

Preučevanje vpliva citozinov na zvižanje z gvanini
bogatih zaporedij

PREUČEVANJE VPLIVA CITOZINOV NA ZVIJANJE Z GVANINI BOGATIH ZAPOREDIJ

KEMIJA

RAZISKOVALNA NALOGA

KLARA LAVRIČ

4.A

MENTOR: dr. MARTINA LENARČIČ ŽIVKOVIĆ

SOMENTOR: TOMAŽ ŽIGON, mag. prof. kem.

2024

BIOTEHNIŠKI IZOBRAŽEVALNI CENTER LJUBLJANA,
GIMNAZIJA IN VETERINARSKA ŠOLA, LJUBLJANA
v sodelovanju s
KEMIJSKI INŠTITUT,
NACIONALNI CENTER ZA NMR SPEKTROSKOPIJO VISOKE
LOČLJIVOSTI, LJUBLJANA

Kazalo

Seznam okrajšav in simbolov	III
Povzetek	IV
Zahvala	V
1. Uvod	1
1.1 Deoksiribonukleinska kislina (DNK)	1
1.2 Sekundarna struktura DNK.....	1
1.3 Sinteza oligonukleotidov	6
1.4 FPLC	8
1.5 NMR spektroskopija	8
2. Cilji in problemi naloge.....	10
3. Hipoteze.....	12
4. Metode dela	13
4.1 Sinteza.....	13
4.2 Cepitev oligonukleotidov s trdnega nosilca in odstranjevanje zaščitnih skupin	13
4.3 Čiščenje oligonukleotidov s kartušami	13
4.4 Čiščenje in razsoljevanje oligonukleotidov s FPLC	14
4.5 Priprava vzorcev za NMR spektroskopijo	14
4.6 NMR spektroskopija	14
5. Rezultati.....	15
6. Razprava	22
7. Zaključek	23
8. Literatura	25

Seznam okrajšav in simbolov

1D.....	enodimenzionalno
A.....	adenin
ACN.....	acetonitril
AMA.....	1:1 v/v raztopina amonijevega hidroksida (NH ₃ (aq)) in metilamina)
C.....	citozin
D ₂ O.....	devterijev dioksid
DMT.....	dimetoksitritil
DNK.....	deoksiribonukleinska kislina
FPLC.....	tekočinska kromatografija za hitro ločevanje proteinov
G.....	gvanin
G4.....	G-kvadrupleks
KCl.....	kalijev klorid
KPi.....	kalijev fosfatni pufer
NMR.....	jedrska magnetna resonanca
ppm.....	število delcev na milijon (<i>ang.</i> parts per million)
RNK.....	ribonukleinska kislina
T.....	timin
TEAA.....	trietilamonijev acetat
TFA.....	trifluorocetna kislina

Povzetek

Ključne besede: DNK, G-kvadrupleks, lasnica, genska regulacija

G-kvadrupleksi (G4) so štiri-vijačne strukture DNK, ki nastanejo na področjih z gvanini bogatih regij DNK. Te so zgoščene predvsem na koncih kromosomov (telomerah) ter v regulatornih regijah genov. Igrajo pomembno vlogo pri prepisovanju in podvojevanju DNK ter ohranjanju stabilnosti genoma. Raziskave kažejo, da lahko spremembe v tvorbi in stabilnosti G4 vplivajo na razvoj različnih bolezni, vključno z rakom. Zato so G4 v zadnjih letih postali pomembna tarča za razvoj zdravil in terapevtskih pristopov, ki ciljajo na genetsko regulacijo.

Na tvorbo in stabilnost G4 poleg gvaninov vplivajo tudi drugi, naravno prisotni ostanki DNK, kot so citozini. Citozini se z gvanini bazno pariyo in tvorijo strukture, ki so v celici konkurenčne G4 – lasnice. V človeškem genomu so našli približno 11000 G/C bogatih zaporedij, v katerih se pojavita vsaj dva trakta s tremi zaporednimi citozini. Poznavanje mehanizma preklopa med G4 in lasnicami je ključnega pomena za razumevanje funkcije sekundarnih struktur DNK v celici. Hkrati nam podrobno poznavanje faktorjev, ki vplivajo na to, katera sekundarna struktura DNK se bo tvorila na določenem mestu v genomu, lahko pomaga razvozlati, zakaj prihaja do nastanka določenih bolezni in kako jih zdraviti.

V okviru projektne naloge smo v modelno zaporedje, ki se zvije v G4, postopoma uvajali citozine. Preučevali smo, kakšno razmerje med gvanini in citozini v izbranem z gvanini bogatem zaporedju še vodi v tvorbo G4. S pomočjo NMR spektroskopije smo ugotovili, da ob dodatku dveh traktov dveh ali več zaporednih citozinov v izbrano zaporedje pride do preklopa iz G4 v lasnico. Lasnica se tvori tudi, če oligonukleotid zvijamo v pogojih, ki vzpodbujajo tvorbo G4 (prisotnost K^+ ionov). Predstavljeni rezultati nudijo nov vpogled v kompleksnost zvijanja z gvanini bogatih zaporedij in nakazujejo na fino regulacijo med različnimi sekundarnimi strukturami DNK v celici.

Zahvala

V največji meri se iskreno zahvaljujem zunanji mentorici dr. Martini Lenarčič Živković z Nacionalnega centra za NMR spektroskopijo visoke ločljivosti na Kemijskem inštitutu, ki me je uvajala v nove metode, pomagala pri eksperimentalnem delu, vedno poskrbela za teoretično podlago in bila na voljo za kakršnekoli nasvete ter diskusijo.

Zahvaljujem se šolskemu mentorju Tomažu Žigonu za možnost opravljanja raziskovalne naloge izven šole, za odzivnost in izkazano podporo.

Iskreno pa se zahvaljujem tudi svoji družini in prijateljem za vso podporo med nastajanjem projektne naloge.

1. Uvod

1.1 Deoksiribonukleinska kislina (DNK)

Odkritje deoksiribonukleinske kisline (DNK) predstavlja enega najpomembnejših mejnikov v zgodovini znanosti. Prvi je prisotnost DNK v celicah opisal Friedrich Miescher leta 1868, vendar sta komaj leta 1953 James Watson in Francis Crick razvozlala njeno strukturo. Njuna odkritja so temeljila na delu Rosalind Franklin, ki je s pomočjo rentgenske kristalografije pridobila ključne podatke o strukturi molekule (Dahm, 2008; Pray, 2008; Watson in Crick, 1953).

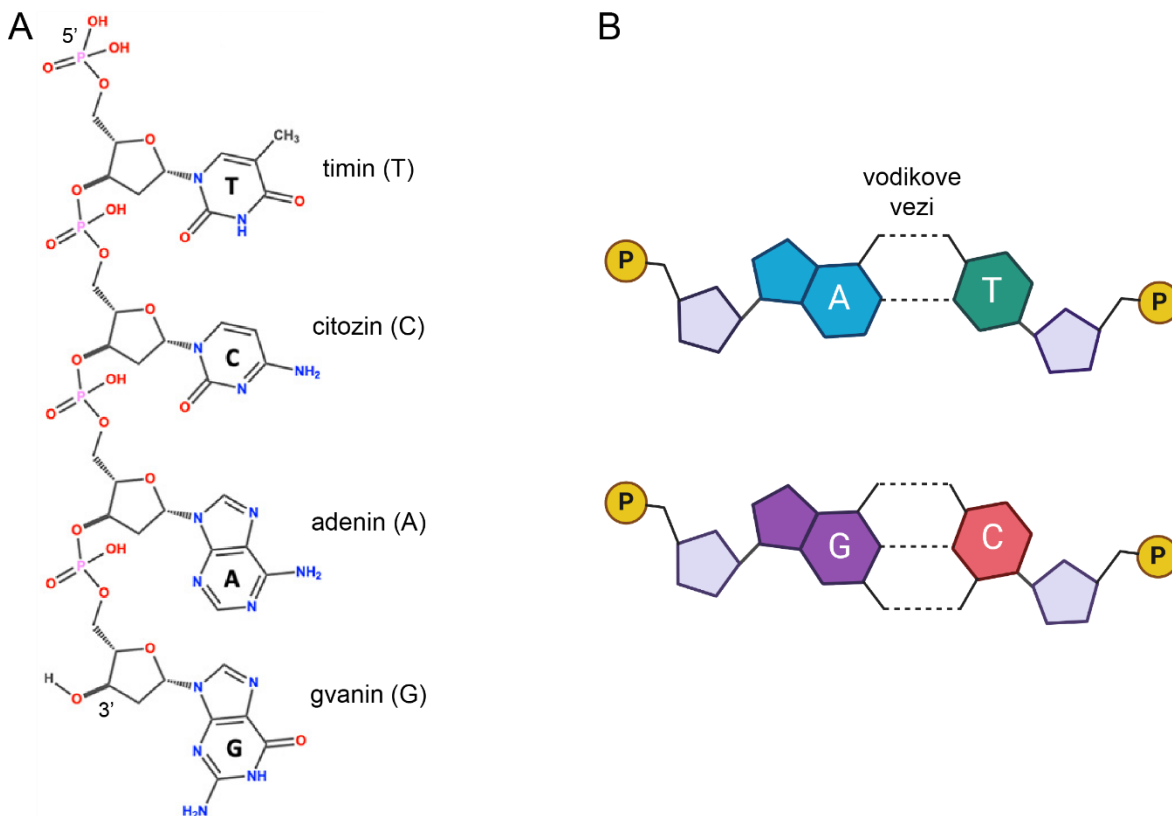
Glavna funkcija DNK je shranjevanje genetskih informacij, ki vodijo razvoj, delovanje in vzdrževanje živih organizmov. Genetski kod, zapisan v primarnem zaporedju DNK, določa, kako se organizem razvija in deluje ter prenaša dedne lastnosti iz generacije v generacijo, njena sekundarna struktura pa igra pomembno vlogo pri uravnavanju izražanja genov. Razumevanje strukture in funkcije DNK je ključnega pomena za razvoj genetskega inženiringa, medicinskih raziskav, terapevtskih pristopov in številnih drugih področij znanosti. Odkritje DNK je tako pustilo trajen pečat na različnih disciplinah, ki se ukvarjajo z življenjskimi znanostmi in je nedvomno prispevalo k nadaljnjemu razumevanju življenja na molekularni ravni (Minchin, 2019; Alberts, 2002; Pray, 2008).

DNK je v celicah z beljakovinami organizirana v goste komplekse, ki jih imenujemo kromosomi. Pri evkariontih se kromosomi nahajajo v jedru, čeprav je DNK prisotna tudi v mitohondrijih in kloroplastih. Prokarionti imajo DNK v obliki enega krožnega kromosoma v citoplazmi. Nekateri prokarionti (bakterije in nekaj evkariontov) imajo zunajkromosomsko DNK, ki jo imenujemo plazmid. Plazmidi so samostojen, genski material, ki ima sposobnost samopodvojevanja. Plazmide uporabljajo v tehnologiji rekombinantne DNK za preučevanje izražanja genov (The Editors of Encyclopædia Britannica, 2024).

1.2 Sekundarna struktura DNK

1.2.1 Dvojna vijačnica

DNK se v celici večinoma nahaja v obliki dvojne vijačnice, ki je sestavljena iz dveh verig nukleotidov. Vsak nukleotid je sestavljen iz deoksiriboze, na katero sta vezani fosfatna skupina in dušikova baza. Dušikove baze delimo na purine (adenin, gvanin) ter pirimidine (citozin, timin) (slika 1A). Nukleotide povezujejo kovalentne vezi med fosfatno skupino enega nukleotida in sladkorjem drugega nukleotida, tako da tvorijo fosfatno-sladkorno ogrodje na katerega so pripete dušikove baze (Minchin, 2019; The Editors of Encyclopædia Britannica, 2024).



Slika 1: (A) Prikaz fosfatno-sladkornega ogrodja DNK, na katerega so pripete dušikove baze. Slika je povzeta po Minchin (2019). (B) Prikaz vodikovih vezi po Watson-Cricku. Shema baznih parov je bila ustvarjena s programom BioRender (BioRender.com).

