

Univerzita Pardubice
Fakulta chemicko-technologická

Korelace plochy pod píkem a koncentrací oligonukleotidu za použití HPLC
v závislosti na sekvenci a modifikaci oligonukleotidu

Diplomová práce

2023

Bc. Jana Marečková

University of Pardubice
Faculty of Chemical Technology

Correlation of Peak Area and Oligonucleotide Concentration Using HPLC
Depending on Oligonucleotide Sequence and Modification

Master's thesis

2023

Bc. Jana Marečková

Univerzita Pardubice
Fakulta chemicko-technologická
Akademický rok: 2020/2021

ZADÁNÍ DIPLOMOVÉ PRÁCE (projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Bc. Jana Marečková**
Osobní číslo: **C19540**
Studijní program: **N0914P360001 Bioanalytická laboratorní diagnostika ve zdravotnictví**
Studijní obor: **Bioanalytická laboratorní diagnostika ve zdravotnictví**
Téma práce: **Korelace plochy pod píkem a koncentrací oligonukleotidu za použití HPLC v závislosti na sekvenci a modifikaci oligonukleotidu**
Zadávací katedra: **Katedra biologických a biochemických věd**

Zásady pro vypracování

Teoretická část:

- 1) Syntéza oligonukleotidů
- 2) HPLC a její modifikace využívané pro analýzu a separaci oligonukleotidů
- 3) Stanovení koncentrace oligonukleotidů pomocí absorbance
- 4) Fyzikálně chemické vlastnosti oligonukleotidů a jejich vliv na HPLC separaci

Experimentální část:

- 1) výběr oligonukleotidů a příprava jejich roztoků pro experimentální porovnání
- 2) optimalizace metody pro stanovení absorbance v destičkovém formátu
- 3) HPLC analýza
- 4) korelace výsledků stanovených na základě absorbance a měření HPLC

Rozsah pracovní zprávy: **35 s.**
Rozsah grafických prací: **dle potřeby**
Forma zpracování diplomové práce: **tištěná**

Seznam doporučené literatury:

Cavaluzzi, M. J., & Borer, P. N. (2004). Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. *Nucleic Acids Research*, 32(1). <https://doi.org/10.1093/nar/gnh015>

Gerstein, A. S. (2003). Nucleotides, Oligonucleotides, and Polynucleotides. In *Molecular Biology Problem Solver* (pp. 267-289). John Wiley & Sons, Inc. <https://doi.org/10.1002/0471223905.ch10>

Hughes, R. A., & Ellington, A. D. (2017). Synthetic DNA Synthesis and Assembly: Putting the Synthetic in Synthetic Biology. *Cold Spring Harb Perspect Biol*, 9(a023812). <https://doi.org/10.1101/cshperspect.a023812>

Tataurov, A. V., You, Y., & Owczarzy, R. (2008). Predicting ultraviolet spectrum of single stranded and double stranded deoxyribonucleic acids. *Biophysical Chemistry*, 133(1-3), 66-70. <https://doi.org/10.1016/j.bpc.2007.12.004>

Vedoucí diplomové práce: **PharmDr. Antonín Libra, Ph.D.**
Generi Biotech s.r.o. Hradec Králové

Konzultant diplomové práce: **Ing. Zuzana Havlínová, Ph.D.**
Generi Biotech s.r.o. Hradec Králové

Datum zadání diplomové práce: **18. prosince 2020**

Termín odevzdání diplomové práce: **7. května 2021**

L.S.

prof. Ing. Petr Kalenda, CSc.
děkan

prof. Mgr. Roman Kandár, Ph.D.
vedoucí katedry

V Pardubicích dne 26. února 2021

Prohlašuji:

Práci s názvem Korelace plochy pod píkem a koncentrací oligonukleotidu za použití HPLC v závislosti na sekvenci a modifikaci oligonukleotidu jsem vypracovala samostatně. Veškeré literární prameny a informace, které jsem v práci využila, jsou uvedeny v seznamu použité literatury.

Byla jsem seznámena s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Beru na vědomí, že v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů, a směrnicí Univerzity Pardubice č. 7/2019 Pravidla pro odevzdávání, zveřejňování a formální úpravu závěrečných prací, ve znění pozdějších dodatků, bude práce zveřejněna prostřednictvím Digitální knihovny Univerzity Pardubice.

V Pardubicích dne 12. 4. 2023

Bc. Jana Marečková v. r.

Poděkování

Ráda bych touto formou poděkovala školitelce diplomové práce Ing. Zuzaně Havlínové, Ph.D. za odborné vedení praktické části, konzultace v laboratoři a realizaci experimentů. Dále mé poděkování patří vedoucímu práce PharmDr. Antonínu Librovi za konzultace při zpracování dat, nekonečnou trpělivost, čas a samotnou realizaci této diplomové práce. Děkuji též laborantkám a celé společnosti Generi Biotech s.r.o. za prostory, nezbytný materiál a především vlídný přístup.

ANOTACE

V diplomové práci řeším průkaz korelace mezi plochou pod píkem oligonukleotidu při HPLC purifikaci a jeho výslednou koncentrací s ohledem na sekvenci a modifikaci, kdy je snaha stanovit koeficient predikce. To by v ideálním případě mohlo vést k eliminaci stanovení obsahu oligonukleotidu po jeho purifikaci. Teoretická část předkládá potřebné pozadí pro provedení experimentu. Součástí práce je dále HPLC analýza a měření absorbance na spektrofotometru, stejně tak statistické zpracování získaných dat.

KLÍČOVÁ SLOVA

oligonukleotidy, duálně značené sondy, purifikace, HPLC, absorbance, koncentrace, Lambert-Beerův zákon

ANNOTATION

The thesis focus on proving a correlation between the peak area at HPLC purification of the oligonucleotide and its final concentration with respect to sequence and modification, which is to determine the coefficient of prediction. Ideally this could lead to the elimination of determining the content of the oligonucleotide after its purification. The theoretical part outlines the necessary background for the experiment. The thesis also includes HPLC analysis and measurement of absorbance via spectrophotometer, as well as statistical processing of the obtained data.

KEY WORDS

oligonucleotides, dual-labeled probes, purification, HPLC, absorbance, concentration, Beer-Lambert's law

SEZNAM ILUSTRACÍ

Obrázek 1: Centrální dogma molekulární biologie.....	17
Obrázek 2: H-fosfonátová metoda syntézy oligonukleotidu	20
Obrázek 3: Fosfodiesterová metoda syntézy oligonukleotidu	20
Obrázek 4: Fosfotriesterová metoda syntézy oligonukleotidu	21
Obrázek 5: Detritylace.	23
Obrázek 6: Kondenzace.	24
Obrázek 7: Capping.	24
Obrázek 8: Oxidace.	25
Obrázek 9: Obecný náčrt kapalínového chromatografu.	30
Obrázek 10: Definice torzních úhlů nukleotidů.....	35

SEZNAM ROVNIC

Rovnice 1: Výpočet metody nejbližšího souseda pro výpočet molární absorpčního koeficientu polynukleotidu.	27
Rovnice 2: Výpočet směrodatné odchylky.	49
Rovnice 3: Výpočet relativní směrodatné odchylky.	49
Rovnice 4: Výpočet parametrů Grubbsova testu.	50
Rovnice 5: Výpočet parametru F-testu	50
Rovnice 6: Rovnice popisující Lambert-Beerův zákon transformovaná pro přímý výpočet koncentrace.	50
Rovnice 7: Rovnice popisující Lambert-Beerův zákon.	51
Rovnice 8: Zjednodušená rovnice Lambert-Beerova zákona.	51
Rovnice 9: Transformovaná rovnice s dosazením koncentrace jako podílu počtu molů a objemu.	51
Rovnice 10: Rovnice lineární regrese.	51
Rovnice 11: Výpočet množství oligonukleotidu či sondy ze získané plochy píku.	51

SEZNAM GRAFŮ

Graf 1: Vliv úpravy křivky gradientu na rozlišení separace oligonukleotidů.....	38
Graf 2: Vliv teploty na rozlišení separace oligonukleotidů.	38
Graf 3: Vliv průtoku a rychlosti dosažení gradientu na rozlišení separace oligonukleotidů....	38
Graf 4: Chromatograf purifikace duálně značené sondy a absorpční spektra látek.....	40
Graf 5: Kalibrační řady oligonukleotidů.....	56
Graf 6: Kalibrační řady sond	56
Graf 7: Závislost absorbance vynásobené objemem na ploše píku pro oligonukleotidy.....	58
Graf 8: Závislost absorbance vynásobené objemem na ploše oligonukleotidů.....	60
Graf 9: Vyřazené oligonukleotidy.	61
Graf 10: Porovnání množství oligonukleotidu získaného purifikací a predikovaného	65
Graf 11: Závislost absorbance vynásobené objemem na ploše všech sond.....	67
Graf 12: Závislost absorbance vynásobené objemem na ploše jednotlivých sond.....	68
Graf 13: Porovnání množství sondy získaného purifikací a predikovaného	71
Graf 14: Zobrazení vlivu značení sond u vybraných predikčních modelů.	72

SEZNAM TABULEK

Tabulka 1: Hodnoty absorpčního koeficientu pro vybrané nukleotidy, sondy a zhášeče.....	28
Tabulka 2: Hodnoty molárního absorpčního koeficientu pro 100-mery.	28
Tabulka 3: Hodnoty pK_A majoritních bází při 20 °C a nulové koncentraci solí.	35
Tabulka 4: Definice torzních úhlů nukleotidů.	36
Tabulka 5: Vybrané oligonukleotidy.	42
Tabulka 6: Vybrané sondy.	43
Tabulka 7: Základní fyzikální vlastnosti vybraných oligonukleotidů.	44
Tabulka 8: Koncentrace zásobních roztoků oligonukleotidů.	53
Tabulka 9: Koncentrace zásobních roztoků sond.	53
Tabulka 10: Naměřené plochy a výšky píku pro validaci opakovatelnosti.	55
Tabulka 11: Výsledky validace opakovatelnosti.	55
Tabulka 12: Přehled výsledků statistických nástrojů oligonukleotidů.	62
Tabulka 13: Přehled výsledků statistických nástrojů sond.	68

SEZNAM ZKRATEK A ZNAČEK

A	adenin
ACN	acetonitril
AdCl	adamantankarbonylchlorid
ADP	adenosindifosfát
AI	artificial intelligence (umělá inteligence)
C	cytosin
CPG	controlled pore glass (porézní sklo)
Cryo-EM	kryoelektronová mikroskopie
DMOCP	5,5-dimethyl-2-oxo-2-chloro-1,3,2-dioxafosforinan
DMT	4,4'-dimethoxytrityl
DNA	deoxyribonukleová kyselina
dNTP	deoxyribonukleotidtrifosfát
DP	deep learning (hluboké učení)
G	guanin
HPLC	high-performance liquid chromatography (vysokoučinná kapalinová chromatografie)
IDMS	isotope dilution mass spektrometry (hmotnostní spektroskopie s izotopovým zředováním)
ML	machine learning (strojové učení)
MS-Cl	mesitylen-2-sulfonylchloridu
NMR	nukleární magnetická rezonance
NP-HPLC	normal-phase HPLC (HPLC s normálními fázemi)
OPC	oligonucleotide purification cartridge (cartridge pro purifikaci oligonukleotidů)
OMP	bis(2-oxo-3-oxoazolidinyl)-fosfinchlorid

PAGE	polyacrylamide gel electrophoresis (elektroforéza v polyakrylamidovém gelu)
PvCl	pivaloylchlorid
PVDF	polyvinylidenfluorid
qPCR	quantitative polymerase chain reaction (kvantitativní polymerázová řetězová reakce)
RNA	ribonukleová kyselina
RP-HPLC	reversed-phase HPLC (HPLC s reverzními fázemi)
SAXS	small-angle X-ray scattering
SE-HPLC	size-exclusion HPLC (rozměrově vylučovací HPLC)
T	thymin
TdT	terminální deoxynukleotidtransferáza
TEAA	triethylamoniumacetát
TPS-Cl	triisopropylbenzensulfonylchloridu
U	uracil
UV	ultrafialové

OBSAH

ÚVOD.....	16
1 TEORETICKÁ ČÁST	17
1.1 Nukleové kyseliny.....	17
1.1.1 Struktura nukleových kyselin	17
1.1.2 Oligonukleotidy	19
1.2 Syntéza oligonukleotidu.....	19
1.2.1 H-fosfonátová metoda.....	19
1.2.2 Fosfodiesterová metoda	20
1.2.3 Fosfotriesterová metoda.....	21
1.2.4 Fosfit-triesterová metoda	21
1.2.5 Fosforamiditová metoda	21
1.2.6 Enzymatická metoda.....	25
1.2.7 Modifikace oligonukleotidů.....	26
1.3 Kvantifikace oligonukleotidů.....	26
1.3.1 UV absorpční spektroskopie.....	26
1.3.2 Další metody stanovení koncentrace nukleotidů	29
1.4 Purifikace	30
1.4.1 Rozdělovací chromatografie	30
1.4.2 HPLC s normálními fázemi	31
1.4.3 Iontově výměnná chromatografie	31
1.4.4 Bioafinitní chromatografie.....	31
1.4.5 Rozměrově vylučovací HPLC	31
1.4.6 HPLC s reverzními fázemi	32
1.4.7 Isokratická a gradientová eluce.....	32
1.4.8 Purifikace oligonukleotidů.....	33
1.5 Fyzikálně chemické vlastnosti oligonukleotidů.....	34
1.5.1 Fyzikální vlastnosti	34
1.5.2 Chemické vlastnosti.....	34

1.5.3	Konformační vlastnosti.....	35
1.5.4	Predikce vlastností	37
1.5.5	Vliv na HPLC separaci	37
2	CÍL PRÁCE	41
3	EXPERIMENTÁLNÍ ČÁST	42
3.1	Metodika	42
3.1.1	Výběr oligonukleotidů a sond.....	42
3.1.2	Definice základních vlastností vybraných vzorků	43
3.1.3	Stanovení koncentrace zásobních roztoků	45
3.1.4	HPLC analýza	46
3.1.5	Měření absorbance purifikátů	48
3.1.6	Zpracování a analýza dat	48
4	VÝSLEDKY	53
4.1	Stanovení koncentrace zásobních roztoků	53
4.2	HPLC analýza	54
4.2.1	Validace opakovatelnosti	54
4.2.2	Výběr vhodných koncentračních bodů:	55
4.3	Analýza oligonukleotidů	57
4.3.1	Hromadná analýza dat	57
4.3.2	Analýza jednotlivých oligonukleotidů.....	59
4.3.3	Porovnání modelu predikce s reálnými výsledky	65
4.4	Výsledky analýzy sond	66
4.4.1	Hromadná analýza dat	66
4.4.2	Porovnání modelu predikce s reálnými výsledky	71
4.4.3	Hodnocení vlivu značení oligonukleotidu	71
5	DISKUZE A ZÁVĚR	73
	POUŽITÁ LITERATURA	77

ÚVOD

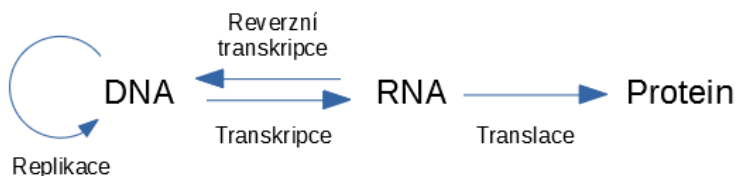
Oligonukleotidy jsou známé, syntetizované a využívány již několik desítek let. Jsou výchozí surovinou velkého množství molekulárně-biologických postupů a v současnosti nalézají uplatnění prakticky v každé větší laboratoři biologického směru. S téměř exponenciálním rozvojem těchto metod roste i poptávka po oligonukleotidech, je tedy cílem zrychlit, zlevnit a celkově zefektivnit jejich syntézu.

Důležitým posyntetizačním krokem pro výslednou dobrou kvalitu produktu je purifikace. V současnosti je několik možností purifikace, ovšem nejvyužívanější je kapalinová chromatografie, která odseparuje produkt od balastu, za níž následuje následuje přesné stanovení koncentrace čistého oligonukleotidu, zpravidla spektrofotometricky. Stanovit koncentraci lze i pomocí velikosti plochy píku chromatografu, avšak je zapotřebí buď kalibrační křivky, nebo standardu. Tato práce si klade za cíl zjistit možnou korelaci mezi měřenou plochou píku kapalinové chromatografie a výslednou koncentrací čistého oligonukleotidu za použití kapalinové chromatografie v závislosti na sekvenci a modifikaci oligonukleotidu.

1 TEORETICKÁ ČÁST

1.1 Nukleové kyseliny

Vedle polypeptidů a polysacharidů patří mezi základní biomakromolekuly života i nukleové kyseliny. Tyto snad nejdůležitější biopolymery, jejichž hlavní úloha je nést genetickou informaci, jsou nezbytnou součástí všech živých organismů. Samotný pojem nukleové kyseliny pak označuje souhrnný název pro deoxyribonukleové kyseliny (DNA) a kyseliny ribonukleové (RNA). Vztah mezi DNA a RNA a jejich roli v organismu popisuje centrální dogma molekulární biologie, které se jako první pokusil zformulovat roku 1957 Francis Crick. Dnešní výklad centrálního dogmatu tvrdí, že DNA je nositel informace, která determinuje transkripcí primární strukturu RNA, a ta je poté určující pro primární strukturu proteinu (viz obr. 1). [1, 2]



Obrázek 1: Centrální dogma molekulární biologie.

1.1.1 Struktura nukleových kyselin

Základním stavebním kamenem nukleových kyselin jsou nukleové báze, purinové a pyrimidinové. DNA obsahuje adenin (A), guanin (G), cytosin (C) a thymin (T), v RNA je thymin nahrazen uracilem (U). Vedle těchto majoritních bází se vyskytují i báze minoritní, často se jedná o metylované majoritní báze, dále pak v DNA hydroxymethylcytosin (u bakteriofágů) a methylcytosin (v eukaryotech) a v RNA např. dimethyladenin, hypoxanthin, isopentenyladenin, kinetin, zeatin, pseudouracil a další. Plnohodnotná monomerová jednotka dále obsahuje cukr-fosfátovou kostru skládající se ze sacharidu ribosy nebo deoxyribosy a zbytku kyseliny fosforečné, kterým se váže k další monomerní jednotce, fosforylovanému nukleosidu, tedy nukleotidu, fosfodiesterovou vazbou za vzniku lineárního polymeru. [3, 4]

Vedle primární struktury, kterou v případě nukleových kyselin určuje pořadí nukleotidů, se tvoří i sekundární, terciární a případně kvartérní struktury, vazby mohou být jak intra-, tak intermolekulární a uplatňují se především vodíkové můstky, hydrofobní

interakce a van der Waalsovy síly. Za samotnou polynukleotidovou konformaci jsou zodpovědně především hrany bází, které obsahují $-\text{NH}_2$, $=\text{N}-$ a $=\text{O}$ skupiny, a mohou tak interagovat s dalšími polárními skupinami nebo molekulami vody. Na rozdíl od hran se plochy heterocyklů tohoto účastnit nemohou a mají naopak tendenci se vodě vyhýbat. Na základě tohoto jevu dochází k tzv. patrové interakci (angl. base stacking), kdy se zmenší hydrofobní plocha a kde se uplatňují π - π interakce. Vedle patrové interakce se pro zvýšení stability polynukleotidu uplatňuje i párování bází dle Watson-Crickovských pravidel komplementarity, tzn. A se páruje s T nebo U dvěma vodíkovými můstky, G s C třemi. Existují i další alternativní párování, např. Hoogsteenovské párování, kdy se A váže N7 dusíkem k T na jeho N3 dusík a NH_2 skupinou na thyminovou ketoskupinu. U guaninu a cytosinu se G váže obdobně N7 dusíkem na N3 atom C. Mohou vznikat i další nepřírodní páry především v souvislosti s výskytem minoritních bází. [1, 5, 6]

DNA je co se týče sekundární struktury méně rozmanitější než RNA. Dvouvláknová DNA téměř vždy zaujímá helikální strukturu, přičemž vlákna jsou antiparalelní. Rozlišujeme tři základní dvojšroubovicové uspořádání a to B-, A- a Z-DNA. Zdaleka nejčastější uspořádání je pravotočivá B-DNA. Levotočivá Z-DNA vzniká z B- formy, hraje pravděpodobně roli v ovlivnění genové exprese a její výskyt je spojován s různými onemocněními jako je Alzheimerova choroba nebo systémový lupus erythematosus. A- forma vzniká dehydratací B-formy. Obecně u dvoušroubovic rozeznáváme velký a malý žlábek, oba jsou důležité např. pro expresi genu. Dále mohou být přítomny ohyby helixu nebo tzv. kruciformní forma. Vedle dvouvláknové se může tvořit i trojvláknová šroubovice, případně čtyřvláknová forma G-kvadruplex. [1, 6–10]

RNA je strukturně mnohem bohatší. Sekundární struktura je odvozena od Watson-Crickovského párování bází s interferujícími nepárujícími oblastmi, dále se mohou objevovat nestandardní páry jako je G-U a další intramolekulární interakce. Podobně jako u proteinů je funkčnost RNA v některých případech podmíněna i její 3D strukturou. Tyto funkční strukturní prvky se nazývají strukturní motivy a je z nich složeno až 40 % vlákna RNA. Mezi nejběžnější prvky strukturních motivů patří helixy (A-forma), smyčky, vlásenky, vícehelixové spoje (angl. multi-helix junctions) a pseudouzle, případně G-kvadruplex. Existuje několik databází strukturních motivů, z nejznámějších např. The RNA 3D Motif Atlas, RNA Bricks, RNAStructuromeDB a další, které poskytují informace např. i ohledně místa vázání proteinů. I v tomto případě se řeší otázka predikce struktury motivu a vznikají různé programy založené zpravidla na deep learningu a neuronových sítích. [11–14]

