 ± 0.2	0.0 ± 0.0	17.0 ± 0.6	0.3 ± 0.2	21.1 ± 1.3
TMPyP ₄	19.0 ± 1.1	7.9 ± 0.8	2.4 ± 1.3	9.1 ± 2.3	23.1 ± 0.8	16.3 ± 1.1	> 30

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_m$ (DNA + ligand) – T_m (DNA)]. The buffer used was 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl. The T_m values for the HPV DNAs are 52.9 ± 0.5 °C for FHPV9T, 52.8 ± 0.2 °C for FHPV16T, 39.0 ± 0.4 °C for FHPV18T, 53.7 ± 0.9 °C for FHPV32T, 49.9 ± 0.5 °C for FHPV52T, 52.3 ± 0.5 °C for FHPV57T, 50.3 ± 0.4 °C for FHPV58T. All experiments were done a minimum of three times, and the values reported correspond to the average of three measurements with the estimated standard deviation.

Table S2 - Ligand-induced structural changes observed in CD titration experiments.

		Ligand						
HPV sequence	Reference ^a	PhenDC ₃	[16]phenN ₂	BRACO-19	C ₈	PDS	360A	TMPyP ₄
HPV9	Parallel	Antiparallel	Parallel	Parallel	Hybrid	Parallel	Antiparallel	Hybrid
HPV16	Mixture ^b	Antiparallel	Mixture	Mixture	Hybrid	Mixture	Antiparallel	Mixture
HPV18	Parallel	Parallel	Parallel	Parallel	Parallel	Parallel	Parallel	Parallel
HPV32	Parallel	Hybrid	Parallel	Parallel	Hybrid	Parallel	Hybrid	Parallel
HPV52	Hybrid	Hybrid	Hybrid	Hybrid	Antiparallel	Parallel	Antiparallel	Parallel
HPV57	Parallel	Antiparallel	Parallel	Parallel	Antiparallel	Hybrid	Antiparallel	Hybrid
HPV58	Hybrid	Hybrid	Hybrid	Hybrid	Hybrid	Hybrid	Parallel	Mixture

^a Three distinct G4 topologies were observed in CD experiments with the following characteristic bands: i) parallel topology with a positive band at ≈264 nm and negative at ≈245 nm; ii) antiparallel topology with positive bands at ≈295 and ≈240 nm, and a negative band at ≈260 nm; iii) hybrid-type (3+1) topologies with a positive band at ≈290 nm, negative band at ≈245 nm and a shoulder peak at ≈260 nm.

^b No distinct topology CD characteristic signals were observed.

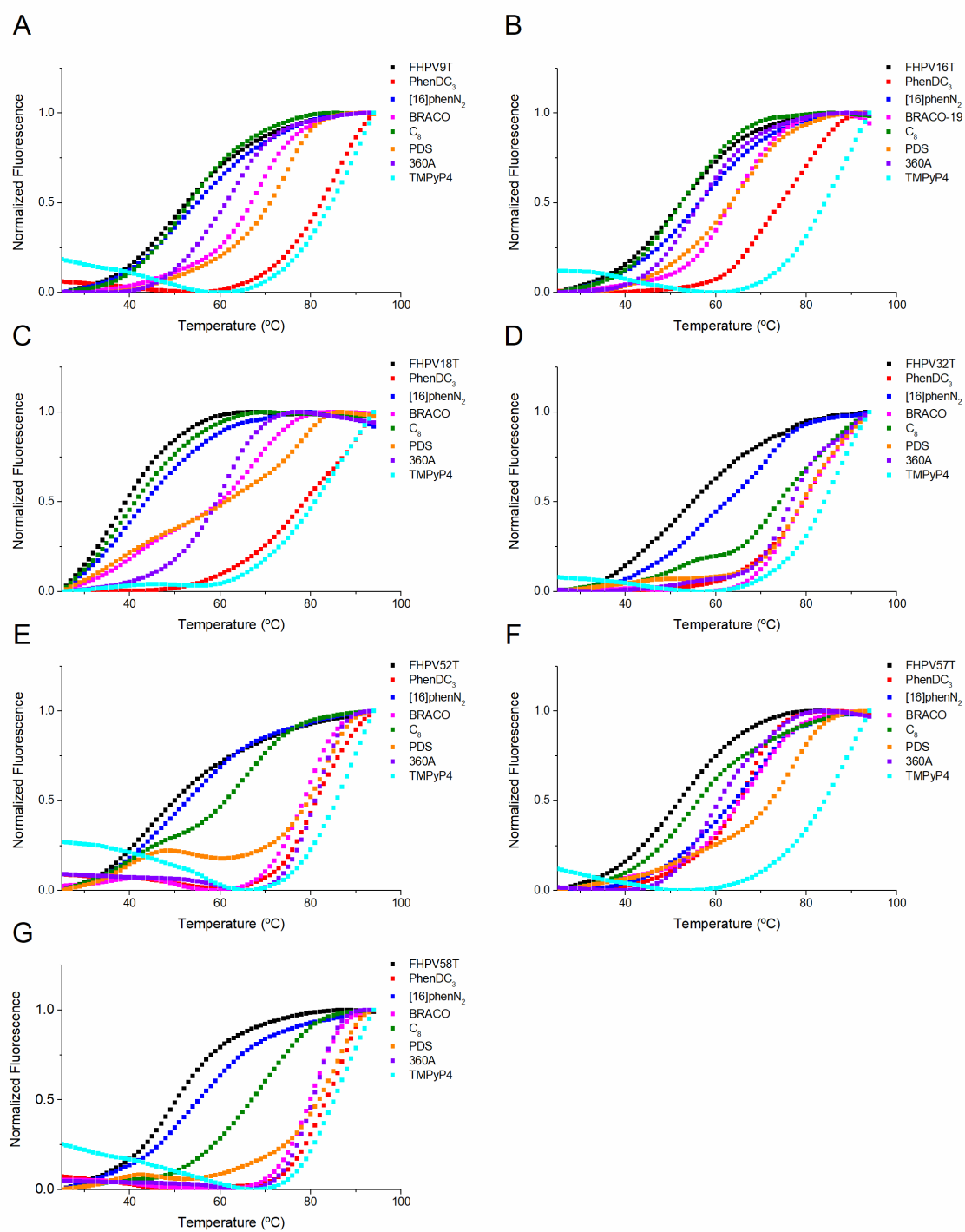


Figure S1 - FRET melting curves in the absence and presence of 5 molar eq. of ligand, obtained in 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl. Determined ΔT_m values are listed in Table 2.

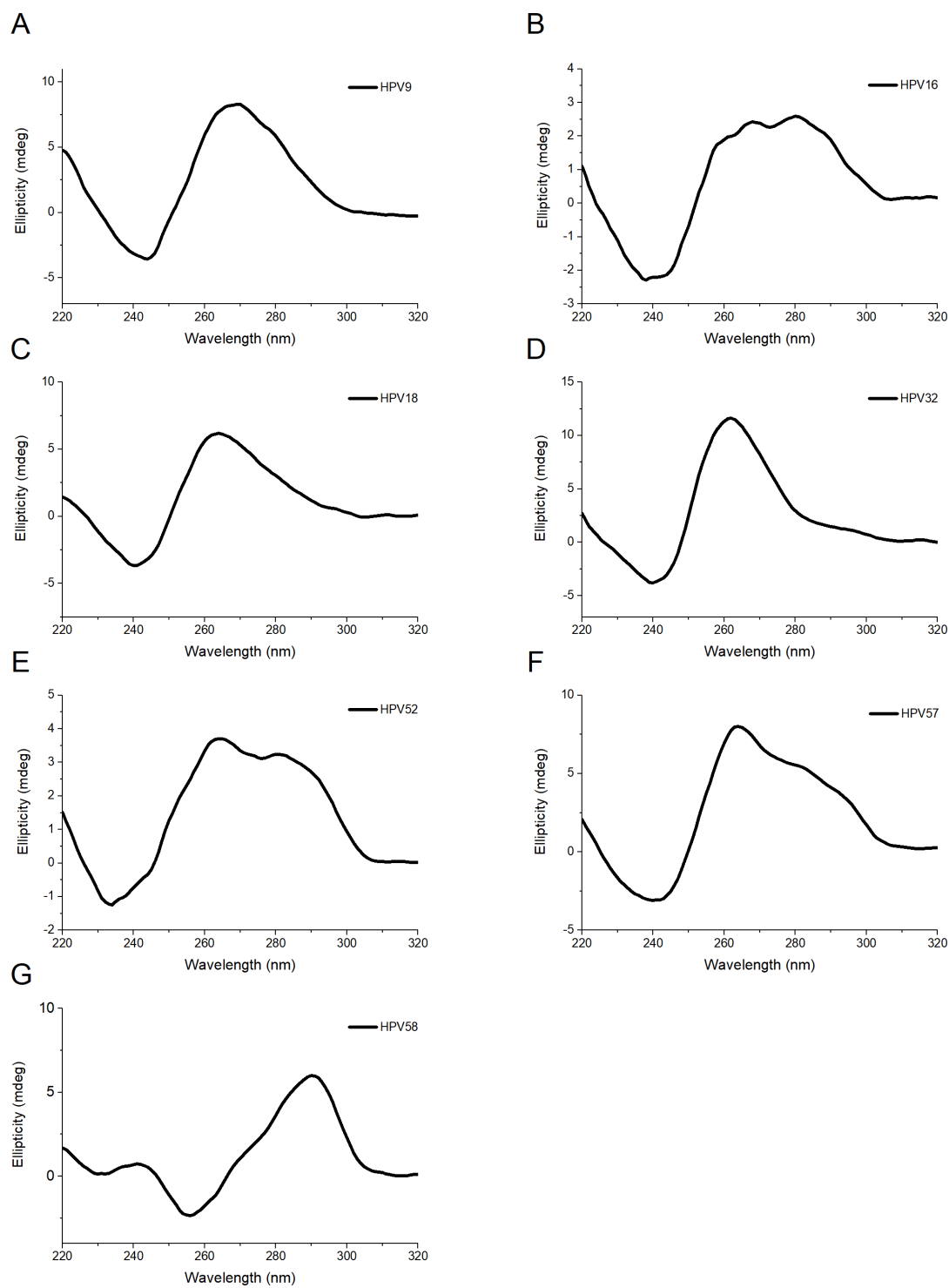


Figure S2 - CD spectra of HPV G4s in buffer mimicking potassium physiological conditions.

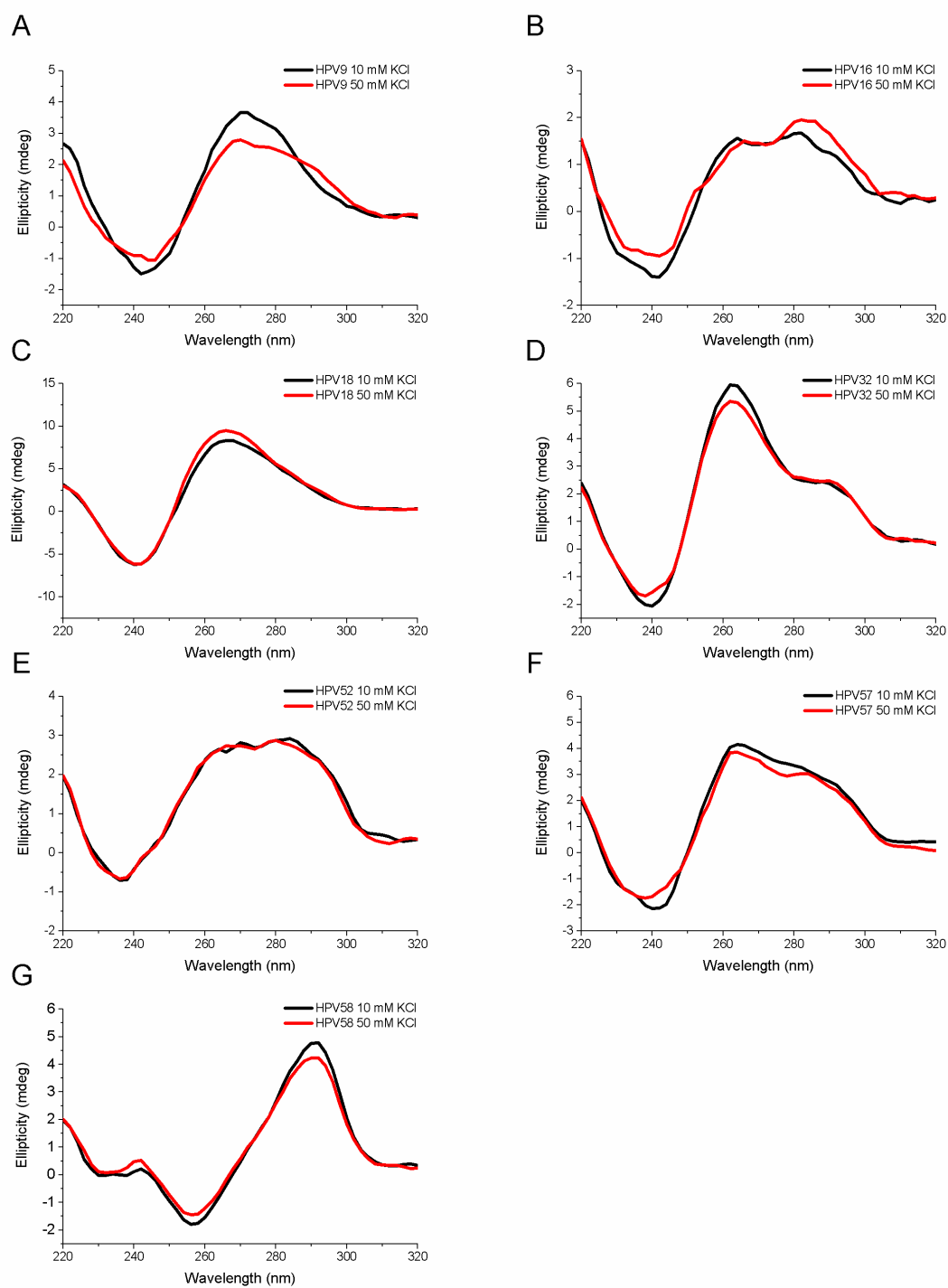


Figure S3 - CD spectra of HPV G4s in buffer containing either 10 or 50 mM KCl concentration.

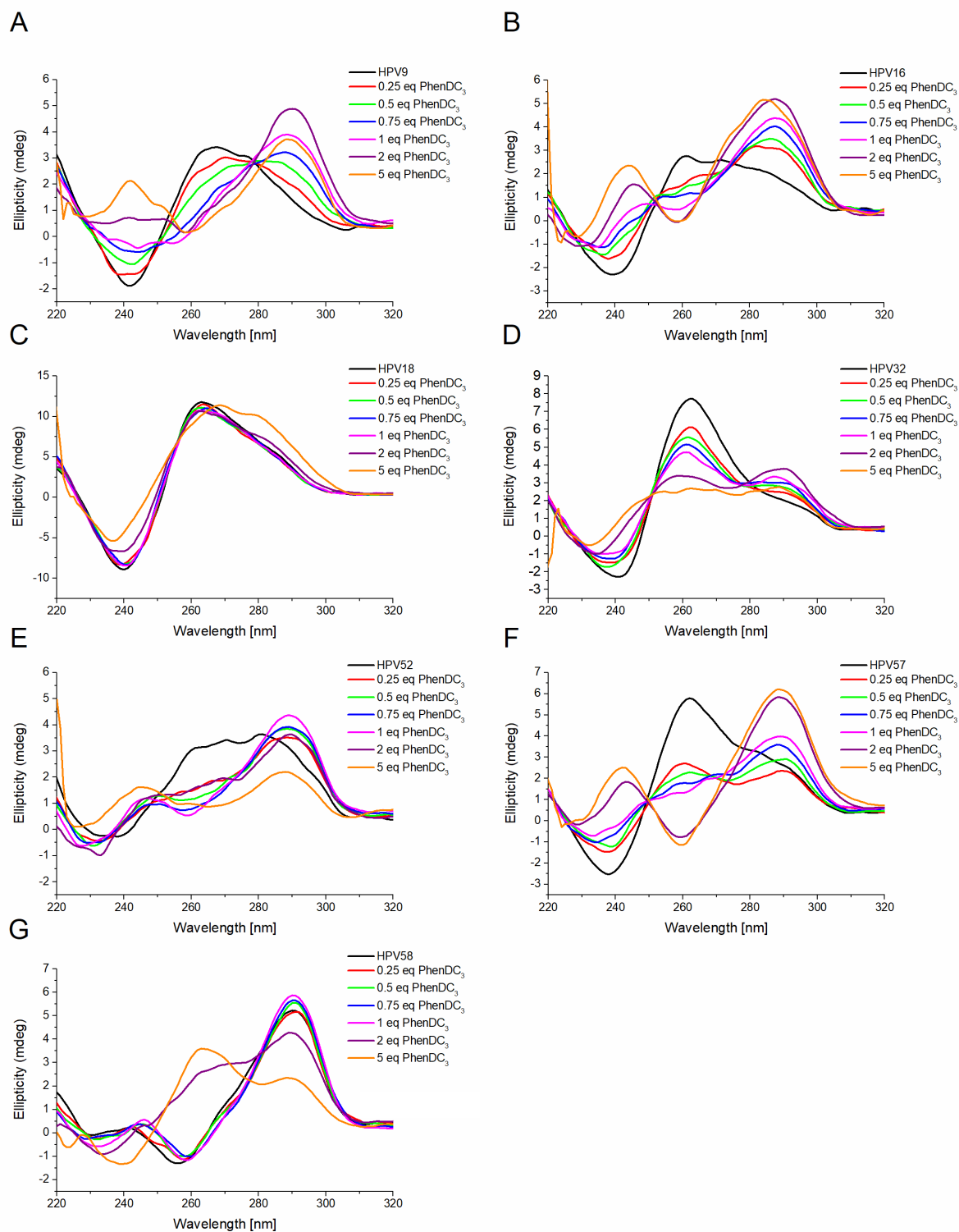


Figure S4 - CD spectra of HPV G4s in the presence and absence of ligand PhendC₃. The buffer used was 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl.