Antiparalelni, komplementarni verigi dvojne vijačnice DNK sta med seboj povezani z vodikovimi vezmi. Dva nukleotida na komplementarnih verigah DNK, ki sta povezana preko vodikovih vezi, se imenujeta bazni par. Pri kanoničnem združevanju baz po Watson–Cricku se v DNK v bazni par z dvema vodikovima vezema povezujeta adenin (A) in timin (T) ter gvanin (G) in citozin (C), ki sta povezana s tremi vodikovimi vezmi (slika 1B) (Cleaves, 2015).

Več desetletij je veljalo splošno prepričanje, da je dvojna vijačnica edina struktura DNK, ki je biološko pomembna, čeprav so podatki difrakcije vlaken Rosalind Franklin jasno pokazali, da lahko molekula DNK prehaja med različnimi strukturami, ki so odvisne od zaporedja baz in okolja (Bansal, 2003).

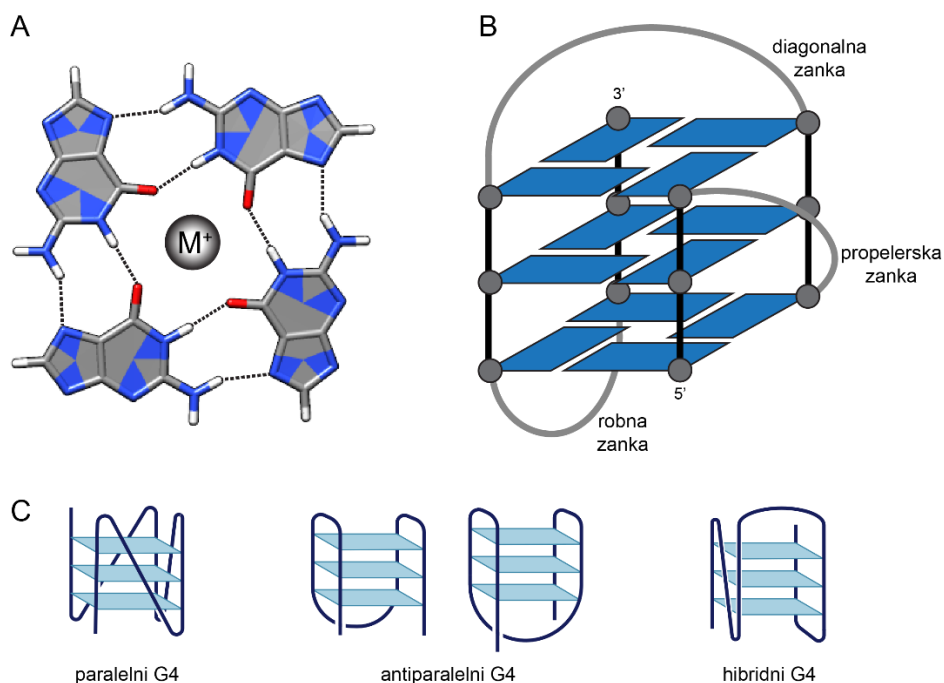
Poleg dvojne vijačnice lahko DNK tvori tudi nekanonične strukture, ki so še posebej stabilne v natrpanem znotrajceličnem okolju, kot so na primer lasnica, križ, DNK tripleks, G-kvadrupleksi in i-motivi. Pojavljajo se predvsem na delih DNK, kjer prihaja do več ponovitev določenih zaporedij (na primer direktne ali zrcalne tandemske ponovitve). Pri nastanku teh od dvojne vijačnice različnih DNK struktur poleg zaporedja ključno vlogo igrajo tudi lastnosti topila in prisotnost ionov (Tateishi-Karimata, 2020; Bansal, 2022; Oh, 2020). Dokazali so, da imajo te nenavadne strukture vlogo pri različnih bioloških procesih, kot so podvajanje, prepisovanje in prevajanje (Dalla Pozza, 2022). Ena izmed najbolj preučevanih skupin nekanoničnih sekundarnih struktur DNK so G-kvadrupleksi.

1.2.2 G-kvadrupleksi

1.2.2.1 Strukturne značilnosti G-kvadrupleksov

Zaporedja DNK, bogata z gvanini, se lahko zložijo v štiriverižne, nekanonične sekundarne strukture, imenovane G-kvadrupleksi (G4). G4 nastanejo z nalaganjem *t.i.* G-kvartetov (slika 2A). Znotraj enega G-kvarteta se štirje gvanini povežejo v planarne strukture preko Hoogsteenovih vodikovih vezi. Za stabilizacijo strukture so potrebni kationi, ki nevtralizirajo odboj karbonilnih skupin znotraj G-kvartetov. Strukturno G4 najboljše stabilizirajo K^+ ioni, sledijo jim NH_4^+ in Na^+ ioni (Spiegel, 2020; Tateishi-Karimata, 2020 in Adrian, 2012).

G4 se tvorijo v z gvanini bogatih regijah DNK (ali RNK), kjer so gvanini razporejeni v določenem vzorcu. Ponavadi zaporedja, ki tvorijo G4 sledijo vzorcu $GG(G)N_{1-7}GG(G)N_{1-7}GG(G)N_{1-7}GG(G)$, kjer so trakti (ponavadi dveh ali treh) zaporednih gvaninov prekinjeni z različno dolgimi, kratkimi segmenti ostalih nukleotidov (A,C,T). Ti ostanki tvorijo *t.i.* zanke, ki med seboj povezujejo G-kvartete (slika 2B). Poznamo tri vrste zank: robna, diagonalna in propelerska. Zaradi variacij v številu naloženih G-kvartetov in ogromnega števila možnih kombinacij zaporedij v zankah, G4 uvrščamo med polimorfne strukturne družine. Glede na način, kako so štiri verige osrčja strukture povezane med seboj ločimo več vrst G-kvadrupleksov. Če so G-kvarteti med seboj povezani s samimi propelerskimi zankami, vse verige od 5' proti 3' koncu zaporedja potekajo v isto smer. G-verige so v takem primeru usmerjene paralelno in zato tako strukturo imenujemo paralelni G4 (slika 2C). V *t.i.* antipara-



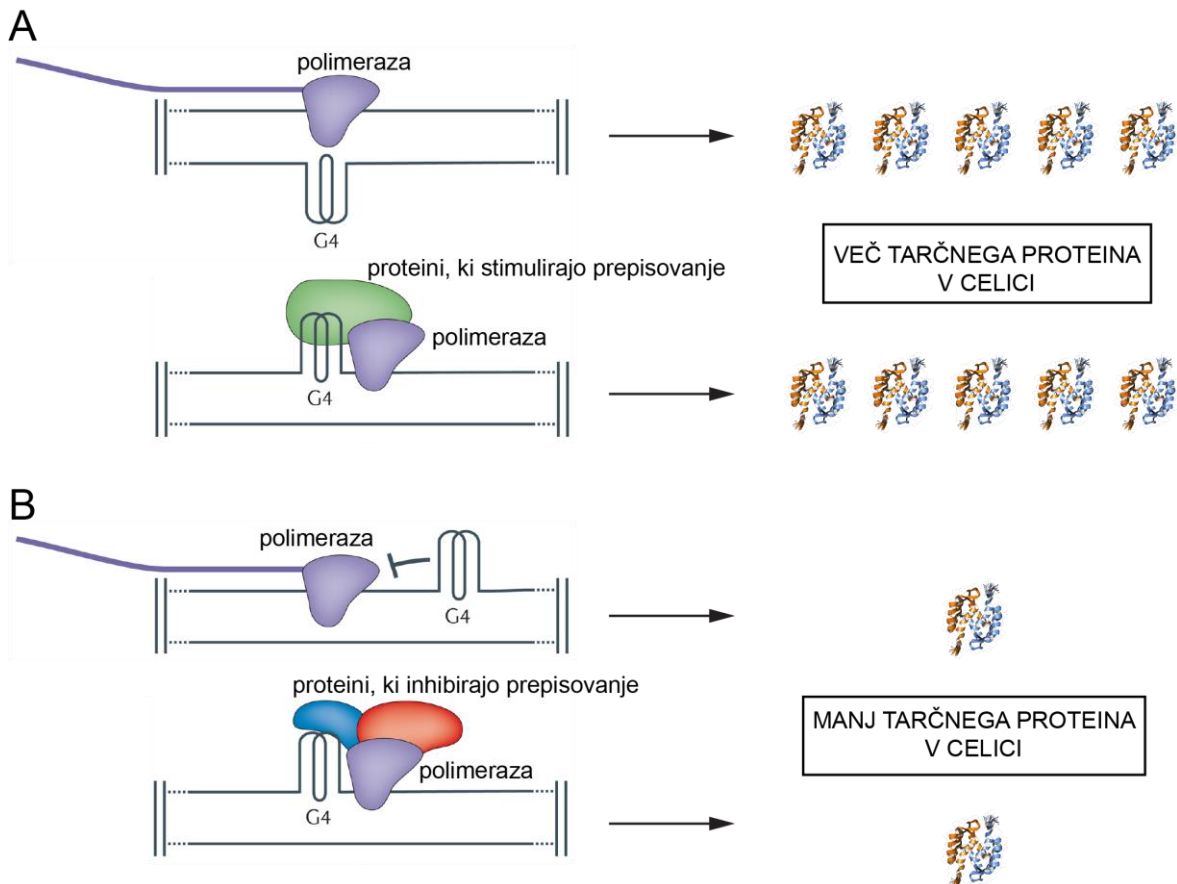
Slika 2. (A) Struktura G-kvarteta, kjer se štirje gvanini povežejo preko Hoogsteen vodikovih vezi. V sredini je koordiniran kation. (B) G-kvadrupleks s tremi naloženimi G-kvarteti, ki jih povezujejo različne zanke. (C) Shematski prikaz različnih topologij zvitja, ki jih lahko zavzamejo intramolekularni G-kvadrupleksi. Slika 2C je bila ustvarjena s programom BioRender (BioRender.com).

lelni strukturi G4 dve verigi potekata v eno smer, dve pa v drugo. Tak G4 nastane, če so G-kvarteti povezani s tremi robnimi ali dvema robnima in sredinsko diagonalno zanko. Kombinacija propelerske zanke z dvema robnima ali eno robno in eno diagonalno zanko vodi v nastanek hibridnega G4, kjer tri verige potekajo v eno smer, ena pa v nasprotno . (Bansal, 2022; Oh ,2020; Adrian, 2012 in Yang, 2019).

1.2.2.2 Z gvanini bogata zaporedja v genomu

Področja DNK, bogata z gvanini, niso naključno razporejena po genomu, ampak so zgoščena na koncih kromosomov ter v regulatornih regijah genov, kot so promotorske regije.

Promotor je nukleotidno zaporedje na 5'-koncu posameznega gena, na katerega se vežejo encimi in transkripcijski faktorji ter ključno vplivajo na proces genskega prepisovanja. Domnevno približno polovica človeških genov v svoji promotorski regiji vključuje zaporedje, bogato z gvanini in ima potencial za tvorbo G4. Vse več je dokazov, da lahko



Slika 3: (A) Tvorba G4 lahko zviša stopnjo prepisovanja, tako da s svojo prisotnostjo drži verigi DNK razprti in polimerazi olajša drsenje po verigi, ki jo prepisuje ali z vezavo proteinov, ki stimulirajo delovanje polimeraze. V obeh primerih v celici zaznamo povešano količino tarčnega proteina. (B) Tvorba G4 v promotorski regiji gena lahko blokira prepisovanje, tako da deluje kot ovira za polimerazo ali z vezavo proteinov, ki znižujejo stopnjo prepisovanja, kar vodi do zmanjšanja količine tarčnega proteina v celici. Slika je povzeta po Bochman (2012).

različni proteini selektivno in tesno vežejo G4 in igrajo bistveno vlogo v regulacijskih procesih. Nekateri od teh proteinov spodbujajo tvorbo in/ali stabilizacijo G4, drugi proteini odvijajo G4. Tvorba G4 znotraj promotorske regije tako lahko deluje kot magnet za vezavo regulatornih proteinov ali pa preprečujejo vezavo encimov in drugih proteinskih partnerjev. Na ta način pozitivno ali negativno vplivajo na izražanje posameznega gena, kar pomeni, da se v celici zaradi nastanka G4 tvori več ali manj tarčnega proteina (slika 3). To pa lahko v posameznih primerih vodi celo do pojavnosti bolezni, kot so različni tipi rakavih obolenj, nevrodegenerativne bolezni in osteoporoza (Bansal, 2022; Oh, 2020; Adrian, 2012 in Yang, 2019).