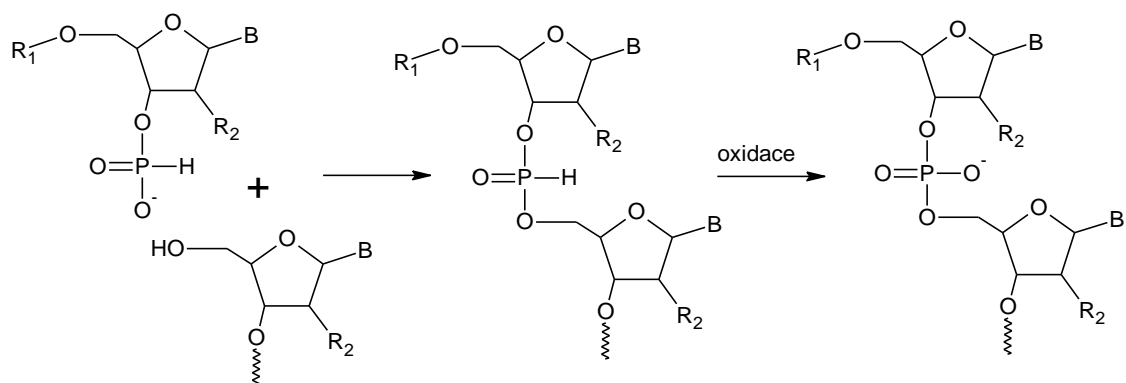
1.1.2 Oligonukleotidy

Oligonukleotidy jsou krátké polymery nukleových kyselin, přičemž jejich délka se pohybuje zhruba od 8 do 50 nukleotidů. V buňce fungují například jako regulátory genové exprese, případně se vyskytují jako zbytky po degradaci delších vláken. Běžně se oligonukleotidy využívají jako primery pro PCR, v DNA čípech a dále například v CRISPR/Cas systému. Mimo jiné se uplatňují i jako terapeutika v medicíně, kdy mohou ovlivňovat genovou expresi, fungují jako inhibitory microRNA, silencers nebo aktivátory transkripčních genů. Konkrétně lze jmenovat léčbu spinální muskulární atrofie, akutní jaterní porfýrie, dědičné tranthyretinové amyloidózy nebo Duchennovy muskulární dystrofie. Zatím je k léčbě schváleno pouze několik málo terapeutik, protože doručení oligonukleotidu do cílové tkáně jiné než jaterní je problémové, ale výzkum stále pokračuje.[15]

1.2 Syntéza oligonukleotidu

1.2.1 H-fosfonátová metoda

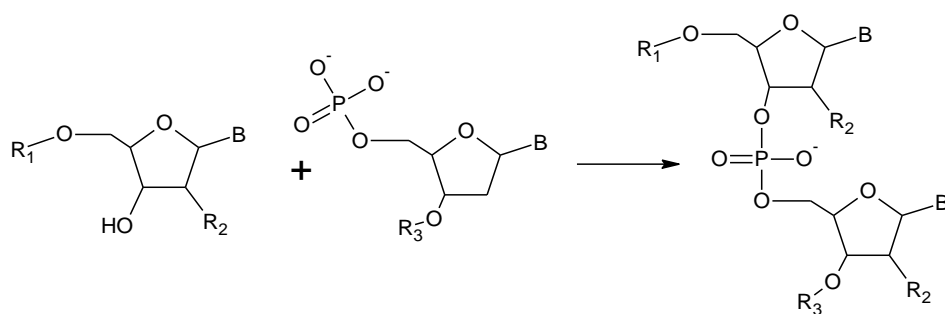
Metoda využívající H-fosfonáty, tedy estery fosfonové kyseliny, byla poprvé představena světu v padesátých letech minulého století sirem Alexanderem Toddem, kdy demonstroval, že zpracování monoesteru benzyl-H-fosfonátu difenylfosfochloridátem vedlo k tvorbě aktivovaného směšného anhydridu. Tuto sloučeninu poté nechal reagovat *in situ* s hydroxylem isopropylidennukleosidu za vzniku diesteru H-fosfonátu, který byl oxidovaný na diester fosfochloridátu. Později byla tato metodologie adaptována na chemickou syntézu 5'-3' spojeného dinukleotidu. Vznik internukleotidové vazby pomocí H-fosfonátové metody spočívá v kondenzaci nukleosidu chráněného monoesterem H-fosfonátu s nukleosidem v přítomnosti vazebného činidla za vzniku odpovídajících diesterů dinukleosidových H-fosfonátů, které lze oxidovat na cílovou fosfodiesterovou skupinu (viz obr. 2). Mezi nejběžněji užívaná kondenzační činidla patří pivaloylchlorid (PvCl), adamantankarbonylchlorid (AdCl), 5,5-dimethyl-2-oxo-2-chloro-1,3,2-dioxafosforinan (DMOCP) a *bis*(2-oxo-3-oxoazolidinyl)-fosfinchlorid (OXF). Největší výhodou H-fosfonátové metody je, že výchozí látka, tedy H-fosfonátový nukleosidový monoester, je hydrolyticky stabilní a jednoduchý na přípravu i manipulaci. Nevýhodou je vznik vedlejších produktů v důsledku reaktivity H-fosfonátu s aktivačním činidlem a tvorba fosfitových anhydridů. [16–18]



Obrázek 2: H-fosfonátová metoda syntézy oligonukleotidu. Převzato a upraveno z [18].

1.2.2 Fosfodiesterová metoda

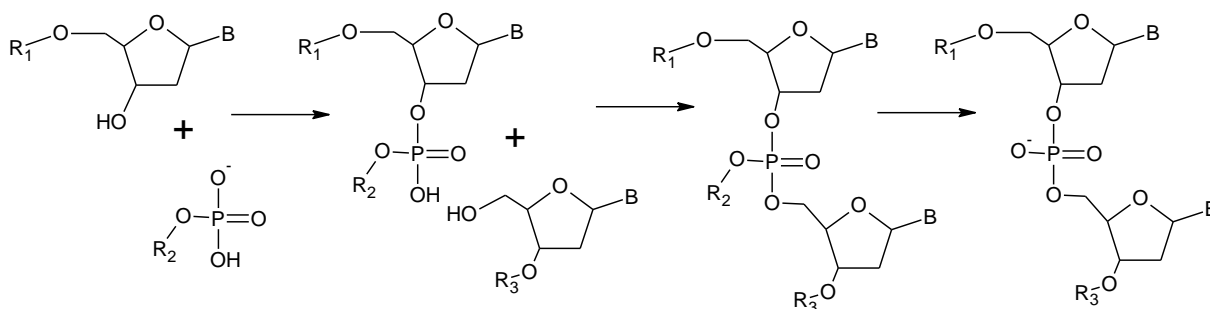
Fosfodiesterová metoda, dříve známá jako Khoranova metoda - pojmenovaná po jejím objeviteli Har Gobindu Khoranovi, byla prvně představena krátce po H-fosfonátové metodě. Na výchozím nukleosidu byla 5'-hydroxylová skupina blokována robustní tritylovou acidosenzitivní skupinou (např. trifenylmethylovou), 3'-hydroxylová skupina byla volná. Druhou reakční komponentou kondenzace je mononukleotid s 3'-hydroxylem blokováným jednoduchou alkalilabilní skupinou. Při syntéze se 5'-fosfomonoestereová skupina nukleotidu aktivovala *p*-toluensulfonylchloridem, nebo preferovanějším dicyklohexylkarbodiimidem a kondensovala s 3'-OH skupinou nukleosidu za vzniku fosfodiesterové vazby (viz obr. 3). Nevýhodou je reaktivita fosfátové skupiny, která soutěží s 3'-OH skupinou za vzniku vedlejších produktů. [19, 20]



Obrázek 3: Fosfodiesterová metoda syntézy oligonukleotidu. Převzato a upraveno z [20].

1.2.3 Fosfotriesterová metoda

Fosfotriesterový přístup k syntéze oligonukleotidů byl zkoumán v druhé polovině 60. let hned v několika laboratořích. Nejlépe na tom byly tým Letsingera a tým Ecksteina a Rizka, které byly schopny syntetizovat delší řetězce. Letsinger nejprve zkoumal metodu syntézy na pevném nosiči, ale nakonec přešel na syntézu v roztoku. V prvním kroku syntézy, fosforylaci, reaguje nukleosid, který má volnou 3'-OH skupinu, s 2-kyanoethylfosfátem v prostředí mesitylen-2-sulfonylchloridu (MS-Cl). Produkt této reakce je poté podroben kondenzaci s nukleosidem, který má volnou 5'-hydroxyskupinu, v prostředí triisopropylbenzensulfonylchloridu (TPS-Cl) viz. obr. 4. Následnou amonolýzou je odstraněna protektivní 2-kyanoethylová skupina. Eckstein a Rizk použili jako zdroj fosfátu trichlorethylfosfát nebo trichlorethylfosforodichloridát. Obdobně postupovala i skupina Reese a Saffhilla. Nevýhodou této metody je časová náročnost kondenzace, nízká výtěžnost a nestabilita reagensů. [21, 22]



Obrázek 4: Fosfotriesterová metoda syntézy oligonukleotidu. Převzato a upraveno z [22].

1.2.4 Fosfit-triesterová metoda

Fosfit-triesterovou metodu vyvinula skupina vedená Marvinem H. Caruthersem. 5'-O-chráněný nukleosid je podroben sekvenci reakcí. Nejprve reaguje s esterem dihalogenfosfitu (ROPX₂), poté s 3'-O-chráněným nukleosidem, jodovou vodou a nakonec s reagentem pro uvolnění 5'-OH- skupiny za vzniku fosfotriesterdinukleosidu. Řetězec lze prodlužovat reakcí vzniklého dinukleosidu s dalším aktivním fosfitem a opakovat tak cyklus. [23]

1.2.5 Fosforamiditová metoda

Fosforamiditová metoda navazuje na fosfit-triesterovou syntézu. Poprvé byla představena roku 1981 týmem Sergem L. Beaucagem a Marvinem H. Caruthersem, který je zároveň autorem výše zmíněné fosfit-triesterové metody. Jejich článek popisoval využití

fosforamiditů jako důležitých meziproductů syntézy deoxypolynukleotidů. Fosforamidity jsou monoamidy fosfitových diesterů s obecnou strukturou $R_3R_4NP(R_2)-O-R_1$. Na rozdíl od předchozích metod neprobíhá syntéza v roztoku, ale nukleosid je vázaný na pevné fázi, která zaručuje, že přebytečné reagenty mohou být odstraněny jednoduchou filtrací. Momentálně se jedná o nejpoužívanější metodu napříč laboratořemi, protože proces je plně automatizovaný, je možné syntetizovat i dlouhé oligonukleotidy (do délky cca 200 bází) a především poskytuje vysoké výtěžky, čímž snižuje i samotnou cenu výroby. [24, 25]

Mezi nejnovější inovace pro tuto metodu, které jsou zatím ve fázi výzkumu, patří možnost syntetizovat fosforamidity z jejich odpovídajících alkoholů, kdy je dosaženo krátkých reakčních časů, výtěžky jsou téměř kvantitativní a není potřeba jejich purifikace před automatizovanou syntézou oligonukleotidu, přičemž cílem je samozřejmě přímá integrace metody do DNA syntetizátorů, kdy se zcela vynechá jak manuální syntéza, tak skladování a manipulace samotná s redoxně nestabilní fosforamidit. [26]

1.2.5.1 Pevná fáze

Během let se studovalo nepřeberné množství druhů pevné fáze pro syntézu oligonukleotidů, mezi nimi např. Sephadex, silikagel, polydimethylakrylamidová pryskyřice, polyakrilmorfolid, celulóza, Kieselguhr-polydimethylakrylamidové kompozity, sklo a různé polystyrenové pryskyřice, přičemž nejvíce se odsvědčily poslední dva jmenované. Konkrétně se používá porézní sklo (CPG, controlled pore glass) a makroporézní polystyrén (MPPS). Momentálně existuje více než 280 pevných fází pro syntézu modifikovaných i nedomodifikovaných oligonukleotidů. [27, 28]

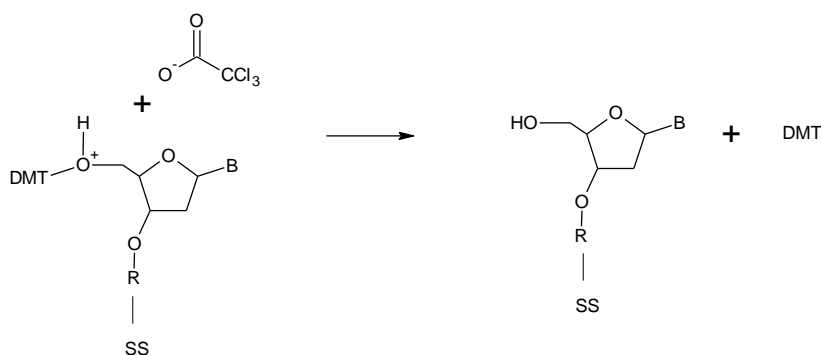
Dále se pro syntézu využívají univerzální supporty (podpěry), u kterých se využívá linkerů (např. 3-amino-1,2-propandiol), jež formují s amidovou skupinou pouze slabý vodíkový můstek, což v zásaditém prostředí vede k amidem asistované defosforylaci, která uvolní 3'-OH-oligonukleotid do roztoku, zatímco produkty postranních reakcí zůstanou vázané na pevné fázi. Výhodou těchto supportů je možnost využití jedné kolony pro jakýkoliv oligonukleotid, nezáleží tedy, jaký je nukleosid na 3'-konci. [29, 30]

1.2.5.2 Syntetický cyklus

Syntéza probíhá směrem od 3' konce k 5' konci a skládá se ze čtyř základních kroků. Na samotném začátku syntézy je první chráněný nukleosid navázán na pevnou fázi. V případě využití univerzálních supportů se první nukleosid váže až během kondenzace. [30]

Detritylace

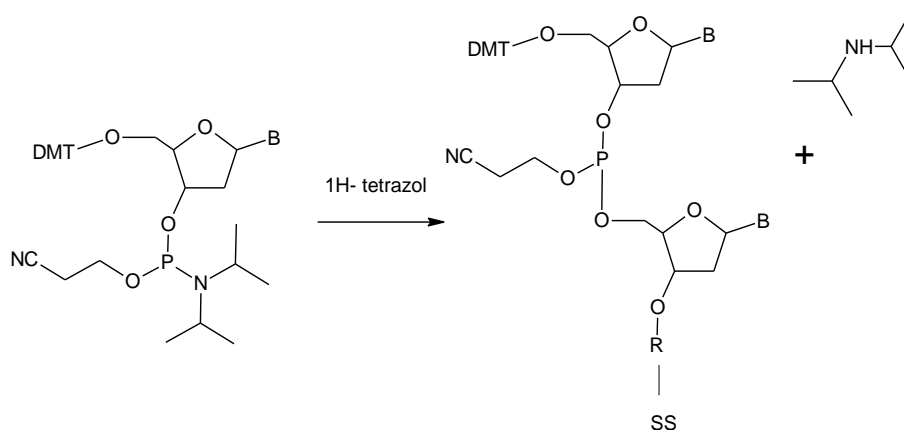
Syntetický cyklus začíná odstraněním acidolabilní 5'-DMT (4,4'-dimethoxytrityl) protektivní skupiny, která má za úkol zabránit polymerizaci nukleosidu při navazování na pevnou fázi. Toho je docíleno přidáním kyselého roztoku jako je 2% trichloroctová nebo 3% dichloroctová kyselina v dichlormethanu. Při použití příliš koncentrovaného činidla či příliš dlouhé době expozice kyselinou může docházet k depurinaci látek, a tedy ke snížení výtěžku. Toto lze monitorovat spektrofotometricky ve VIS oblasti, neboť DMT-karbokationt absorbuje při 495 nm a je tudíž oranžový. Výsledkem detritylace je odkrytí reaktivní 5'-hydroxy skupiny, na kterou se může vázat nukleosidový fosforamiditový monomer (viz obr. 5). [16, 31–33]



Obrázek 5: Detritylace.

Coupling (kondenzace)

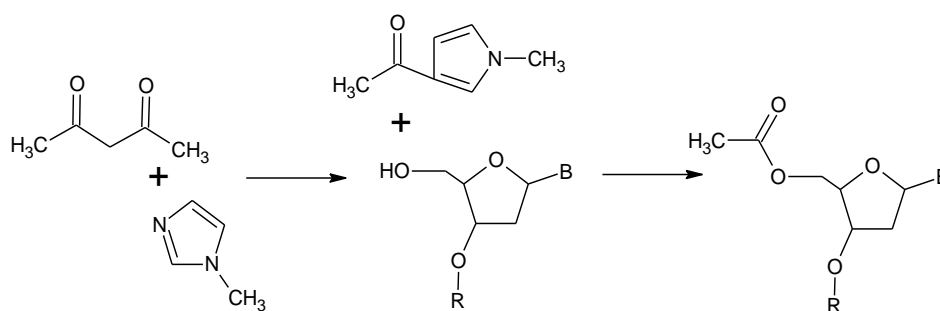
Při kondenzaci je nejprve adovaný monomer upraven protonací kondenzačním činidlem. K velkému přebytku příslušného nukleosidového fosforamiditu je přidán aktivátor 1H-tetrazol (nebo jeho deriváty) rozpuštěný v acetonitrilu jakožto vhodném rozpouštědle pro nukleofilní vytěšňovací reakce, čímž dojde k protonaci aminoskupiny. Tímto způsobem protonovaný amin dá vzniku meziprojektu tetrazolylfosforamiditu, který následně vytvoří internukleotidový triesterfosfit s volným 5'-hydroxyem nukleosidu vázaného k pevné fázi (viz obr. 6). Diisopropylaminoskupina běžně využívaných fosforamiditů poskytuje díky sterickému efektu ideální rovnováhu mezi stabilitou a snadností aktivace. [16, 31–33]



Obrázek 6: Kondenzace.

Capping (acylace)

Dalším krokem syntézy je capping. Přesto, že kondenzační krok je velmi přesný co se reakčního mechanismu a kvantitativního průběhu reakce týče, zůstává malé procento (řádově desetiny až jednotky) 5'-hydroxylové skupiny nezreagované. Navolný aktivní hydroxyl by se v dalším cyklu při kondenzaci mohl standardně vázat fosforamidit, což by mělo za následek zkrácení o jeden nukleotid oproti předpokladu. Pokud by tyto volné hydroxylové skupiny nebyly po každé kondenzaci ošetřeny acylací, vedlo by to ve výsledku ke směsi oligonukleotidů lišících se délkou o několik málo nukleotidů. Takovou směs by bylo velmi těžké purifikovat a samozřejmě by byla snižována samotná hodnota produktu. Zablokování těchto volných hydroxylů se provádí acetanhydridem a N-methylimidazolem v tetrahydrofuranu s pyridinem, který zajišťuje bazicitu roztoku, aby nedocházelo k případné detritylaci nukleosidového fosforamiditu (viz obr. 7). [16, 31–33]

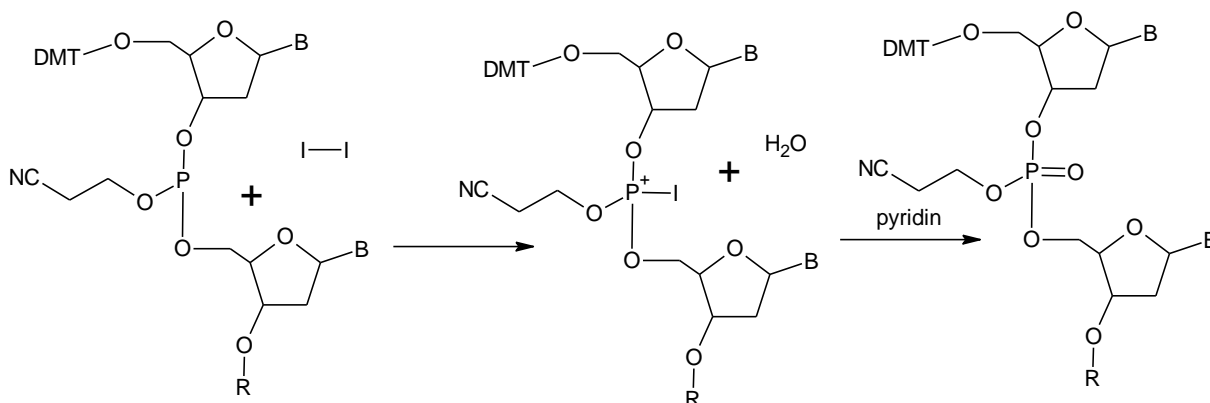


Obrázek 7: Capping.

Oxidace

Finální reakcí cyklu je oxidace fosfitriesteru na fosfát. Vzniklá internukleosidová fosfitriesterová vazba je acidolabilní a kvůli možnosti opakování syntetického cyklu je třeba ji převést na stabilní fosfátovou vazbu. Toho je docíleno oxidačním činidlem, zpravidla jodem

v přítomnosti pyridinu (viz obr. 8). Takto upravený dinukleotid je připraven na další detrylacii a opakování cyklu. [16, 31–33]



Obrázek 8: Oxidace.