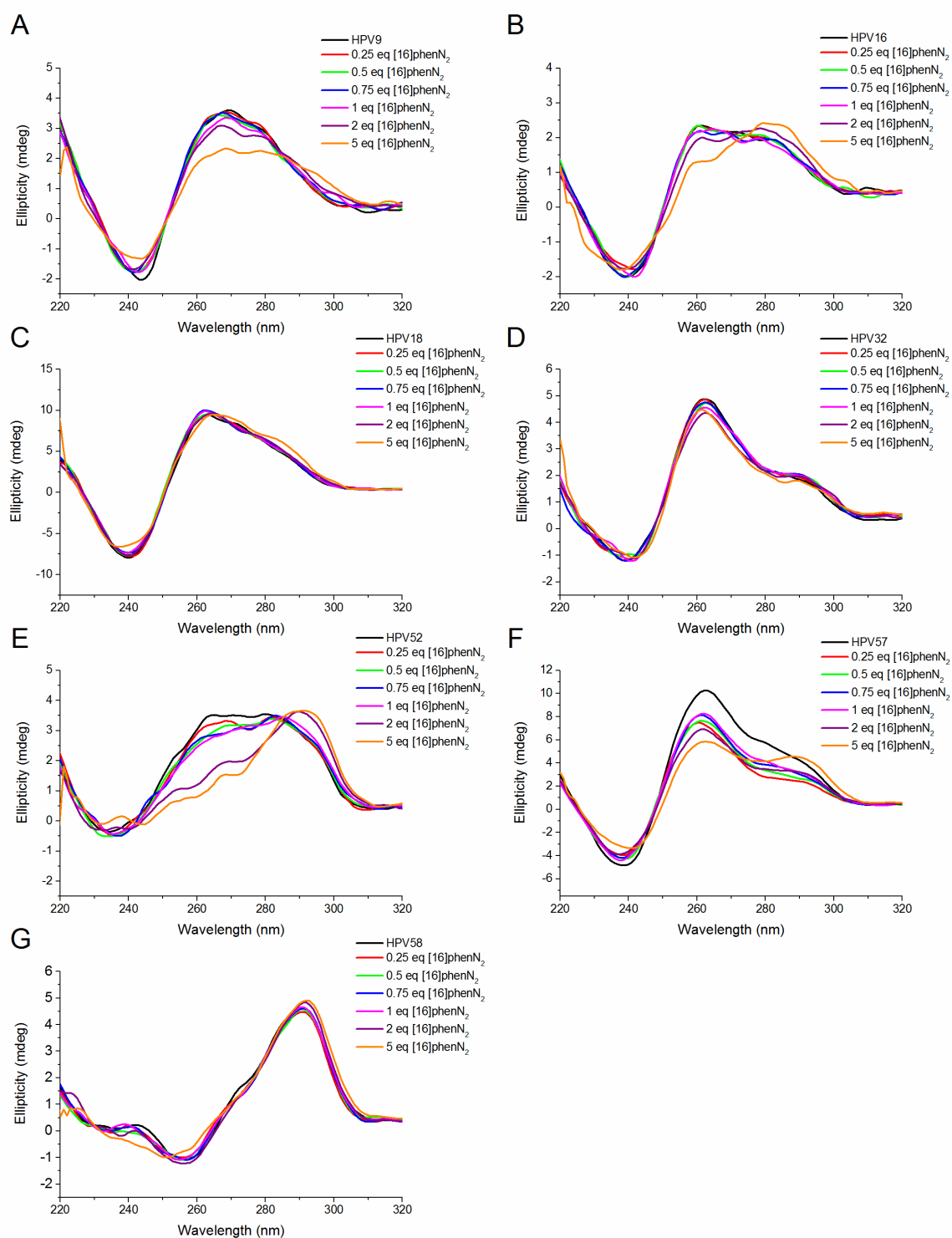


Figure S5 - CD spectra of HPV G4s in the presence and absence of ligand [16]phenN₂. The buffer used was 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl.

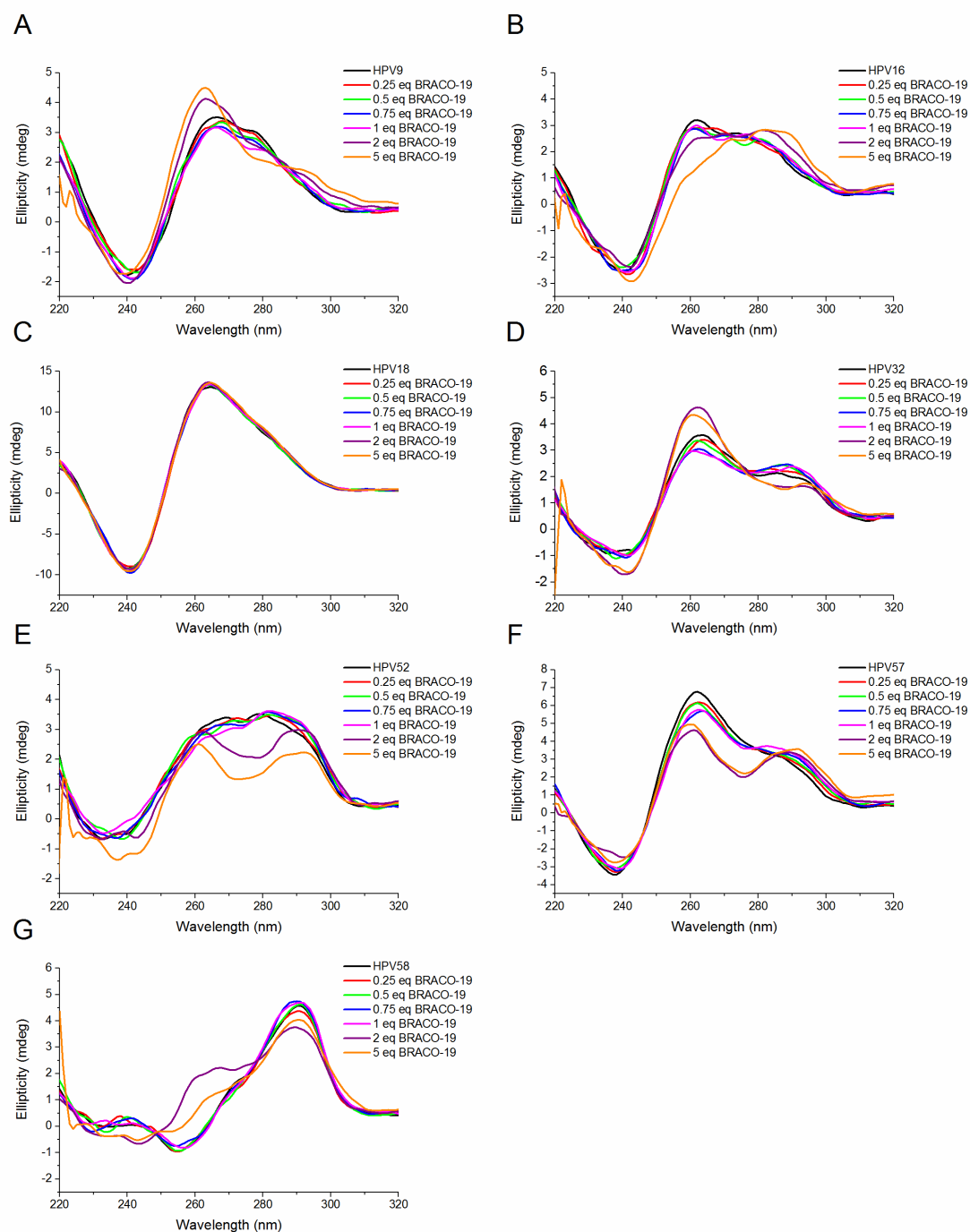


Figure S6 - CD spectra of HPV G4s in the presence and absence of ligand BRACO-19. The buffer used was 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl.

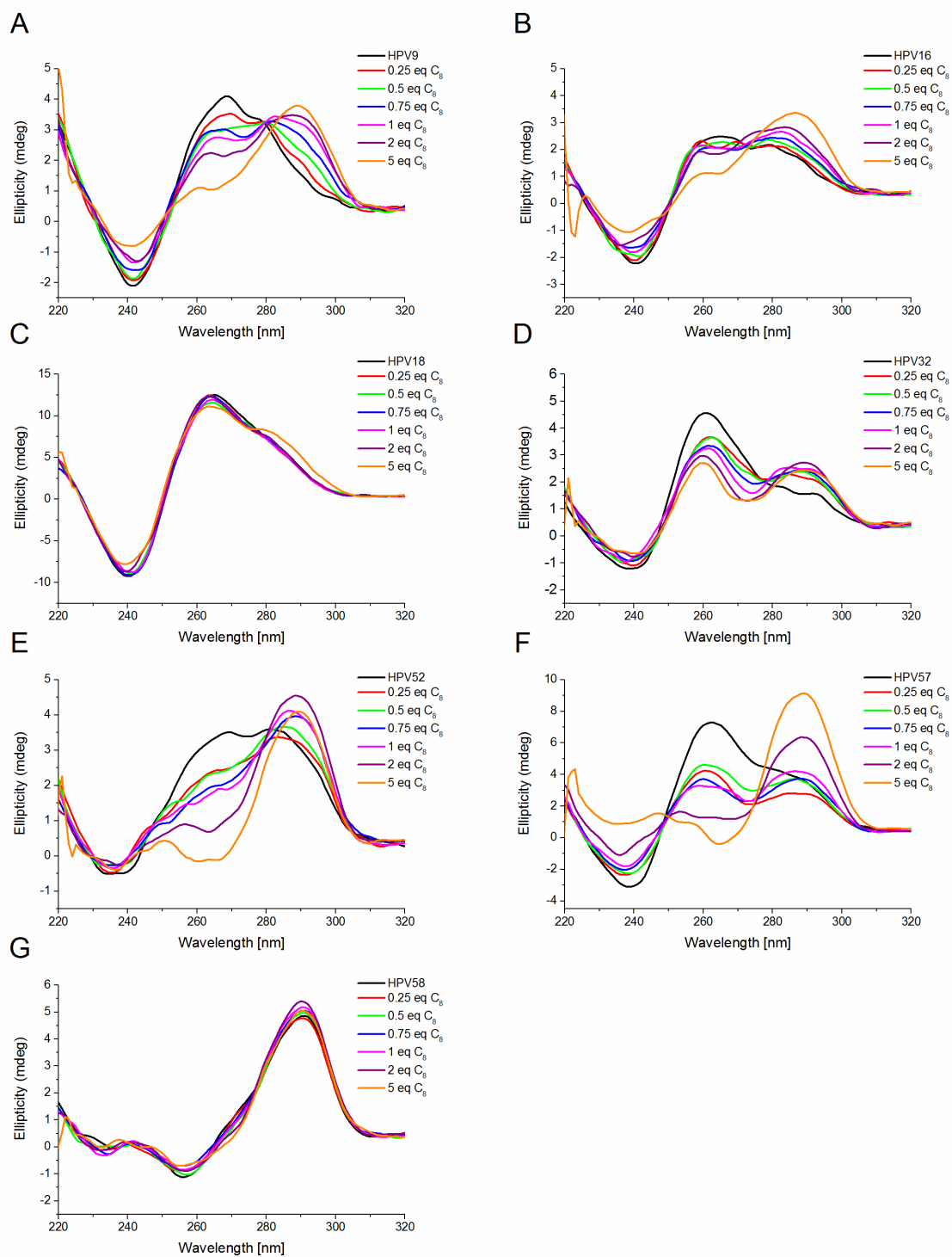


Figure S7 - CD spectra of HPV G4s in the presence and absence of ligand C₈. The buffer used was 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl.

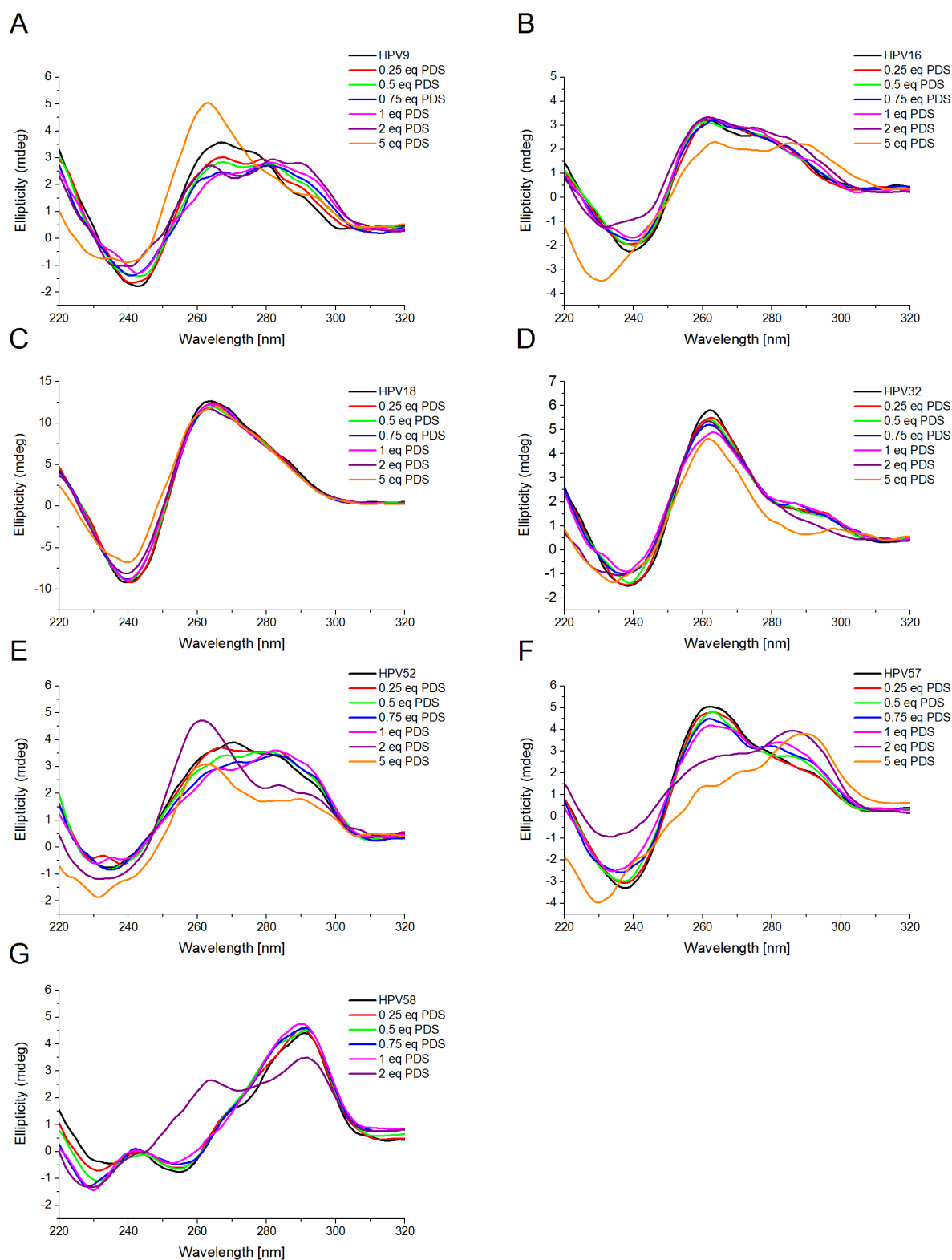


Figure S8 - CD spectra of HPV G4s in the presence and absence of ligand PDS. The buffer used was 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl.