Eden temeljnih razlogov, da so DNK strukture, ki vsebujejo G-kvartete zanimive, je ta, da so zaporedja, bogata z gvanini, široko zastopana v vseh genomih, ki so bili do danes odkriti. Številne študije so pokazale, da se G4 tvorijo *in vivo* in igrajo pomembno vlogo v bioloških procesih, kot so podvojevanje DNK, transkripcija, rekombinacija, epigenetska regulacija, mejoza in vzdrževanje stabilnosti telomer. Zaradi njihove vpletenosti v ključne funkcije genoma danes pospešeno raziskujemo mehanizme tvorbe G4 in odkrivamo morebitne priložnosti za njihovo uporabo v terapevtske namene. (Spiegel, 2020). Z vidika odkrivanja zdravil so številni G4-vezavni proteini in/ali njihovi kompleksi z G4 potencialna zdravila. (Qiu, 2015).

1.2.2.3 Vpliv citozinov na tvorbo G4

Stabilnost posameznega G4 je odvisna od števila G-kvartetov, dolžine zank ter nastanka morebitnih dodatnih stabilizirajočih strukturnih elementov, ki jih tvorijo ostanki iz zank (na primer nalaganje ostankov iz zank na G-kvartete in tvorba baznih parov ali triad). Znanih je več primerov, kjer so pokazali, da interakcije ostankov iz zank občutno vplivajo na (termično) stabilnost posameznega G4 (Ngoc Nguyen, 2020; Lim, 2009).

Poleg razporeditve gvaninov pa lahko na tvorbo in stabilnost G4 vplivajo tudi citozini, ki so naravno prisotni v teh zaporedjih. V človeškem genomu so našli več kot 11000 z gvanini bogatih zaporedij, ki vsebujejo vsaj dva trakta s tremi zaporednimi citozini. Takšna mešana G/C bogata zaporedja lahko poleg G4 tvorijo tudi konkurenčne strukture, kot so lasnice na podlagi baznih parov G-C. Relativne stabilnosti obeh konkurenčnih struktur (G4 in lasnice) niso odvisne le od zaporedja, ampak tudi od ionskih pogojev. Medtem ko je tvorba G4 pogojena s prisotnostjo kationov v raztopini, tvorba lasnic načeloma ni odvisna od ionov. Raziskovalci so že uspeli dokazati, da je koncentracija K^+ ionov v rakastih celicah nižja, zaradi povečanega delovanja $[K^+]$ -kanalčkov v membranah (Tateishi-Karimata, 2018). To naj bi v rakastih celicah vodilo v destabilizacijo G4 in povečano transkripcijo onkogenov (to so geni, ki sprožijo nastanek rakaste celice). Višja koncentracija K^+ ionov v normalnih, zdravih celicah pa naj bi inhibirala transkripcijo onkogenov preko stabilizacije G4. Prisotnost citozinov bi lahko ključno vplivala na hitrejši preklon iz G4 v lasnico, kar bi lahko vodilo v povečano transkripcijo določenih genov.

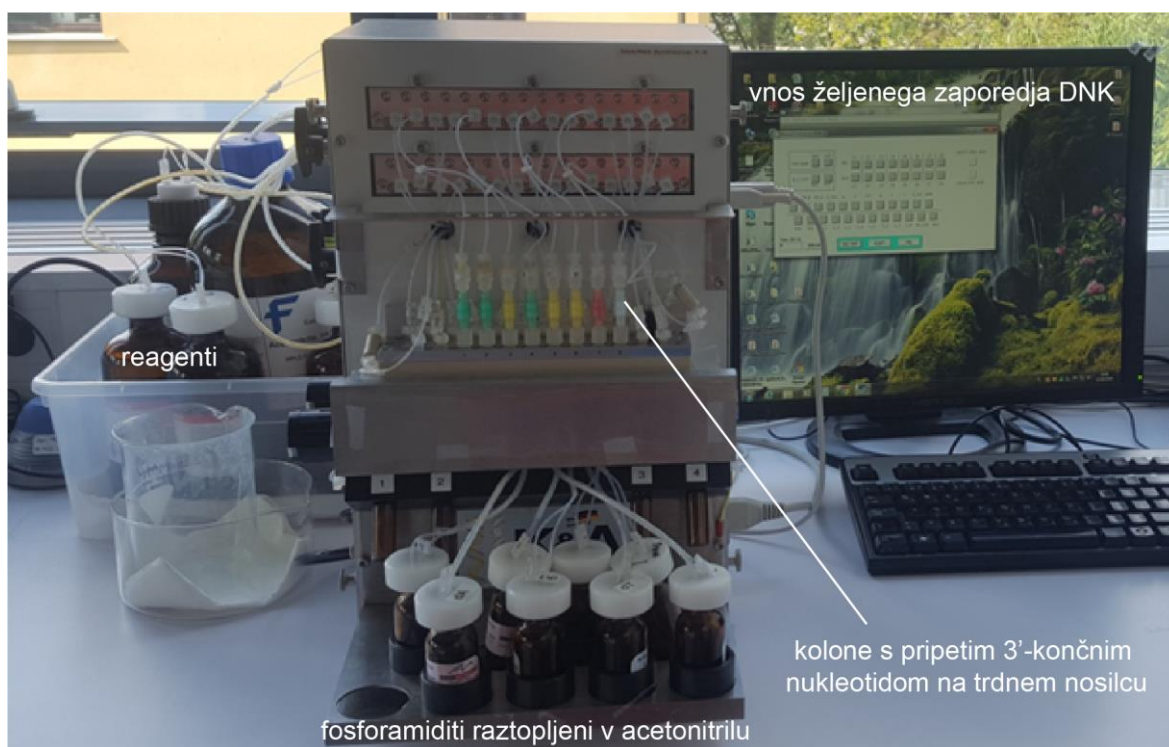
Za natančno poznavanje regulacije tvorbe G4 v primerjavi z lasnicami v celicah je ključnega pomena preučevanje modelnih sistemov, pri katerih lažje nadzorujemo vpliv različnih spremenljivk (na primer koncentracija in razmerje ionov, topologija preučevanega G4 in število citozinov). Pred kratkim so raziskovalci pokazali, da večanje števila citozinov v zankah modelnega G4 s samimi propelerskimi zankami (*t.i.* paralelni G4) ne prepreči tvorbe G4 in ne vpliva na stabilnost G4, ampak zviša možnost nastanka konkurenčne strukture –

lasnice (Luo, 2024). Zavedanje, da na tvorbo in stabilnost G4 vplivajo tudi drugi ostanki (predvsem citozini) se v zadnjih letih veča, vendar je naše znanje na tem področju še vedno pomanjkljivo.

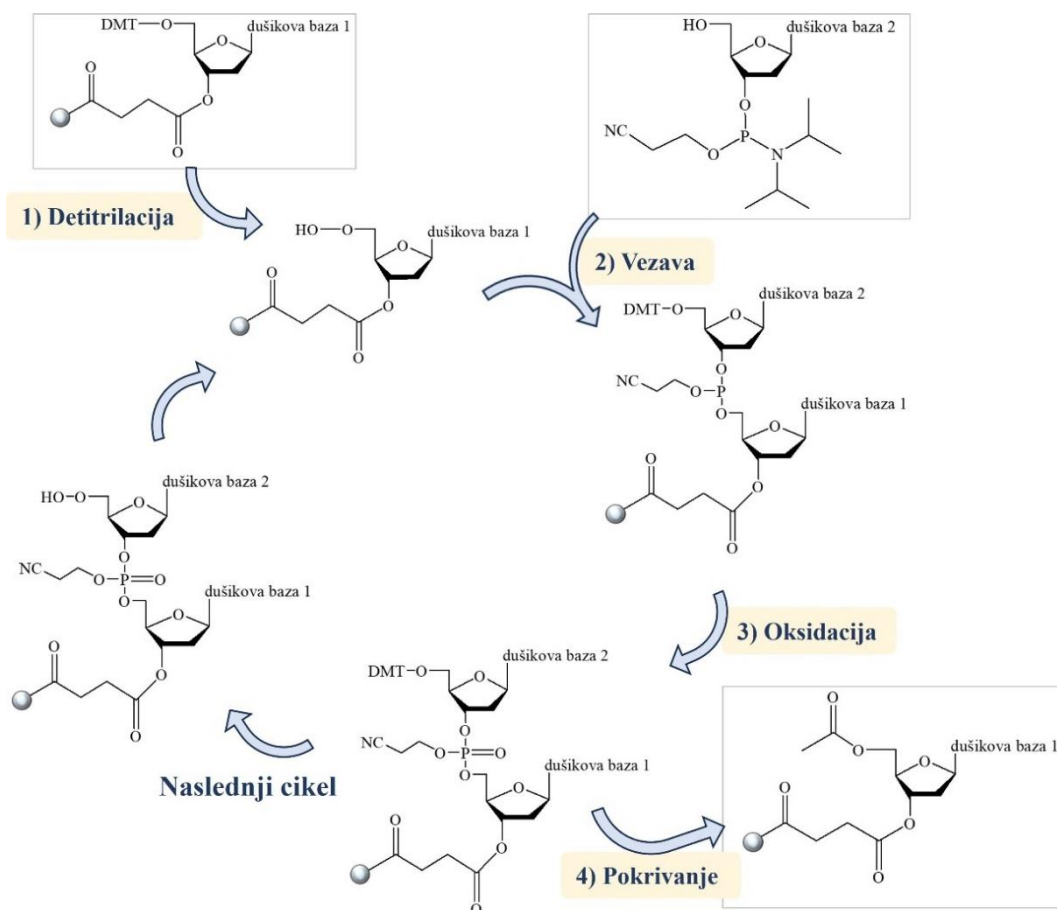
1.3 Sinteza oligonukleotidov

Zaporedja DNK, ki smo jih preučevali v projektni nalogi smo pripravili s kemijsko sintezo, ki temelji na fosforamiditni metodi na trdnem nosilcu. Kemijska sinteza oligonukleotidov, ki poteka na sintetizatorju (slika 4) se od encimske sinteze DNK v živih celicah razlikuje po tem, da gre za ciklično dodajanje fosforamiditov – nukleozidov z zaščitnimi skupinami, na trdni nosilec od 3'- proti 5'-koncu DNK (slika 5). Metoda je primerna za pripravo oligonukleotidov z dolžino do največ 200 nukleotidov. Ta tip kemijske sinteze fragmentov DNK uporabljamo med ostalim za določevanje strukture oligonukleotidov, za sinteza začetnih oligonukleotidov (primerjev) ali oligonukleotidov, ki jih uporabljamo kot sonde.

Serijski ventilni in črpalni postopoma uvajajo pravilne nukleotidne monomere in reagente, potrebne za nastajanje izbranega zaporedja DNK. Že pred začetkom kemijske sinteze DNK je začetni nukleozidni fosforamidit aktiviran s kislino in pritrjen na kroglico iz stekla s kontroliranimi porami ali polistirena (nosilec) (Masaki, 2022).



Slika 4: DNK/RNK sintetizator, ki ga uporabljajo na Nacionalnem centru za NMR spektroskopijo visoke ločljivosti na Kemijskem inštitutu v Ljubljani in temelji na fosforamiditni metodi na trdnem nosilcu.



Slika 5: Postopek sinteze oligonukleotida DNK s fosforamiditno metodo. Sinteza se začne z deprotekcijo (detritylacija) nukleotida na 3'-koncu, ki je preko povezovalca vezan na trdni nosilec (moder krogec). Sledi vezava naslednjega fosforamidita na prosto 5'-OH skupino ter oksidacija nastalega oligonukleotida. Nato poteče še acetilacija (pokrivanje) oligonukleotidov, na katere se fosforamidit ni vezal (Masaki, 2022; Merck, 2023).