1.2.5.3 Posyntetické úpravy

Jakmile je syntéza kompletní a oligonukleotid dosáhne požadované délky, je uvolněn hydrolyzací z pevné fáze. K níse používá vodný roztok amoniaku, který zároveň odstraní protektivní amidové skupiny na nukleových bázích a případně i protektivní kyanoethylové skupiny na fosfátu. Následuje purifikace za použití elektroforézy v polyakrylamidovém gelu (PAGE, polyacrylamide gel electrophoresis) či dnes spíše vysokoúčinné kapalinové chromatografie (HPLC, high-performance liquid chromatography). [16, 33]

1.2.6 Enzymatická metoda

Ačkoliv fosforamiditová metoda je stále zlatý standard mezi syntézami nukleotidů a aktivně se využívá již 40 let, má i své limity, mezi které patří především možná syntetizovaná délka, která se pohybuje do cca 200 bází. Tato délka je sice naprosto dostačující pro syntézu primerů či sond, pro potřeby např. RNA vakcín či pro CRISPR technologii už zcela nestačí. Myšlenka enzymatické syntézy nukleových kyselin přišla už začátkem 60. let minulého století, poté co byl identifikován enzym terminální deoxynukleotidtransferáza (TdT), což je enzym schopný syntézy DNA řetězce bez nutnosti templátu. Jeho konkrétní funkce zahrnuje zavádění nukleotidů do genové sekvence kódující receptory T a B-lymfocytů pro rozšíření rozmanitosti rozpoznávaných epitopů. Při enzymatické syntéze TdT zprostředkovává postupné přidávání upravených nukleotidů obsahujících terminátory, které syntézu zastaví. Syntéza tedy probíhá krokově po jednotlivých nukleotidech. Vedle TdT se objevují i další enzymy, případně se uvažuje o jejich kombinaci. Oproti fosforamiditové metodě se vedle možnosti delších nukleotidů ukazuje i o něco vyšší výtěžnost. Zatím touto metodou

syntetizuje nukleové kyseliny pouze několik málo firem především z USA a Velké Británie a je stále předmětem výzkumu. [34–36]

1.2.7 Modifikace oligonukleotidů

Samotný oligonukleotidový řetězec je často syntetizován pouze za účelem modifikace, přičemž úpravy mohou být jak během samotné syntézy, tak po ní.

1.2.7.1 Enzymatické značení

Pro enzymatické techniky se využívá často polymeráz, transferáz a ligáz, při kterých dochází k přidání určitého množství modifikovaných nukleotidů k samotnému oligonukleotidovému řetězci. Z konkrétních enzymů se využívá např. TdT nebo T4 RNA ligáza 1. [37]

1.2.7.2 Chemické modifikace

Chemické modifikace nukleových kyselin umožňují především inkorporaci různých značek do DNA a RNA práb (sond). Mezi běžné modifikace oligonukleotidů patří přidání spaceru s funkční skupinou, např. aminem, přes který se může vázat další molekula jako je fluorescenční značka či nějaké zesíťovací činidlo. V některých případech lze značku umisťovat přímo na oligonukleotid bez potřeby spaceru. Jedná se o různé fluorescenční molekuly nebo biotinové značky. Velmi využívané jsou duálně značené sondy, které na 5'-konci obsahují fluorescenční značku a na 3'-konci pak zhášec, a slouží především pro qPCR (kvantitativní polymerázová řetězová reakce). [37]

Dalším důvodem pro chemickou modifikaci syntetického oligonukleotidu může být příprava řetězce odolávajícího endonukleázám, dále zvýšení stability, hydrofobicity a teploty tání. Úpravy často zahrnují inkorporaci fluoro- nebo methoxy-skupiny na pozici 2'cukerné složky. Dále to může být fosforthioátová modifikace na fosfátové kostře řetězce, tedy náhrada nevezbného atomu kyslíku za atom síry. [38]

1.3 Kvantifikace oligonukleotidů

Kvantifikace oligonukleotidů nebo obecně nukleových kyselin využívá především jejich fyzikálně-chemických vlastností a samozřejmě se jedná o jednu z nejdůležitějších informací o produktu.

1.3.1 UV absorpční spektroskopie

Zcela jistě nepoužívanější a nejjednodušší metodou pro kvantifikaci oligonukleotidu je UV absorpční spektroskopie, která využívá schopnosti nukleotidů absorbovat záření

v ultrafialové (UV) oblasti. Za tuto vlastnost může přítomnost heterocyklů, konkrétně rezonanční struktury purinových a pyrimidinových bází. Absorpční spektrum se měří většinou mezi 200 a 320 nm, přičemž absorpční maximum je zpravidla při 260 nm a odráží tak vlastní koncentraci nukleové kyseliny, je však závislé na pH roztoku a poměru obsažených bází. Dále se počítají poměry absorbancí A_{260}/A_{280} a A_{260}/A_{230} pro zjištění čistoty vzorku; absorbance při 280 nm značí přítomnost proteinů, při 230 nm se pak jedná o další kontaminanty, jako jsou organická rozpouštědla (ethanol), EDTA, fenol, soli či detergenty (TritonTM X-100) v závislosti na předchozím nakládání s měřeným vzorkem. Ideální hodnoty pro čisté látky se pak pohybují do ~ 1.8 u DNA a ~ 2,0 RNA v případě prvního poměru, druhý poměr by pak měl vycházet v rozmezí přibližně od 2,0 do 2,2. [39–41]

Zajímavou vlastností nukleotidů je hyperchromicita, tedy nárůst absorbance materiálu. Konkrétně dvouvláknová DNA má nižší absorbanci než její denaturovaná forma, tedy 2 řetězce jednovláknové DNA. Díky tomuto jevu je možné měřit bod tání DNA spektrofotometricky za konstantního zvyšování teploty, kdy T_m odpovídá přesně polovině křivky vzestupu absorbance. [42]

1.3.1.1 Molární absorpční koeficient

Molární absorpční (někdy též extinkční) koeficient ϵ je číslo vyjadřující schopnost látky tlumit, resp. absorbovat procházející světlo při dané vlnové délce. SI jednotkou tohoto koeficientu je m^2/mol , ale v praxi se používá spíše $M^{-1}\cdot cm^{-1}$. Tato veličina je závislá na délce nukleotidu, jeho sekvenci a pH roztoku a může být spočítána buď pomocí prostého součtu extinkčních koeficientů jednotlivých nukleotidů, které se násobí s ohledem na base-stacking faktorem 0,9, nebo pomocí o něco přesnější metody nejbližšího souseda, která se vypočítá podle rovnice 1.

$$\epsilon_{260} = \sum_1^{N-1} \epsilon_{Nearest\ Neighbor} - \sum_2^{N-1} \epsilon_{Individual\ Bases} + \sum_1^N \epsilon_{Modification}$$

Rovnice 1: Výpočet metody nejbližšího souseda pro výpočet molární absorpčního koeficientu polynukleotidu.

Při využití druhého postupu se na řetězec pro výpočet dívá spíše jako na soubor několika dinukleotidů ve směru od 5'-konce k 3'-konci a hodnoty jsou tedy dané pro jednotlivé sousední dvojice. Hodnoty pro oba druhy výpočtů jsou uvedeny v tabulce 1, porovnání teoretických hodnot s naměřenou experimentální hodnotou pak v tabulce 2. Pro dlouhé nukleotidové řetězce se extinkční koeficient určuje experimentálně porovnáním

absorbance vzorku před a po totální enzymatické digesci případně se využívá pouze aproximace s typickými hodnotami pro jednotlivé nukleové kyseliny. [40, 41, 43]

Tabulka 1: Hodnoty absorpčního koeficientu pro vybrané nukleotidy, sondy a zhášecí (A) a pro dvojice metody nejbližšího souseda (B). Hodnoty jsou uvedeny v $\text{mM}^{-1}\text{cm}^{-1}$. Převzato a upraveno z [44, 45].

A

Báze	Sondy	Zhášecí
A	15,4 FAM	20,9 BHQ1 8
C	7,3 HEX	31,58 BHQ2 8
G	11,7 Cy5	10 BBQ 15,07
T	8,8	

B

		3'-pozice				
		Báze	A	C	G	T
5'-pozice	A		27,4	21,2	25	22,8
	C		21,2	14,6	18	15,2
	G		25,2	17,6	21,6	20
	T		23,4	16,2	19	16,8

Tabulka 2: Hodnoty molárního absorpčního koeficientu pro 100-mery získané prostým součtem ($\epsilon - \text{BC}$), metodou nejbližšího souseda ($\epsilon - \text{NN}$) a experimentálně ($\epsilon - \text{EX}$). Hodnoty jsou uvedeny v $\text{mM}^{-1}\text{cm}^{-1}$. Hodnoty převzaty z [46].

Měřený vzorek	$\epsilon - \text{BC}$	$\epsilon - \text{NN}$	$\epsilon - \text{EX}$
AT (100-mer)	1 140,75 \pm 5,7	1 082,7 \pm 5,4	927,4 \pm 4,6
GC (100-mer)	841,5 \pm 4,2	846 \pm 4,2	896,8 \pm 4,9

1.3.1.2 Lambert-Beerův zákon

Lambert-Beerův nebo Beer-Lambert-Bouguerův zákon je beze sporu nejdůležitějším zákonem optické spektroskopie. Beerův nebo Bouguer-Beerův zákon jakožto rozšíření Bouguerova zákona na roztoky říká, že při průchodu světla skrz vrstvu absorbujícího roztoku o koncentraci c , intenzita procházejícího světla I_{trS} se při lineárním růstu koncentrace snižuje exponenciálně. Toto lze zapsat rovnicí jako $I_{trS} = I_0 e^{-kc}$, kdy I_0 je intenzita dopadajícího světla a k je konstanta. Do jaké míry je elektromagnetické záření absorbováno vzorkem, se vyjadřuje pomocí absorbance A nebo transmittance T , přičemž vztah těchto veličin k intenzitě záření lze vyjádřit jako $T(\%) = 100 \cdot \frac{I_{trS}}{I_0}$, resp. $A = -\log_{10} \left(\frac{I_{trS}}{I_0} \right) = -\log(T)$. Kombinací s poznatky německého vědce Johanna Heinricha Lamberta vznikl Beer-Lambert-Bouguerův

zákon, který je běžně zapisován jako $A = \epsilon cl$, kdy l je tloušťka vrstvy, resp. délka průchodu paprsku vzorkem. [25, 47]

1.3.2 Další metody stanovení koncentrace nukleotidů

1.3.2.1 Fluorescenční značení

Další možnou spektrofotometrickou metodou pro kvantifikaci nukleotidů je využití fluorescenčních barviv, které se specificky vážou na určitý typ nukleové kyseliny, na jedno- či dvouvláknovou DNA nebo RNA, což je výhoda oproti UV absorpční spektroskopii. K dispozici je široké množství komerčně dodávaných fluorescenčních barviv, přičemž se vždy volí vhodné barvivo pro daný účel a instrumentální možnosti. Barviva se liší jak svými excitačními a emisními spektry, tak i senzitivitou či stabilitou. Mezi nejběžněji používaná fluorescenční barviva se řadí ethidium bromid, PicoGreen[®] či SYBR Green[®]. Fluorescenčních sond je využíváno i při kvantitativní real-time PCR či při gelové nebo kapilární elektroforéze, kdy je v obou případech možnost kvantifikace či alespoň semikvantifikace nukleové kyseliny. [40]

1.3.2.2 Izotopová zředovací hmotnostní spektrometrie

Hmotnostní spektrometrie s izotopovým zředováním (IDMS) je pravděpodobně nejpřesnější metoda pro kvantifikaci oligonukleotidů a využívá se proto u referenčních materiálů či primárních standardů. Vzorek je smíchán s izotopicky značeným analogem, software pak pomocí plochy pod píkem a známého množství značeného analogu dopočítá množství vzorku. [40, 48]

1.3.2.3 Alternativní metody

Enzymatická metoda

Alternativním přístupem ke kvantifikaci nukleové kyseliny je využití enzymu luciferázy. Nejprve je na 3'-konec DNA navázán pyrofosfát za vzniku deoxyribonukleotidtrifosfátu (dNTP), který následně přenesení terminální fosfát na adenosindifosfát (ADP) dodaný do reakce. Luciferáza poté spotřebovává ATP za vzniku světla, jehož množství je přímo úměrné množství přítomné DNA. Tohoto se využívá především ve forenzní analýze. [40]

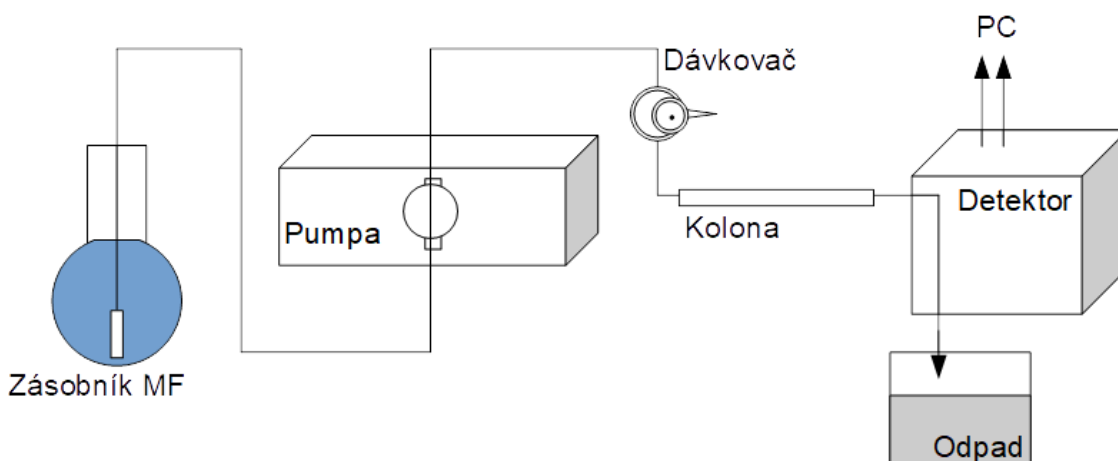
Stanovení pomocí difenylaminu

Deoxyribóza purinových bází v DNA v kyselém prostředí tvoří β -hydroxylevulinálddehyd, který při reakci s difenylaminem vytváří modrý produkt, který má

absorpční maximum při 595 nm. Výsledná absorbance odpovídá polovičnímu množství přítomné deoxyribózy. [49]

1.4 Purifikace

Nejrozšířenější metodou pro purifikaci oligonukleotidů je chromatografie, která se beze sporu řadí mezi základní separační metody; jednou z nejvíce využívaných je vysokoúčinná kapalinová chromatografie (HPLC, high-performance liquid chromatography). Principem metody je zadržování analytu unášeného mobilní fází na stacionární fázi kolony, kam je vhnán pod tlakem, přičemž zadržování probíhá na základě fyzikálně-chemických vlastností analyzovaných složek. Dávkování vzorku probíhá pomocí vícecestného (většinou šesticestného) dávkovacího ventilu. Součástí zařízení je i optický, hmotnostní nebo vodivostní detektor. Obecný náčrt kapalinového chromatografu je zobrazen na obr. 9. V průběhu let došlo k vývoji mnoha podob této metody.



Obrázek 9: Obecný náčrt kapalinového chromatografu.

1.4.1 Rozdělovací chromatografie

Rozdělovací chromatografie, někdy zvaná i jako liquid-liquid chromatografie, patří mezi historicky nejstarší metody HPLC. Fáze jsou tvořeny dvěma kapalinami, přičemž stacionární fázi tvoří kapalina imobilizovaná v pórech kolony. Samotné dělení složek tak probíhá na základě rozdílné distribuce látek v jednotlivých, navzájem nemísitelných kapalinách. Obrovskou výhodou této metody je téměř neomezený počet fázových systémů a dále možnost velmi dobře termodynamicky definovat chování systému; mezi nevýhody se pak zcela jistě řadí nemožnost použít kontinuální gradient. [50]

1.4.2 HPLC s normálními fázemi

HPLC s normálními fázemi (normal-phase HPLC, NP-HPLC) označuje metodu, kdy stacionární fáze je polární a mobilní fáze méně polární či nepolární. Tato metoda byla také vyvinuta jako jedna z prvních a označení normální slouží především pro odlišení od chromatografie s reverzními fázemi, která vznikla o něco později. Kolona je naplněna zpravidla silikagelem či aluminou, případně se silikagel modifikuje amino-, nitro- nebo diolovou skupinou. Jako mobilní fáze se využívá *n*-hexan, diethylether, aceton, ethanol, isopropanol a další. Nevýhodou metody je z chemické podstaty vysoká citlivost k přítomnosti vody, kdy i stopové množství se váže na kolonu a má za následek nereprodukovatelnost měření, proto se metoda příliš nevyužívá. Hlavní oblastí použití metody je separace lipidů, vitamínů a polyaromatických uhlovodíků. [51]

1.4.3 Iontově výměnná chromatografie

Mechanismus separace při iontové výměnné, někdy nazývané též ionexové, chromatografii závisí na elektrostatické interakci iontů vzorku a opačně nabitých iontů funkčních skupin stacionární fáze. Mobilní fází je vždy pufovaný vodný roztok o fixním pH, přičemž při použití více mobilních fází o rozdílných pH může separace probíhat i v pH gradientu. Náplň kolony bývají chemicky upravené pryskyřice např. vysoce zesíťovaný polystyren s chemicky navázanými ionizovanými či ionizovatelnými funkčními skupinami, kdy velikost separovaných iontů odpovídá míře zesíťování. Jako další náplň kolon se využívá i celulóza či Sephadex (zesíťovaný dextran). [38, 52]

1.4.4 Bioafinitní chromatografie

Bioafinitní chromatografie je významnou technikou v purifikaci sloučenin široce využívanou pro svou specifitu. První bioafinitní chromatografie byla provedena roku 1910 a sloužila k purifikaci α -amylázy imobilizovaným škrobem. Principem tohoto typu chromatografie je vazba biologického prvku na stacionární fázi; u purifikace enzymu může být ligandem koenzym, kofaktor, substrát nebo inhibitor enzymu. Příkladem může být imobilizovaná DNA nebo RNA pro purifikaci nukleáz. Dále se jako ligandy využívají lecitiny nebo v případě imunoafinitní chromatografie protilátky. [38, 53]

1.4.5 Rozměrově vylučovací HPLC

Běžně využívanou metodou pro analýzu polymerů je rozměrově vylučovací chromatografie (size-exclusion HPLC, SE-HPLC), která pro separaci využívá kombinaci velikosti, difúzního koeficientu a povrchových vlastností molekuly, přičemž metoda vychází

z gelové permeační chromatografie. Při průchodu vzorku kolonou s inertní porézní náplní (nejčastěji agaróza, dextran či polyakrylamid na silice) dochází k většímu zadržování menších molekul v pórech stacionární fáze, větší molekuly procházejí snáz, eluce probíhá tedy od největší molekuly po nejmenší. Na rozdíl od předchozích metod neprobíhá žádná chemická interakce mezi analytem a stacionární fází, separace probíhá čistě na základě velikosti a tvaru, rozměrově vylučovací chromatografie je tak vhodná pro separaci multimerů nebo při použití vhodného detektoru, k determinaci molekulové hmotnosti. [38, 54, 55]

1.4.6 HPLC s reverzními fázemi

Vedle NP-HPLC zmíněné výše existuje HPLC s reverzními fázemi (reversed-phase HPLC, RP-HPLC), někdy nazývaná též HPLC na nepolárních absorbentech, kdy stacionární fázi tvoří hydrofobní látka navázaná na nosiči kolony a mobilní fázi pak zpravidla vodný roztok, pufr, nebo polární organický roztok, např. acetonitril, methanol, isopropanol nebo tetrahydrofuran. Analyt bude eluovat od nejpolárnější látky po méně polární. Běžně je tato metoda doplněna o iontově-párové interakce. Nejčastěji se metoda využívá pro separaci peptidů nebo krátkých oligonukleotidů, kdy poskytuje vysokou míru rozlišení a reprodukovatelnosti. [56, 57]

1.4.6.1 Kolony a stacionární fáze

Typickou náplň kolony pro RP-HPLC tvoří silika s kovalentně navázaným oktylem (C8), oktadecylem (C18), methylem (C1), butylem (C4), nebo fenylem, kdy delší řetězec bude mít za následek vyšší hydrofobicitu. Pro iontově-párové interakce mezi kolonou a analytem je navíc navázán i protiion. Rozdíly v kolonách budou velmi patrné při potřebě preparativní a analytické HPLC. Zatímco u analytické chromatografie je žádoucí vysoký počet píků a nízký objem nástřiku, u preparativní HPLC je důležitý především pík žádoucího produktu a vyšší průtok pro co možná nejvyšší a nejrychlejší možný výtěžek na kolektoru frakcí. Z tohoto důvodu jsou preparativní kolony o mnoho větší s menší rozlišovací schopností, přičemž preparativní kolona může přesahovat i metrovou délku. Kompromisem bývají semipreparativní kolony. Samozřejmostí u preparativní HPLC je nedestruktivní detektor. [57, 58]

1.4.7 Isokratická a gradientová eluce

Dávkování mobilní fáze je možné provádět ve dvou různých módech; isokratickou či gradientovou elucí. Při isokratické eluci se složení mobilní fáze nemění a po dobu celého měření tak zůstává stejná. Používá se především při malém množství vzorku, u něhož se

nepředpokládá velký počet píků. Retenční časy bývají delší a reprodukovatelnost metody je velmi jednoduchá. Gradientová eluce, kdy se v čase mění složení mobilní fáze, se může potýkat s větším šumem na základní linii, čemuž se lze alespoň částečně vyvarovat použitím velmi čistých chemikálií. Další nevýhodou je také větší náročnost na instrumentaci. Co tyto nedostatky ovšem předčí, je o mnoho větší kapacita píků, která umožňuje separovat i několikasožkové vzorky a i zkrácení času separace v porovnání s isokratickou metodou. Z těchto důvodů je metoda hojně využívána. [53, 59]

1.4.8 Purifikace oligonukleotidů

Pro další využití oligonukleotidů je po syntéze potřeba produkt zpurifikovat, tedy odstranit z roztoku vše nežádoucí, jako jsou soli či nedosyntetizované zbytky. Prvním krokem bývá odsolení, při kterém je roztok zbavován organických sloučenin, jako jsou např. benzamidy, isobutyramidy, octan amonný nebo akrylonitril, které vznikají jako vedlejší produkty syntézy. Nejjednodušší metodou odsolování je dialýza. Při odsolování nedochází k odstranění nedosyntetizovaných zbytků, lze tedy takto vypurifikované oligonukleotidy použít pouze v robustních technikách, jakou je např. PCR. [60]

Další možností je využití cartridge (oligonucleotide purification cartridge, OPC), která navíc vedle solí odstraňuje i krátké nedosyntetizované řetězce na základě rozdílné hydrofobicity mezi kompletním oligonukleotidem a krátkými řetězci, kdy je kompletní řetězec zachycen na koloně pomocí 5'-DMT skupiny a ostatní produkty jsou eluovány. Výsledkem je negarantovatelná, avšak zpravidla dostatečná čistota oligonukleotidu pro většinu aplikací. [61]

Pokud je potřeba získat oligonukleotid o vyšší čistotě, nabízí se hned několik metod HPLC, nejčastěji se jedná o RP-HPLC a iontově výměnnou HPLC. Při RP-HPLC se využívá stejného principu jako při OPC, tedy separace závislé na hydrofobních vlastnostech DMT skupiny. Nejběžněji využívané kolony jsou typu C18 naplněny 2,5 μ m porézním sorbentem a jako mobilní fáze slouží TEAA (triethylamoniumacetátový) pufr. Iontově výměnná HPLC pro purifikaci oligonukleotidů využívá nejčastěji jako náplň kolony pórezní materiál s navázanou terciární či kvartérní amoniou solí. V obou případech je vhodné předejít tvorbě sekundárních a terciárních struktur či dimerů denaturací teplotou (50 – 65 °C) nebo přidáním organického rozpouštědla (ethanol, acetonitril). V obou případech klesá selektivita s délkou oligonukleotidu. [62–64]

Další metodou purifikace, která nezahrnuje HPLC, je PAGE, jež je též hojně využívána. Možnou novou metodou je využití kombinace interakce biotin-streptavidin a fotoštěpitelné skupiny. [63, 64]

1.5 Fyzikálně chemické vlastnosti oligonukleotidů

Fyzikálně chemické vlastnosti oligonukleotidů vyplývají jako v případě ostatních chemických látek primárně z jejich molekulového složení. Charakterizovat vlastnosti daného oligonukleotidu je důležité především pro určitou možnost predikce chování, např. při HPLC analýze nebo i designu primerů.