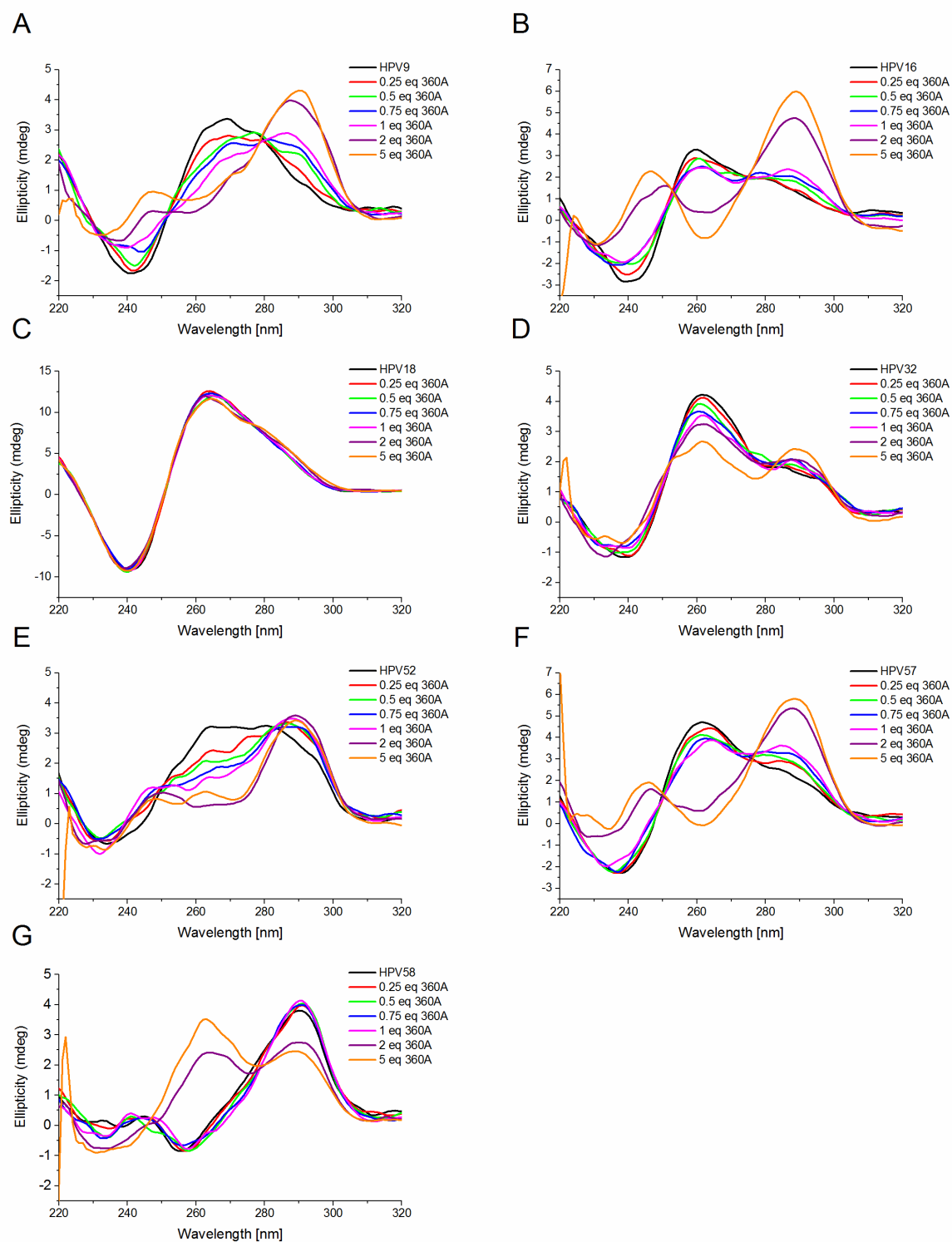


Figure S9 - CD spectra of HPV G4s in the presence and absence of ligand 360A. The buffer used was 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl.

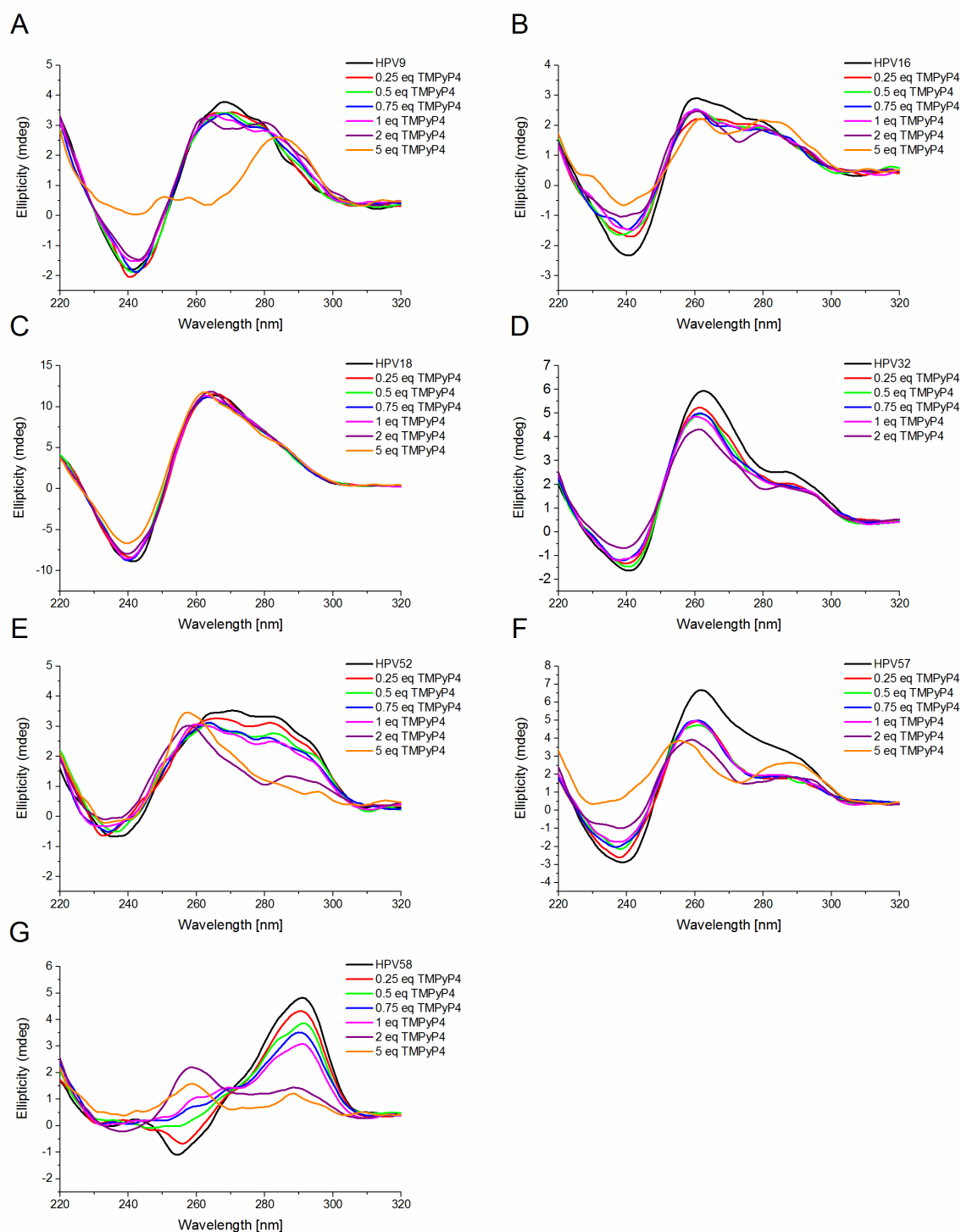


Figure S10 - CD spectra of HPV G4s in the presence and absence of ligand TMPyP4. The buffer used was 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl.

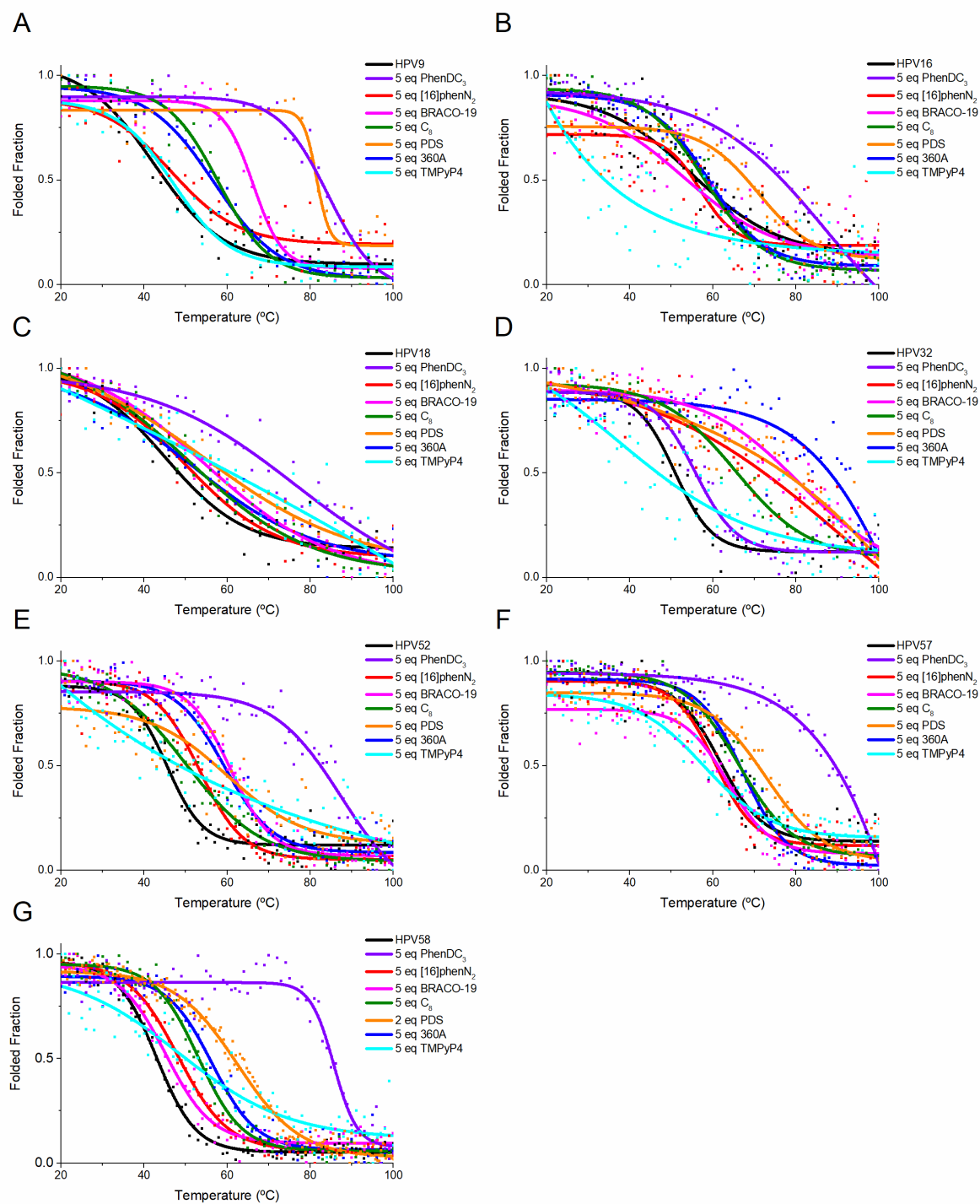


Figure S11 - CD melting curves of HPV G4s in the absence and presence of 5 molar eq. of ligand, obtained in 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl. Determined ΔT_m values are listed in Table 3.

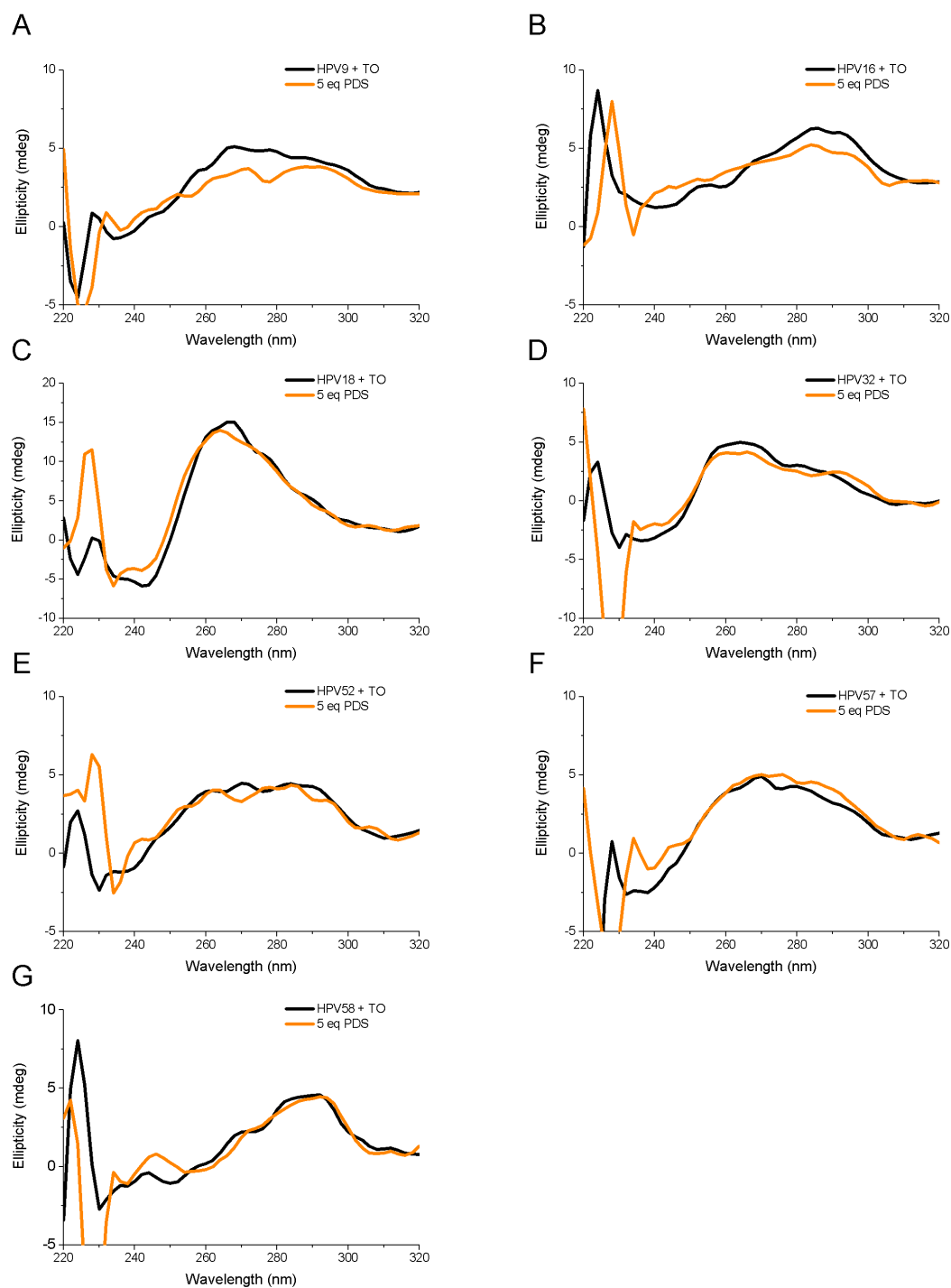


Figure S12 - CD spectra of selected HPV G4s incubated with 2 molar eq. of TO probe, in the presence and absence of PDS. The buffer used was 10 mM K-phosphate buffer, pH 7.2, with 50 mM KCl.

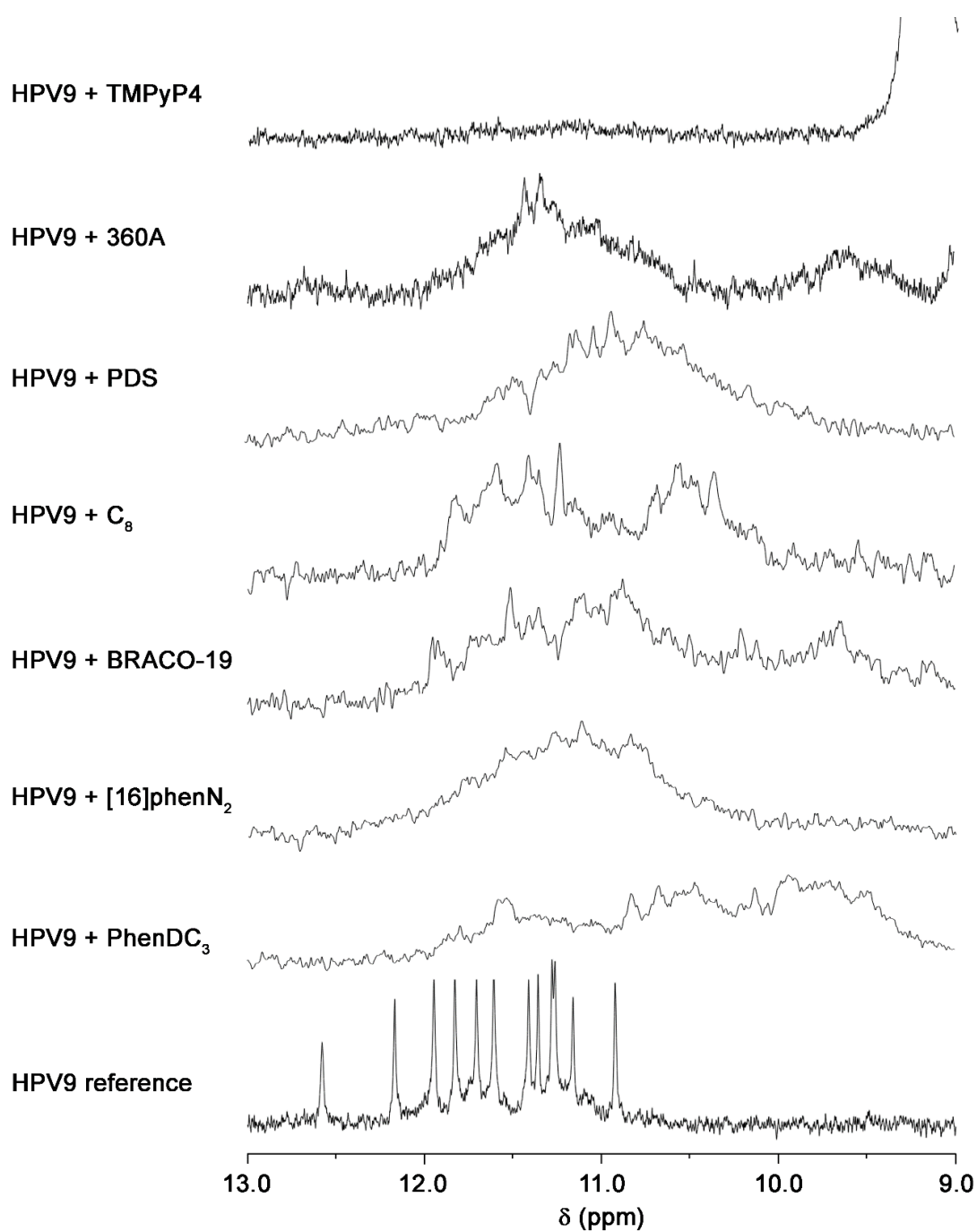


Figure S13 - ¹H NMR spectra (imino region) of HPV9 G4 in the absence and presence of the tested ligands. Ligands were titrated in the NMR tube at 0.5, 1, 2 and 4 molar equivalents. Titrations were stopped once no imino signals were observed or the chemical shifts remained unchanged. The buffer used was 10 mM K-phosphate buffer, pH 7.2, with 50 mM KCl.