Ciklično dodajanje dodatnih monomerov k obstoječi oligonukleotidni verigi poteka v štirih korakih: deprotekcija, vezava, oksidacija in pokrivanje. Vsak fosforamidit, ki ga dodamo nastajajoči verigi, ima vezano zaščitno skupino dimetoksitritil (DMT). Z odstranitvijo zaščitne skupine DMT izpostavimo 5'-OH skupino, kamor se v naslednjem koraku cikla veže naslednji fosforamiditni monomer. V naslednjem koraku nestabilni fosfitni triester z oksidacijo pretvorimo v stabilni fosfatni triester. V zadnjem koraku vse 5'-OH skupine, na katerih ni potekla reakcija, acetiliramo (pokrijemo), da preprečimo sintezo krajših, neželenih produktov. Opisani postopek ponavljamo dokler ne dosežemo željenega zaporedja. Končni rezultat tega procesa je izbrano oligonukleotidno zaporedje, vezano na nosilec (Masaki, 2022).

Po sintezi oligonukleotid z uporabo amonijevega hidroksida in metilamina odcepimo s trdnega nosilca in mu odstranimo zaščitne skupine. Vzorce dodatno očistimo in razsolimo s hitro proteinsko tekočinsko kromatografijo (FPLC)

1.4 FPLC

FPLC ali tekočinska kromatografija za hitro ločevanje proteinov je vrsta kromatografije, ki poteka pod srednje visokim tlakom. Stacionarno fazo sestavljajo majhne kroglice, ki so prečno povezane z agarozo in se nahajajo v steklenih ali plastičnih kolonah. Mobilno fazo predstavljajo vodne raztopine. Pretok mobilne faze skozi stacionarno uravnava črpalka. FPLC uporabljamo za ločevanje različnih makromolekul. Med drugim tudi za razsoljevanje DNK oligonukleotidov po sintezi in deblokiranju. Ker je od molekul v raztopini DNK oligonukleotid po velikosti največji, se izloči kot prva frakcija, ki jo zbiramo (Points, 2017).

1.5 NMR spektroskopija

Jedrska magnetna resonanca (NMR) je spektroskopska metoda, ki jo uporabljamo za določevanje struktur makromolekul. Uporabna je za vzorce v raztopini in trdnem stanju. Temelji na magnetnih lastnostih jeder. Elementi, ki jih zaznavamo z NMR morajo imeti jedrsko kvantno število (I) oziroma spin različen od 0. Najprimernejša so jedra z $I = \frac{1}{2}$, kot so na primer ^1H , ^{13}C , ^{15}N , ^{19}F in ^{31}P . Za večino meritev bioloških molekul zadostuje 600 MHz spektrometer (slika 6) (Yang., 2019; Marion, 2013).

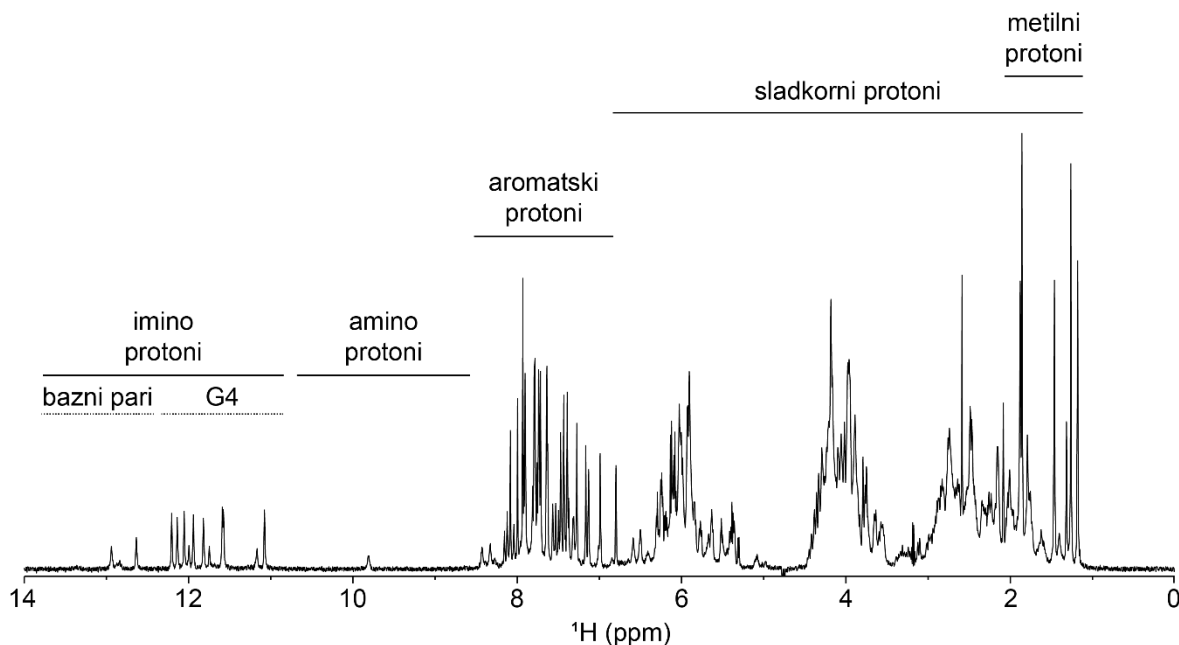


Slika 6: 600 MHz NMR spektrometer, postavljen na Nacionalnem centru za NMR spektroskopijo visoke ločljivosti na Kemijskem inštitutu v Ljubljani. Uporabljajo ga za strukturne študije in študije dinamike proteinov, DNK/RNK molekul, kompleksov proteinov in nukleinskih kislin ter majhnih molekul.

Iz NMR spektrov lahko razberemo informacije o kemijskih premikih (δ). Kemijski premik se nanaša na spremembo v frekvenci jedra atoma v molekuli, ki jo povzroča vpliv okolice tega atoma. Vsako jedro ima svojo naravno frekvenco, ki je odvisnega od magnetnega polja, v katerem se nahaja. Ko je atom v molekuli obkrožen z drugimi atomi, elektroni teh atomov vplivajo na magnetno okolje jedra, kar povzroči premik frekvence. Kemijski premik, ki se izraža v ppm (*ang.* parts per million), nam pove, koliko je frekvenca določenega jedra odmaknjena od referenčne frekvence. Na podlagi kemijskega premika torej pridobimo informacije o kemijskem okolju, v katerem se nahaja jedro (Yang, 2019; Liu, 2021).

S posebnim setom 1D in 2D (včasih tudi 3D) NMR eksperimentov lahko določimo tridimenzionalno strukturo opazovane molekule (Marion, 2013). NMR-spektroskopija makromolekul (nukleinskih kislin in proteinov) je tako uporabna za strukturne študije, interakcijske študije, študije dinamike in zvijanja ter druge (Yang, 2019 in Liu, 2021).

1D ^1H NMR eksperiment je najbolj osnoven NMR eksperiment, s katerim ocenimo, ali je vzorec v raztopini čist, hkrati pa z njim dobimo prve informacije o zvitju biomolekule. Tako lahko na podlagi protonskega NMR spektra hitro ocenimo, ali se je preučevani oligonukleotid zvil v G4 ali ne. Pri tem tipu NMR eksperimenta opazujemo signale v območju kemijskih premikov 0–14 ppm (slika 7). Imino protoni, ki sodelujejo v tvorbi G-kvartetov in so zato zaščiteni pred izmenjavo s topilom, resonirajo med 10 in 12 ppm. Iz števila signalov v imino območju lahko tudi sklepamo na število G-kvartetov, ki tvorijo G-kvadrupleks (Adrian, 2012; Yang, 2019).

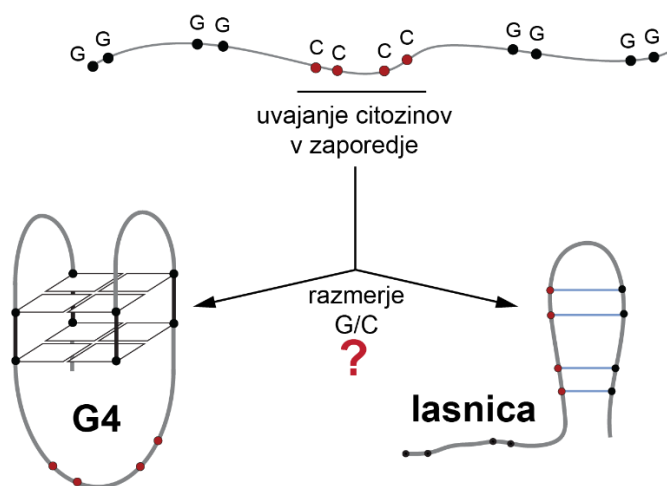


Slika 7: 1D ^1H NMR spekter. Spekter DNK oligonukleotida s koncentracijo 0,3 mM v vodni raztopini 10 mM kalijevega fosfatnega pufra, 50 mM KCL in 10 % D2O smo posneli na 600 MHz NMR spektrometru pri 25 °C. Nad spektrom so označena značilna območja, ki pripadajo določeni skupini protonov.

2. Cilji in problemi naloge

G4 so nekanonične, sekundarne strukture DNK, ki jih tvorijo z gvanini bogata zaporedja v promotorskih regijah genov. S svojo prisotnostjo direktno ali indirektno vplivajo na izražanje gena, kar občutno vpliva na (pato)fiziologijo celice. Vpliv posameznega G4 na izražanje gena je nemalokrat direktno povezano z njegovo stabilnostjo, ki je odvisna od števila G-kvartetov, dolžine zank ter dodatnih strukturnih elementov v zankah, ki lahko stabilizirajo G4. Poleg naštetega na tvorbo G4 (*in vitro* in v celici) vplivajo tudi ostanki, ki lahko znižujejo možnost nastanka teh štiriverižnih struktur – tak primer so na primer citozini, ki so naravno prisotni v teh zaporedjih. V človeškem genomu so našli več kot 11000 z gvanini bogatih zaporedij, ki vsebujejo vsaj dva trakta s tremi zaporednimi citozini. Takšna mešana G/C bogata zaporedja lahko namesto v tvorbo G4 vodijo v nastanek z G-C baznimi pari stabiliziranih lasnic.

Zaradi izredne raznolikosti struktur za vsak G4 posamezno preučujemo potek njegovega zvijanja, stabilnost ter vezavo na različne proteine. Zaradi želje po njihovi uporabi v terapevtske ali diagnostične namene se veliko pozornosti posveča tudi razvoju molekul z manjšo molekulsko maso (ligandov), ki bi specifično vezale G4 ter v celici sprožile željen odziv (stabilizacija G4 ali njegovo razvijanje). Vendar pa moramo za razvoj ustrezne učinkovine, ki bo v celici delovala specifično (oziroma bi bila specifična za en tip topologije G4) poznati čim več faktorjev, ki lahko vplivajo na njihovo tvorbo. Zavedanje, da na tvorbo in stabilnost G4 ključno vplivajo tudi drugi ostanki (predvsem citozini) se v zadnjih letih večja, vendar je naše znanje na tem področju še vedno pomanjkljivo.



Slika 8. Shema poteka dela projektne naloge.

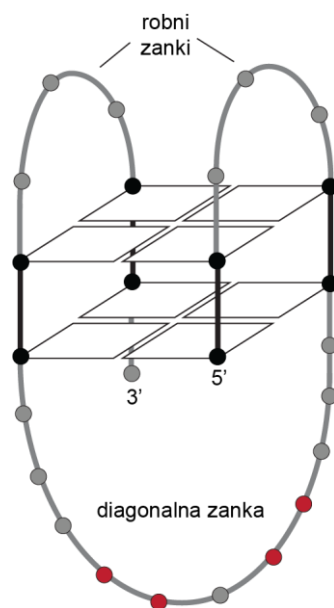
Pred kratkim raziskovalci pokazali, da večanje števila citozinov v zankah modelnega G4 s samimi propelerskimi zankami (*t.i.* paralelni G4) ne vpliva na stabilnost G4 in da tudi še ob razmerju G:C 12:6 v 40-140 mM KCl še vedno vodi do tvorbe G4. Že ob razmerju G:C 12:9 pa se je ne glede na uporabljeno koncentracijo K⁺ ionov v raztopini tvorila izključno lasnica,

stabilizirana z G-C baznimi pari (Luo, 2024). Paralelni G4 so termično najbolj stabilni med vsemi znanimi G4 strukturami.