1.5.1 Fyzikální vlastnosti

Mezi základní vlastnosti z fyzikálního hlediska patří zcela jistě délka, tedy absolutní počet bází daného oligonukleotidu, a obsah GC bází udávaný v procentech, kdy vyšší procenta značí lepší stabilitu duplexu, což se projeví na teplotě tání T_m , která je důležitá především pro PCR primery a jejich teplotu nasedání. Dále se u oligonukleotidů determinuje jejich relativní molekulová hmotnost pomocí prostého součtu relativních hmotností všech obsažených atomů. Mezi další fyzikální vlastnost patří i molární absorpční koeficient zmíněný v samostatné podkapitole (kapitola 1.3.1.1 *Molární absorpční koeficient*) či termodynamické konstanty. [65–67]

1.5.2 Chemické vlastnosti

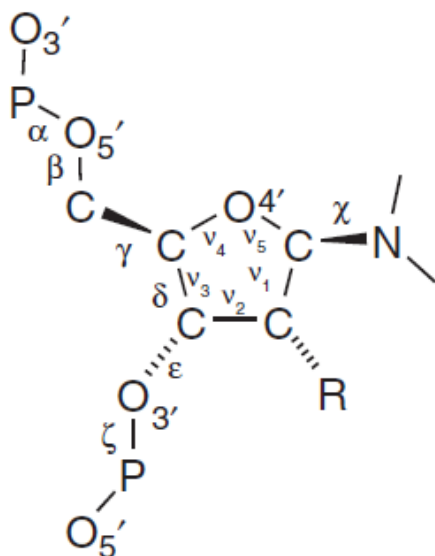
Chemické vlastnosti nukleotidů jsou spjaty především s tvorbou vodíkových můstků, resp. amino- a ketoskupiny mohou působit jako donory či akceptory díky tendenci disociovat atomy a vytvořit tak na dané skupině náboj. Toto chování popisuje disociační konstanta pK_A . Při fyziologickém rozhraní pH jsou báze i cukerná složka nenabitě, z fosfátových skupin má náboj pouze terminální fosfát. Hodnoty pK_A pro jednotlivé majoritní báze jsou uvedeny v tabulce 3. [68]

Tabulka 3: Hodnoty pK_A majoritních bází při 20 °C a nulové koncentraci solí. Převzato z [68].

Báze (místo protonace)	Nukleosid	3'-nukleotid	5'-nukleotid
Adenin (N1)	3,52	3,7	3,88
Cytosin (N3)	4,17	4,43	4,56
Guanin (N7)	3,3	3,5	3,6
Guanin (N1)	9,42	9,84	10
Thymin (N3)	9,93	-	10,47
Uracil (N3)	9,38	9,96	10,06

1.5.3 Konformační vlastnosti

Obdobně jako všechny molekuly mají i oligonukleotidy tendence zaujímat co nejstabilnější konformaci v prostoru. Stabilitu je dosaženo pomocí hned několika faktorů a sil, jako je vertikální stohování sousedních bází a elektrostatické interakce způsobené negativně nabitými fosfátovými skupinami. Pyrimidiny i puriny, resp. heterocykly, jsou planární molekuly, konformace je tedy definována sérií torzních úhlů vazeb mimo tyto heterocykly, kdy jsou preferovány takové, které vyhovují sterickému omezení. Definice torzních úhlů nukleotidů je uvedena na obrázku 9 a v tabulce 4. V některých případech mohou oligonukleotidy tvořit i sekundární struktury či strukturální motivy. [68, 69]



Obrázek 10: Definice torzních úhlů nukleotidů. Převzato z [68].

Tabulka 4: Definice torzních úhlů nukleotidů. Převzato z [68].

Torzní úhel	Atomy definující úhel
α	$(n - 1)O3' - P - O5' - C5'$
β	$P - O5' - C5' - C4'$
γ	$O5' - C5' - C4' - C3'$
δ	$C5' - C4' - C3' - O3'$
ε	$C4' - C3' - O3' - P$
ζ	$C3' - O3' - P - O5'(n + 1)$
χ	$O4' - C1' - N1 - C2$ (pyrimidiny)
χ	$O4' - C1' - N9 - C4$ (puriny)
ν_0	$C4' - O4' - C1' - C2'$
ν_1	$O4' - C1' - C2' - C3'$
ν_2	$C1' - C2' - C3' - C4'$
ν_3	$C2' - C3' - C4' - O4'$
ν_4	$C3' - C4' - O4' - C1'$

Konformace nukleotidu se projeví na výsledné struktuře řetězce, je tedy důležitou vlastností pro predikci vlastností DNA a má dále i vliv na separaci. Mezi konkrétní příklady vlivu konformace patří např. levotočivá Z-DNA, která je výsledkem střídavé konformace GC dinukleotidu, kdy báze zaujímají antikonformaci pro cytosinové nukleotidy s cukry v C2'-endo ($\chi = 208^\circ$) konformaci a synkonformaci pro guaninové nukleotidy s cukry v C3'-endo ($\chi = 67^\circ$). Za příklad vlivu na separaci lze uvést base-flip izomery, kdy báze v B-DNA zaujímá extrahelikální postavení, u kterého byly při iontové párové chromatografii s reverzními fázemi prokázány rozdílné retenční časy. [70, 71]

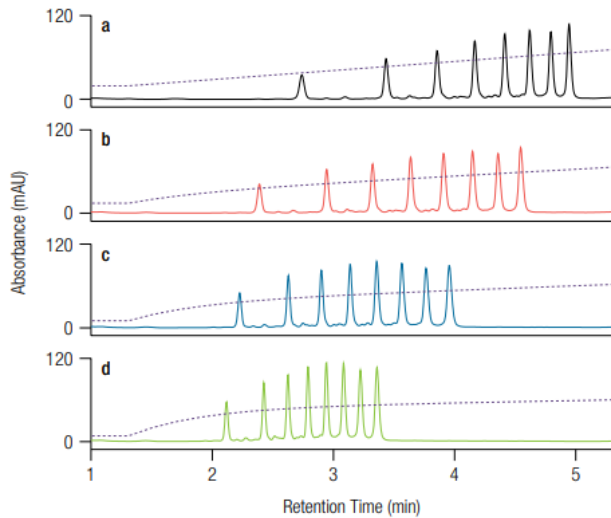
Determinace konformace daného oligonukleotidu lze provést experimentálně pomocí nukleární magnetické rezonance (NMR), při které se využívá interakce atomových jader s nenulovým jaderným spinem a magnetickým polem, nebo pomocí techniky SAXS (small-angle X-ray scattering), která využívá difrakce rentgenových paprsků na molekule, přičemž vzorek není potřeba mít v krystalické formě, což je benefitem oproti běžné rentgenové krystalografii. Tato výhodna je v určitém smyslu ovšem vykompenzována o něco nižším rozlišením. Alternativou k těmto metodám je pak kryoelektronová mikroskopie (cryo-EM). [72–74]

1.5.4 Predikce vlastností

Pro běžné použití se výše zmíněné vlastnosti nestanovují experimentálně, ale predikují se pomocí softwaru. Jedním z volně dostupných programů na webu je OligoCalc, který je schopný na základě primární struktury DNA či RNA, případně jejich modifikace a množství, zobrazit délku, procentuální obsah GC bází, molekulovou hmotnost, teplotu tání, absorpční koeficient nebo intermolekulární vazby a struktury či termodynamické konstanty. Teplota tání je v tomto případě určena třemi způsoby výpočtů. Mezi další softwary pak patří Biomath Calculator či OligoEvaluator. K predikci sekundární struktury a následného zobrazení výsledné kompozice se používá program RNAstructure, který využívá čtyři algoritmy pro predikci zahrnující výpočet stavby s nejnižší volnou energií a hledání struktury s maximální předpokládanou přesností. Pro predikci terciární struktury DNA či RNA slouží UNAFold. Komplexním zdrojem pro biologická data je iniciativa INSDC (International Nucleotide Sequence Database Collaboration), která funguje mezi NCBI (National Center for Biotechnology Information), DDBJ (DNA Data Bank of Japan) a EMBL-EBI (European Molecular Biology Laboratory's European Bioinformatics Institute). [45, 65, 75–77]

1.5.5 Vliv na HPLC separaci

Pro úspěšnou separaci na HPLC je nejdůležitější veličinou rozlišení píků. Obecně lze rozlišení ovlivnit třemi faktory: účinností, selektivitou a retencí. Účinnost lze zvýšit délkou kolony, zmenšením částic náplně kolony, odstraněním či redukcí chvostování píku (např. úpravou pH, zvýšením mrtvého objemu, snížením dávkovaného objemu atd.), zvýšením teploty či zvýšením extra-kolonového objemu. Selektivitu je možno upravit vhodným výběrem stacionární fáze, úpravou pH mobilní fáze či její výměnou za jiný roztok. Úpravu samotné retence lze provést především změnou polarity, ať už použitím slabšího rozpouštědla, silnější stacionární fáze nebo změnou ionizace analytu pomocí pH. Příklady ovlivnění separace oligonukleotidů změnou různých faktorů jsou zobrazeny na grafu 1, 2 a 3. [78]

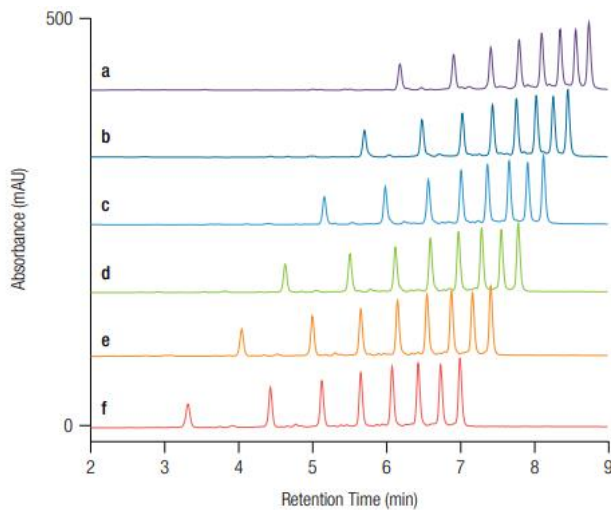


Column: **DNAPac RP**, 4 μ m
 Format: 2.1 \times 50 mm
 Mobile Phase A: 0.1 M TEAA, pH 7.0
 Mobile Phase B: 0.1 M TEAA in water/acetonitrile (25:75 v/v)
 Gradient:

Time (min)	%A	%B
-3.0	85	15
0.0	85	15
4.0	61	39
4.1	10	90
6.0	10	90

Gradient Curve: a) 5 (linear)
 b) 4
 c) 3
 d) 2
 Flow Rate: 0.60 mL/min
 Inj. Volume: 2 μ L
 Temp.: 60 $^{\circ}$ C
 Detection: UV (260 nm)
 Sample: 8-Combo DNA (5 μ M)

Graf 1: Vliv úpravy křivky gradientu na rozlišení separace oligonukleotidů. Převzato z [75].

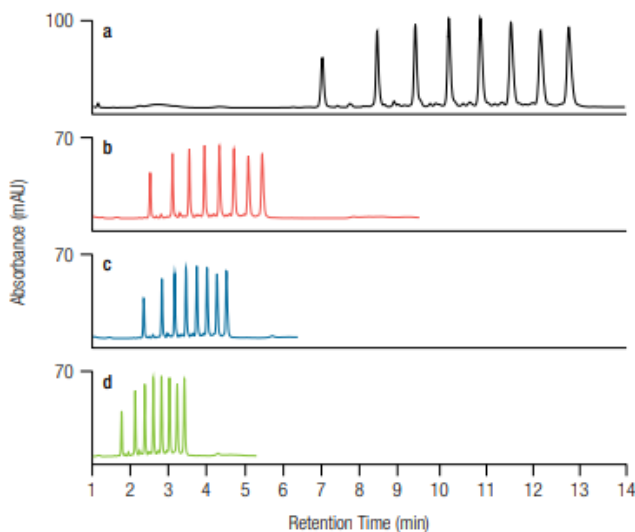


Column: **DNAPac RP**, 4 μ m
 Format: 2.1 \times 50 mm
 Mobile Phase A: 0.1 M TEAA, pH 7.0
 Mobile Phase B: 0.1 M TEAA in water/acetonitrile (25:75 v/v)
 Gradient:

Time (min)	%A	%B
-3.0	90	10
0.0	90	10
8.0	54	46
8.1	10	90
10.0	10	90

Flow Rate: 0.40 mL/min
 Inj. Volume: 2 μ L
 Temp.: a) 30 $^{\circ}$ C
 b) 40 $^{\circ}$ C
 c) 50 $^{\circ}$ C
 d) 60 $^{\circ}$ C
 e) 70 $^{\circ}$ C
 f) 80 $^{\circ}$ C
 Detection: UV (260 nm)
 Sample: 8-Combo DNA (5 μ M)

Graf 2: Vliv teploty na rozlišení separace oligonukleotidů. Převzato z [75]



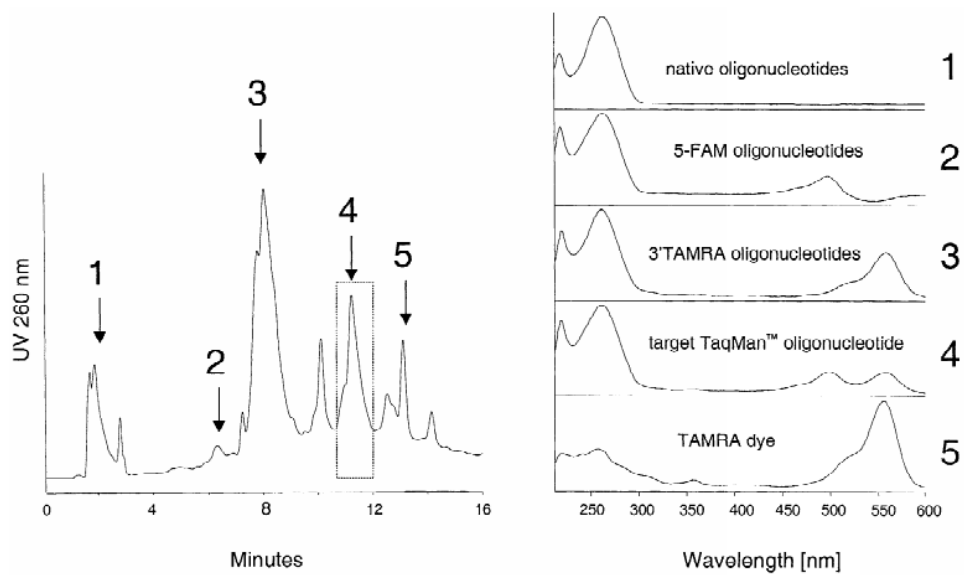
Column: **DNAPac RP**, 4 μ m
 Format: 2.1 \times 50 mm
 Mobile Phase A: 0.1 M TEAA, pH 7.0
 Mobile Phase B: 0.1 M TEAA in water/acetonitrile (75:25 v/v)
 Gradient:

Time (min)	%A	%B
15	37	63

Gradient Curve: a) 15 to 37%B in 12 min
 b) 15 to 37%B in 6 min
 c) 15 to 37%B in 4 min
 d) 15 to 37%B in 3 min
 Gradient Curve: 3
 Flow Rate: a) 0.20
 b) 0.40
 c) 0.60
 d) 0.80
 Inj. Volume: 2 μ L
 Temp.: 60 $^{\circ}$ C
 Detection: UV (260 nm)
 Sample: 8-Combo DNA (5 μ M)

Graf 3: Vliv průtoku a rychlosti dosažení gradientu na rozlišení separace oligonukleotidů. Převzato z [75].

HPLC separaci ovlivňuje z pohledu oligonukleotidů hned několik faktorů. Primárně je to samotná délka - delší oligonukleotidy mají vyšší retenční čas než ty kratší. Dále hraje roli pH mobilní fáze, které ovlivňuje náboj báze i kostry a vhodně zvolený pufr tedy může separaci zlepšit. I při teplotě separace okolo 50–60 °C nemusí docházet k denaturaci vodíkových vazeb a mohou zůstat zachovány sekundární struktury; nejběžnější typ vlásenky má vyšší distribuci náboje než lineární řetězec. Největší vliv má ovšem primární struktura a to, zda se jedná o RNA nebo DNA. Obecně se RNA sorbuje dříve kvůli extra hydroxylovým skupinám, které mají za následek o trochu větší polaritu řetězce. Samotné báze se vyznačují rozdílnou hydrofobicitou: $C < G < U < A < T$. To znamená, že oligonukleotid s větším obsahem GC se bude sorbovat dříve než oligonukleotid se stejnou délkou a nižším procentem GC. Celková hydrofobicita je dána součtem hydrofobicity jednotlivých bází a bude ji ovlivňovat i případné značení oligonukleotidu. Poněkud obtížnější je separace duálně značených sond, kdy se ve vzorku nachází nežádoucí částečně značené sondy a absolutně neznačené oligonukleotidy, které vznikly jako nechtěné produkty syntézy. V tomto případě se využívá detekce při dvou vlnových délkách (absorpční maxima obou značek) nebo duální HPLC purifikace (iontově výměnná následovaná reverzní). Příklad separace duálně značené sondy je na grafu 4. Pro predikci výsledného retenčního času je možno využít existujících modelů, jež mohou vycházet pouze ze samotné sekvence oligonukleotidu. Základní model počítá se sumou retenčních příspěvků jednotlivých nukleotidů, které byly získány experimentálně analýzou homooligonukleotidů. Vylepšení pak spočívá v započítání i vzniklé sekundární struktury. [71, 79–81]



Graf 4: Chromatograf purifikace duálně značené sondy a absorpční spektra jednotlivých látek. Převzato z [71].

2 CÍL PRÁCE

Cílem této diplomové práce je analyzovat možnou závislost mezi plochou pod píkem a koncentrací stanovovaných oligonukleotidů a duálně značených sond s ohledem na jejich délku, obsah GC a případnou modifikaci, tuto závislost definovat a pokusit se o stanovení koeficientu predikce pro další oligonukletidy a sondy. Tohoto cíle bude dosaženo pomocí následujících dílčích kroků:

- výběr oligonukleotidů a sond a definice jejich základních vlastností, jako je délka, obsah GC, molární absorpční koeficient a teplota tání,
- stanovení koncentrace zásobních roztoků výpočtem z naměřené absorbance při 260 nm,
- optimalizace replikátů nástřiků,
- výběr vhodných koncentračních bodů pro analýzu,
- provedení HPLC analýzy s odběrem frakcí,
- měření absorbance purifikátů,
- zpracování dat,
- stanovení závislosti mezi výstupem HPLC analýzy a molárním množstvím stanoveným pomocí měření absorbance.