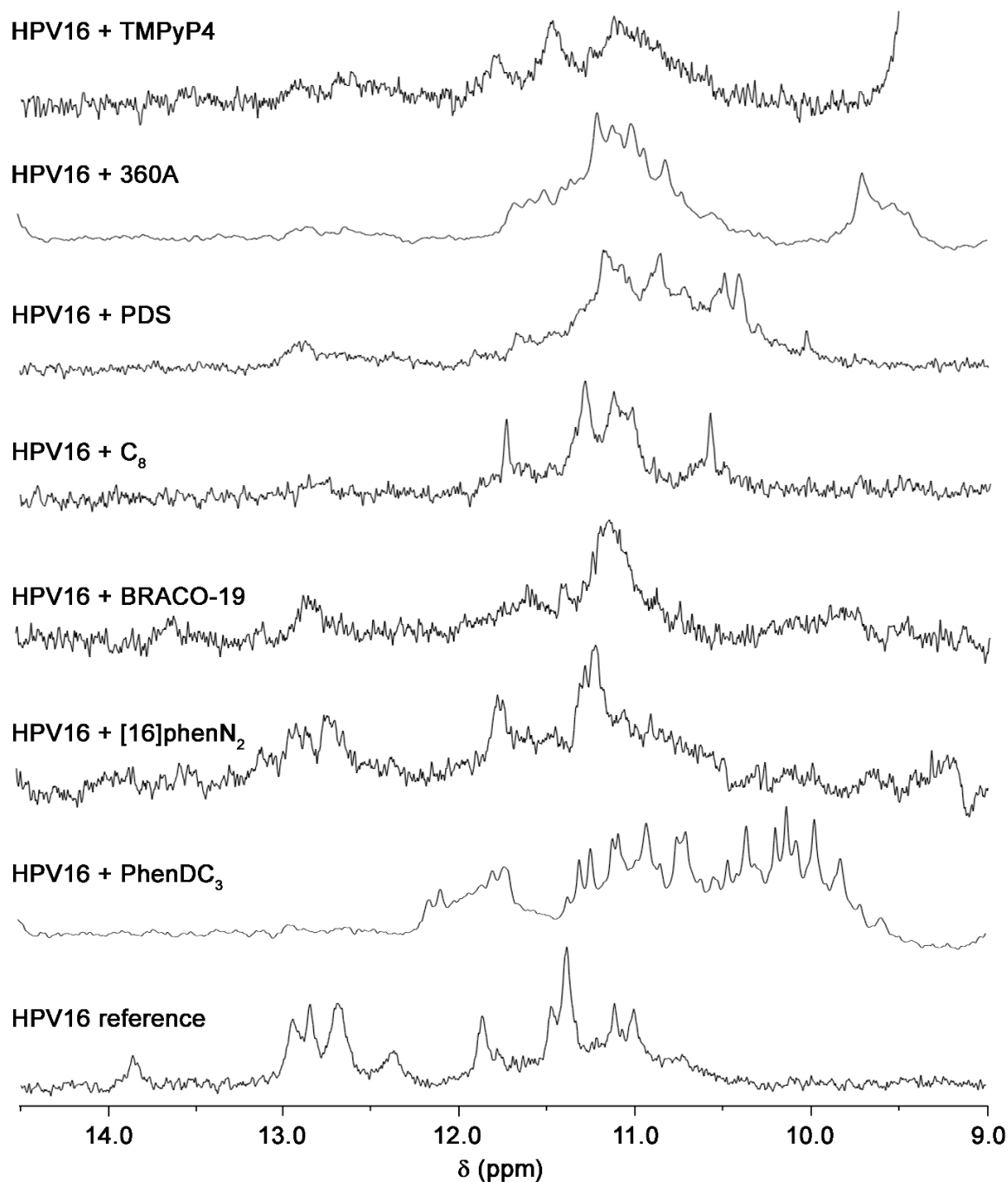


Figure S14 - ¹H NMR spectra (Watson-Crick H-bonds and imino regions) of HPV16 G4 in the absence and presence of the tested ligands. Ligands were titrated in the NMR tube at 0.5, 1, 2 and 4 molar equivalents. Titrations were stopped once no imino signals were observed or the chemical shifts remained unchanged. The buffer used was 10 mM K-phosphate buffer, pH 7.2, with 50 mM KCl.

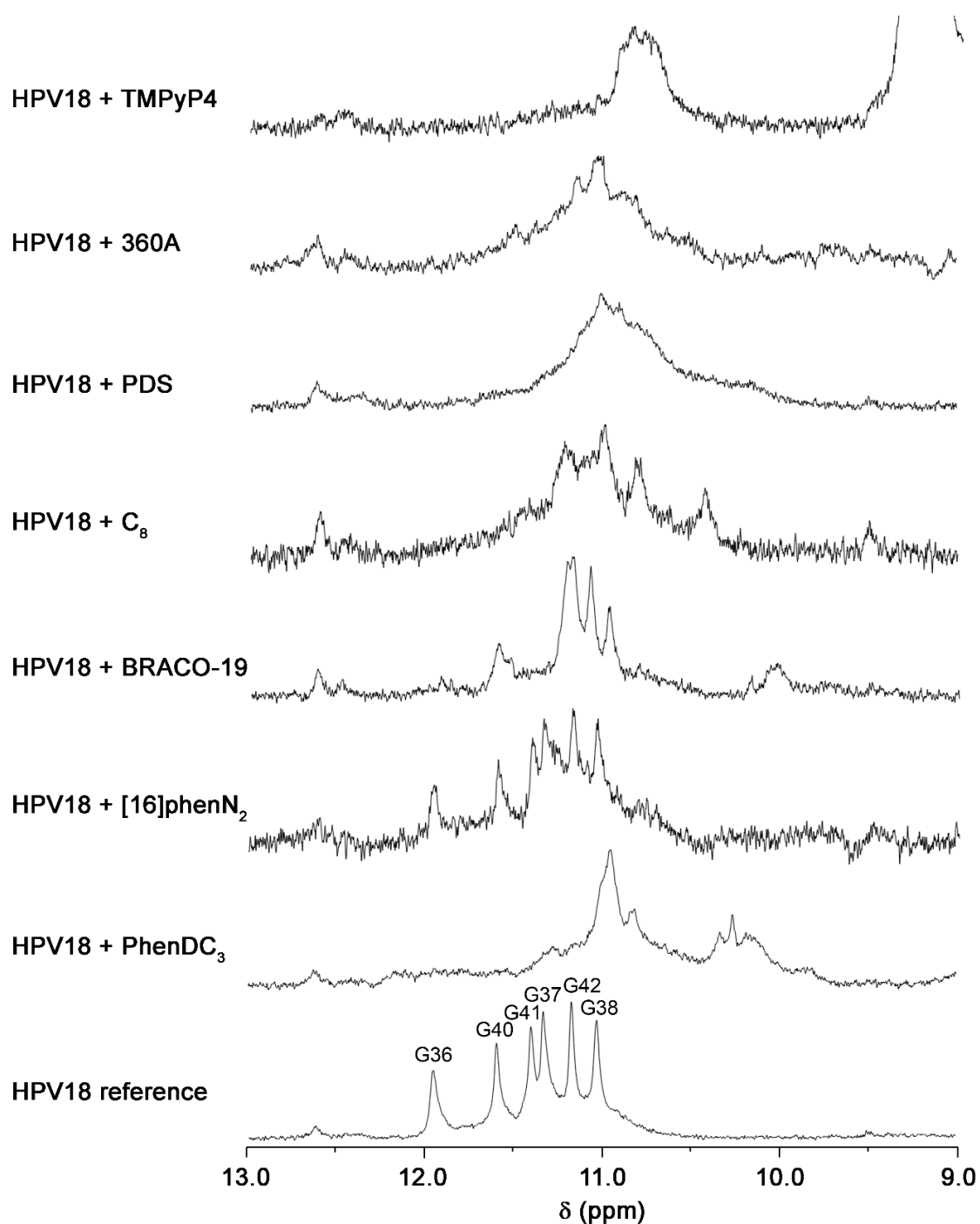


Figure S15 - ^1H NMR spectra (imino regions) of HPV18 G4 in the absence and presence of the tested ligands. Ligands were titrated in the NMR tube at 0.5, 1, 2 and 4 molar equivalents. Titrations were stopped once no imino signals were observed or the chemical shifts remained unchanged. Guanines involved in G4 formation are assigned at the top. The buffer used was 10 mM K-phosphate buffer, pH 7.2, with 50 mM KCl.

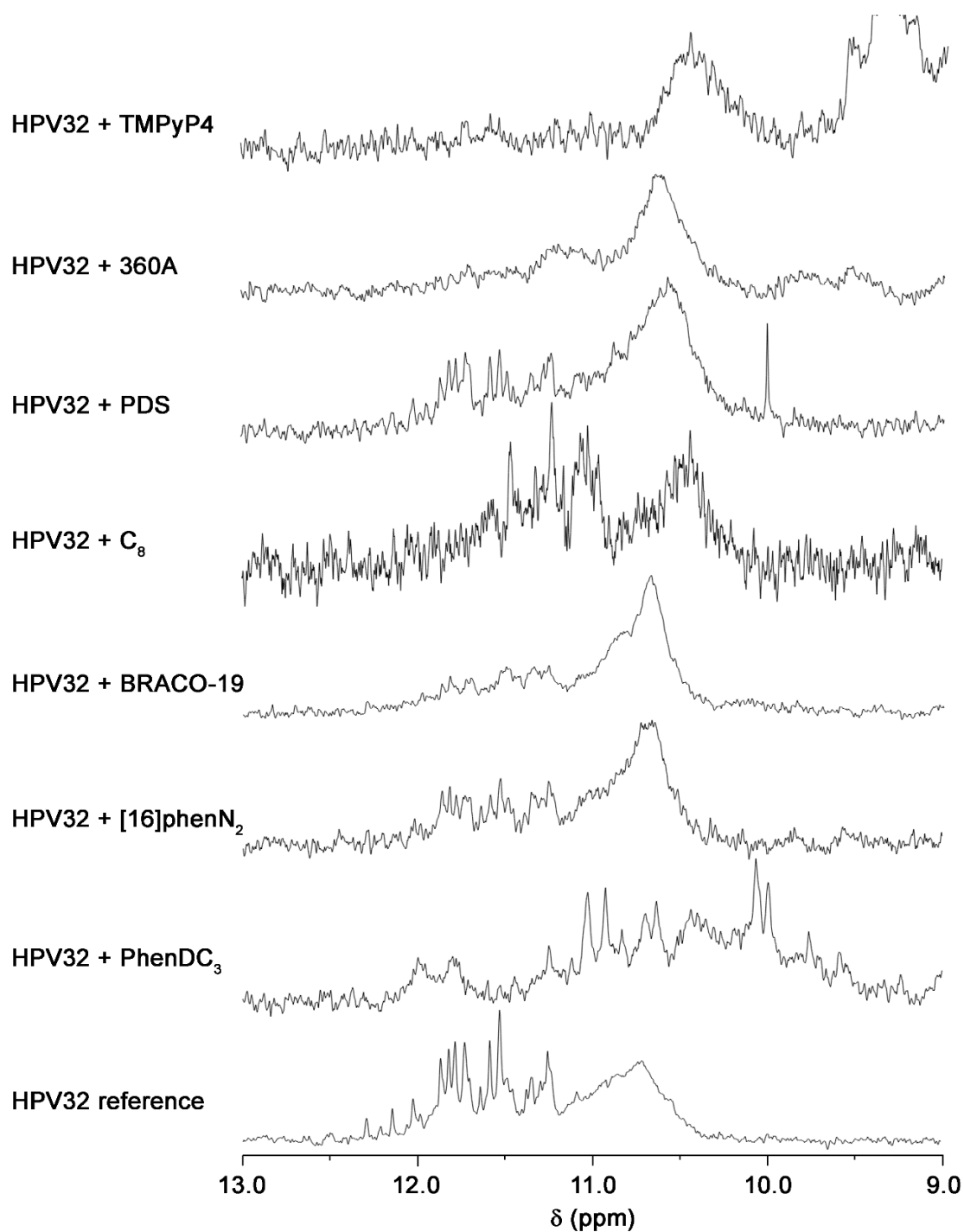


Figure S16 - ^1H NMR spectra (imino regions) of HPV32 G4 in the absence and presence of the tested ligands. Ligands were titrated in the NMR tube at 0.5, 1, 2 and 4 molar equivalents. Titrations were stopped once no imino signals were observed or the chemical shifts remained unchanged. A broad signal indicative of aggregation can be observed between 10.5 and 11 ppm. The buffer used was 10 mM K-phosphate buffer, pH 7.2, with 50 mM KCl.

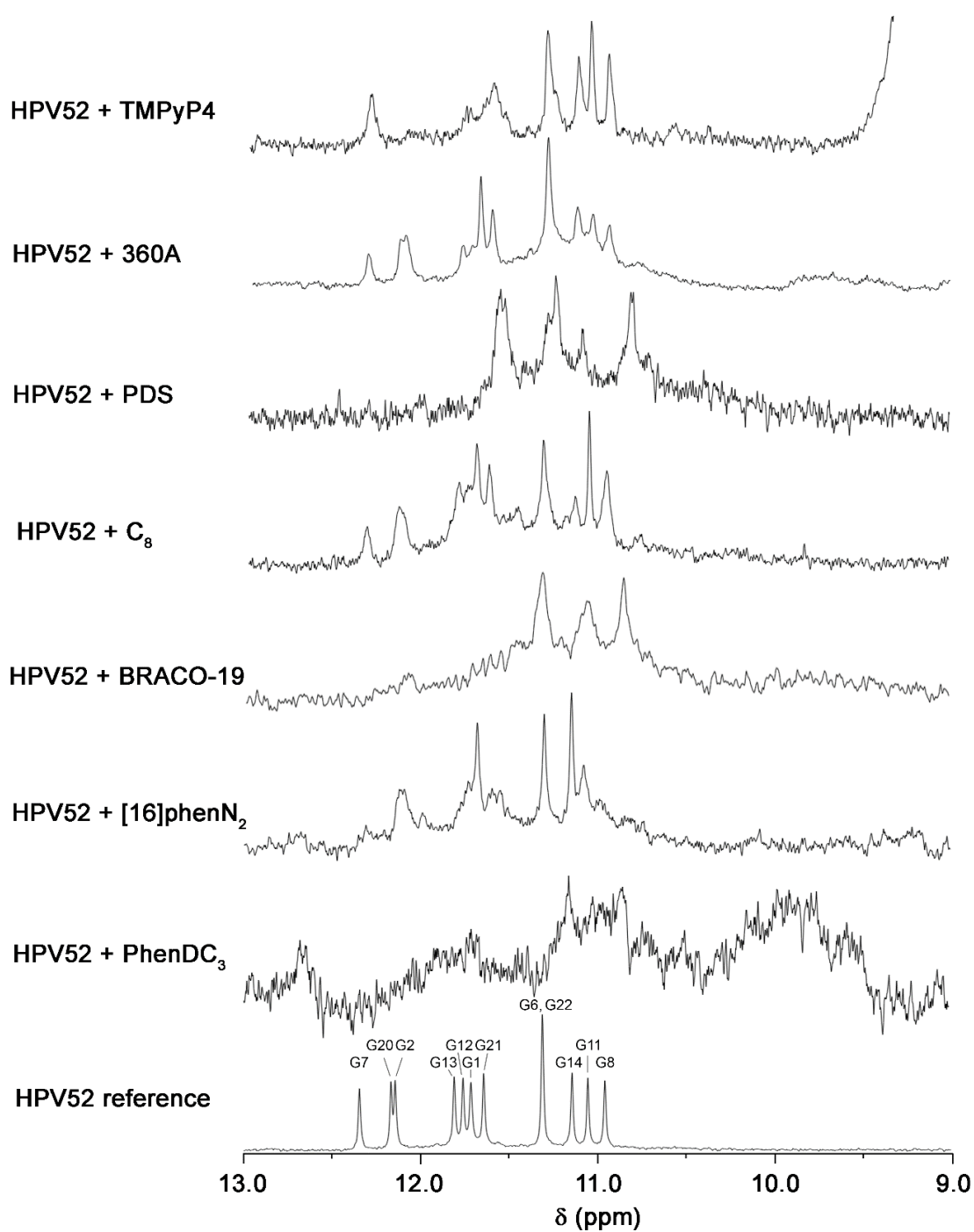


Figure S17 - ¹H NMR spectra (imino regions) of HPV52 G4 in the absence and presence of the tested ligands. Ligands were titrated in the NMR tube at 0.5, 1, 2 and 4 molar equivalents. Titrations were stopped once no imino signals were observed or the chemical shifts remained unchanged. Guanines involved in G4 formation are assigned at the top. The buffer used was 10 mM K-phosphate buffer, pH 7.2, with 50 mM KCl.