Cilj projektne naloge je bil v modelno, z gvanini bogato zaporedje, ki v prisotnosti K^+ ionov tvori dvo-kvartetni antiparalelni G4 z dvema robnima in eno diagonalno zanko postopoma uvajati citozine ter z NMR spektroskopijo spremljati njihov vpliv na tvorbo G4 (slika 8). Citozine smo uvajali v dolgo, diagonalno zanko. Želeli smo preveriti, kakšno razmerje G:C (še) vodi v tvorbo G4. Dodatno nas je zanimalo, ali uvajanje citozinov v termično manj stabilno G4 strukturo (temperatura polovice prehoda $48\text{ }^{\circ}\text{C}$ v primerjavi s paralelnim G4, ki znaša okoli $80\text{ }^{\circ}\text{C}$) vpliva na hitrost preklapljanja v lasnico.

3. Hipoteze

V projektni nalogi smo se osredotočili na z gvanini bogato zaporedje, ki tvori dvo-kvartetni, antiparalelni G4 z dvema robnima in eno dolgo, diagonalno zanko (slika 9). V diagonalno zanko smo postopoma uvajali citozine in z NMR spektroskopijo spremljali tvorbo G4 in konkurenčne strukture – lasnice.



Slika 9: Topologija, ki jo zavzame modelno zaporedje v prisotnosti K^+ ionov, brez dodatnih citozinov, ki smo ga poimenovali KL.

Zastavili smo si tri sledeče hipoteze:

1. Število na novo uvedenih citozinov v zaporedje bo vplivalo na pretvorbo G4 v lasnico. Pretvorbo bomo opazili v imino regiji 1D 1H NMR spektra.
2. Prisotnost ostankov, ki v diagonalni zanki tvorijo dva G-C bazna para, bo vplivala na zvijanje konstruktov, vendar ne bo igrala ključne vloge pri pretvorbi G4 v lasnico.
3. Gvanin na 3'-koncu preučevanega zaporedja bo vplival na tvorbo sekundarnih struktur DNK v vzorcu, vendar ne bo odločilno vplival na prekop med G4 in lasnico.

4. Metode dela

4.1 Sinteza

Izbrane oligonukleotide smo s pomočjo programa vnesli v DNK/RNK sintetizator H-8 (K&A Laborgeräte). Izbrane DNK oligonukleotide smo pripravili z uporabo fosforamiditne metode sinteze na trdnem nosilcu. Modelno zaporedje KL ima na 3'-koncu gvanin, zato smo za sintezo uporabljali trdne nosilce z vezanim začetnim nukleozidom gvanozinom. Kot gradnike smo uporabili fosforamidite. Raztopili smo jih v predpisanem volumnu acetonitrila, da smo dobili ustrezne koncentracije. Dodali smo tudi ostale potrebne kemikalije (acetonitril, mešanico za deblokiranje, kemikalije za pokrivanje in oksidant).

Po dodajanju kemikalij smo na sintetizator namestili ustrezne kolone s trdnimi nosilci, in vnesli zaporedja oligonukleotidov. Po sintezi je na 5'-koncu oligonukleotida ostala vezana zaščitna skupina DMT.

4.2 Cepitev oligonukleotidov s trdnega nosilca in odstranjevanje zaščitnih skupin

Za cepitev sintetiziranih oligonukleotidov z nosilca smo najprej pripravili raztopino mešanice amonijevega hidroksida in metilamina (AMA), v razmerju 1:1. Z uporabo igle in brizge smo odmerili 1,5 mL sveže pripravljene mešanice. Na kolono z vzorcem smo na eno stran namestili prazno brizgo, na drugo pa napolnjeno brizgo. Vzorce smo s pritiski na brizgi premešali in pustili 15 minut. Po pretečenem času smo postopek ponovili ter jih namakali v raztopini AMA dodatnih 15 minut pri sobni temperaturi. Zatem smo jih 2–3x premešali in prenesli v epruvete ter jih 30 minut segrevali na 65 °C. Tako smo odstranili zaščitne skupine vezane na fosfatne skupine in dušikove baze.

4.3 Čiščenje oligonukleotidov s kartušami

Ohlajenim vzorcem iz prejšnjega koraka smo dodali po 1,5 mL raztopine NaCl s koncentracijo 100 mg/mL (končna koncentracija NaCl v vzorcih je znašala 50 mg/mL). Tako smo jih pripravili za čiščenje z Glen-Pak kartušami, ki smo jih namestili na vakuumski razdelilnik. Skupina DMT na oligonukleotidih interagira z materialom membran in tako omogoča vezavo vzorcev.

Z nanosom 0,5 mL ACN na kolono smo omočili membrane in z njih sprali ostanke organskih molekul. Z 1 mL TEAA smo olajšali vezavo oligonukleotidov na membrano. Po dodatku vzorca so se oligonukleotidi z DMT skupino na 5'-koncu vezali na kartuše, krajše pa smo s 5 % raztopino ACN v 100 mg/mL NaCl sprali v odpad. V naslednjem koraku smo z 2 % trifluoroocetno kislino (TFA) z vzorcev odcepili skupine DMT ter kartuše sprali z MiliQ vodo. Oligonukleotide smo s kartuš eluirali z nanosom 50 % acetonitrila v vodi z 0,5 % amonijevim hidroksidom, ki nevtralizira morebitne ostanke TFA.

Raztopine vzorcev smo nato zamrznili in jih čez noč sušili v vakuumskem liofilizatorju.

4.4 Čiščenje in razsoljevanje oligonukleotidov s FPLC

Liofilizirane vzorce iz prejšnjega koraka smo raztopili v 1 ml 100 mM raztopine LiCl. Pripravljene vzorce smo čistili in razsolili s hitro tekočinsko kromatografijo (FPLC), ki je prilagojena za delo z večjimi molekulami, kot so proteini in nukleinske kisline. Frakcije z vzorcem smo zbirali v epruvete in jih shranili v zamrzovalnik. Zamrznjene smo čez noč sušili v vakuumskem liofilizatorju.

4.5 Priprava vzorcev za NMR spektroskopijo

Posamezen vzorec smo raztopili v 300 μ L MiliQ vode. Koncentracijo oligonukleotidov smo določili z merjenjem absorbance pri valovni dolžini 260 nm (A_{260}) na UV spektrometru. Za merjenje koncentracije smo posamezen vzorec 500-krat redčili (2 μ L vzorca dodali v 998 μ L MiliQ). S spletnim orodjem Oligoanalyzer smo določili ekstinkcijske koeficiente izbranih zaporedij. Na podlagi tega smo z Beer-Lambertovim zakonom preračunali molsko koncentracijo vzorcev.

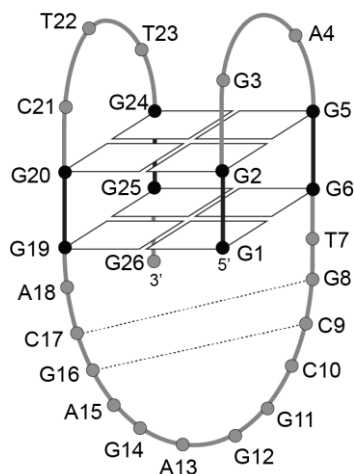
Vzorce smo zvalili v 10 mM kalijevem fosfatnem pufri (KPi, pH 7), 50 mM KCl in 10 % D₂O s končnim volumnom 600 μ L. Koncentracije oligonukleotidov v vzorcih so bile 0.3 mM. Pripravljene raztopine smo 5 minut segrevali v vodni kopeli na 95 °C in jih čez noč postopoma ohlajali do sobne temperature ($T = 25$ °C). Po zvijanju smo vzorce prenesli v 5 mm cevke za NMR spektroskopijo.

4.6 NMR spektroskopija

Za določanje sekundarne strukture sintetiziranih oligonukleotidov smo uporabili NMR spektroskopijo. Meritve smo izvajali na Bruker AVANCE NEO 600 MHz spektrometrih s 5 mm hladno sondo. Spektre smo referirali glede na zunanji NMR standard 2,2-dimetil-2-silapentan-5-sulfonsko kislino (DSS).

5. Rezultati

Kot izhodno zaporedje našega dela smo izbrali z gvanini bogato zaporedje, ki se nahaja v promotorski regiji gena *sost*. Delovanje gena *sost* je povezano z metabolizmom kosti, saj igra pomembno vlogo pri inhibiciji nastajanja kosti. Omenjeno z gvanini bogato zaporedje je dolgo 26 nukleotidov (5'-GGG AGG TGC CGG AGA GCA GGC TTG GG-3', SOST26).



Slika 10: Topologija zvitja izhodnega z gvanini bogatega zaporedja SOST26. Topologijo so predhodno določili na Nacionalnem centru NMR spektroskopije visoke ločljivosti v Ljubljani kot del diplomskega dela Anje Moškrič.

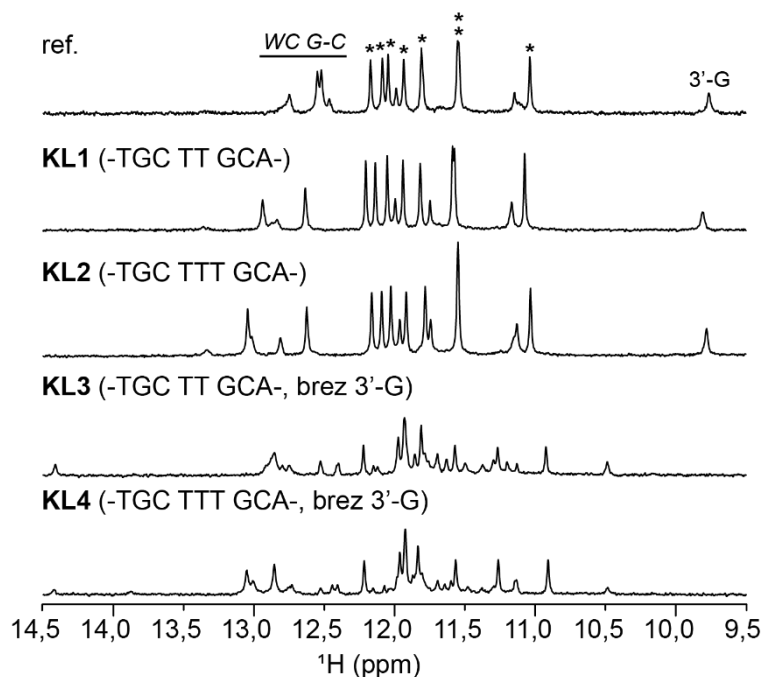
Strukturna študija, ki so jo predhodno izvedli na Nacionalnem centru za NMR spektroskopijo na Kemijskem inštitutu v Ljubljani, je pokazala, da oligonukleotid SOST26 tvori dvo-kvartetni, antiparalelni G4 (slika 10) (Moškrič, 2023). Čeprav v zaporedju najdemo 15 gvaninov, pri tvorbi dveh G-kvartetov sodeluje le osem od njih (odebeljeni v zaporedju zgoraj in prikazani s črno barvo na sliki 10). Prva in tretja zanka sta robni, vmesna zanka pa je diagonalna. Preučili so tudi vpliv ostankov iz 12 nukleotidov dolge diagonalne zanke (-TGC CGG AGA GCA-) na nastanek G4. Opazili so, da se v zanki tvorita dva G-C bazna para (G8-C17 in C9-G16), ki sta ključna za stabilizacijo osrednje strukture. Na zvitje vpliva tudi G26 s 3'-konca zaporedja, ki se najverjetneje naloži pod spodnji G-kvartet.