3 EXPERIMENTÁLNÍ ČÁST

Experimentální část diplomové práce byla realizována ve společnosti GENERI BIOTECH s.r.o. v Hradci Králové. Termínem oligonukleotid je v následující části zamýšlen neznačený oligonukleotidový řetězec, termín sonda pak označuje duálně značený oligonukleotid.

3.1 Metodika

3.1.1 Výběr oligonukleotidů a sond

Pro experimentální část bylo vybráno celkem 27 oligonukleotidů v délce od 18 do 43 nukleotidů s obsahem GC bází od 23 do 92 %. Pro zjednodušené označení jednotlivých oligonukleotidů se využíval formát X/Y, kdy X značilo procentuální obsah GC bází a Y celkovou délku. V případě shody délky a množství GC bylo za značení přidáno písmeno. Přehled jednotlivých oligonukleotidů, jejich sekvence a označení je v tabulce 5.

Tabulka 5: Vybrané oligonukleotidy.

Sekvence	% GC	Délka	ID
AAAAAAGGCAAGATGCTG	39	18	39/18
GATGGTTCGGATAATGCG	50	18	50/18
GCCACCCTGCGGGACTTC	72	18	72/18
TACTCACAAAGCATTCTTA	35	20	35/20
CGCCAACACAGAAATTATTG	40	20	40/20
GGTCCGCTGCGAGGCTTGC	70	20	70/20
AAAAATAGCTTATCAGACTGATG	30	23	30/23
GCTGGAAGCCTGTCAAAGAGAGA	52	23	52/23
GGGGTACCGTGTAGACCTGGAGC	65	23	65/23
GATCCAAGACTCTGTTTTCTAAATC	36	25	36/25
ACTCGCTCAACGGCCCGCTGCACGC	72	25	72/25
CCGCCGCCGTCGCACCCGGCCCCGC	92	25	92/25
GTTTTTCTCTTTCTGGTAGGACAAATA	32	28	32/28
ATCTTAAGTCACACTGGCAGTCCAGAC	50	28	50/28
ACGCTAGCCGCCATGACAACCAAGTCACC	61	28	61/28
AAGCATTTGGTTTTAAATTATGGAGTATATT	23	31	23/31
GAAGATCTACCACCATCACCATCACCATCAC	48	31	48/31
GCTAGCCACCATGCGGCTCTGGAGTTGGGTG	65	31	65/31
AAAAACAGGAAAAATGTAAGTACTTAGACCCTACAATG	32	34	32/34
GCTCTAGAGCCATGGAAGTCTGGGTGTCTGCATT	53	34	53/34
GGCCCAGCCGGCCGACCTGGGCCCGCAGATGCTTC	77	35	77/35
AAAAAAAAAAAAAAAAAAAAATGTAAACGACGGCCAGT	24	38	24/38
CAGGAAACAGCTATGACCACTGTAAGTGGTTTCTCAGGA	46	39	46/39
ACCCAAGCTGGCTAGCCACCATGGTTCGCCTCGGGGCTC	67	39	67/39

<i>Sekvence</i>	<i>% GC</i>	<i>Délka</i>	<i>ID</i>
TGTAAAACGACGGCCAGTAATTATTGTTTTCTCTTAGATATGC	35	43	35/43
GCGATCGTCACTGTTCTCCACTCCTGATATTTTTGCCTTGATC	47	43	47/43/A
CAGGAAACAGCTATGACTAGACTTCTGGGTGGCCAAAGAATCA	47	43	47/43/B

Pro analýzu sond byly vybrány duálně značené sondy s fluorescenční značkou FAM, HEX a Cy5 v délce od 16 do 33 nukleotidů a obsahu GC v rozmezí 35–82 %. Identifikační kód sondy byl tvořen obdobně jako u oligonukleotidů s přidáním značení. Přehled vybraných sond je uveden v tabulce 6.

*Tabulka 6: Vybrané sondy.*Pro HPLC měření byla z důvodu nedostatečného množství původní sondy využita i sonda se stejnou značkou i sekvencí lišící se pouze ve zhášeči, proto jsou údaje k tomuto ID dvoji.*

<i>Sekvence</i>	<i>Sonda</i>	<i>% GC</i>	<i>Délka</i>	<i>ID</i>
FAM				
<i>TCTATGATCATGAGAGT</i>	FAM-BHQ1	35	17	35/17_FAM
<i>ATCATGAGAGTCGCCGT</i>	FAM-BHQ1	53	17	53/17_FAM
<i>CCCCACCAGCCTGCCC</i>	FAM-BHQ1	81	16	81/16_FAM
HEX				
<i>CCTCCAATGGGTCTCCAGGG</i>	HEX-BHQ1	67	21	67/21_A_HEX
<i>ACACTGGCAAAACAATGCAGAC</i>	HEX-BHQ1	45	22	45/22_HEX
<i>CCGACACCACGCCCCC</i>	HEX-BHQ1	82	17	82/17_HEX*
	HEX-BBQ			
Cy5				
<i>GTTCTGGTTACTATAATGATCCACTGGTCTGT</i>	Cy5-BHQ2	41	32	41/32_Cy5
<i>TTCTGGTTACTATAATGATCCACTGGTCTGTA</i>	Cy5-BHQ2	38	32	38/32_Cy5
<i>GTTCTGGTTACTATAATGATCCACTGGTCTGTA</i>	Cy5-BHQ2	39	33	39/33_Cy5

3.1.2 Definice základních vlastností vybraných vzorků

Molární absorpční koeficient byl definován prostým součtem absorpčních koeficientů jednotlivých nukleotidů řetězce, kdy se pro měření při 260 nm počítalo s hodnotami uvedenými v kapitole 1.3.1.1 *Molární absorpční koeficient*. Prostý součet byl využit především díky dobrým zkušenostem s touto metodou v laboratořích, kde výzkum k diplomové práci probíhal. Teplota tání byla stanovena pomocí kalkulace vycházející ze sekvence. Přehled vlastností oligonukleotidů a sond je v tabulce 7.

Tabulka 7: Základní fyzikální vlastnosti vybraných oligonukleotidů.

ID	T_M [°C]	ε [mM ⁻¹ cm ⁻¹]
39/18	58,6	229,3
50/18	61,1	202,1
72/18	72,2	174,1
35/20	58,2	216,1
40/20	60,1	230,0
70/20	74,9	196,8
30/23	58,2	275,5
52/23	68,8	272,4
65/23	73,5	250,3
36/25	61,5	265,9
72/25	81,8	245,8
92/25	90,8	222,9
32/28	63,1	294,5
50/28	71,7	300,2
61/28	77,7	295,7
23/31	61,4	354,7
48/31	71,7	331,7
65/31	80,7	322,0
32/34	66,7	416,9
53/34	76,5	355,8
77/35	89,6	346,7
24/38	67,3	514,5
46/39	74,8	439,4
67/39	86	392,2
35/43	70,9	470,3
47/43/A	76,1	418,8
47/43/B	76,4	495,8
35/17_FAM	52,3	220,1
35/17_FAM	52,3	220,1
53/17_FAM	62,8	213,4
81/16_FAM	73,2	172,2
67/21_A_HEX	72,4	249,58
45/22_HEX	66,6	301,78
82/17_HEX	75,4	196,78
	75,4	203,85
41/32_Cy5	68,6	350,5
38/32_Cy5	67,4	354,2
39/33_Cy5	68,2	365,9

3.1.3 Stanovení koncentrace zásobních roztoků

Použité roztoky a chemikálie:

- zásobní roztoky oligonukleotidů a sond,
- IWA voda (voda pro analytické účely).

Laboratorní vybavení:

- laminární box,
- automatické mikropipety, kalibrované, různé objemy,
- sterilně balené špičky pro mikropipety, různé objemy,
- mikrozkušavky, různé objemy,
- centrifuga pro mikrozkušavky,
- mikrotitrační destička s plochým dnem, 96 jamek,
- třepačka vortex,
- absorpční spektrofotometr Tecan Infinite 200Pro.

Pracovní postup:

1. Pro měření bylo do jamek napipetováno po 200 μl IWA vody a 2 μl vzorku oligonukleotidu, přičemž každý vzorek byl měřen v duplikátu. Jako slepý vzorek bylo použito 200 μl IWA vody.
2. Destička s napipetovanými a promíchanými vzorky byla proměřena na přístroji Tecan Infinite 200Pro při vlnové délce 260 nm.
3. Byl proveden manuální odečet výsledných hodnot naměřených absorbancí roztoků a slepých vzorků.

3.1.4 HPLC analýza

Použité roztoky a chemikálie:

- zásobní roztoky oligonukleotidů a sond,
- mobilní fáze pro analýzu oligonukleotidů
 - 0,1M TEAA,
 - acetonitril,
- mobilní fáze pro analýzu sond
 - směs methanolu a acetonitrilu v poměru 1:1,
 - 50mM TEAA.

Laboratorní vybavení:

- běžné laboratorní vybavení viz kapitola 3.1.3 *Stanovení koncentrace zásobních roztoků*,
- filtrační aparatura s 0,45 μ m PVDF (polyvinylidenfluoridovou) membránou Durapore[®],
- Agilent 1260 Infinity II LC systém,
- kolona Clarity[®] 5 μ m Oligo-RP, LC Column 150 x 4,6 mm – stacionární fáze C18 na plně porézní sílice (velikost částic 5 μ m, velikost pórů 110 Å) pro analýzu oligonukleotidů,
- kolona Arion[®] Plus C18 HPLC Column 100 x 4,6 mm – stacionární fáze C18 na plně porézní sílice (velikost částic 5 μ m, velikost pórů 100 Å) pro analýzu sond,
- centrifugační koncentrátor.

Software:

- analýza byla sledována za využití softwaru Agilent Instrument Control Framework dodaného výrobcem přístroje,
- vyhodnocení bylo provedeno pomocí tabulkového procesoru Microsoft Excel.

Pracovní postup – validace opakovatelnosti:

1. pro validaci opakovatelnosti výsledků byl zvolen oligonukleotid, který nebyl zařazen do testovaných vzorků. Sekvence: ATCATTGCTATGTCCACCACGGGAG, GC = 52 %, $T_M = 70,66$ °C.
2. Bylo připraveno celkem šest zásobních roztoků; tři o teoretické koncentraci 100 μM a tři o koncentraci 50 μM .
3. Byla připravena mobilní fáze pro analýzu oligonukleotidů, která byla přefiltrována pod tlakem na 0,45 μm PVDF membráně Durapore®.
4. Z každého roztoku byly provedeny tři nástříky o objemu 15 μl . Analýza probíhala v gradientu mobilní fáze 0,1M TEAA a čistého acetonitrilu. Teplota na koloně byla 55 °C, průtok 1,1 ml/min, doba analýzy 13 min, signál byl detekován při 260 nm - tedy stejné nastavení jako při samotné analýze vybraných vzorků.

Pracovní postup – výběr vhodných koncentračních bodů:

1. byly zvoleny 2 oligonukleotidy se středním obsahem GC a nejnižším a nejvyšším molárním absorpčním koeficientem. V případě sond byl zvolen jeden zástupce daného značení.
2. Ze zvolených vzorků byla připravena dvanáctibodová kalibrační řada s ředícím faktorem 1,5 s nejvyšším teoretickým bodem koncentrace 100 μM a nejnižším 0,77 μM (bod o koncentraci 1,16 μM byl vyjmut).
3. Byly připraveny mobilní fáze pro analýzu oligonukleotidů a sond, které byly přefiltrovány pod tlakem na 0,45 μm PVDF membráně Durapore®.
4. Připravené kalibrační řady byly proměřeny na HPLC, nástřík o objemu 15 μl probíhal vždy v duplikátu. Podmínky pro měření oligonukleotidů jsou popsány výše. Pro analýzu sond byl zvolen gradient mobilní fáze směsi methanolu a acetonitrilu (v poměru 1:1) a 50mM roztoku TEAA. Nastavení analyzátoru pro sondy se měnilo dle jednotlivého značení následovně:
 - FAM: teplota na koloně 50 °C, průtok 0,9 ml/min, doba analýzy 12,5 min,

- HEX: teplota na koloně 50 °C, průtok 1,7 ml/min, doba analýzy 16,5 min,
- Cy5: teplota na koloně 50 °C, průtok 1,5 ml/min, doba analýzy 19,1 min.

Pracovní postup – HPLC analýza:

1. byly připraveny roztoky vzorků oligonukleotidů a sond ve zvolených teoretických koncentracích, přičemž každý koncentrační bod byl připraven ve dvou alikvotech a příprava byla provedena vždy těsně před samotným měřením.
2. Z každého alikvotu byly provedeny dva nástřiky na kolonu.
3. Výsledky měření byly zaznamenány do tabulky a dále zpracovány.
4. Z každého nástřiku byl automatickým kolektorem odebrán do sběrné zkumavky purifikovaný vzorek. Tyto purifikáty byly vysušeny v centrifugačním koncentrátoru a následně rozpuštěny ve 400 µl IWA vody, čímž bylo dosaženo shodného objemu pro všechny získané vzorky.

3.1.5 Měření absorbance purifikátů

Měření absorbance získaných purifikátů probíhalo obdobně jako při stanovení koncentrace zásobních roztoků. Z popsanych sběrných zkumavek bylo do dvou jamek 96jankové destičky s plochým dnem pipetováno po 200 µl vzorku, měření tedy bylo v duplikátu. Absorbance na destičce byla změřena na přístroji Tecan při 260 nm s korekcí na jeden centimetr. Výsledná absorbance vzorku je uvedena jako průměr dvou jamek.

3.1.6 Zpracování a analýza dat

Pro prvotní zpracování dat a jejich základní vizualizaci byl využit software Excel z balíčku Microsoft Office; šlo především o výpočty koncentrací z absorbancí. Pro následné zpracování dat byl zvolen programovací jazyk R v grafickém rozhraní RStudio, který je primárně zaměřen na statistické a grafické zpracování dat a ve kterém byla tvořena většina grafů a náročnější statistické výpočty včetně modelu predikce. Vedle základního balíčku basic bylo dále užito balíčků readxl a tidyverse, které sloužily pro import dat z MS Excel, jejich úpravu a analýzu. Grafické zpracování dat bylo provedeno za pomoci balíčků scales, gridExtra, ggplot2, ggforce a nlme. [82–90]

3.1.6.1 Obecná a popisná statistika

Jako jednoduchý orientační nástroj pro ověření normality dat, byl využit hrubý odhad, v rámci kterého by se aritmetický průměr neměl lišit oproti mediánu o více než 10 %.

Pro validaci opakovatelnosti nástřiků byl využit nástroj Popisná statistika z doplňku Analytické nástroje pro Excel, která poskytuje informace o celkovém trendu a variabilitě dat vybrané oblasti, pro níž byla zvolena hladina spolehlivosti 95 %. Výpočty byly provedeny v Microsoft Excel.

3.1.6.2 Směrodatná odchylka

Základním matematickým nástrojem, ze kterého vycházely i pokročilejší statistické analýzy, byla směrodatná odchylka, což je odmocnina z rozptylu využívaná pro popis souborů. Pro výpočet byl použit následující vztah:

$$SD = \sqrt{\frac{\sum(x - \bar{x})^2}{(n - 1)}}$$

SD směrodatná odchylka

x hodnota

\bar{x} aritmetický průměr

n počet hodnot souboru

Rovnice 2: Výpočet směrodatné odchylky.

Relativní směrodatná odchylka, někdy též variační koeficient, porovnává směrodatnou odchylku s průměrem a zpravidla se udává v procentech.

$$rSD = \frac{SD}{\bar{x}} \cdot 100$$

rSD relativní směrodatná odchylka

Rovnice 3: Výpočet relativní směrodatné odchylky.

Jako SDR je v práci označována směrodatná odchylka reziduí. Reziduum je definováno jako rozdíl mezi pozorovanou a predikovanou hodnotou.

Výpočty *SD* a *rSD* byly provedeny v Microsoft Excel, výpočet SDR byl součástí regresní analýzy základního balíčku RStudio.

3.1.6.3 Grubbsův test odlehlosti bodů

Grubbsův test přítomnosti odlehlých bodů slouží k vyloučení extrémních hodnot u souborů s Gaussovým rozložením dat. Pro výpočet testovacích kritérií byly použity tyto rovnice:

$$T_1 = \frac{\bar{x} - x_1}{SD} \quad T_n = \frac{\bar{x} - x_n}{SD} \quad T \dots\dots \text{testovací kritérium}$$

Rovnice 4: Výpočet parametrů Grubbsova testu.

Jako kritická hodnota pro porovnání při hladině spolehlivosti 99 % a velikosti výběru 9 byla stanovena na 2,387. Tato byla převzata z tabulky kritických hodnot pro Grubbsův test. Výpočty byly provedeny v Microsoft Excel. [91]

3.1.6.4 F-test

F-test byl využit k testování významnosti rozdílu mezi soubory. Výpočet testovacího kritéria se prováděl podle vzorce:

$$F = \frac{SD_1^2}{SD_2^2} \quad F \dots\dots \text{testovací kritérium}$$

Rovnice 5: Výpočet parametru F-testu

Ze získaného testovacího kritéria byla následně s ohledem na stupně volnosti a hladinu spolehlivosti vypočítána p-hodnota. V případě, kdy je p-hodnota nižší než koeficient spolehlivosti α , je nulová hypotéza zamítnuta a platí hypotéza alternativní. [92]

Současně byl uveden odhad síly hypotézy koeficientem determinace R^2 a adjustovaným koeficientem determinace \bar{R}^2 . Koeficient determinace R^2 nabývá hodnot od 0 do 1, zároveň s hodnotou roste míra závislosti veličin, tedy jak dobře data odpovídají predikci průběhu funkce a udává tedy informaci o kvalitě modelu. Adjustovaný koeficient determinace \bar{R}^2 se pak využívá u vícenásobných regresních modelů, který na rozdíl od R^2 , zohledňuje vliv počtu prediktorů.

3.1.6.5 Stanovení koncentrace zásobních roztoků

Pro výpočet koncentrace byl využit Lambert-Beerův zákon, viz rovnici 6.

$$c = \frac{Abs}{\varepsilon_{260} \cdot l}$$

$c \dots\dots$ koncentrace
 $Abs \dots\dots$ absorbance
 $l \dots\dots$ optická vzdálenost
 $\varepsilon_{260} \dots\dots$ molární absorpční koeficient

Rovnice 6: Rovnice popisující Lambert-Beerův zákon transformovaná pro přímý výpočet koncentrace.

Výsledná koncentrace oligonukleotidů byla stanovena průměrem hodnot z obou přístrojů. Výsledná koncentrace zásobních roztoků sond byla stanovena průměrem získaných hodnot.

3.1.6.6 Lineární regrese a konečná úprava dat

Předpokladem pro správnou interpretaci výsledků je postupná úprava rovnic vycházející z rovnice Lambert-Beerova zákona (rovnice 7) v obecném stavu.

$$Abs = \varepsilon \cdot c \cdot l$$

Rovnice 7: Rovnice popisující Lambert-Beerův zákon.

S ohledem na to, že při měření absorbance proběhla vždy korekce tloušťky absorbující vrstvy na jeden centimetr, lze pro další úpravu vzorce veličinu l zanedbat, čímž lze získat upravenou rovnici 8.

$$Abs = \varepsilon \cdot c$$

Rovnice 8: Zjednodušená rovnice Lambert-Beerova zákona.

Dalším krokem úpravy vzorce je vyjádření koncentrace jako podílu počtu molů ku objemu. Mírnou transformací vznikne rovnice 9.

$$n = \frac{Abs \cdot V}{\varepsilon}$$

Rovnice 9: Transformovaná rovnice s dosažením koncentrace jako podílu počtu molů a objemu.

Obecnou rovnici pro lineární regresi získaných hodnot, tedy vlivu koncentrace na výsledné ploše píku, lze vyjádřit rovnicí 10, kde přímka prochází počátkem a k je směrnice.

$$Abs \cdot V = k \cdot Area$$

Rovnice 10: Rovnice lineární regrese.

Výsledným spojením rovnice 9 a 10 je poté závislost, která vede k výpočtu molárního množství oligonukleotidu či sondy na základě jeho naměřené plochy a molárního absorpčního koeficientu, viz rovnici 11, dle které lze pak spočítat i koncentraci.

$$n = \frac{k \cdot Area}{\varepsilon}$$

Rovnice 11: Výpočet množství oligonukleotidu či sondy ze získané plochy píku.

Vzhledem k tomu, že se pro lineární regresi vycházelo z teze, že nulová koncentrace bude vykazovat nulovou absorbanci, je pro výpočty aproximace přímkou vedena skrz počátek. Výpočty byly provedeny v jazyku R s využitím základního balíčku funkcí. Míra závislosti regrese přímkou byla stanovena koeficientem determinace R^2 . [93, 94]

V absolutních číslech bylo pro jeden vzorek o čtyřech různých koncentracích, kdy každý koncentrační bod byl po dvou alikvotech a každý alikvot byl aplikován ve dvou nástricích, získáno celkem 16 údajů o ploše píku a jeho výšce. Pro všech 27 vzorků oligonukleotidů a 9 sond to dělá 576 různých píků jakožto výchozích hodnot pro datovou analýzu a tvorbu predikčního modelu. Z důvodu větší přesnosti informace o píku byla zpracována pouze získaná data ploch, nikoliv výšek.

Protože odběr purifikované frakce vzorku byl především u nižších koncentrací méně přesný, byla provedena manuální korekce plochy tak, aby odpovídala reálné odebrané frakci.

4 VÝSLEDKY

4.1 Stanovení koncentrace zásobních roztoků

Výsledky měření z přístroje Tecan jsou zobrazeny v tabulce 8 a 9.

Tabulka 8: Koncentrace zásobních roztoků oligonukleotidů.