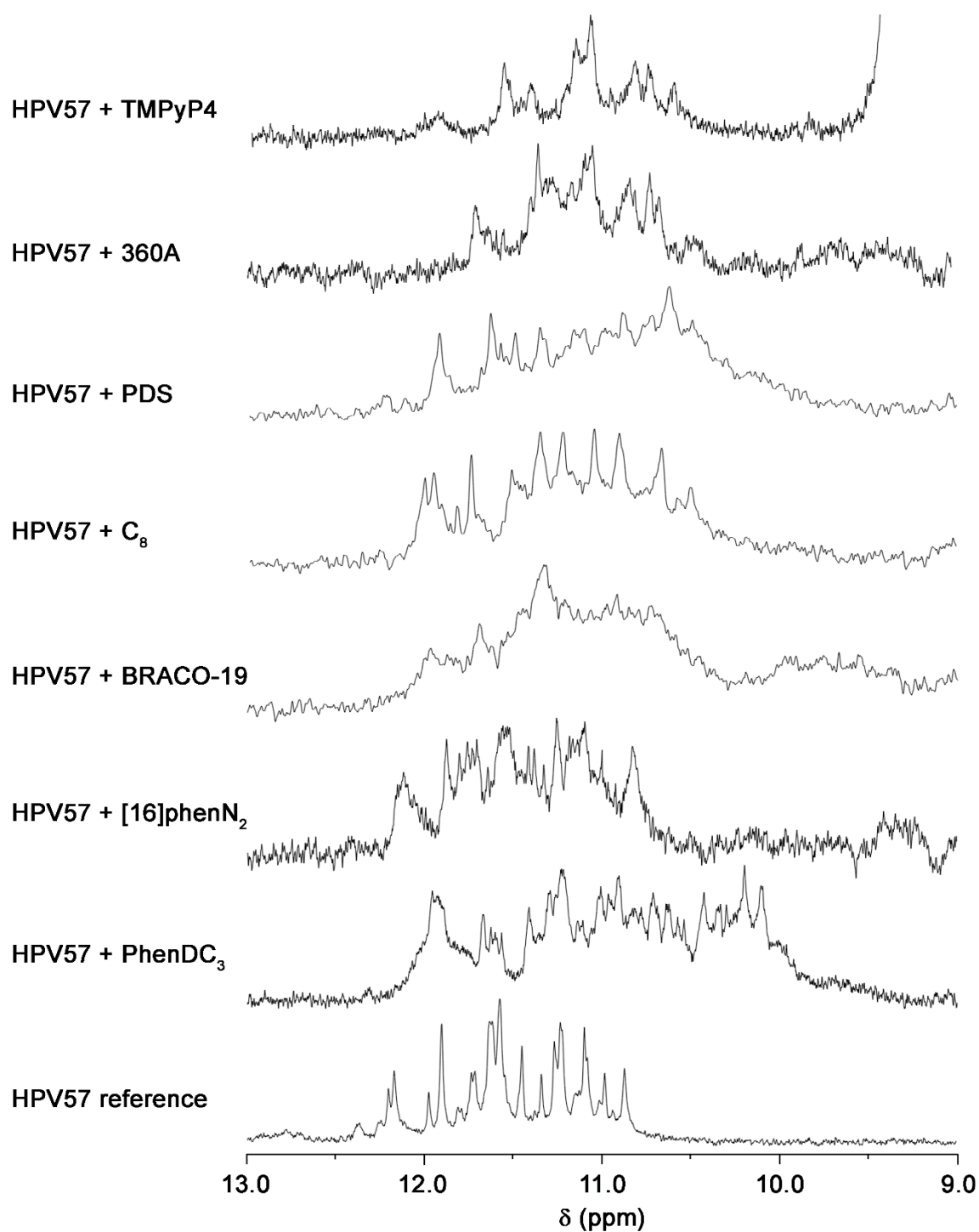


Figure S18 - ^1H NMR spectra (imino regions) of HPV57 G4 in the absence and presence of the tested ligands. Ligands were titrated in the NMR tube at 0.5, 1, 2 and 4 molar equivalents. Titrations were stopped once no imino signals were observed or the chemical shifts remained unchanged. The buffer used was 10 mM K-phosphate buffer, pH 7.2, with 50 mM KCl.

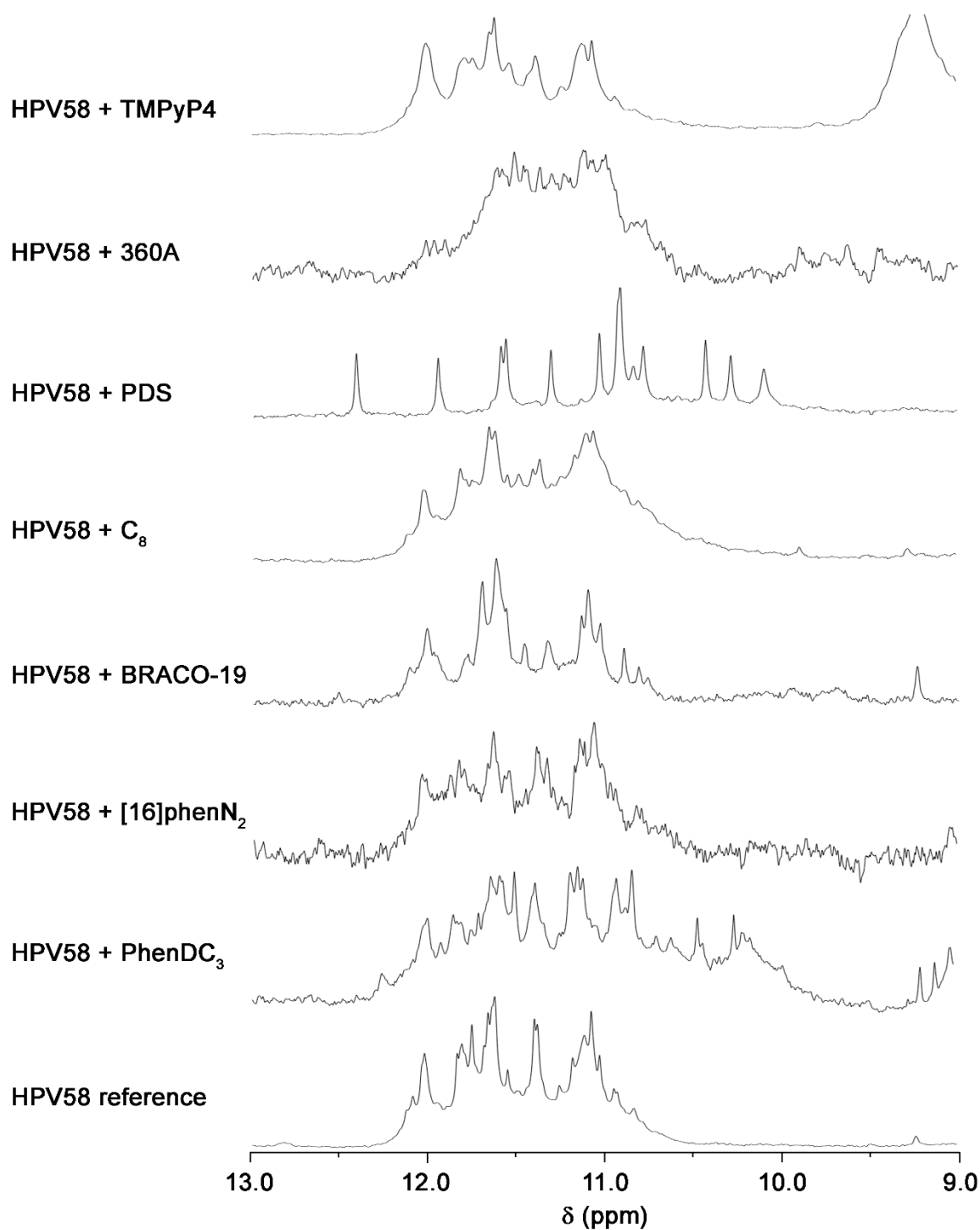


Figure S19 - ^1H NMR spectra (imino regions) of HPV58 G4 in the absence and presence of the tested ligands. Ligands were titrated in the NMR tube at 0.5, 1, 2 and 4 molar equivalents. Titrations were stopped once no imino signals were observed or the chemical shifts remained unchanged. The buffer used was 10 mM K-phosphate buffer, pH 7.2, with 50 mM KCl.

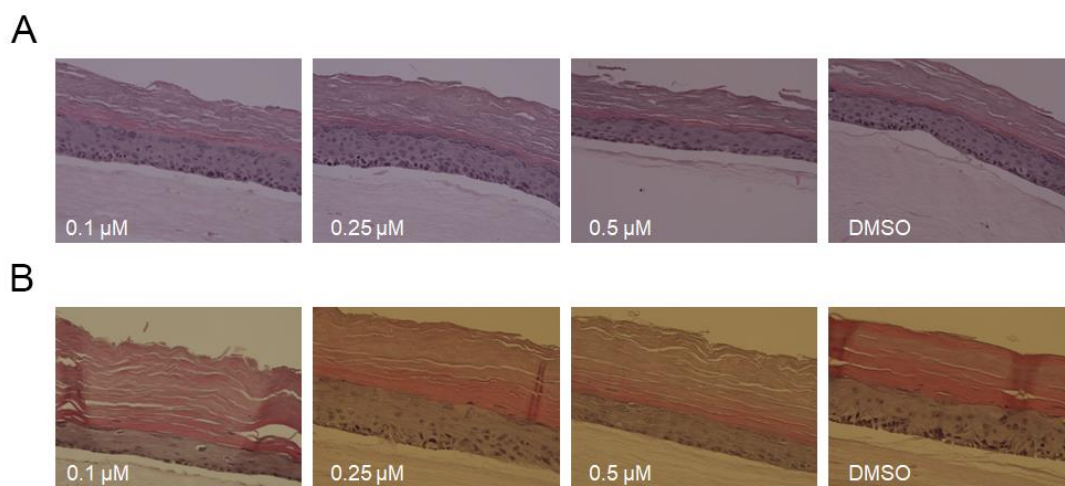


Figure S20 - H&E staining and histological analysis of the effect of PhenDC3 treatment on (A) HCK16-5 and (B) HCK18C raft culture tissues. Raft cultures treated with different concentrations of PhenDC3 (0.1, 0.25 and 0.5 μ M) were grown for 20 days, harvested, fixed in 10% neutral buffered formalin, embedded in paraffin, cut into sections and stained with H&E. DMSO was used as solvent control.

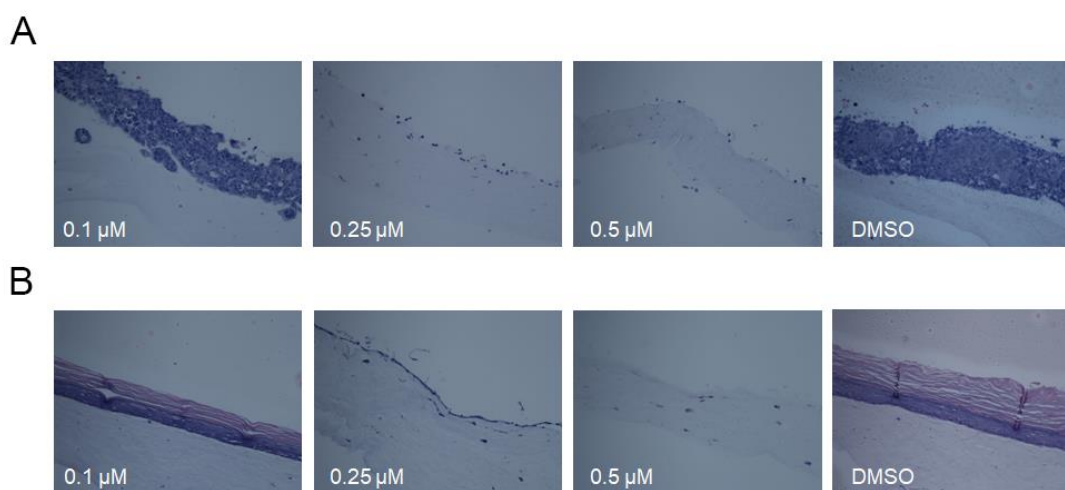


Figure S21 - H&E staining and histological analysis of the effect of C₈ treatment on (A) HCK16-5 and (B) HCK18C raft culture tissues. Raft cultures treated with different concentrations of C₈ (0.1, 0.25 and 0.5 μ M) were grown for 20 days, harvested, fixed in 10% neutral buffered formalin, embedded in paraffin, cut into sections and stained with H&E. DMSO was used as solvent control.

Appendix E

Supporting Information for Chapter 8

Table S1 - Ligand-induced thermal stabilization (ΔT_m) measured by CD melting experiments.

	ΔT_m (°C) ^a			
Aptamer	0.5 eq C ₈	1 eq C ₈	2 eq C ₈	4 eq C ₈
AS1411	12.7 ± 0.3	20.5 ± 0.2	28.4 ± 0.2	>30
LNA-AS1411	6.4 ± 0.2	12.0 ± 0.2	17.5 ± 0.1	22.6 ± 0.2
U-AS1411	5.1 ± 0.5	14.9 ± 0.5	24.5 ± 0.4	>30

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_m$ (DNA + ligand) – T_m (DNA)]. The buffer used was 10 mM lithium cacodylate, pH 7.2, supplemented with 10 mM KCl and 90 mM LiCl. The T_m values for the aptamers are 56.3 ± 0.1 °C for AS1411, 51.6 ± 0.2 °C for LNA-AS1411, 52.6 ± 0.2 °C for U-AS1411. Values are reported as $\Delta T_m \pm$ SE.

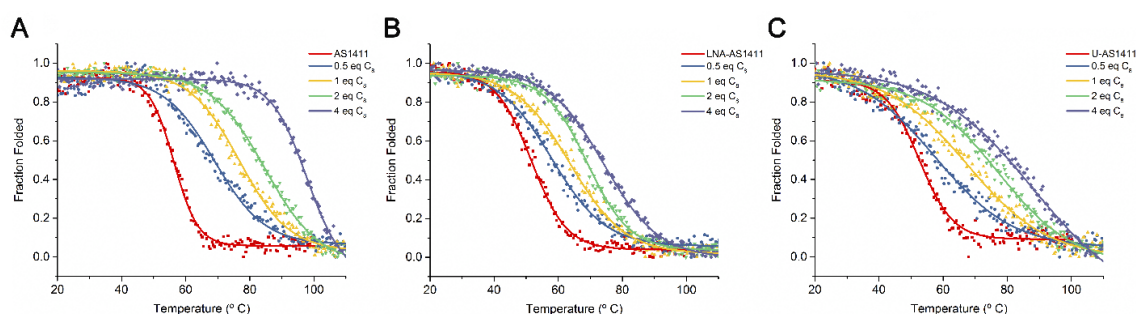


Figure S1 - CD melting curves of (A) AS1411, (B) LNA-AS1411 and (C) U-AS1411 in the presence of increasing amounts of ligand C₈. Melting curves were obtained by monitoring the ellipticity at 262 nm between 20 and 110 °C.

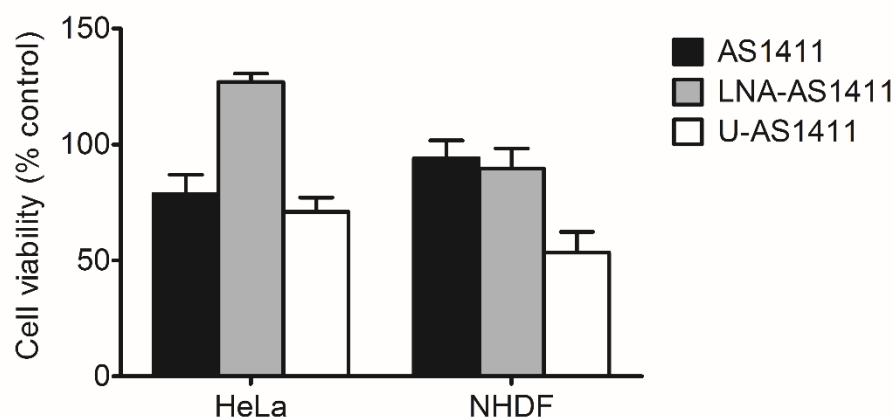


Figure S2 - Relative cell viability of HeLa and NHDF cells incubated for 7 days with the aptamers at 15 μM.

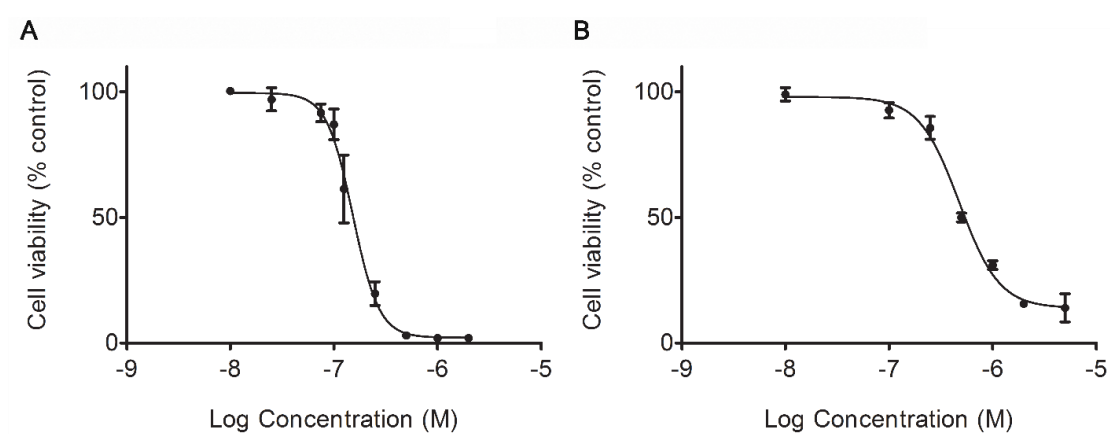


Figure S3 - Dose-response data of cell viability measured by MTT assay after incubation of (A) HeLa and (B) NHDF cells with C_8 for 7 days.