Vsa zaporedja, ki smo jih preučevali v projektni nalogi, so zbrana v tabeli 1. Ker so predhodni rezultati pokazali, da na zvijanje SOST26 v G4 vplivajo le prvi in zadnji trije ostanki iz diagonalne zanke (to so ostanki, ki so odebeljeni: -TGC CGG AGA GCA-), smo v prvem koraku svojega dela 6 nukleotidov, ki niso odebeljeni zamenjali s TT in TTT (tabela 1). Zaporedji smo poimenovali KL1 oziroma KL2. Dodatno smo pripravili še dva konstrukta, ki se od KL1 in KL2 ločita po tem, da smo jima odstranili 3'-končni gvanin (KL3 in KL4) in s tem preverili še njegov vpliv na zvijanje.

Tabela 1: Seznam z gvanini bogatih zaporedij, ki smo jih preučevali v projektni nalogi. Konstrukti, ki izhajajo iz drugih konstruktov so barvno usklajeni (na primer: KL9, KL10 in KL11 izhajajo iz KL2 in so zato obarvani z enako barvo).

Ime konstrukta	Zaporedje (5'-3')			Število nukleotidov	Število uvedenih citozinov	Razmerje G/C*
	5'-konec	diagonalna zanka	3'-konec			
KL1	GGG AGG	TGC TT GCA	GGC TTG GG	22	0	/
KL2	GGG AGG	TGC TTT GCA	GGC TTG GG	23	0	/
KL3	GGG AGG	TGC TT GCA	GGC TTG G	21	0	/
KL4	GGG AGG	TGC TTT GCA	GGC TTG G	22	0	/
KL5	GGG AGG	TT	GGC TTG GG	16	0	/
KL6	GGG AGG	TTT	GGC TTG GG	17	0	/
KL7	GGG AGG	TT	GGC TTG G	15	0	/
KL8	GGG AGG	TTT	GGC TTG G	16	0	/
KL9	GGG AGG	TGC TCTCT GCA	GGC TTG GG	25	2	3,00
KL10	GGG AGG	TGC TCCTCCT GCA	GGC TTG GG	27	4	2,00
KL11	GGG AGG	TGC TCCCTCCCCT GCA	GGC TTG GG	29	6	1,50
KL12	GGG AGG	TGC TCTCT GCA	GGC TTG G	24	2	2,75
KL13	GGG AGG	TGC TCCTCCT GCA	GGC TTG G	26	4	1,83
KL14	GGG AGG	TGC TCCCTCCCCT GCA	GGC TTG G	28	6	1,37
KL15	GGG AGG	TCTCT	GGC TTG GG	19	2	3,33
KL16	GGG AGG	TCCTCCT	GGC TTG GG	21	4	2,00
KL17	GGG AGG	TCCCTCCCCT	GGC TTG GG	23	6	1,43
KL18	GGG AGG	TCTCT	GGC TTG G	18	2	3,00
KL19	GGG AGG	TCCTCCT	GGC TTG G	20	4	1,80
KL20	GGG AGG	TCCCTCCCCT	GGC TTG G	22	6	1,28

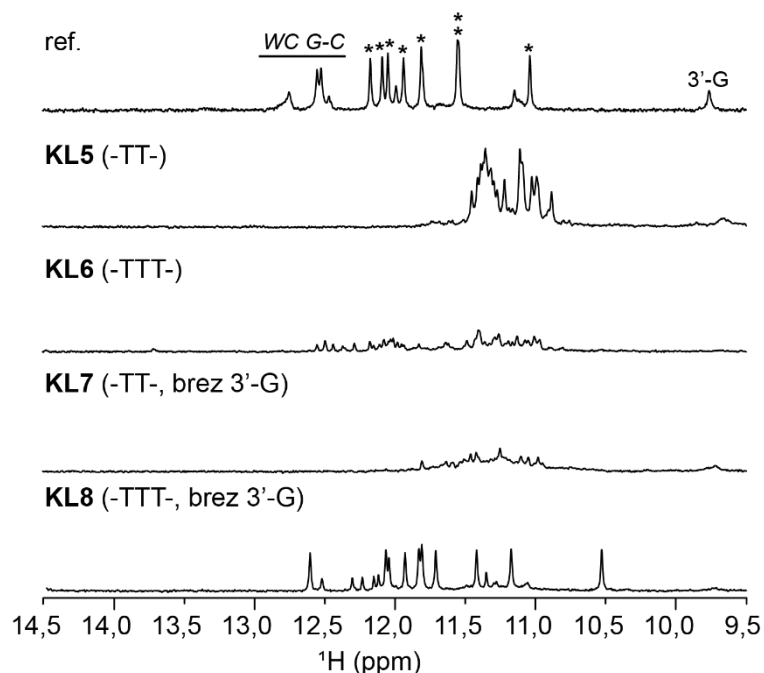
*Pri izračunu razmerja med gvanini in citozini smo upoštevali vse gvanine iz posameznega zaporedja (gvanin iz G-traktov, vključno s 3'-končnim gvaninom ter gvanina iz dveh G-C baznih parov) ter vse citozine (na novo uvedene citozine, citozina iz dveh G-C baznih parov ter citozin iz prve robne zanke). Četudi citozin iz prve robne zanke ne sodeluje pri tvorbi G4, lahko vpliva na pretvorbo G4 v lasnico.



Slika 11: Imino regije 1D ^1H NMR spektrov konstruktov KL1, KL2, KL3 in KL4 v primerjavi z referenčnim spektrom SOST26 (označen z ref.). Imino signali gvaninov, ki pripadajo dvema G-kvartetoma so označeni z zvezdico, signal pri približno 9,8 ppm pripada 3'-končnemu G (označen kot 3'-G), signali nad 12,5 ppm so značilni za Watson-Crick G-C bazne pare (označeni kot WC G-C).

Osredotočili smo se predvsem na analizo imino regij 1D ^1H NMR spektrov konstruktov KL1, KL2, KL3 in KL4 (slika 11). Medtem ko imata konstrukta KL1 in KL2 primerljivi imino regiji izhodnemu, referenčnemu spektru SOST26 (označen z ref.), se regiji KL3 in KL4 od njega močno razlikujeta. Pri KL1 in KL2 se ohrani *t.i. fingerprint* spektra, kar pomeni, da se konstrukta kljub uvedenim spremembam v zaporedju, zvijeta v enak G4. V primerjavi z referenčnim spektrom, opazimo enako število imino signalov, ki pripadajo gvaninom iz G-kvartetov in signale, ki so značilni za tvorbo G-C baznih parov. Po drugi strani spremembe v imino regiji NMR spektrov KL3 in KL4 nakazujejo, da 3'-končni gvanin poglavitno vpliva na njuno zvijanje in da delecija 3'-končnega gvanina ni ugodna, v kolikor imamo v zaporedju še vedno prisotne ostanke iz originalne diagonalne zanke (ostanki TGC in GCA).

Da bi preverili njihov vpliv na zvijanje modificiranih konstruktov, smo pripravili štiri konstrukte brez teh ostankov, kjer smo predpostavili, da bodo zanko tvorili le TT oziroma TTT z ali brez 3'-končnega gvanina. Konstrukte smo poimenovali KL5, KL6, KL7 in KL8 (tabela 1, slika 12). NMR analiza je razkrila, da uvedba -TT- zanke (konstrukt KL5) sicer vodi v tvorbo G4, vendar se njegova topologija po uvedbi sprememb močno spremeni. Predvidevamo, da se KL5 zvije v paralelni G4, kot je to značilno za G4 s krajšimi zankami in imino signali, ki resonirajo okoli 11 ppm. Uvedba zanke TTT (konstrukt KL6) podre antiparalelno G4 strukturo SOST26 in vodi v tvorbo več različnih G4, ki so v vzorcu zastopani v manjših odstotkih. Zato tudi v imino regiji vidimo več prekrivajočih signalov nizke intenzitete. Podobno opazimo tudi v primeru konstrukta KL7. Konstrukt KL8 z zanko TTT in brez 3'-končnega G ima podoben imino *fingerprint* kot referenčni spekter, zato predvidevamo, da ohrani antiparalelno G4 topologijo, ki je primerljiva s SOST26. Dodatni



Slika 12: Imino regije 1D ^1H NMR spektrov konstruktov KL5, KL6, KL7 in KL8 v primerjavi z referenčnim spektrom SOST26 (označen z ref.). Imino signali gvaninov, ki pripadajo dvema G-kvartetoma so označeni z zvezdico, signal pri približno 9,8 ppm pripada 3'-končnemu G (označen kot 3'-G), signali nad 12,5 ppm so značilni za Watson-Crick G-C bazne pare (označeni kot WC G-C).

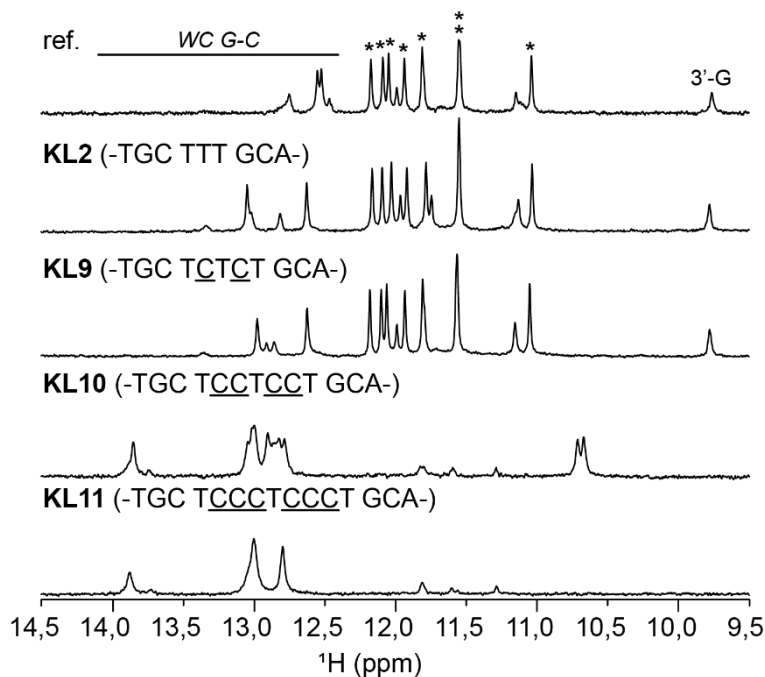
imino signali nižjih intenzitet, ki jih opazimo pri KL8 med 11,0 in 12,5 ppm najverjetneje pripadajo dodatni G4 strukturi, ki je v vzorcu manj zastopana.

V naslednjem koraku smo pričeli z uvajanjem citozinov v zaporedje. Kot izhodni zaporedji smo si izbrali KL2 in KL8, medtem ko sta nam zaporedji KL4 in KL6 služili kot kontroli. Z njima smo želeli preveriti, ali do tvorbe lasnice pride kljub zvitju, ki ne pripada izhodnemu G4.

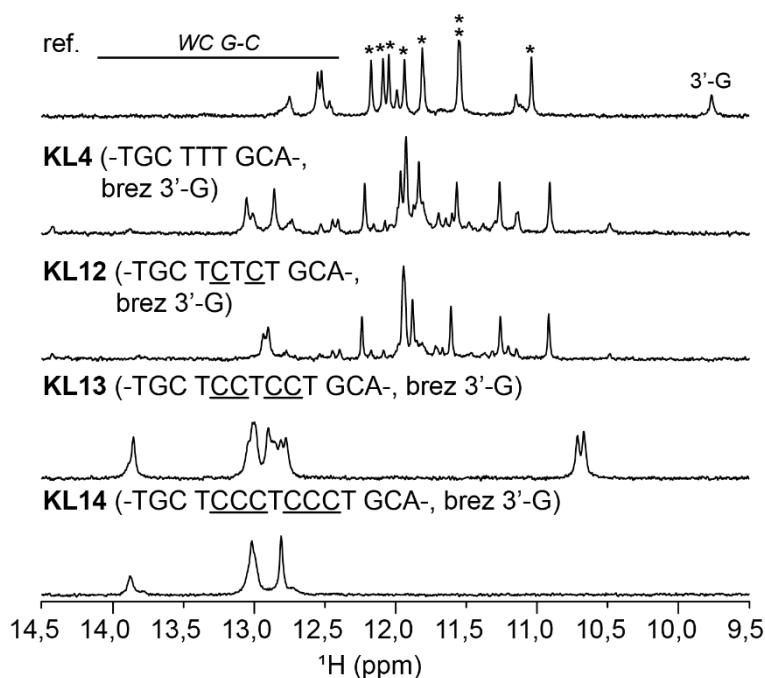
Citozine smo uvajali med tri timine v zanki v prisotnosti ali odsotnosti TGC in GCA ostankov ter v prisotnosti ali odsotnosti 3'-končnega gvanina. Na tak način smo pripravili dvanajst dodatnih konstruktov, ki so prikazani v tabeli 1. Barvna shema prikazuje, iz katerega konstrukta izhajajo posamezna zaporedja.