<i>ID</i>	<i>Koncentrace[μM]</i>	<i>ID</i>	<i>Koncentrace [μM]</i>
39/18	115,44	61/28	100,95
50/18	124,22	23/31	102,71
72/18	109,64	48/31	116,59
35/20	118,76	65/31	102,88
40/20	106,72	32/34	107,19
70/20	109,95	53/34	111,43
30/23	105,55	77/35	119,73
52/23	112,35	24/38	112,21
65/23	100,54	46/39	104,59
36/25	109,98	67/39	107,28
72/25	103,38	35/43	113,98
92/25	110,85	47/43/A	100,04
32/28	105,32	47/43/B	91,79
50/28	109,65		

Tabulka 9: Koncentrace zásobních roztoků sond.

<i>ID</i>	<i>Koncentrace [μM]</i>
35/17_FAM	216,31
53/17_FAM	166,54
81/16_FAM	83,25
67/21_A_HEX	81,91
45/22_HEX	94,81
82/17_HEX	111,46
	122,79
41/32_Cy5	94,81
38/32_Cy5	93,91
39/33_Cy5	90,00

4.2 HPLC analýza

4.2.1 Validace opakovatelnosti

Validace opakovatelnosti byla provedena především kvůli zavádění metody, tedy pro ověření míry shody ploch a výšek píků jednotlivých nástřiků a možnou další optimalizaci samotné analýzy. Získaná data byla podrobena popisné statistice (tabulka 10 a 11), kdy procentuální rozdíl průměru a mediánu byl menší než 10 % (menší než 1 %), tudíž nejsou předpokládány asymetrie a odlehlé hodnoty. Dále byl proveden i Grubbsův test odlehlosti bodů, který opět potvrdil nepřítomnost odlehlých hodnot, a test významnosti rozdílu. Veškeré provedené statistické testy naznačují, že míra shody mezi jednotlivými nástřiky je vysoká (hladina spolehlivosti byla 95 %). Celková relativní směrodatná odchylka získaná z průměrů odchylek z tabulky 10 u roztoků o teoretické vstupní koncentraci 100 μM je 1,36 pro plochu a 0,24 pro výšku. U koncentrace 50 μM pak 1,57, resp. 1,26. Celková míra variability vyjádřená relativní směrodatnou odchylkou je 1,11, tedy nižší než 5 %. Na základě těchto dat lze stanovit, že pro spolehlivou analýzu stačí využít pouze nástřiků v duplikátu, jelikož míra shody mezi jednotlivými nástřiky je pro potřeby práce vyhovující.

Tabulka 10: Naměřené plochy a výšky piku pro validaci opakovatelnosti.

Plocha			
<i>c</i> [μM]	<i>Průměr A</i>	<i>SD</i>	<i>rSD</i> [%]
100	1,74E+04	238,16	1,36
50	1,02E+04	161,59	1,57

Výška			
<i>c</i> [μM]	<i>Průměr A</i>	<i>SD</i>	<i>rSD</i> [%]
100	2,62E+03	6,32	0,24
50	1,78E+03	22,50	1,26

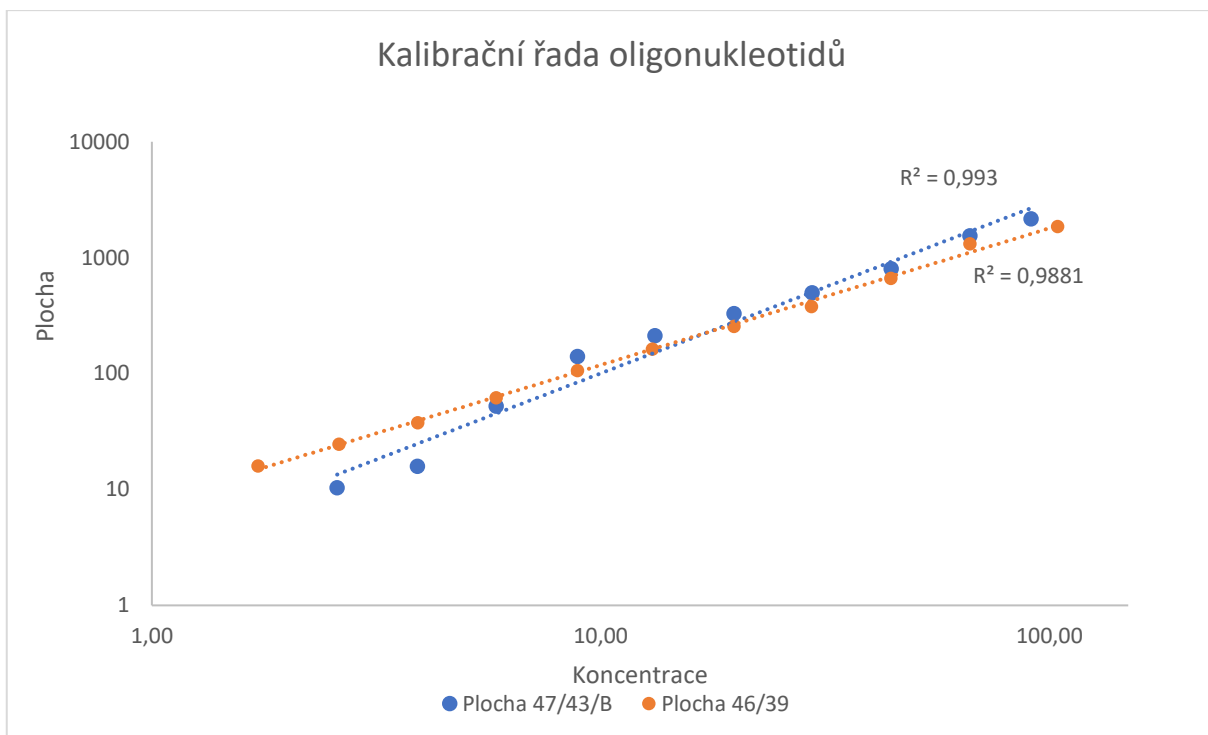
Tabulka 11: Výsledky validace opakovatelnosti.

Orientační test		Plocha		Výška		
Koncentrace	Průměr	Medián	Rozdíl [%]	Průměr	Medián	Rozdíl [%]
100 μM	17535,5	17558,4	0,13	2624,4	2635,52	0,42
50 μM	10245,6	10319,1	0,72	1778,44	1830,63	2,93

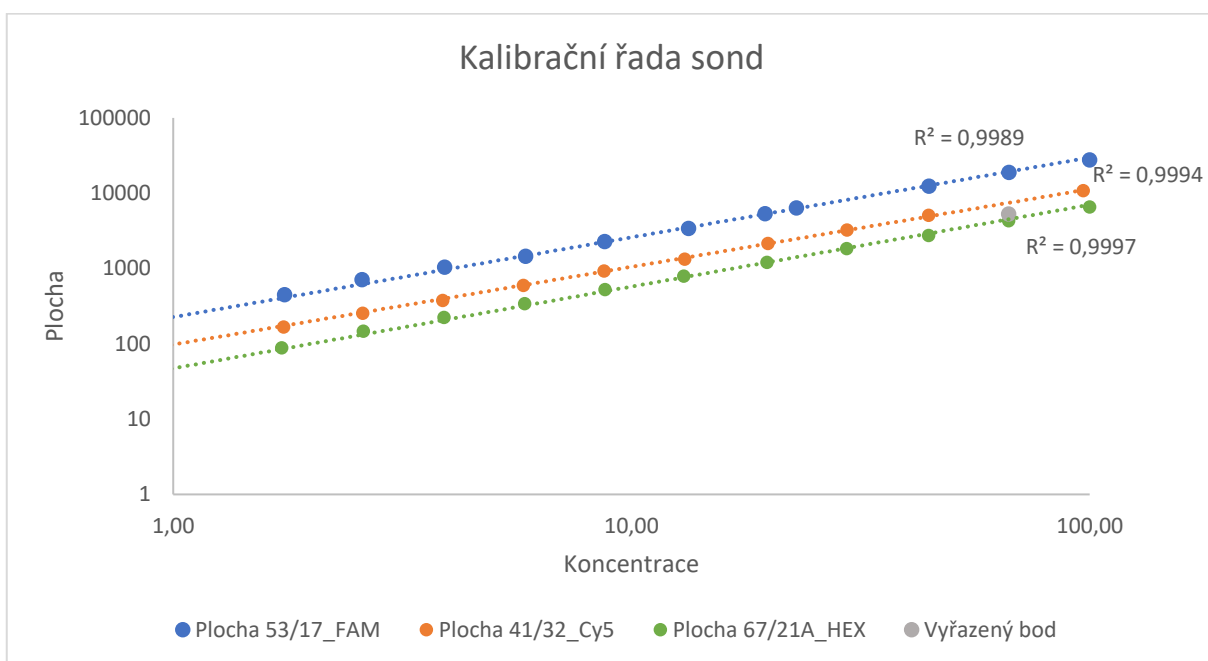
Grubbsův test		Plocha		Výška	
Koncentrace	100 μM	50 μM	100 μM	50 μM	
Aritm. průměr	17535,50	2624,40	10245,60	1778,44	
SD	260,43	24,35	313,48	101,31	
Minimum	17149,20	2582,67	9749,77	1625,59	
T_1	1,48	1,71	1,58	1,51	
Maximum	17925,30	2648,09	10685,20	1877,39	
T_n	1,50	0,97	1,40	0,98	

4.2.2 Výběr vhodných koncentračních bodů:

Výsledky měření jsou zobrazeny na grafu 5 a 6. Pro nižší vstupní koncentrace, ač byly detekovatelné, nebylo možno odebrat z důvodu nízkého signálu frakci, proto nebyly zařazeny do hodnocení kalibrační řady. Na základě těchto výsledků byly pro samotnou analýzu vzorků vybrány body o teoretické koncentraci 5,85 μM , 16,16 μM , 29,60 μM a 100 μM pro oligonukleotidy a 5,85 μM , 19,73 μM , 44,4 μM a 100 μM pro sondy, kdy první bod kalibrační řady je nejnižší odezva pro odběr frakce, poslední bod je nominální koncentrace zásobních roztoků. Prostřední dva body jsou poté rozloženy napříč škálou. U všech těchto vybraných bodů byla zjištěna lineární odezva absorbance na koncentraci. Jelikož se jedná o vstupní koncentrace před nástřikem, není nutné, aby korelace procházely bodem 0.



Graf 5: Kalibrační řady oligonukleotidů, logaritmické měřítko. Uvedené koncentrace odpovídají reálným koncentracím připravených roztoků.



Graf 6: Kalibrační řady sond, logaritmické měřítko. Uvedené koncentrace odpovídají reálným koncentracím připravených roztoků.

4.3 Analýza oligonukleotidů

4.3.1 Hromadná analýza dat

Naměřené hodnoty ploch a absorbancí vynásobených výsledným objemem (400 μ l) byly vyneseny do grafu (graf 7) s logaritmickým měřítkem a rozlišením jednotlivých vzorků a alikvotů. Z tohoto grafu je jasně patrná lineární závislost měřených hodnot. Byly provedeny následující statistické analýzy:

F-test oboustranný:

Cílem testu bylo zamítnutí nulové hypotézy a potvrzení hypotézy alternativní.

H_0 = regresní koeficient je roven nule

H_A = regresní koeficient není roven nule

$\alpha = 0,05$

hladina jistoty = 95 %

F-test: 7233 a 431 stupňů volnosti, p-hodnota $< 2,2 \cdot 10^{-16}$

Nulová hypotéza se zamítá, platí alternativní hypotéza.

Pro výpočet byl využit jazyk R, základní balíček, funkce `lm()`.

Lineární regrese (procházející počátkem):

Cílem testu bylo získat rovnici lineární regrese a její koeficient determinace.

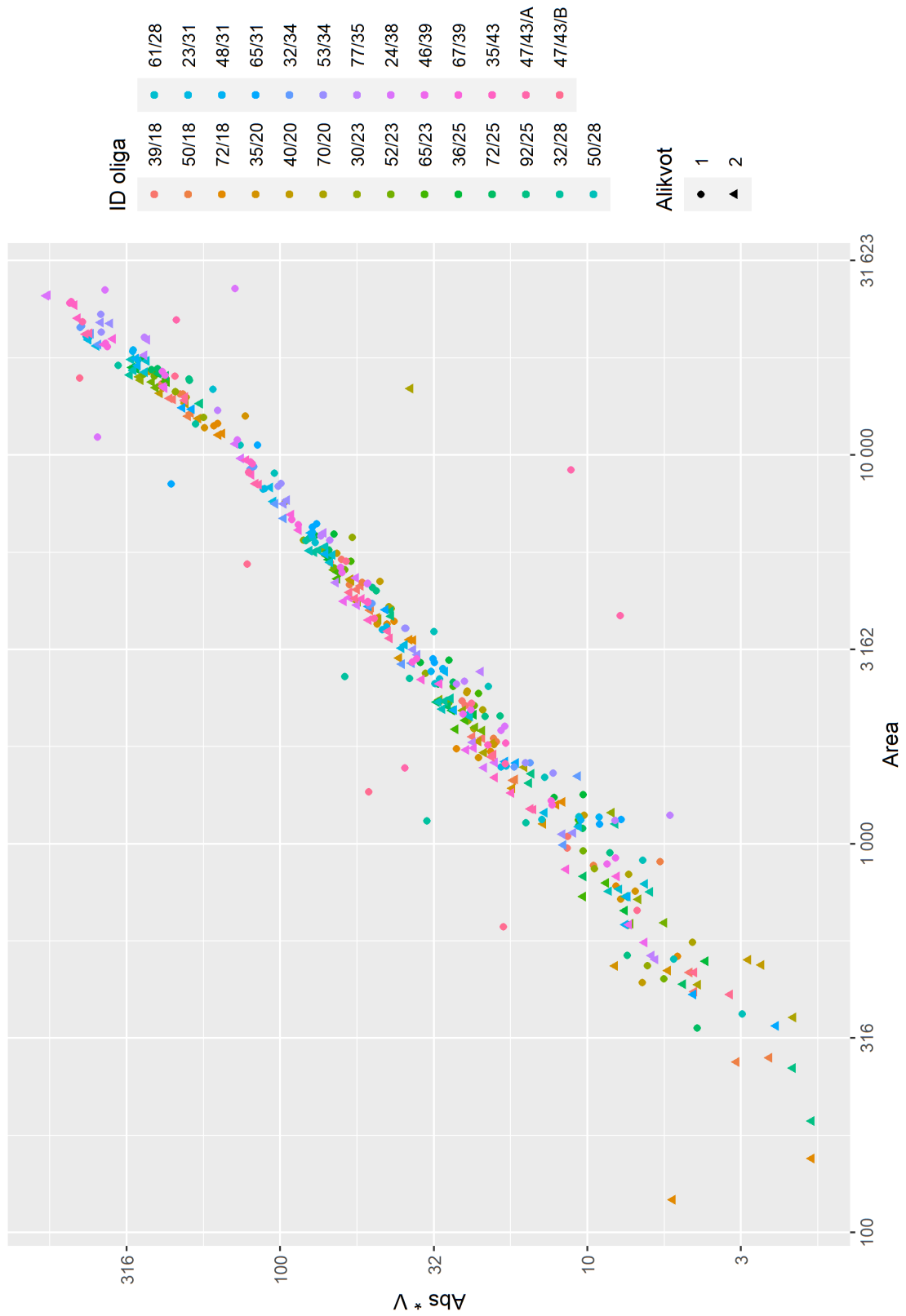
$Abs \cdot 400 = 0,0163 \cdot Area$

koeficient determinace $R^2 = 0,9438$

adjustovaný koeficient determinace $\bar{R}^2 = 0,9436$

Pro výpočet byl využit jazyk R, základní balíček, funkce `lm()`.

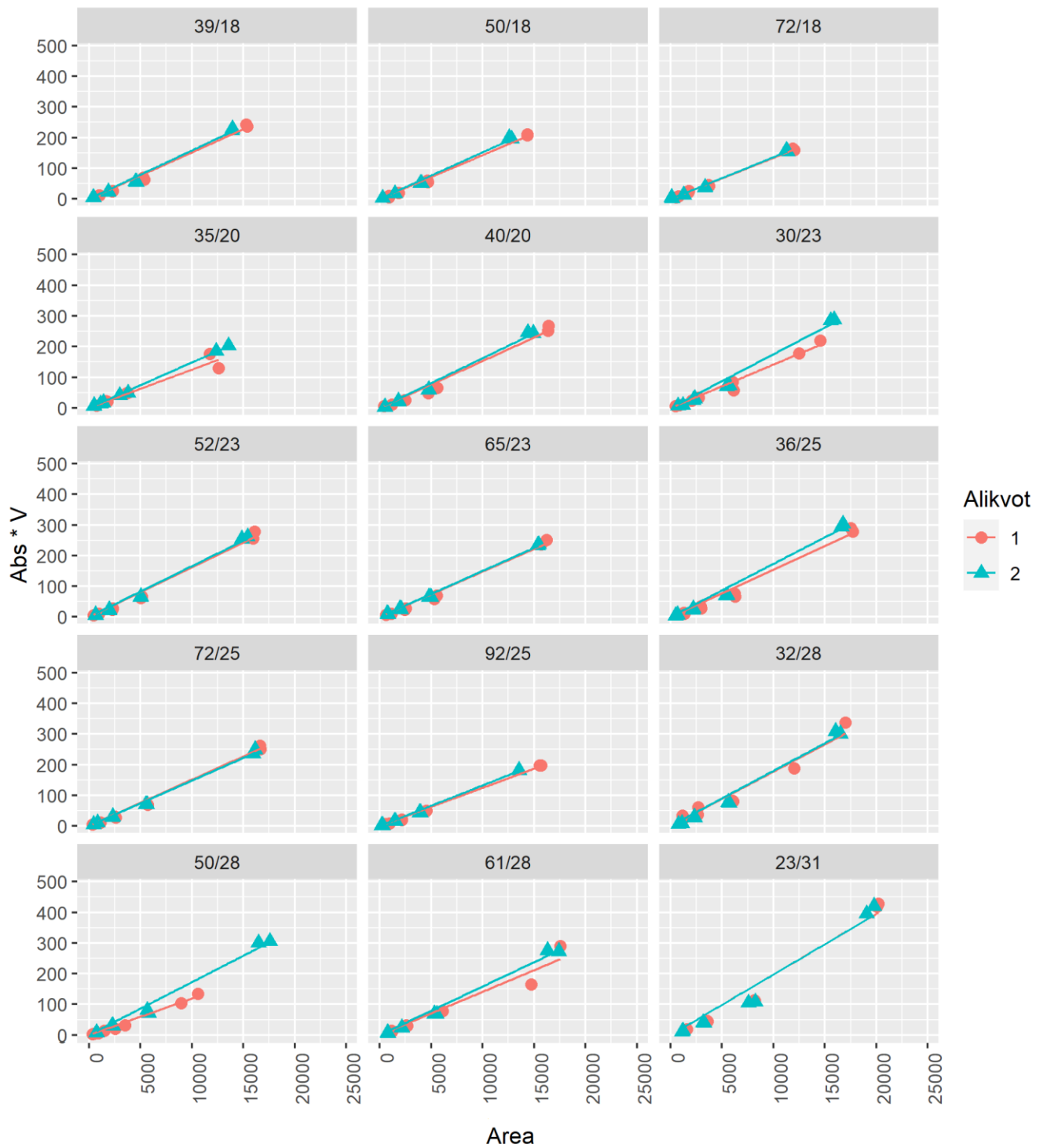
Z výsledků analýz lze usuzovat, že závislost je opravdu lineární, hodnoty korelují a lze tedy přímku lineární regrese použít k hrubému odhadu výsledné absorbance, resp. množství.



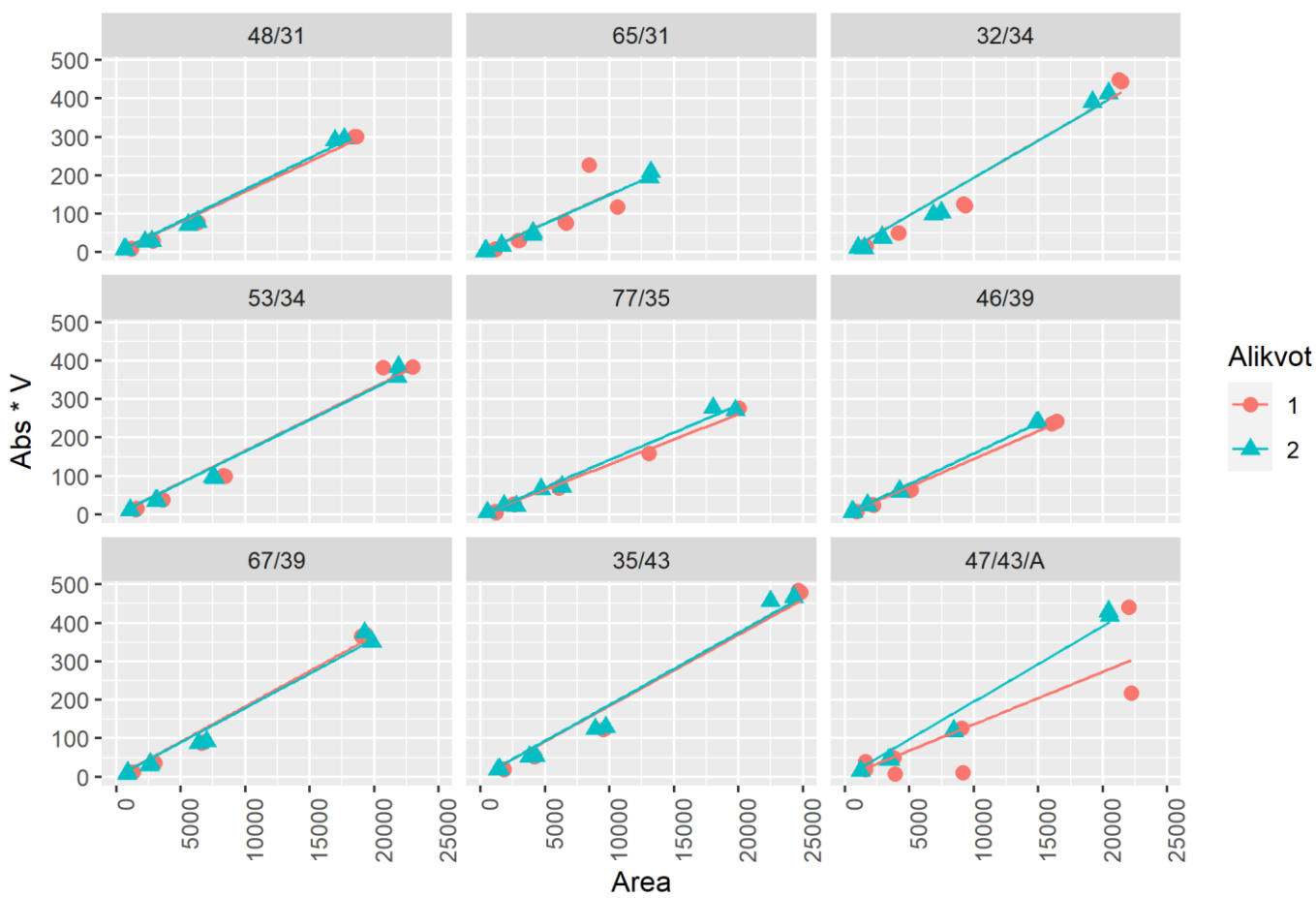
Graf 7: Závislost absorbance vynásobené objemem na ploše píku pro všechny oligonukleotidy. Jednotlivé vzorky jsou odlišeny barvou, jednotlivé alikvoty pak tvarem.