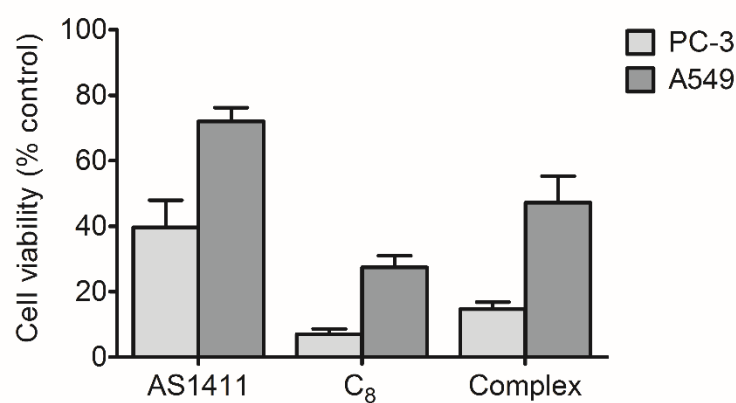


Figure S4 - Relative cell viability of PC-3 and A549 cancer cells incubated for 7 days with free AS1411, free C_8 or the pre-formed aptamer-ligand complex at a C_8 concentration of $1 \mu\text{M}$.

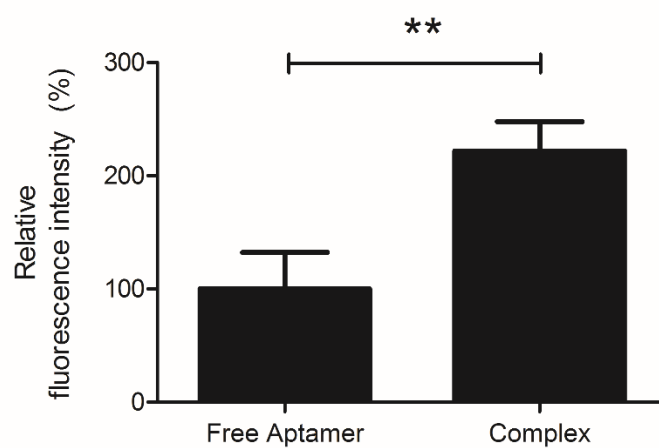


Figure S5 - CTCF of the cells incubated with free Cy5-AS1411 or Cy5-AS1411- C_8 complex. Data is represented as fluorescence intensity relatively to free aptamer images (positive control). Data represent the average of three experiments $\pm$ SEM, having considered at least 100 Cy5-AS1411-positive cells; ** $p < 0.01$.

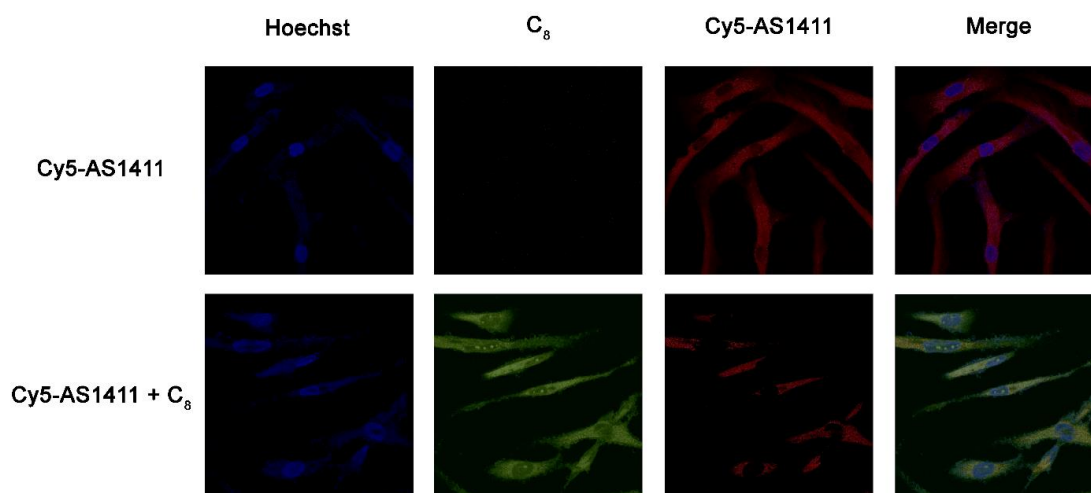


Figure S6 - Confocal microscopy images of NHDF cells incubated with Cy5-AS1411 and Cy5-AS1411-C₈ complex for 7 days. Cell nuclei are stained with Hoechst 33342 (blue), C₈ emits green fluorescence and Cy5-AS1411 is shown in red. Overlapping of the two stains, observed as yellow regions can be seen in the merge images. Brightness was adjusted to facilitate analysis.

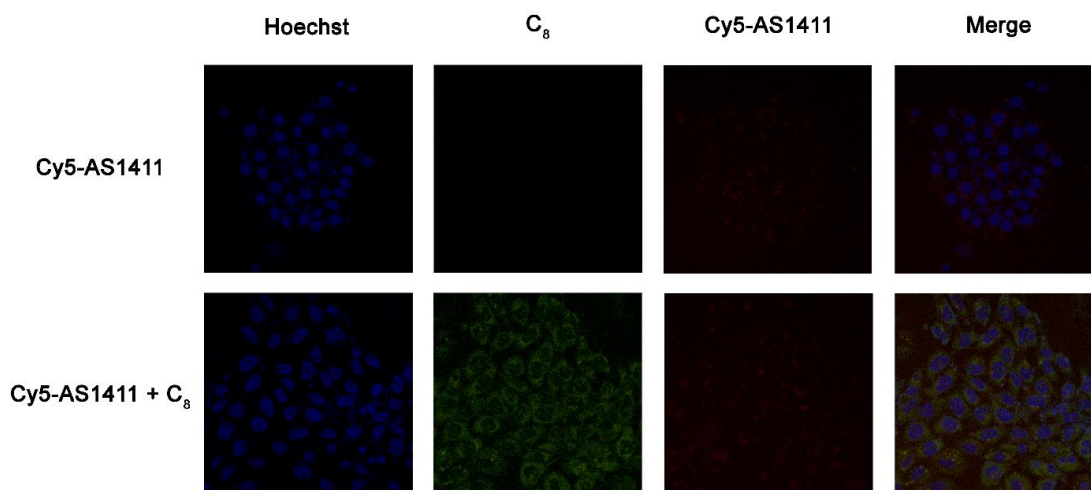


Figure S7 - Confocal microscopy images of HeLa cells incubated with Cy5-AS1411 and Cy5-AS1411-C₈ complex for 6 days followed by 1-day incubation in fresh media. Cell nuclei are stained with Hoechst 33342 (blue), C₈ emits green fluorescence while Cy5-AS1411 fluorescence is showed in red. Brightness was adjusted to facilitate analysis.

Appendix F

Supporting Information for Chapter 9

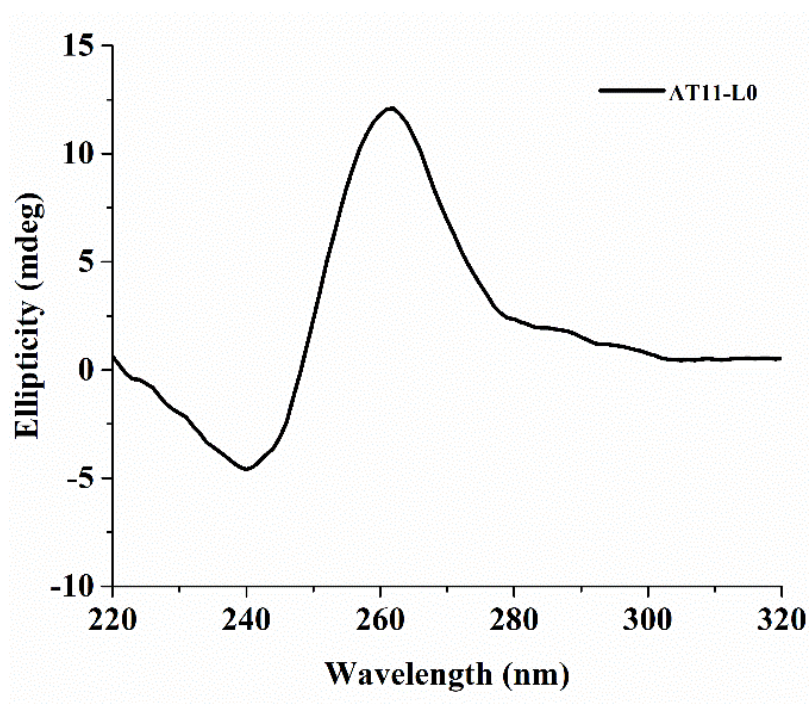


Figure S1 - CD spectrum of AT11-Lo acquired in 20 mM potassium phosphate buffer containing 65 mM KCl.

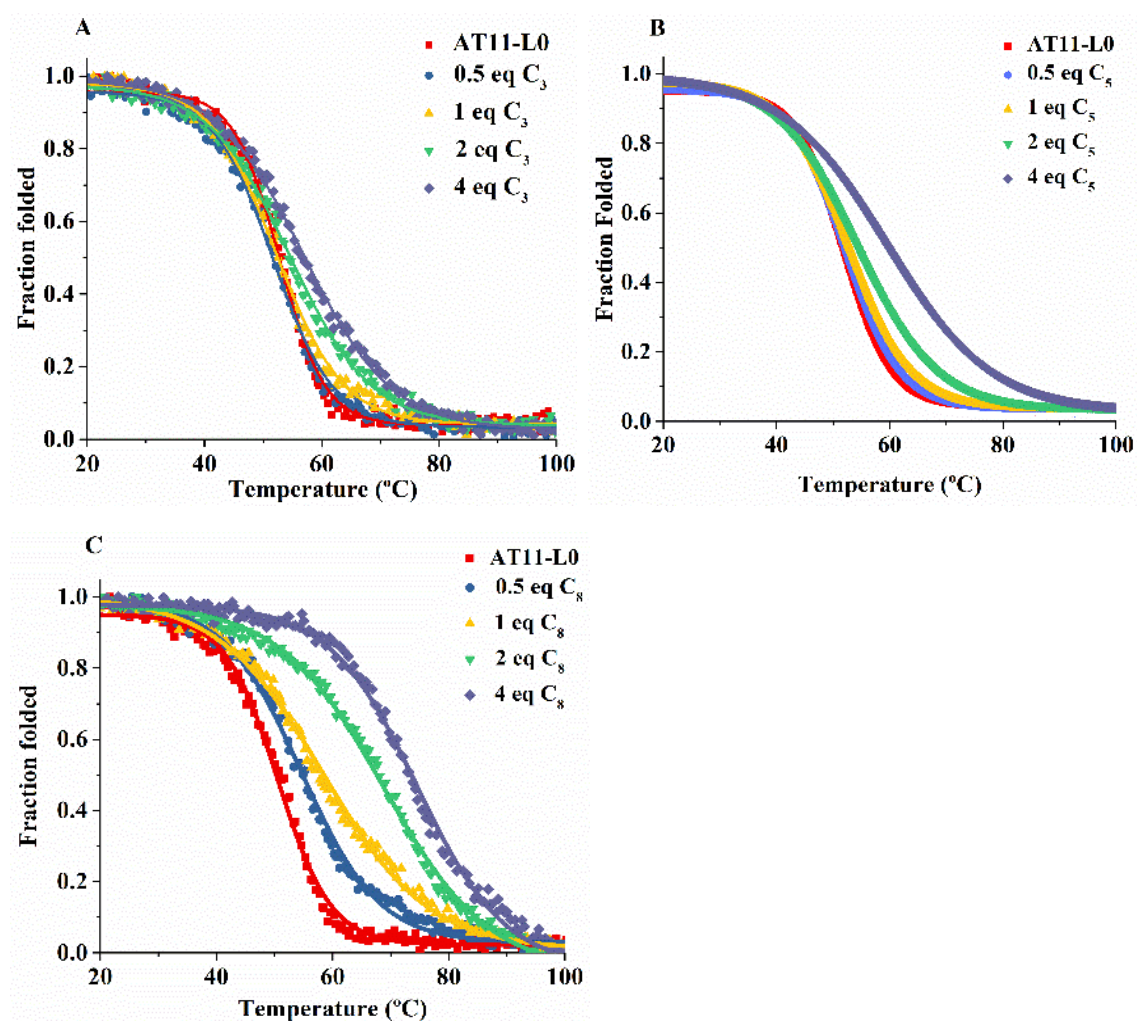


Figure S2 - CD melting curves of AT11-L0 with increasing amounts of ligands (A) C_3 (B) C_5 and (C) C_8 . Melting curves were obtained by monitoring the ellipticity at 262 nm between 20 and 100 $^{\circ}\text{C}$.

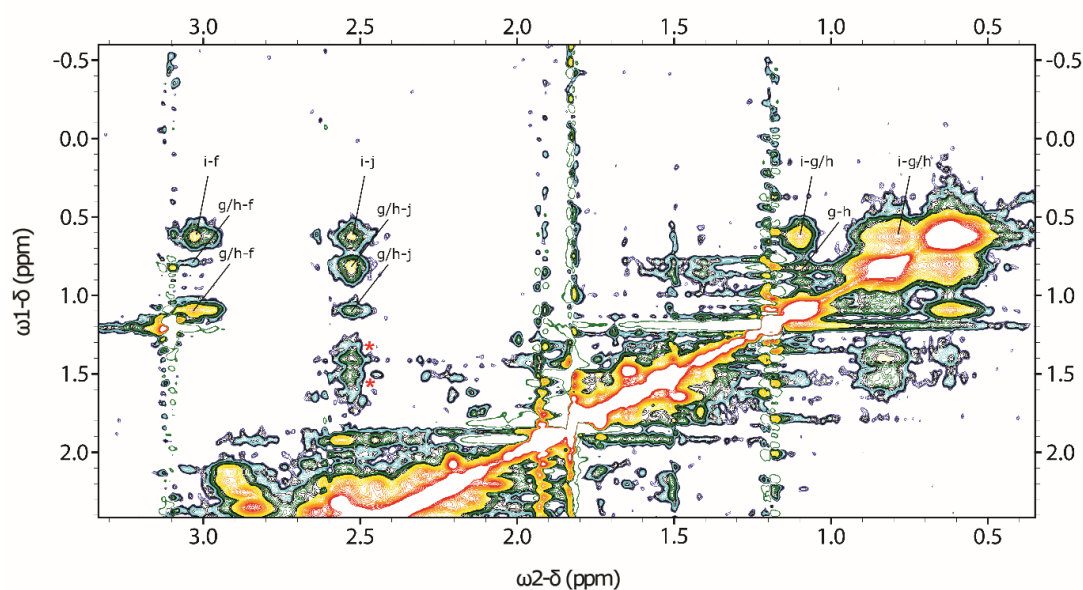


Figure S3 - Expanded region of the 2D NOESY spectrum showing the intramolecular NOEs of ligand C₈, particularly the aliphatic protons f, g, h, i and j. Intermolecular NOEs are also detected for proton j (highlighted with asterisks).

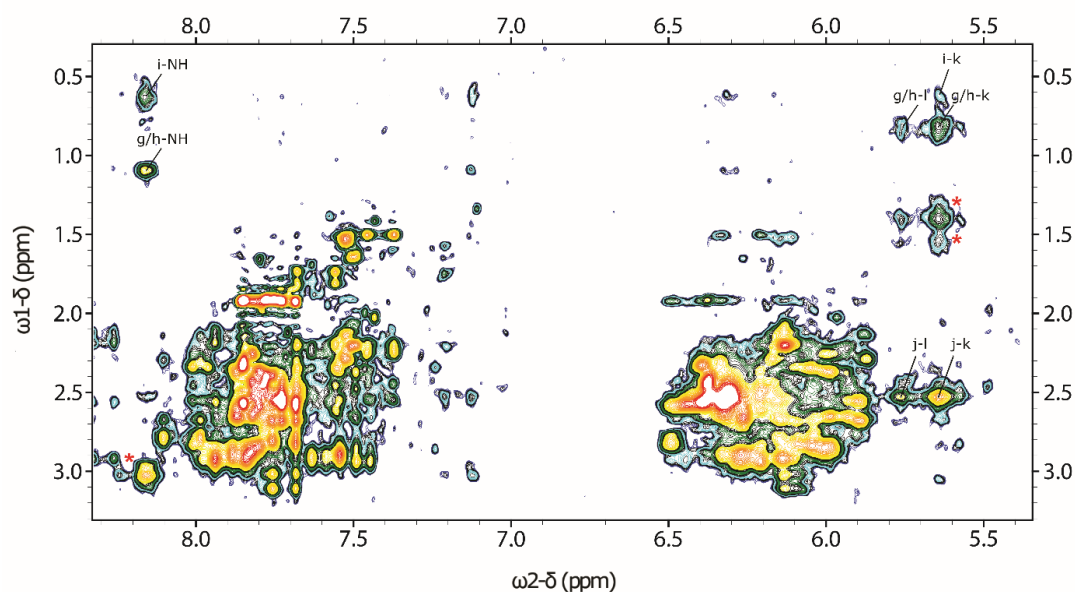


Figure S4 - Expanded region of the 2D NOESY spectrum showing the intramolecular NOEs of ligand C₈, namely between aliphatic protons and the iodobenzene moiety protons k and l. This suggests a crescent-shape conformation of C₈. Intermolecular NOEs are also detected for protons k and l with thymine's CH₃, presumably the capping T1 residue (highlighted with asterisks).