Konstrukt KL2 tvori antiparalelni G4 po enakem principu kot SOST26. Ko smo v zaporedje uvedli dva posamezna citozina (KL9), se je struktura G4 ohranila (slika 13). Že pri uvedbi dveh traktov dveh zaporednih citozinov (KL10) pa se je konstrukt zvil v lasnico. Tvorbo lasnice smo opazili tudi pri konstruktu KL11, kjer smo v zaporedje uvedli dva trakta treh zaporednih citozinov.

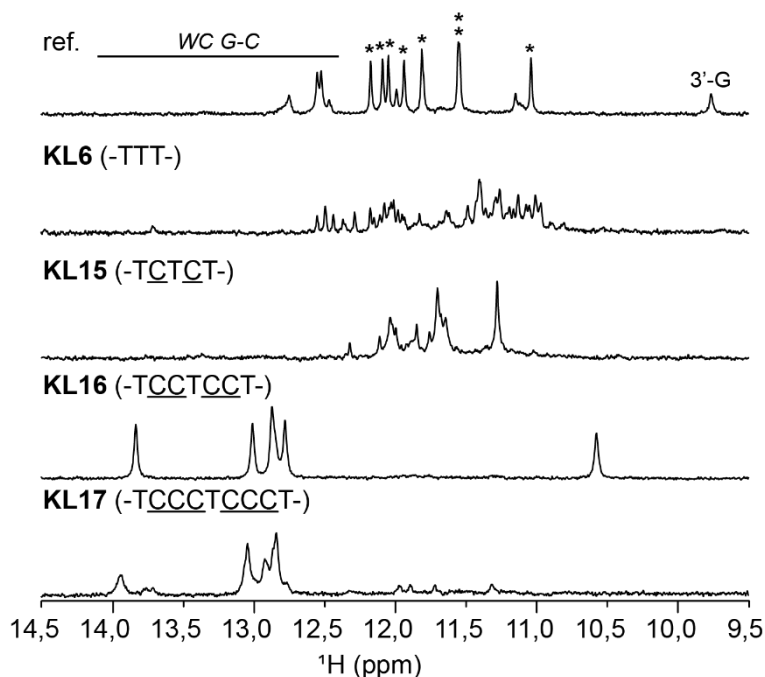
Preverimo smo tudi, ali na tvorbo lasnice vpliva odsotnost 3'-G oziroma odsotnost lepo definirane antiparalelne strukture G4. Podobno kot pri konstruktih na sliki 13, smo tudi tukaj pretvorbo med G4 in lasnico opazili, ko smo v zaporedje uvedli dva trakta dveh zaporednih citozinov (slika 14). Ne glede na zaporedje iz katerega smo izhajali (KL2 ali KL4), torej ali je v vzorcu bil prisoten izhodni G4, je uvedba treh traktov dveh ali treh citozinov privedla do tvorbe enake lasnice.



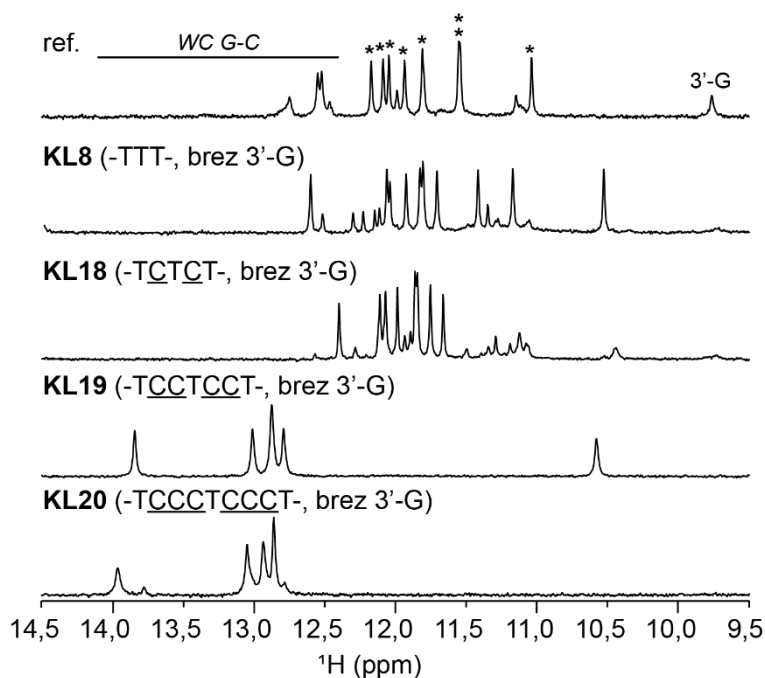
Slika 13: Imino regije 1D ¹H NMR spektrov konstruktov KL9, KL10 in KL11 v primerjavi s KL2 brez citozinov in z referenčnim spektrom SOST26 (označen z ref.). Imino signali gvaninov, ki pripadajo dvema G-kvartetoma so označeni z zvezdico, signal pri približno 9,8 ppm pripada 3'-končnemu G (označen kot 3'-G), signali nad 12,5 ppm so značilni za Watson-Crick G-C bazne pare (označeni kot WC G-C).



Slika 14: Imino regije 1D ¹H NMR spektrov konstruktov KL12, KL13 in KL14 v primerjavi s KL4 brez citozinov in z referenčnim spektrom SOST26 (označen z ref.). Imino signali gvaninov, ki pripadajo dvema G-kvartetoma so označeni z zvezdico, signal pri približno 9,8 ppm pripada 3'-končnemu G (označen kot 3'-G), signali nad 12,5 ppm so značilni za Watson-Crick G-C bazne pare (označeni kot WC G-C).



Slika 15: Imino regije 1D ^1H NMR spektrov konstruktov KL15, KL16 in KL17 v primerjavi s KL6 brez citozinov in z referenčnim spektrom SOST26 (označen z ref.). Imino signali gvaninov, ki pripadajo dvema G-kvartetoma so označeni z zvezdico, signal pri približno 9,8 ppm pripada 3'-končnemu G (označen kot 3'-G), signali nad 12,5 ppm so značilni za Watson-Crick G-C bazne pare (označeni kot WC G-C).



Slika 16: Imino regije 1D ^1H NMR spektrov konstruktov KL18, KL19 in KL20 v primerjavi s KL8 brez citozinov in z referenčnim spektrom SOST26 (označen z ref.). Imino signali gvaninov, ki pripadajo dvema G-kvartetoma so označeni z zvezdico, signal pri približno 9,8 ppm pripada 3'-končnemu G (označen kot 3'-G), signali nad 12,5 ppm so značilni za Watson-Crick G-C bazne pare (označeni kot WC G-C).

Delo smo nadaljevali z analizo konstruktov, pri katerih smo odstranili ostanke iz originalne zanke, torej brez ostankov –TGC- in –GCA-. Preučili smo šest konstruktov, kjer smo citozine uvajali med tri timine, ki so se na 5'- in 3'- koncu direktno dotikali G-traktov. Konstrukti so se med seboj razlikovali v številu uvedenih citozinov ter prisotnosti oziroma odsotnosti 3'-končnega G.

Konstrukt KL6 s prisotnim gvaninom na 3'-koncu zaporedja tvori več G4 struktur (slika 15). Uvedba dveh posameznih citozinov (KL15) vodi do tvorbe G4, vendar imino *fingerprint* ni podoben antiparalelnemu G4, iz katerega smo izhajali. Čeprav se KL6 in KL15 ne zvijeta v G4, ki bi bil podoben SOST26, uvedba dveh traktov dveh ali treh zaporednih citozinov (KL16 in KL17) vodi v tvorbo lasnice.

Za podobne konstrukte s krajšimi zankami, ampak brez 3'-končnega G, smo opazili, da uvajanje dveh posameznih citozinov (KL18) v večji meri vodi do tvorbe dvo-kvartetnega G4 (slika 16). V manjši meri se tvori(jo) tudi dodatni G4 (signali med 11,0 in 11,5 ppm). Zelo podobno kot pri ostalih konstruktih, uvedba dveh traktov dveh ali treh zaporednih citozinov (KL19 in KL20) vodi v tvorbo lasnice. Še enkrat smo potrdili, da ne glede na strukturo konstrukta brez uvedenih citozinov (KL6 ali KL8), dodatek dveh traktov dveh ali treh zaporednih citozinov vodi v tvorbo strukturno zelo podobne lasnice.

6. Razprava

Preučevali smo, kakšen vpliv imajo citozini na zvijanje izbranega, z gvanini bogatega zaporedja. Pripravili smo 20 različnih konstruktov, ki so se med seboj razlikovali v dolžini srednje, diagonalne zanke ter prisotnosti 3'-končnega gvanina. Vse konstrukte smo zvijali pod pogoji, ki vzpodbujajo tvorbo G4 – v prisotnosti K⁺ ionov. Izhajali smo iz 26 nukleotidov dolgega zaporedja SOST26, ki se zvije v dvo-kvartetni, antiparalelni G4 (slika 10) (Moškrič, 2023).

Predhodne študije so pokazale, da prvi in zadnji trije ostanki iz diagonalne zanke vplivajo na tvorbo željenega G4. Zato smo študijo začeli s konstrukti, kjer smo ohranili šest ostankov (TGA in GCA), dodali tri timine ter med njih dodajali citozine. Rezultati so pokazali, da se struktura G4 ohrani, če v zaporedje dodamo dva posamezna citozina (konstrukt KL9, slika 13). Dodatek novih dveh citozinov pa že vodi v tvorbo lasnice (KL10). Enako se zgodi, če v zaporedje uvedemo še dva citozina (KL11). Predvidevamo, da do spremembe pride zaradi spremembe v razmerju med gvanini in citozini v zaporedju. Za izračun razmerja G/C teh konstruktov smo upoštevali gvanine iz G-traktov vključno s 3'-končnim gvaninom in dva gvanina iz G-C baznih parov. V izračun smo upoštevali citozina iz G-C baznih parov ter na novo uvedene citozine. Tako smo za KL9 določili, da je razmerje med G/C = 12/4 = 3,00, za KL10 12/6 = 2,00 in za KL11 12/8 = 1,50. Iz tega lahko zaključimo, da do pretvorbe iz G4 v lasnico pride, ko razmerje G/C pade na 2 oziroma ko imamo v zaporedju 2 gvanina na posamezni citozin.

Podobno smo opazili za konstrukte, pri katerih smo sicer ohranili ostanke iz originalne zanke, vendar v zaporedje nismo vključili 3'-končnega gvanina. Uvedba dveh citozinov (KL12) še vodi v tvorbo G4, medtem ko pride do pretvorbe v lasnico, ko v zaporedje uvedemo še dva (KL13) ali štiri (KL14) dodatne citozine. Pri tem spreminjamo razmerje G/C z 2,75 (KL12) na 1,83 (KL13) oziroma 1,37 (KL14). Ti rezultati kažejo, da do pretvorbe iz G4 v lasnico pride, ko razmerje med gvanini in citozini pade pod 2.

Primerjava konstruktov, ki v zanki vsebujejo ostanke TGC in GCA, vendar se med seboj razlikujejo le po odsotnosti 3'-končnega gvanina, je pokazala, da 3'-končni gvanin odločilno vpliva na zvijanje izbranega z gvanini bogatega zaporedja (KL2 in KL4, slika 11). Hkrati pa rezultati kažejo, da to ne vpliva na pretvorbo zaporedja iz G4 v lasnico. Lasnica se torej tvori, ko v zaporedju dosežemo določeno razmerje med gvanini in citozini, neodvisno od izhodne strukture.