4.3.2 Analýza jednotlivých oligonukleotidů

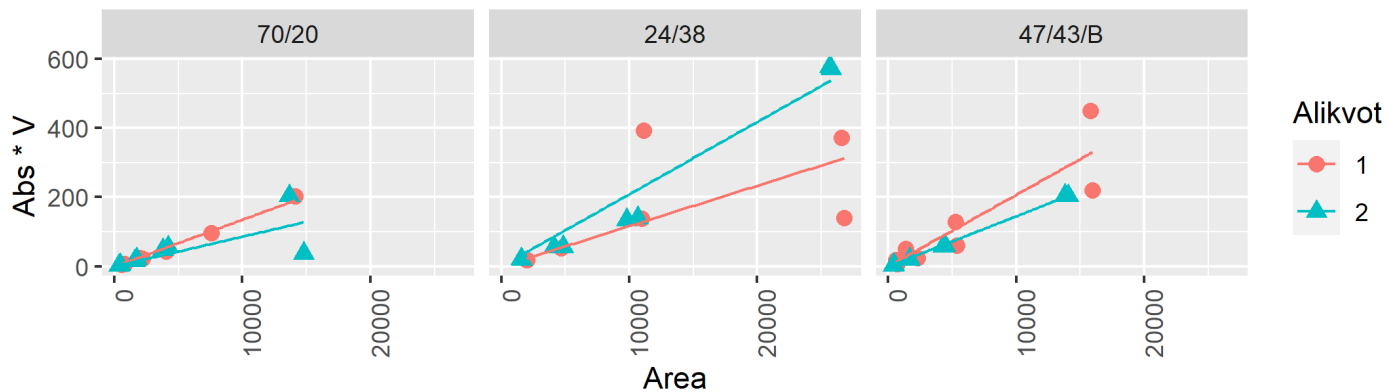
Pro každý oligonukleotid byl vytvořen graf závislosti absorbance vynásobené celkovým objemem a plochy (graf 8). Dále byla provedena lineární regrese spolu s oboustranným F-testem při zachování stejných výchozích podmínek a hodnot jako u souhrnného grafu 7. Trojice oligonukleotidů s nejnižším korelačním koeficientem byla z dalšího statistického zpracování vyřazena. Přehled výsledků statistických analýz je zobrazen v tabulce 12. Ve všech případech byla p-hodnota nižší než hladina významnosti, platí tedy hypotéza alternativní.



Graf 8: Závislost absorbance vynásobené objemem na ploše jednotlivých oligonukleotidů – část 1.



Graf 8: Závislost absorbance vynásobené objemem na ploše jednotlivých oligonukleotidů – část 2.



Graf 9: Vyřazené oligonukleotidy.

Tabulka 12: Přehled výsledků statistických nástrojů oligonukleotidů. $\alpha = 0,05$ * Vyřazené oligonukleotidy.

ID	Směrnice k	R^2	Adj. R^2	F-test	P-hodnota
39/18	0,0154	0,9929	0,9924	2103	<2,2E-16
50/18	0,0147	0,9949	0,9946	2925	<2,2E-16
72/18	0,0135	0,9967	0,9965	4553	<2,2E-16
35/20	0,0137	0,98	0,9786	733,4	3,76E-14
40/20	0,0157	0,9891	0,9883	1358	3,96E-16
70/20 *	0,0106	0,7992	0,7858	59,69	1,32E-06
30/23	0,0160	0,9776	0,9761	653,8	8,75E-14
52/23	0,0164	0,9923	0,9918	1944	<2,2E-16
65/23	0,0149	0,995	0,9947	2999	<2,2E-16
36/25	0,0163	0,9859	0,985	1052	2,63E-15
72/25	0,0150	0,9956	0,9953	3388	<2,2E-16
92/25	0,0128	0,9965	0,9963	4290	<2,2E-16
32/28	0,0179	0,9836	0,9826	901,9	8,18E-15
50/28	0,0159	0,9712	0,9693	505,9	5,72E-13
61/28	0,0150	0,9791	0,9777	701,4	5,22E-14
23/31	0,0198	0,9797	0,9784	725,5	4,07E-14
48/31	0,0160	0,9918	0,9912	1804	<2,2E-16
65/31	0,0150	0,9179	0,9125	167,8	1,51E-09
32/34	0,0194	0,9807	0,9794	761	2,87E-14
53/34	0,0165	0,9866	0,9857	1101	1,87E-15
77/35	0,0137	0,9913	0,9907	1701	<2,2E-16
24/38 *	0,0151	0,8174	0,8053	67,16	6,39E-07
46/39	0,0151	0,9957	0,9954	3475	<2,2E-16
67/39	0,0181	0,9877	0,9869	1206	9,53E-16
35/43	0,0186	0,9839	0,9828	917,8	7,19E-15
47/43/A	0,0164	0,8965	0,8896	129,9	8,69E-09
47/43/B *	0,0180	0,8883	0,8809	119,3	1,55E-08

Získané hodnoty směrnice k byly využity k tvorbě predikčního modelu. Za prediktory byly zvoleny tři ze základních vlastností oligonukleotidu, tedy jeho délka, procentuální obsah GC bází a molární extinkční koeficient.

Na základě získaných hodnot směrnice k lineárních regresí s ohledem na přítomné prediktory (obsah GC, délka a molární extinkční koeficient) byly vytvořeny čtyři predikční modely: A, B, C a D. Tyto modely byly porovnány z hlediska jejich p-hodnoty, adjustovaného R^2 a směrodatné odchylky reziduí (SDR).

Nejkomplexnějším predikčním modelem je model A, který využívá všech třech prediktorů a jejich vzájemných interakcí. Model lze zapsat rovnicí:

$$k_A = 3,580 \cdot 10^{-4} \cdot GC + 1,050 \cdot 10^{-4} \cdot ext + 7,221 \cdot 10^{-4} \cdot lng - 1,457 \cdot 10^{-6} \cdot GC \cdot ext - 1,176 \cdot 10^{-5} \cdot GC \cdot lng - 2,539 \cdot 10^{-6} \cdot ext \cdot lng + 3,996 \cdot 10^{-8} \cdot GC \cdot ext \cdot lng - 1,065 \cdot 10^{-2}$$

Rezidua:

Min	1Q	Medián	3Q	Max
$-2,472 \cdot 10^{-3}$	$-4,844 \cdot 10^{-4}$	$1,505 \cdot 10^{-5}$	$5,730 \cdot 10^{-4}$	$2,315 \cdot 10^{-3}$

$$p\text{-hodnota} = 9,759 \cdot 10^{-4}$$

$$SDR = 1,107 \cdot 10^{-3} \text{ při 16 stupních volnosti}$$

$$\text{Adj. } R^2 = 0,626$$

Model B vznikl odstraněním interakcí mezi prediktory:

$$k_B = -4,937 \cdot 10^{-5} \cdot GC + 1,640 \cdot 10^{-5} \cdot ext - 5,575 \cdot 10^{-5} \cdot lng - 1,500 \cdot 10^{-2}$$

Rezidua:

Min	1Q	Medián	3Q	Max
$-2,642 \cdot 10^{-3}$	$-3,329 \cdot 10^{-4}$	$3,204 \cdot 10^{-5}$	$4,078 \cdot 10^{-4}$	$2,177 \cdot 10^{-3}$

$$p\text{-hodnota} = 7,81 \cdot 10^{-5}$$

$$SDR = 1,143 \cdot 10^{-3} \text{ při 20 stupních volnosti}$$

$$\text{Adj. } R^2 = 0,601$$

Následné vyloučení nejméně specifického prediktoru, délky, dá vzniku modelu C:

$$k_C = -5,464 \cdot 10^{-5} \cdot GC + 1,103 \cdot 10^{-5} \cdot ext + 1,500 \cdot 10^{-2}$$

Rezidua:

Min	1Q	Medián	3Q	Max
$-2,536 \cdot 10^{-3}$	$-2,991 \cdot 10^{-4}$	$3,265 \cdot 10^{-5}$	$4,928 \cdot 10^{-4}$	$2,140 \cdot 10^{-3}$

$$p\text{-hodnota} = 1,614 \cdot 10^{-5}$$

$$SDR = 1,119 \cdot 10^{-3} \text{ při 21 stupních volnosti}$$

$$\text{Adj. } R^2 = 0,617$$

Posledním vytvořeným modelem, je model D, který operuje pouze s délkou a obsahem GC bázi:

$$k_D = -6,558 \cdot 10^{-5} \cdot GC + 1,074 \cdot 10^{-4} \cdot lng + 1,619 \cdot 10^{-2}$$

Rezidua:

Min	1Q	Medián	3Q	Max
$-2,323 \cdot 10^{-3}$	$-2,547 \cdot 10^{-4}$	$-2,366 \cdot 10^{-5}$	$5,734 \cdot 10^{-4}$	$2,141 \cdot 10^{-3}$

$$p\text{-hodnota} = 2,786 \cdot 10^{-5}$$

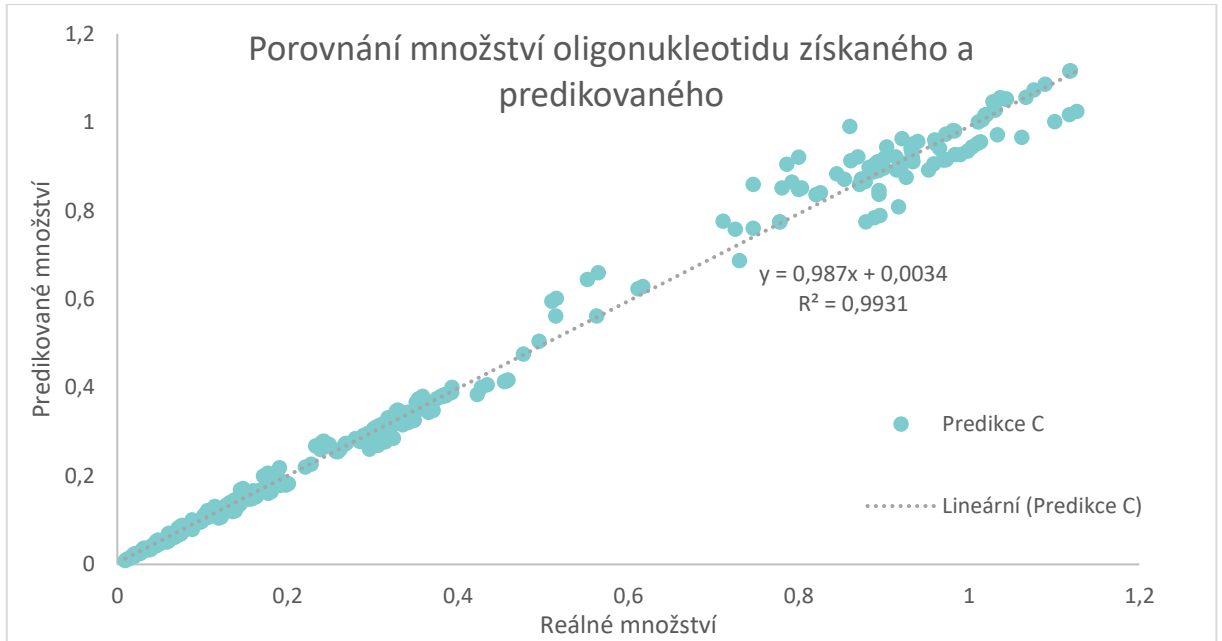
$$SDR = 1,149 \cdot 10^{-3} \text{ při 21 stupních volnosti}$$

$$\text{Adj. } R^2 = 0,597$$

Při porovnání p-hodnot lze označit za nejlépe vyhovující model C, z pohledu R^2 sedí lineární křivka nejpřesněji u modelu A. Modely B a D vykazují horší hodnoty než předešlé zmíněné. Model A, přestože je oproti modelu C mnohem komplexnější, nevykazuje velké rozdíly. Na základě těchto údajů byl pro porovnání s reálnými daty vybrán predikční model C.

4.3.3 Porovnání modelu predikce s reálnými výsledky

Na základě predikčního modelu C bylo vypočítáno predikované množství oligonukleotidu a následně porovnáno s reálně získaným množstvím (graf 10). Na základě dvouvýběrového párového t-testu na střední hodnotu byla potvrzena nulová hypotéza, tedy že střední hodnoty datových souborů jsou shodné.



Graf 10: Porovnání množství oligonukleotidu získaného purifikací a predikovaného dle predikčního modelu C.

4.4 Výsledky analýzy sond

4.4.1 Hromadná analýza dat

Obdobně jako v případě oligonukleotidů byla provedena grafická a statistická analýza měřených hodnot duálně značených sond. Celkové grafické zobrazení absorbance vynásobené objemem vzorku a plochy píku všech duálně značených sond je v grafu 11.

Lineární regrese (procházející počátkem):

Cílem testu bylo získat rovnici lineární regrese a její koeficient determinace.

$$Abs \cdot 400 = 0,0135 \cdot Area$$

koeficient determinace $R^2 = 0,9065$

upravené $R^2 = 0,9059$

Pro výpočet byl využit jazyk R, základní balíček, funkce `lm()`.

F-test oboustranný:

Cílem testu bylo zamítnutí nulové hypotézy a potvrzení hypotézy alternativní.

H_0 = regresní koeficient je roven nule

H_A = regresní koeficient není roven nule

$\alpha = 0,05$

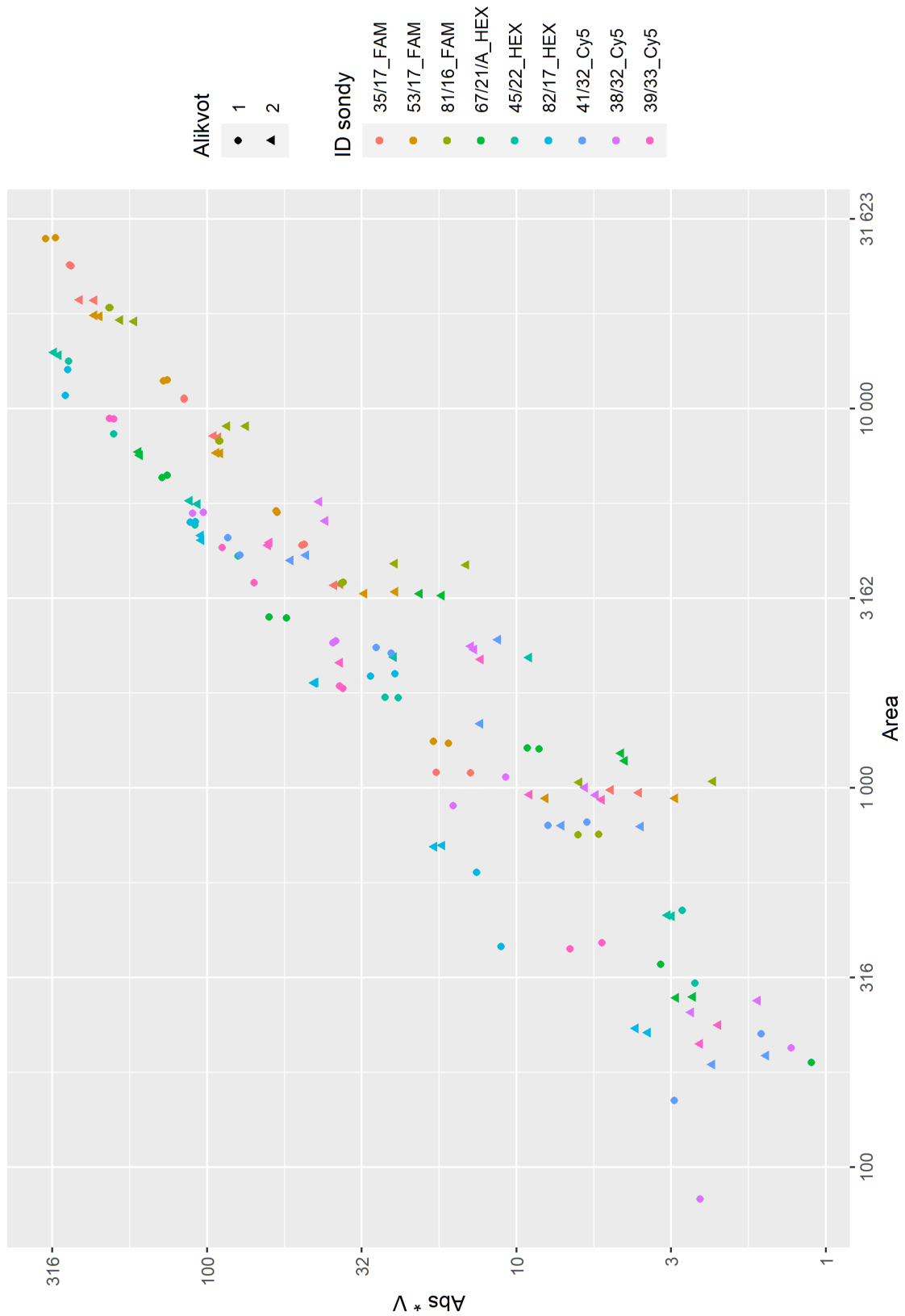
hladina jistoty = 95 %

F-test: 1387 a 143 stupňů volnosti, p-hodnota $< 2,2 \cdot 10^{-16}$

Nulová hypotéza se zamítá, platí alternativní hypotéza.

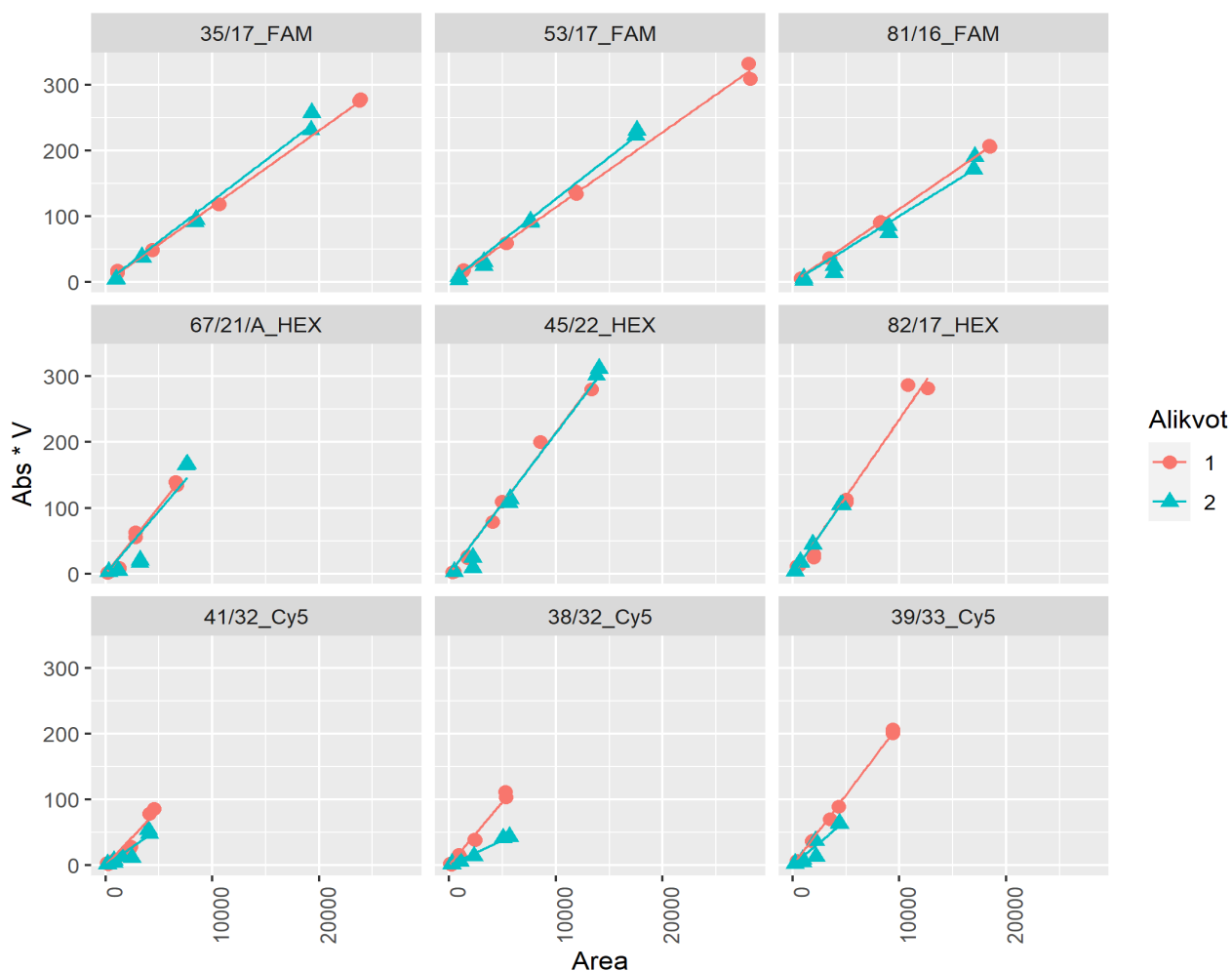
Pro výpočet byl využit jazyk R, základní balíček, funkce `lm()`.