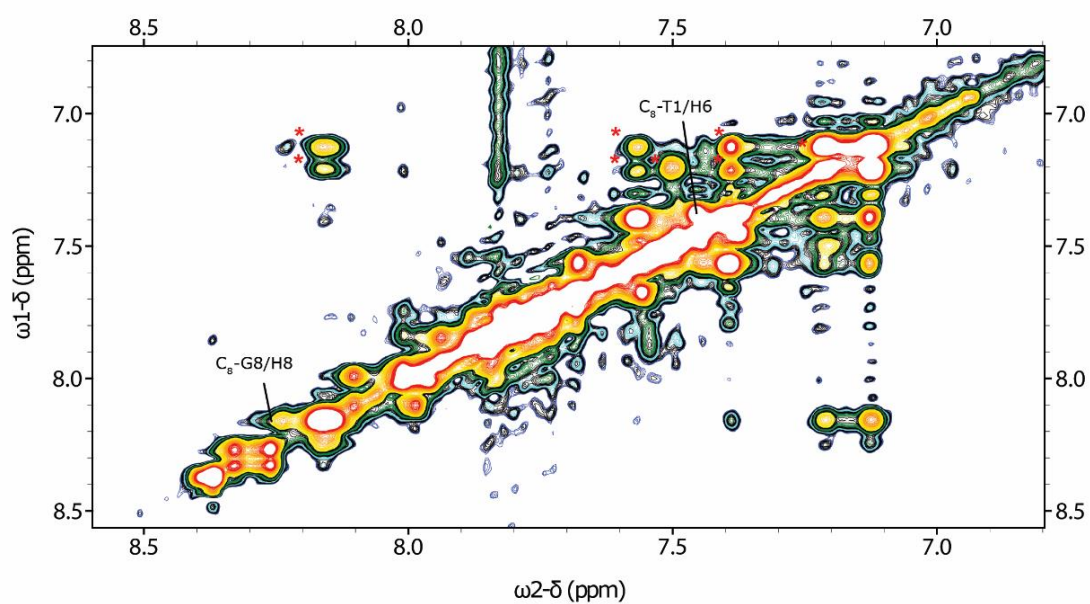


Figure S5 - Expanded region of the 2D NOESY spectrum showing the intermolecular NOEs between the AT11-Lo G8-H8, T1-H6 protons and ligand C₈ acridine moiety. Ligand C₈ intramolecular NOEs are highlighted with asterisks.

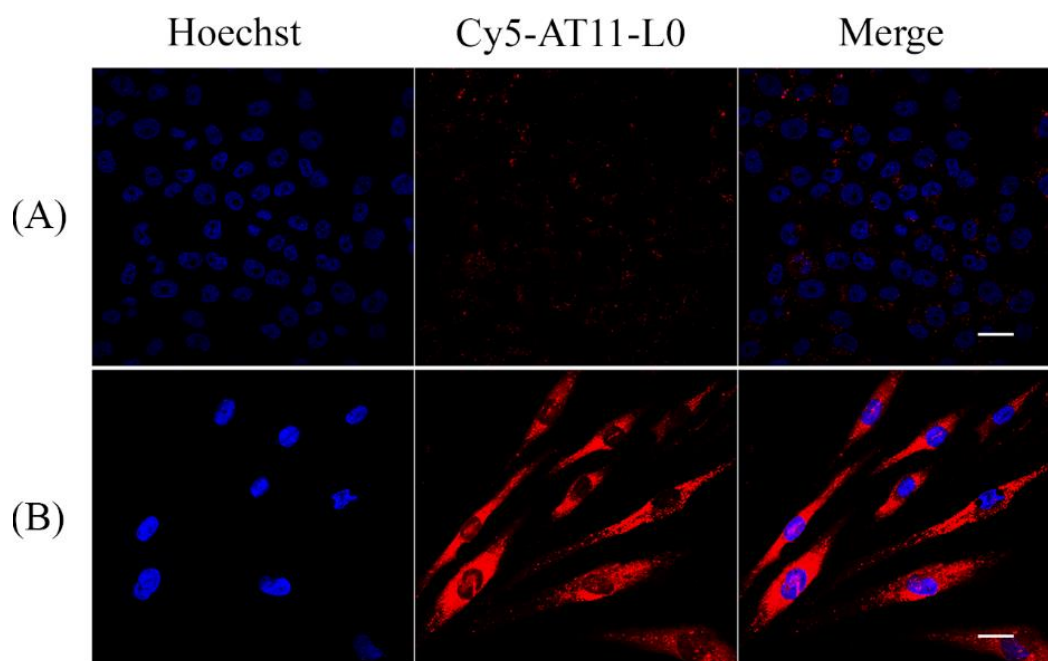


Figure S6 - Confocal microscopy images of (A) HeLa cells and (B) NHDF cells incubated with Cy5-AT11-Lo for 5 days. Cell nuclei are stained with Hoechst 33342 (blue), C₈ emits green fluorescence and Cy5-AT11-Lo is shown in red. Scale bar: 25 μ m.

Appendix G

Supporting Information for Chapter 10

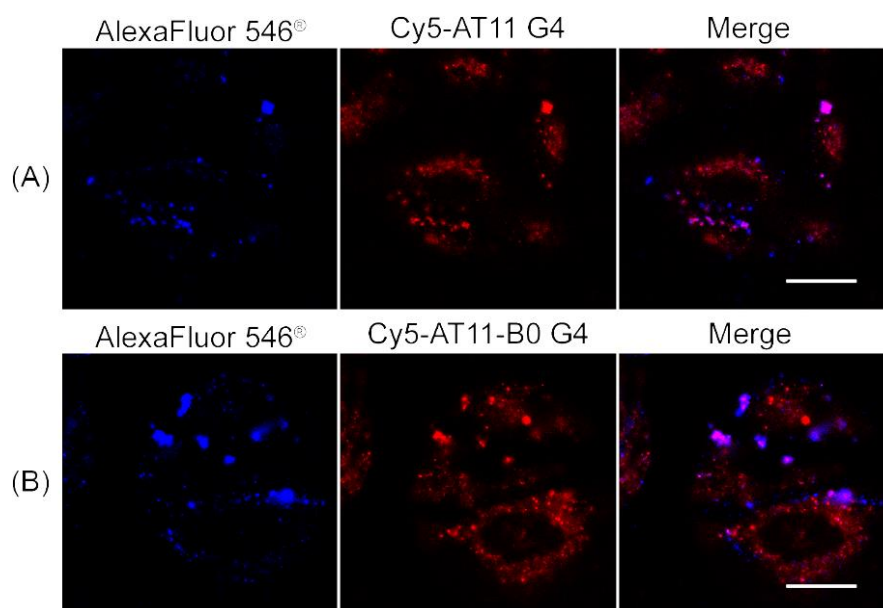


Figure S1 - Confocal fluorescence images of HeLa cells incubated with (A) Cy5-AT11 and (B) Cy5-AT11-Bo after 2 h at 37°C. Nucleolin is stained with AlexaFluor 546®, Cy5-AT11 and Cy5-AT11-Bo are shown in red. Scale bar: 25 μ m. Brightness was equally adjusted in all images

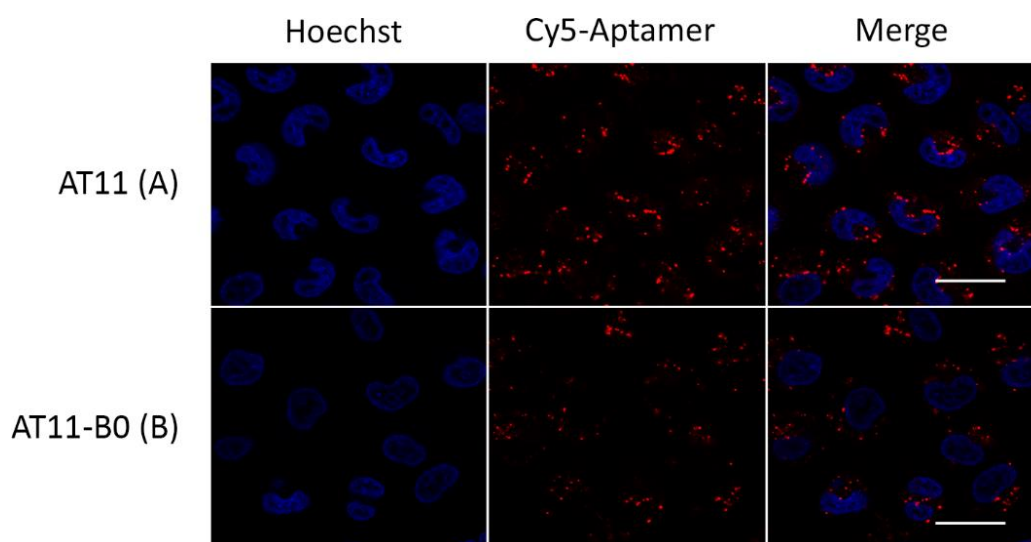


Figure S2 - Confocal microscopy images of HeLa cells incubated for 3 days with (A) AT11 G4 or (B) AT11-Bo G4. Cell nuclei are stained with Hoechst 33342 (blue), C₈ emits green fluorescence and Cy5- AT11 or AT11-Bo G4s are shown in red. Scale bar: 25 μ m

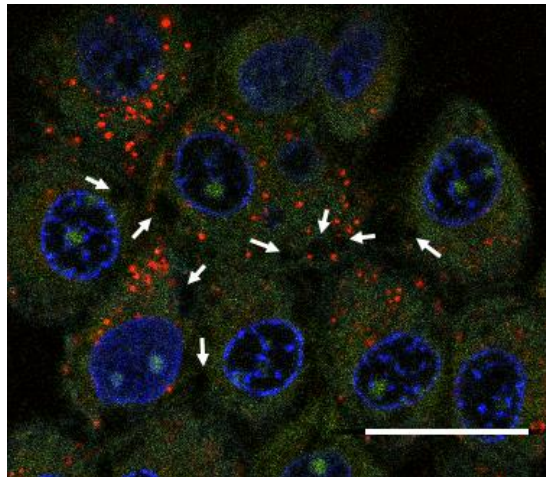


Figure S3 - Representative confocal microscopy image of HeLa cells incubated with G4-C8 complexes showing vacuolization (highlighted by white arrows) indicative of methuosis. Cell nuclei are stained with Hoechst 33342 (blue), C8 emits green fluorescence and Cy5-G4 sequence is shown in red. Scale bar: 25 μ m

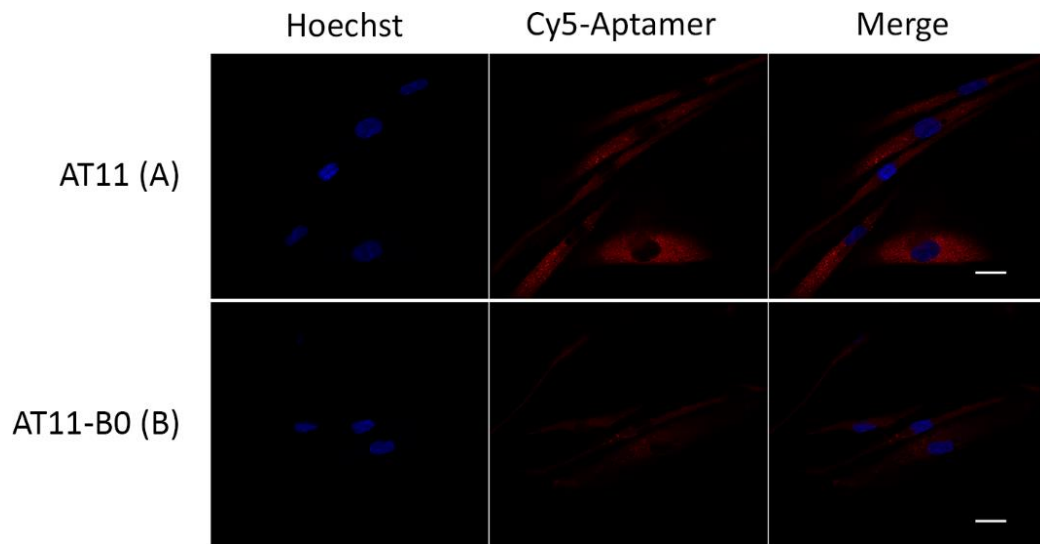


Figure S4 - Confocal microscopy images of NHDF cells incubated for 3 days with (A) AT11 G4 or (B) AT11-Bo G4. Cell nuclei are stained with Hoechst 33342 (blue), C8 emits green fluorescence and Cy5- AT11 or AT11-Bo G4s are shown in red. Scale bar: 25 μ m

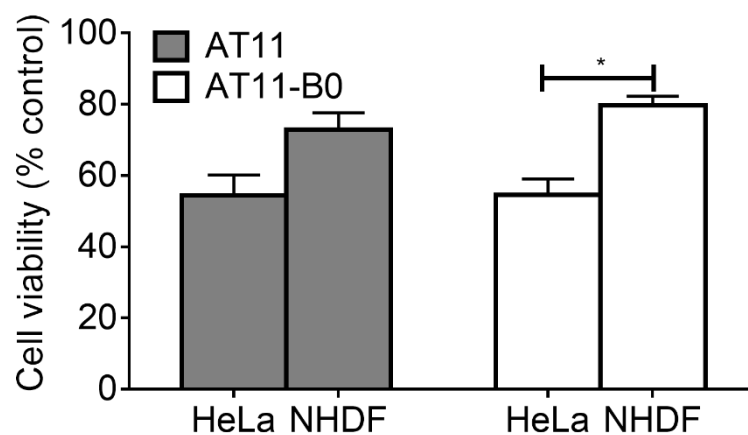


Figure S5 - Percentage of viable HeLa and NHDF cells after incubation with free AT11 G4 and AT11-Bo G4 at 15 μ M for 7 days

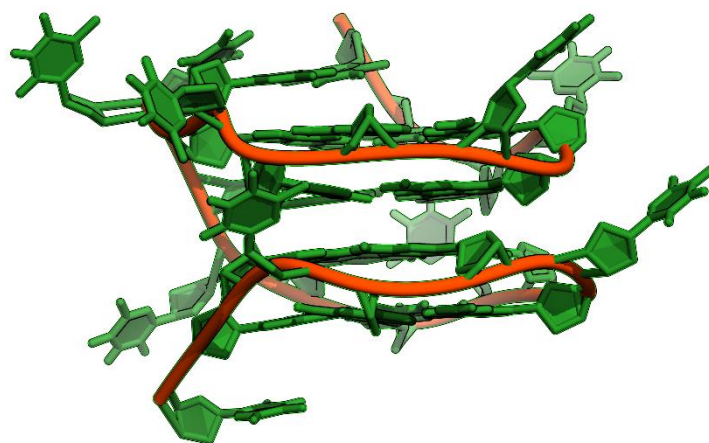


Figure S6 - NMR derived 3D structure of AT11 G4 (PDB: 2N3M) . G4 structure is depicted in green with backbone colored in orange.

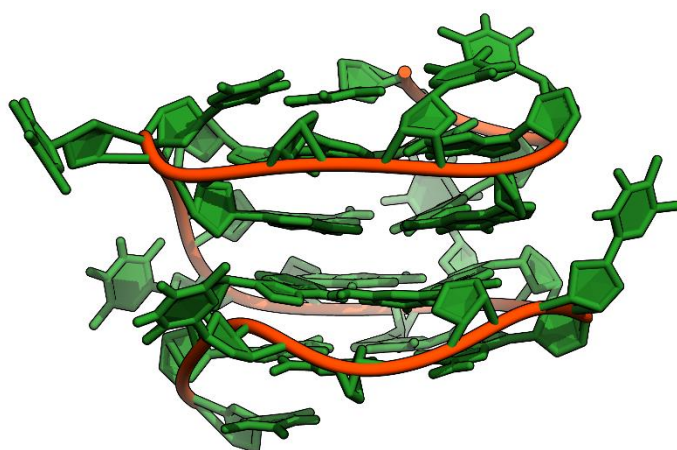


Figure S7 - Predicted structure of AT11-Bo G4, extracted from the final snapshot of the 200 ns MD simulation. G4 structure is depicted in green with backbone colored in orange.

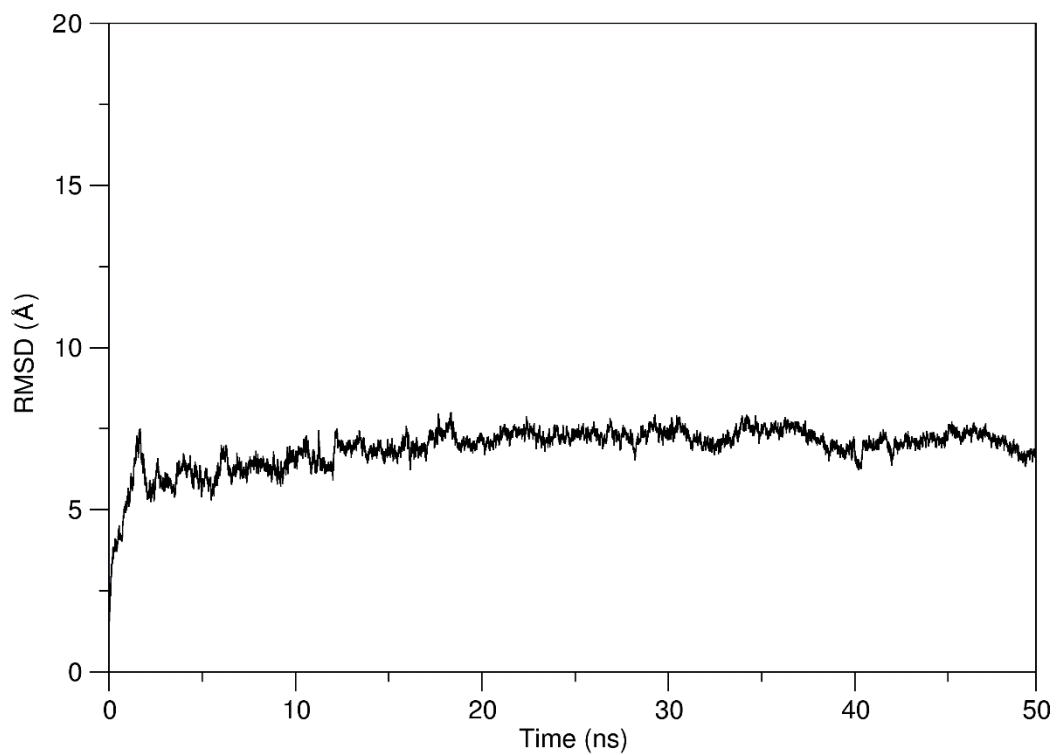


Figure S8 - RMSD plot for the AT11 G4- RBD1,2 complex structure during full 50 ns MD simulation.