Preverili smo tudi, ali do pretvorbe prihaja enako hitro, če v zaporedju ni prisotnih ostankov iz diagonalne zanke, za katere so predhodno pokazali, da odločilno vplivajo na nastanek in stabilizacijo G4 (Moškrič, 2023). Iz NMR spektrov smo razbrali, da prisotnost 3'-končnega gvanina neugodno vpliva na tvorbo izbranega G4, če v zaporedju hkrati ni prisotnih še ostankov TGA in GCA iz diagonalne zanke (KL6 in KL8, slika 12). Hkrati je NMR analiza pokazala, da do pretvorbe iz G4 v lasnico pride takoj, ko v zaporedje uvedemo dva trakta dveh zaporednih citozinov, neodvisno od prisotnosti 3'-končnega gvanina (sliki 15 in 16). Opaženo še enkrat potrjuje dejstvo, da 3'-končni gvanin vpliva na tvorbo G4 v vzorcu, vendar ne igra ključne vloge pri pretvorbi med G4 in lasnico.

Izračunali smo razmerje med gvanini in citozini za konstrukte s krajšimi zankami (torej brez TGA in GCA ostankov). Z uvajanjem citozinov v zaporedje s 3'-končnim gvaninom razmerje G/C pada z 10/3 = 3,33 za KL15, v katerega smo uvedli dva posamezna citozina na 10/5 = 2,00 za KL16 s štirimi uvedenimi citozini in 10/7 = 1,43 za KL17 s šestimi uvedenimi citozini. Za zaporedja brez 3'-končnega gvanina in krajšimi zankami smo ugotovili, da razmerje G/C pada z 9/3 = 3,00 za KL18 na 9/5 = 1,80 za KL19 in 9/7 = 1,28

za KL20. Tudi ti rezultati kažejo, da do pretvorbe med G4 in lasnico pride, ko razmerje med gvanini in citozini pade pod določeno vrednost – v primeru konstruktov med KL15 in KL20 je to na 2,00 oziroma 1,80.

Primerjava NMR spektrov konstruktov z daljšimi (-TGC TC_nTC_nT GCA-) in krajšimi (-TC_nTC_nT-) zankami je razkrila, da prisotnosti ostankov TGC in GCA igra ključno vlogo pri zvijanju izbranega z gvanini bogatega zaporedja v G4 (sliki 11 in 12). Istočasno pa lahko iz rezultatov zaključimo, da preklapljanje iz G4 v lasnico poteka neodvisno od prisotnosti/odsotnosti teh ostankov. Na preklapljanje torej ne vplivajo potencialni G-C bazni pari, ki jih v zanki G4 izhodnega zaporedja tvorijo nekateri od teh ostankov in naj bi vplivali na stabilnost antiparalelnega G4, ki ga tvori SOST26. Ne glede na to, ali se v zaporedju nahajajo ostanki TGC in GCA, do pretvorbe iz G4 v lasnico pride, ko v zaporedju razmerje med gvanini in citozini pade pod določeno mejo in tudi strukturni elementi, ki stabilizirajo G4 (kot na primer G-C bazni pari) ne vplivajo na preklap.

V predstavljeni projektni nalogi smo dokazali, da prisotnost citozinov vpliva na tvorbo antiparalelnega G4 s sredinsko diagonalno zanko. Ne glede na dolžino zanke, prisotnost 3'-končnega gvanina ali zvijanje v pogojih, ki vzpodbujajo tvorbo G4, je izbrano, z gvanini bogato zaporedje po uvedbi dveh traktov dveh zaporednih citozinov tvorilo lasnico. Ugotovili smo, da do tvorbe lasnice prihaja, ko razmerje med gvanini in citozini pade na približno 2- *t.j.* ko imamo na približno vsaka dva gvanina, ki sodelujeta pri tvorbi G4, v zaporedju prisoten en citozin. Predvidevamo, da pomembno vlogo igra tudi to, da so uvedeni citozini v zaporedju skupaj in niso naključno razpršeni po zaporedju.

Eden izmed možnih razlogov, zakaj do preklopa med G4 in lasnico pri preučevanem zaporedju prihaja približno že pri razmerju G:C = 2:1, čeprav smo konstrukte zvijali v pogojih, ki vzpodbujajo tvorbo G4, je relativno nizka termična stabilnost preučevanega sistema. Predhodne študije so pokazale, da temperatura polovice prehoda (temperatura, pri kateri je polovica oligonukleotida v vzorcu zvitega, polovica pa razvitega) za SOST26 znaša približno 48 °C (Moškrič, 2023). Podobna študija, kjer so citozine uvajali v propelersko zanko termično mnogo bolj stabilnega paralelnega G4 (temperatura polovice prehoda čez 80 °C), je pokazala, da do preklopa ne glede na koncentracijo K⁺ ionov pride pri razmerju G/C = 1,33 (Luo, 2024). Pri razmerju G/C = 2,00 pa so lahko s spreminjanjem koncentracije K⁺ ionov še ujeli tvorbo G4. Ti rezultati kažejo, da z gvanini bogata zaporedja, ki so podobna SOST26 in imajo možnost zvitja v antiparaleln G4 z diagonalno zanko ter imajo v zaporedju vsaj dva trakta dveh zaporednih citozinov, v celici najverjetneje ne bi zavzela strukture G4, ampak bi se zvila v lasnico, kar lahko poglobitno vpliva na njihovo funkcijo v celičnih procesih.

7. Zaključek

Potrdili smo, da prisotnost dveh traktov dveh ali več zaporednih citozinov vpliva na tvorbo antiparalelnih G4 s srednjo, diagonalno zanko. Preklapljanje med G4 in lasnicami smo spremljali s pomočjo 1D ¹H NMR spektrov. Ugotovili smo, da prisotnost ostankov, ki v diagonalni zanki tvorijo dva G-C bazna para, vpliva na zvijanje konstruktov, vendar ne igra ključne vloge pri pretvorbi G4 v lasnico. Pokazali smo, da gvanin s 3'-konca zaporedja vpliva na tvorbo G4, vendar ne vpliva na preklap med G4 in lasnico. Do preklopa med G4 in lasnico pride približno pri razmerju G/C = 2, ne glede na dolžnico zanke ali prisotnost 3'-končnega gvanina. S tem smo potrdili vse tri hipoteze, ki smo si jih zastavili na začetku študije.

Ker je zvijanje G4 zelo kompleksno in so G4 strukturno polimorfna družina, bi lahko v prihodnosti izvedli podobno študijo še za z gvanini bogato zaporedje, ki se zvije v antiparalelni G4 s tremi robnimi zankami. S tem bi sistematično preučili, ali se dovzetnost za tvorbo lasnice razlikuje med različni tipi srednjih zank antiparalelnih G4. Delo bi lahko nadaljevali tudi v smeri iskanja z gvanini in citozini bogatega zaporedja, ki bi v prisotnosti K^+ ionov tvoril G4, v prisotnosti Na^+ ionov pa lasnico. Na ta način bi lahko razvili Na^+/K^+ senzor in G4 približali biotehnološki rabi.

8. Literatura

Dahm, R. (2008). Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. *Hum. Genet.*, 122 (6), 565-581.

Pray, L. (2008). Discovery of DNA structure and function: Watson and Crick. *Nature Education*, 1(1), 100.

Watson, J. D. in Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*, 171, 737-738.

Minchin, S., & Lodge, J. (2019). Understanding biochemistry: structure and function of nucleic acids. *Essays in biochemistry*, 63(4), 433-456.

Alberts, A. J. B. (2002). Julian Lewis, Martin Raff, Keith Roberts and Peter Walter. *Molecular Biology of the Cell*.

Britannica, T. Editors of Encyclopaedia (2024, February 1). DNA. *Encyclopedia Britannica*. <https://www.britannica.com/science/DNA>

Cleaves, H. J. J. (2015). Watson-Crick Pairing. *Encyclopedia of Astrobiology*, 2650-2650.

Bansal, M. (2003). DNA structure: Revisiting the Watson–Crick double helix. *Current Science*, 1556-1563

Tateishi-Karimata, H., & Sugimoto, N. (2020). Chemical biology of non-canonical structures of nucleic acids for therapeutic applications. *Chem. Comm.*, 56(16), 2379-2390.

Bansal, A., Kaushik, S., & Kukreti, S. (2022). Non-canonical DNA structures: Diversity and disease association. *Frontiers in Genetics*, 13, 959258.

Oh, K. I., Kim, J., Park, C. J., & Lee, J. H. (2020). Dynamics studies of DNA with non-canonical structure using NMR spectroscopy. *Int. J. Mol. Sci.*, 21(8), 2673.

Dalla Pozza, M., Abdullrahman, A., Cardin, C. J., Gasser, G., & Hall, J. P. (2022). Three's a crowd—stabilisation, structure, and applications of DNA triplexes. *Chemical Science*.

Spiegel, J., Adhikari, S., & Balasubramanian, S. (2020). The structure and function of DNA G-quadruplexes. *Trends in Chemistry*, 2(2), 123-136.

M. Adrian, B. Heddi, A. T. Phan (2012). NMR spectroscopy of G-quadruplexes. *Methods*, 57, 11–24.

D. Yang, C. Lin: G-Quadruplex Nucleic Acids: Methods and Protocols. 2035. *Methods in Molecular Biology* 2019.

Bochman, M. L., Paeschke, K., & Zakian, V. A. (2012). DNA secondary structures: stability and function of G-quadruplex structures. *Nature Reviews Genetics*, 13(11), 770-780

Qiu, J., Wang, M., Zhang, Y., Zeng, P., Ou, T. M., Tan, J. H., Huang, S. L., An, L. K., Wang, H., Gu, L. Q., Huang, Z. S., & Li, D. (2015). Biological Function and Medicinal Research Significance of G-Quadruplex Interactive Proteins. *Current topics in medicinal chemistry*, 15(19), 1971–1987.

T. Q. Ngoc Nguyen, K. W. Lim, A. T. Phan (2020) Duplex formation in a G-quadruplex bulge. *Nucleic Acids Res.*, 48, 10567–10575.

K. W. Lim, P. Alberti, A. Guédin, L. Lacroix, J. F. Riou, N. J. Royle, J. L. Mergny, A. T. Phan (2009). Sequence variant (CTAGGG)_n in the human telomere favors a G- quadruplex structure containing a G·C·G·C tetrad. *Nucleic Acids Res.*, 37, 6239–6248.

H. Tateishi-Karimata, K. Kawauchi, N. Sugimoto (2018). Destabilization of DNA G-quadruplexes by chemical environment changes during tumor progression facilitates transcription. *J. Am. Chem. Soc.* 140 (2), 642-651.

Luo, Y., Živković, M. L., Wang, J., Ryneš, J., Foldynová-Trantírková, S., Trantírek, L., & Mergny, J. L. (2024). A sodium/potassium switch for G4-prone G/C-rich sequences. *Nucleic Acids Research*, 52(1), 448-461.

Y. Masaki, Y. Onishi, K. Seio (2022). Quantification of synthetic errors during chemical synthesis of DNA and its suppression by non-canonical nucleosides. *Sci. Rep*, 12.

Merck KGaA: DNA Oligonucleotide Synthesis <https://www.sigmaaldrich.com/SI/en/technical-documents/technical-article/genomics/pcr/dna-oligonucleotide-synthesis> (pridobljeno 10. maja 2023).

H. G. Pontis (2017) Protein and Carbohydrate Separation and Purification. Methods for Analysis of Carbohydrate Metabolism in Photosynthetic Organisms. *Elsevier*, str. 45–63.

D. Marion (2013). An introduction to biological NMR spectroscopy. *Mol. Cell Proteomics*, 12, 3006–3025.

B. Liu, H. Shi, H. M. Al-Hashimi (2021). Developments in solution-state NMR yield broader and deeper views of the dynamic ensembles of nucleic acids. *Curr. Opin. Struct. Biol.*, 70, 16–25.

A. Moškrič (2023). Študija z gvanini bogatega zaporedja promotorske regije gena sost, povezanega z metabolizmom kosti: diplomsko delo. Ljubljana.