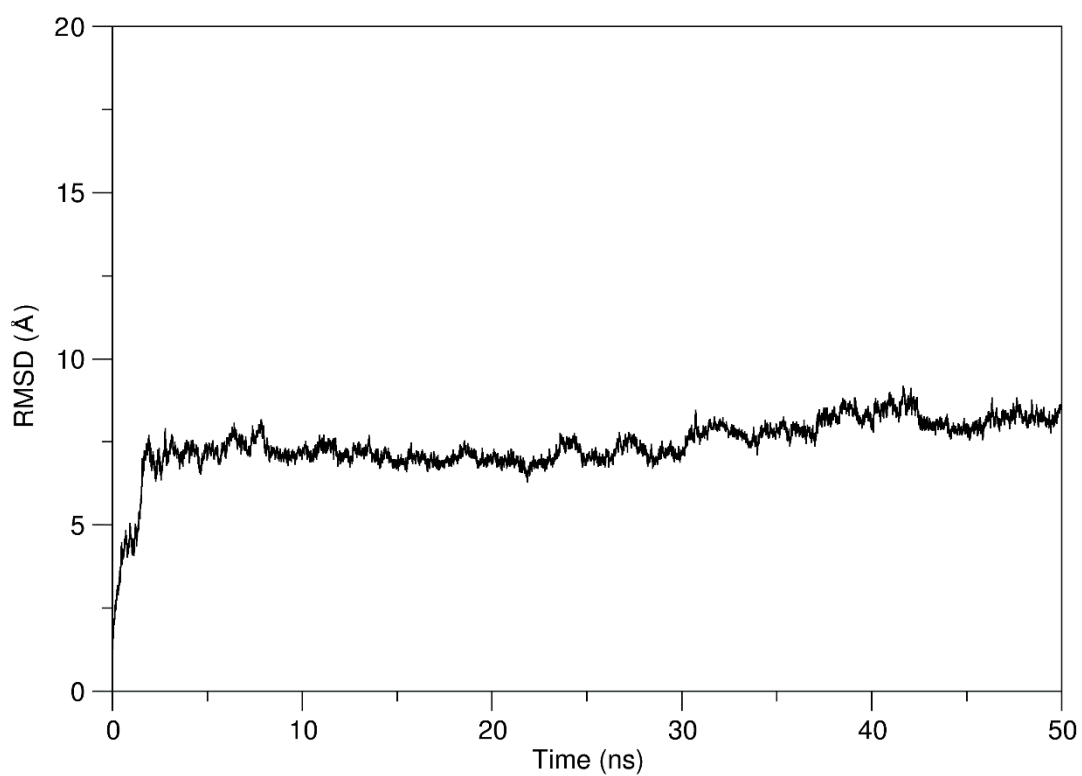


Figure S9 - RMSD plot for the AT11-Bo G4- RBD1,2 complex structure during full 50 ns MD simulation.

Table S1 - Pearson coefficients for the colocalization of the AT11-C₈ and AT11-Bo-C₈ in HeLa and NHDF cell lines. The values are represented as mean $\pm$ SD and were obtained from 3 different confocal images (Figures 10 and 11).

	HeLa	NHDF
AT11-C₈	0.67 $\pm$ 0.06	0.86 $\pm$ 0.02
AT11-Bo-C₈	0.60 $\pm$ 0.06	0.83 $\pm$ 0.03

Appendix H

Supporting Information for Chapter 11

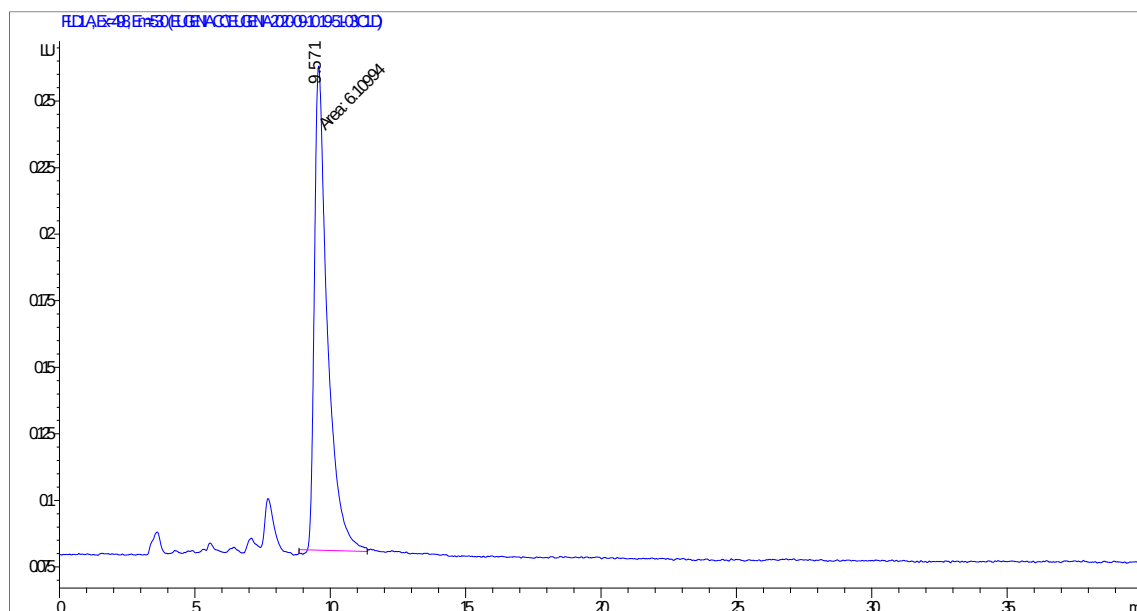


Figure S1 - Chromatogram of C8 at the concentration of 0.98 ng/mL.

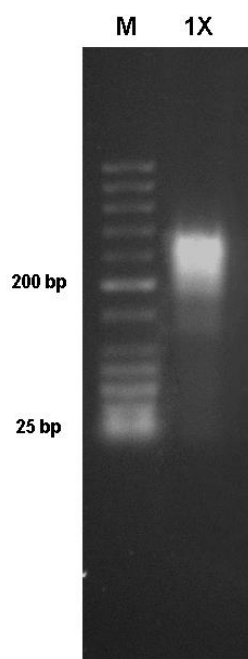


Figure S2 - Native agarose gel electrophoresis of 5'-lipid-AS1411 in 1× buffer (20 mM potassium phosphate buffer pH 6.9 containing 50 mM NaCl, 5 mM KCl and 5 mM MgCl₂).

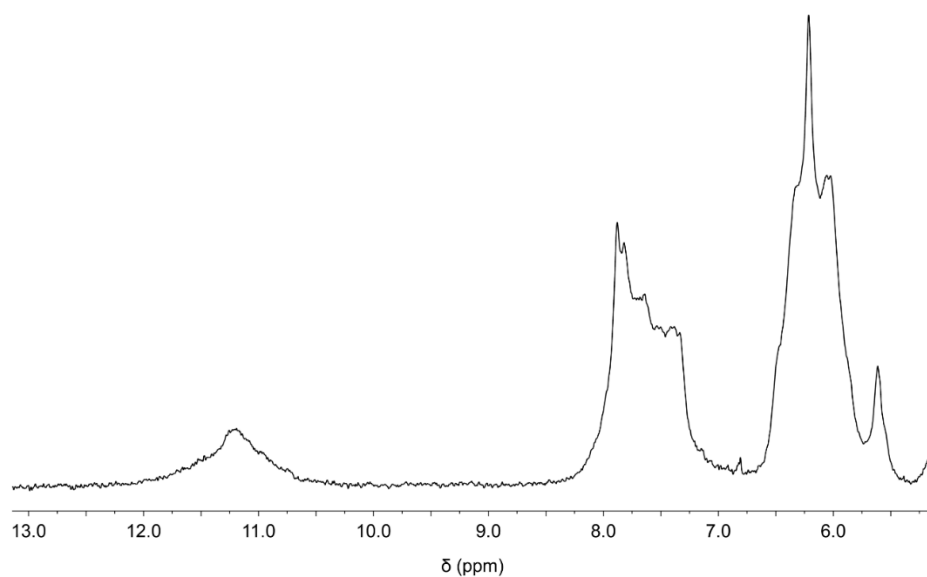


Figure S3 - Expanded imino (10-12 ppm) and aromatic regions (6.5-8.5 ppm) of the ^1H NMR spectra of 5'-lipid-AS1411 in $1\times$ buffer indicating the formation of a G4 structure.

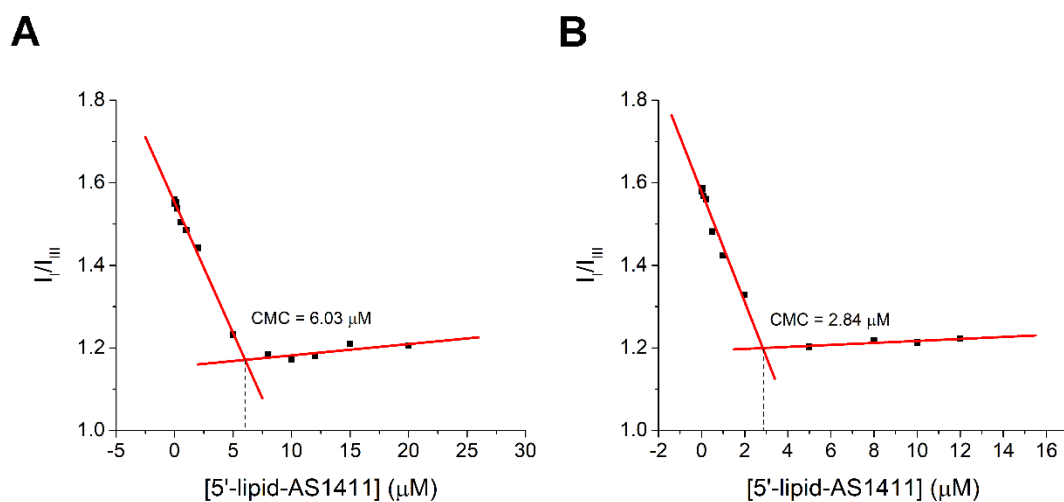


Figure S4 - Critical micellar concentration (CMC) determination by the pyrene fluorescence method. CMC was determined from the inflexion point in the plot representing the quotient of vibrational band intensities (I_I/I_{III}) as a function of 5'-lipid-AS1411 concentration.

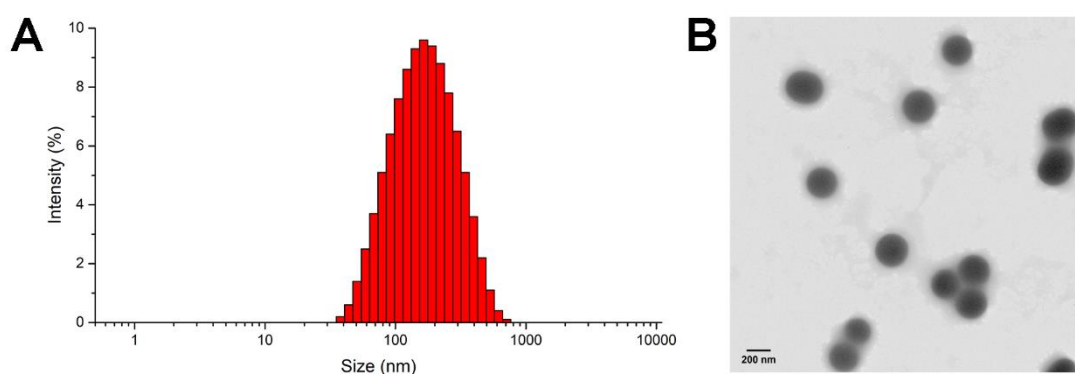


Figure S5 - (A) DLS and (B) TEM (12,000x magnification) analysis of nanoaggregates formed by 5'-lipid-AS1411 in 1X buffer, after 7 days at room temperature.

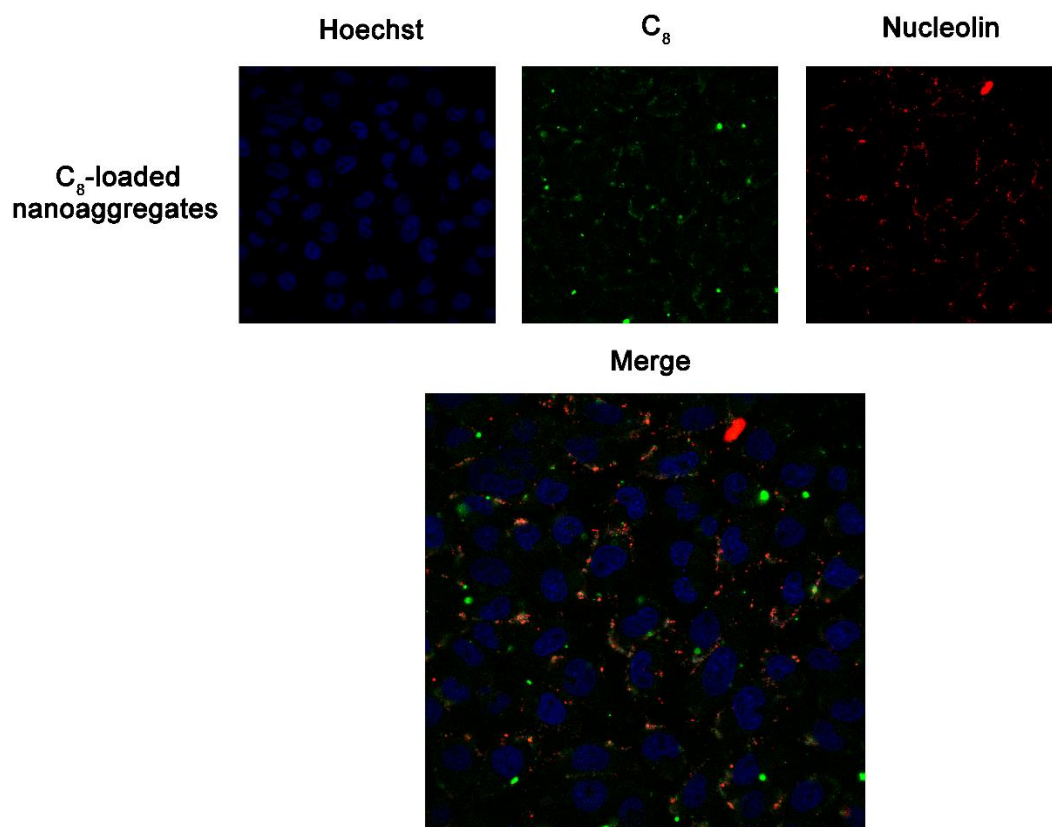


Figure S6 - Fluorescence confocal microscopy images of HeLa cells incubated with C_8 -loaded nanoaggregates for 2h. Cell nuclei are stained with Hoechst 33342 in the blue channel, C_8 emits green fluorescence and nucleolin was stained with Alexa Fluor™ 647 antibody. Overlapping of the green and red stains, observed as yellow regions, can be seen in the merged images of the cells.

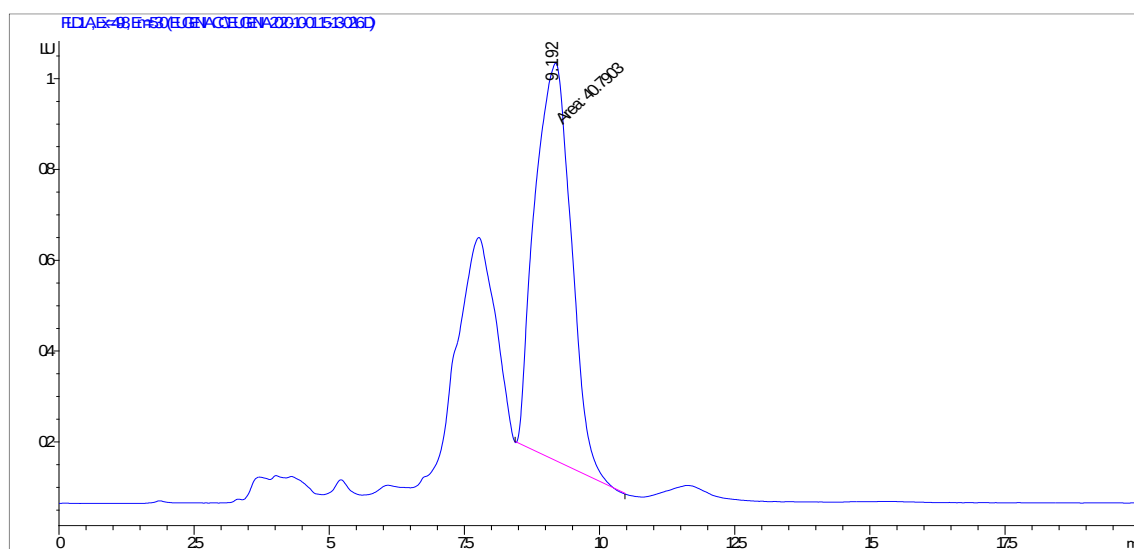


Figure S7 - Chromatogram of C₈-loaded nanoaggregates (3.64 ng/mL).