Z výsledků analýz lze usuzovat, že závislost je opravdu lineární, hodnoty korelují a přímku lineární regrese lze tedy použít k hrubému odhadu výsledné absorbance, resp. množství.



Graf 11: Závislost absorpance vynásobené objemem na ploše všech sond. Jednotlivé vzorky jsou odlišeny barvou, jednotlivé alikvoty pak tvarem.

Pro každou sondu byl sestaven samostatný graf (graf 12), vypočítán korelační koeficient a provedena lineární regrese procházející počátkem. Přehled výsledných hodnot je uveden v tabulce 13. Ve všech případech byla p-hodnota nižší než hladina významnosti, platí tedy hypotéza alternativní.



Graf 12: Závislost absorpance vynásobené objemem na ploše jednotlivých sond.

Tabulka 13: Přehled výsledků statistických nástrojů sond.

ID	Směrnice	R^2	Adj. R^2	F-test	P-hodnota
35/17_FAM	0,0119	0,9962	0,996	3972	<2,2E-16
53/17_FAM	0,0118	0,9952	0,9949	3095	<2,2E-16
81/16_FAM	0,0106	0,9891	0,9884	1363	3,84E-16
67/21/A_HEX	0,0196	0,9406	0,9367	237,6	1,32E-10
45/22_HEX	0,0214	0,9911	0,9905	1664	<2,2E-16
82/17_HEX	0,0234	0,9897	0,989	1442	2,53E-16
41/32_Cy5	0,0143	0,9155	0,9099	162,6	1,88E-09
38/32_Cy5	0,0134	0,8299	0,8186	73,19	3,73E-07
39/33_Cy5	0,0200	0,9726	0,9707	531,7	3,98E-13

Na závěr byly hodnoty směrnice k využity k tvorbě predikčního modelu obdobně jako v případě oligonukleotidů, v úvahu byla vzata délka sondy, obsah GC a molární extinkční koeficient.

Predikční model A_s počítá se všemi prediktory a jejich vzájemnými interakcemi a lze vyjádřit rovnicí:

$$k_{A_s} = -5,655 \cdot 10^{-3} \cdot GC - 6,731 \cdot 10^{-4} \cdot ext - 2,730 \cdot 10^{-2} \cdot lng + 1,477 \cdot 10^{-5} \cdot GC \cdot ext + 3,849 \cdot 10^{-4} \cdot GC \cdot lng + 6,610 \cdot 10^{-5} \cdot ext \cdot lng - 1,018 \cdot 10^{-6} \cdot GC \cdot ext \cdot lng + 3,635 \cdot 10^{-1}$$

Rezidua:

1	2	3	4	5
$1,741 \cdot 10^{-3}$	$-2,186 \cdot 10^{-3}$	$-1,887 \cdot 10^{-3}$	$-1,221 \cdot 10^{-3}$	$-1,989 \cdot 10^{-4}$
6	7	8	9	
$3,441 \cdot 10^{-3}$	$1,732 \cdot 10^{-3}$	$-7,637 \cdot 10^{-4}$	$6,591 \cdot 10^{-4}$	

p-hodnota = 0,7045

SDR = $5,363 \cdot 10^{-3}$ při 1 stupni volnosti

Adj. $R^2 = -0,235$

Predikční model B_s je oprostěn od vzájemných interakcí prediktorů:

$$k_{B_s} = 3,178 \cdot 10^{-4} \cdot GC + 2,389 \cdot 10^{-4} \cdot ext - 1,801 \cdot 10^{-3} \cdot lng - 2,365 \cdot 10^{-2}$$

Rezidua:

1	2	3	4	5
$2,420 \cdot 10^{-3}$	$-1,794 \cdot 10^{-3}$	$-3,774 \cdot 10^{-3}$	$1,563 \cdot 10^{-4}$	$-1,707 \cdot 10^{-3}$
6	7	8	9	
$4,639 \cdot 10^{-3}$	$1,161 \cdot 10^{-3}$	$-2,022 \cdot 10^{-3}$	$3,243 \cdot 10^{-3}$	

p-hodnota = 0,121

SDR = $3,570 \cdot 10^{-3}$ při 5 stupních volnosti

Adj. $R^2 = 0,453$

Model Cs bere v potaz pouze obsah GC a molární absorpční koeficient:

$$k_{Cs} = 1,892 \cdot 10^{-4} \cdot GC + 4,639 \cdot 10^{-5} \cdot ext - 6,345 \cdot 10^{-3}$$

Rezidua:

Min	1Q	Medián	3Q	Max
$-6,333 \cdot 10^{-3}$	$-3,355 \cdot 10^{-3}$	$1,368 \cdot 10^{-3}$	$1,956 \cdot 10^{-3}$	$5,246 \cdot 10^{-3}$

p-hodnota = 0,3582

SDR = $4,696 \cdot 10^{-3}$ při 6 stupních volnosti

Adj. $R^2 = 0,053$

Predikce Ds počítá s délkou a obsahem GC:

$$k_{Ds} = 1,154 \cdot 10^{-4} \cdot GC - 2,555 \cdot 10^{-4} \cdot lng + 4,222 \cdot 10^{-3}$$

Rezidua:

Min	1Q	Medián	3Q	Max
$-7,019 \cdot 10^{-3}$	$-2,918 \cdot 10^{-3}$	$-7,467 \cdot 10^{-4}$	$2,812 \cdot 10^{-3}$	$6,381 \cdot 10^{-3}$

p-hodnota = 0,6443

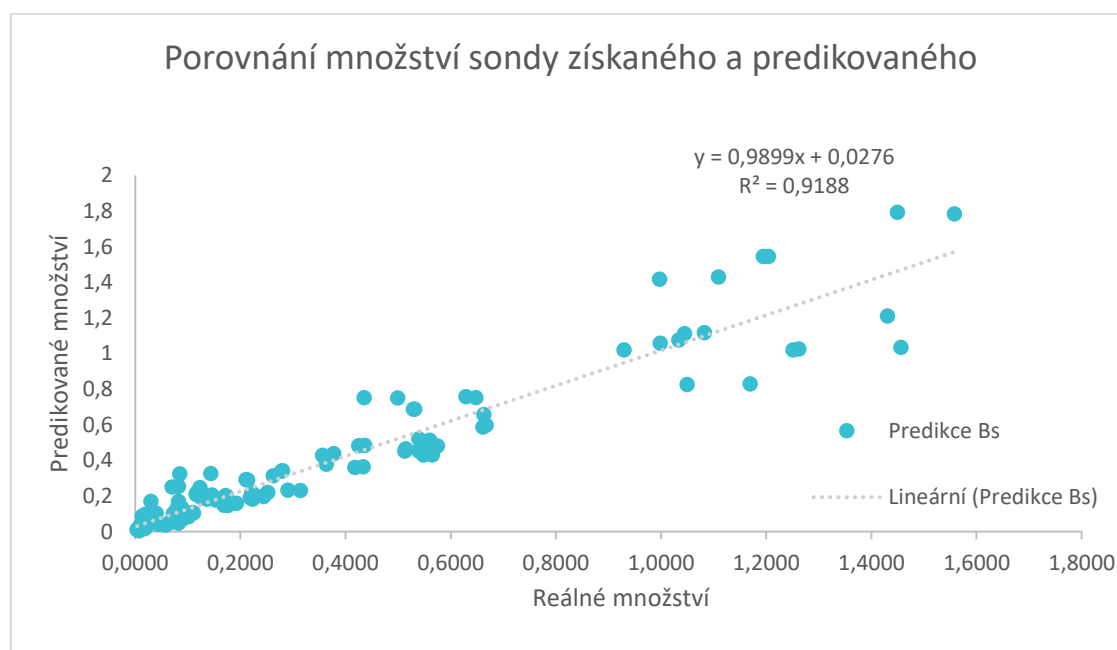
SDR = $5,178 \cdot 10^{-3}$ při 6 stupních volnosti

Adj. $R^2 = -0,152$

Hodnocení sond analogicky k neznačeným oligonukleotidům označuje za preferovaný predikční model B, který má nejnižší p-hodnotu a zároveň i nejvyšší adjustované R^2 , avšak je patrné, že ani jeden z predikčních modelů sond nedosahuje kvalit modelů pro oligonukleotidy. Je zde velká interference použité značky a ani jeden z modelů nesplňuje předpoklady pro využití.

4.4.2 Porovnání modelu predikce s reálnými výsledky

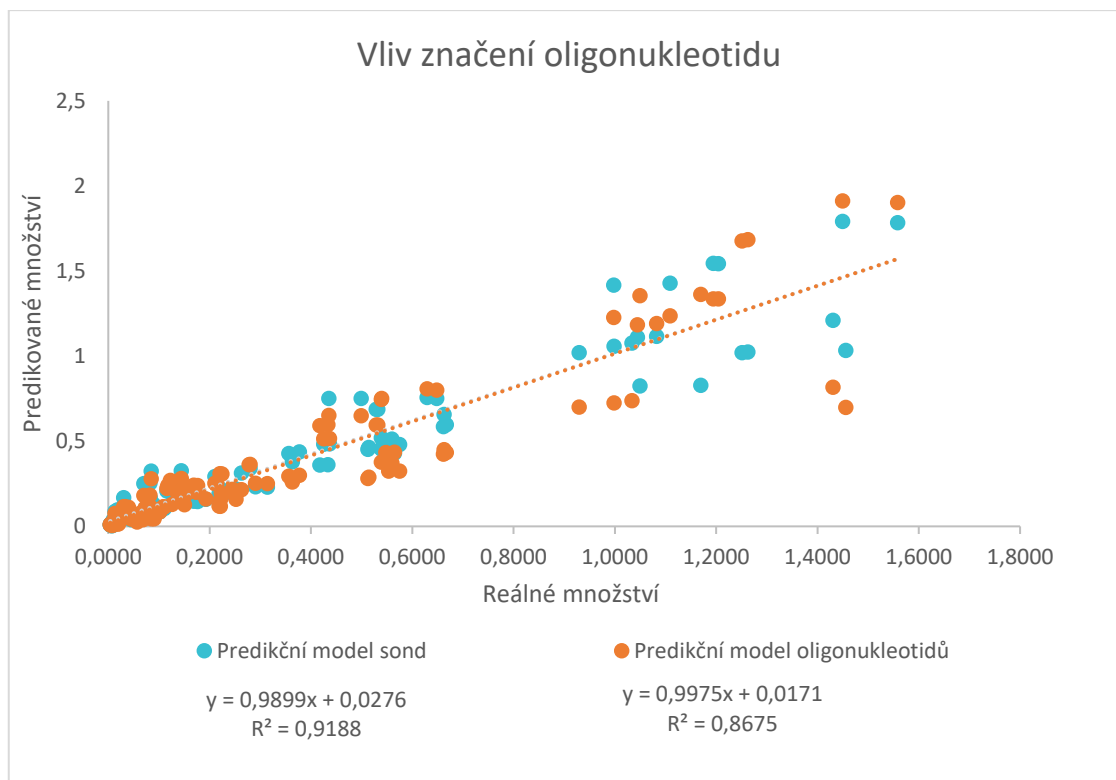
Podle vybraného modelu Bs bylo vypočítáno predikované množství sond, které bylo následně porovnáno s reálně odebraným množstvím (graf 13).



Graf 13: Porovnání množství sondy získaného purifikací a predikovaného dle predikčního modelu Bs.

4.4.3 Hodnocení vlivu značení oligonukleotidu

Pro zhodnocení vlivu značky na sondě byly zobrazeny hodnoty predikce množství dané sondy vybraným modelem pro obecné oligonukleotidy (model C) a pro samotné sondy (model Bs). Graf 14 potvrzuje, že pro predikci sond není vhodné používat predikční model pro oligonukleotidy, protože dosahuje nižšího koeficientu determinace. Nižší výkonnost modelu je dána zejména tím, že vznikl na jiných datech než na posuzovaných sondách, z tohoto důvodu nelze ani vyvodit vliv samotného značení.



Graf 14: Zobrazení vlivu značení sond u vybraných predikčních modelů.

5 DISKUZE A ZÁVĚR

Cílem této diplomové práce bylo analyzovat možnou závislost mezi plochou píku a koncentrací stanovovaných oligonukleotidů a duálně značených sond s ohledem na jejich délku, procentuální zastoupení GC bází a modifikaci fluorescenční značkou. V případě výsledků svědčících pro korelaci poté stanovit korelační koeficient, na základě kterého by bylo možno predikovat výsledky z naměřené plochy. Tohoto mělo být dosaženo analyzováním 27 vybraných oligonukleotidů a 9 duálně značených sond. Tyto vytyčené cíle byly splněny tak, že ze získaných predikčních modelů byl pro každou skupinu vybrán jeden nejlépe vyhovující určeným parametrům.

Prvním krokem diplomové práce bylo stanovení koncentrace zásobních roztoků. Před samotnou analýzou vzorků byla provedena optimalizace replikátů nástříků, která měla za cíl výzkumu k diplomové práci testovat přesnost chromatografu napříč nástříky, resp. ověřit vhodný počet nástříků pro další analýzu. Shoda mezi naměřenými hodnotami byla velmi vysoká, z čehož bylo vyvozeno optimální množství dvou nástříků na roztok. Následně byly připraveny pro každý vzorek 4 kalibrační body v rozmezí od cca 5,85 μM do cca 0,1mM, které byly měřeny ve dvou alikvotech a dvou nástříkách pro každý alikvot. Tyto roztoky byly purifikovány na HPLC, purifikát byl vysušen, rozpuštěn ve 400 μl IWA vody a měřen v dubletu na absorbanci při 260 nm. Získaná data byla vložena do tabulky a zpracována v prostředí R, což je vhodný nástroj pro statistickou analýzu. Výsledkem je prokázána korelace mezi plochou píku a absorbcí oligonukleotidů, kde platí lineární závislost. Výstupem práce bylo objasnění vlivu obsahu GC, délky a molárního absorpčního koeficientu oligonukleotidu na směrnici lineární regrese plochy píku a absorbcí. Zjištěným závislostem odpovídají nejlépe rovnice $k_C = -5,464 \cdot 10^{-5} \cdot GC + 1,103 \cdot 10^{-5} \cdot ext + 1,500 \cdot 10^{-2}$ pro oligonukleotidy a $k_{BS} = 3,178 \cdot 10^{-4} \cdot GC + 2,389 \cdot 10^{-4} \cdot ext - 1,801 \cdot 10^{-3} \cdot lng - 2,365 \cdot 10^{-2}$ pro sondy; k lze označit jako koeficient predikce. Obě tyto rovnice poměrně dobře zpětně predikují získaná data.

Přestože byl výběr vzorků oligonukleotidů dostatečně široký pro potřeby práce, byl limitován jak délkou, tak poměrem GC. Nejsou tedy přítomna data pro nukleotidy kratší než 18 bází, delší než 43 bází a pro obsah GC pod 23 % a nad 92 %. Dále by bylo vhodné pokrýt širší pásmo obsahu GC bází pro jednu délku oligonukleotidu, tedy ideálně 0-100 % GC pro každou analyzovanou délku, ačkoliv zvolené vzorky spadají parametry mezi běžně využívané primery.

Výběr vzorků sond byl oproti oligonukleotidům třetinový, přičemž ve výběru měly zastoupení tři druhy značení. Počet sond bylo limitováno jak vysokou časovou a materiální, tak i finanční náročností. Toto množství však nebylo dostatečné. Rozšíření škály vzorků by jistě vedlo ke komplexnějším datům.

Pro optimalizaci replikátů nástřiků byl využit vzorek oligonukleotidu, který parametry zapadal mezi výběr vzorků, avšak tento proces nebyl proveden i s využitím sondy. Přesto se nástřiky navzájem mezi sebou shodovaly ve stejné míře jako u oligonukleotidů.

Vliv na HPLC analýzu mohlo mít i stáří vzorků či opakované rozmrazování a manipulace s nimi. V některých případech byly přítomny další píky, pravděpodobně zčásti rozpadlé oligonukleotidy. Zároveň byla velmi nízká návratnost vzorku po purifikaci oproti předpokladu. V případě nižších koncentrací roztoků nebyl vždy proveden správný odběr frakce a bylo tedy zapotřebí plochu píku manuálně integrovat tak, aby odpovídala odebrané frakci. Vhodnější by bylo využít vyšší koncentrace roztoků a tomuto problému předejít.

Statistické zpracování dat bylo rozděleno na hodnocení oligonukleotidů a sond, dále pak do podkapitol s hromadným stanovením získaných dat a s jednotlivými vzorky. Ve všech případech byla prokázána lineární závislost mezi plochou píku a množstvím oligonukleotidu. U této lineární závislosti byl stanoven Pearsonův korelační koeficient, přičemž tři oligonukleotidy s jeho nejnižší hodnotou byly z další analýzy vyřazeny. Následně byla směrnice lineární regrese využita k tvorbě predikčních modelů vztahujících se k obsahu GC, délce a molárnímu absorpčnímu koeficientu. Ze získaných predikčních modelů byl vybrán jeden parametrově nejvhodnější.

Vzhledem k rozdílným rovnicím pro výpočet koeficientu predikce lze usuzovat interferenci použitého značení, avšak s ohledem na nízké množství dat nelze stanovit míru ovlivnění značkou v obecné rovině, natož s ohledem na konkrétní fluorescenční značení. Opět by bylo vhodné provést analýzu pro každou použitou značku v množství alespoň obdobném, jaké bylo vybráno pro oligonukleotidy, ideálně rozšířeném o širší pokrytí škály především GC bázi, eventuálně i délky. Pro konkrétní zhodnocení vlivu samotné značky by byla třeba analýza sond o stejné sekvenci jako u samotných oligonukleotidů. Právě rozdíl mezi výsledky těchto dat by mohl dát požadovanou informaci.

Predikce jakožto hlavní cíl regresní analýzy zaostává v několika ohledech. Především nelze extrapolovat mimo rozsah použitých dat, jsou platné, s ohledem na jejich nepřesnost, jen v takovém rozmezí dat, ze kterých se původně vycházelo. Nelze je tedy uplatnit např. pro

oligonukleotidy s nižším či vyšším obsahem GC nebo v jiných délkách, než které byly v práci analyzovány. To značně snižuje celkový význam predikcí. Dále bylo z důvodu rozsahu práce zcela zanedbáno případné skládání do sekundárních struktur, které by mohly mít vliv i u takhle krátkých řetězců na jejich HPLC analýzu, čemuž by se dalo předejít zvýšením teploty na koloně na 60 °C. Vycházelo se však z kladné zkušenosti pracoviště. Pravděpodobně však největším nedostatkem, který by bylo vhodné řešit v rámci návaznosti práce, je validace predikčních modelů na nezávislých datech, tedy na vzorcích jiných, než které byly užity pro tvorbu modelu. Ověření nebylo provedeno s ohledem na jeho časovou náročnost a bylo by vhodné tuto oblast pokrýt v rámci návaznosti práce. [94]

Tvorbou predikčních modelů v oblasti HPLC analýzy oligonukleotidů se zabýval Gilar M. a kol. v roce 2002, jehož cílem bylo vytvoření matematického modelu pro predikci retenčního času syntetického oligonukleotidu. Za základní prediktory zvolil délku a sekvenci a vycházel z předpokladu, že různé typy nukleotidů přispívají při retenci různým dílem. Na jeho práci navázal o pět let později Sturm M. a kol. s modelem počítajícím s regresním parametrem, který vychází ze sekvence a predikované sekundární struktury, čímž se mu podařilo dosáhnout lepší předpovědi pro nižší teploty, a tím demonstrovat vliv sekundární struktury. [81, 95]

S ohledem na datové zpracování, jednoduchou explorativní datovou analýzu a následnou tvorbu predikce z lineárního modelu, jednalo se spíše o klasický, možná až konzervativní, ač stále funkční způsob. V ideálním případě pro vhodnější predikci by bylo lépe využít umělé inteligence (AI z ang. *artificial intelligence*) a strojového učení (ML z ang. *machine learning*). Předpokladem pro aplikaci těchto nástrojů je použití velkého množství dat v jejich surové podobě. ML, případně subkategorie hluboké učení (DL z ang. *deep learning*), je v oblasti bioinformatiky značně rozšířeným fenoménem, který je schopný si s vysokou pravděpodobností poradit i s komplexnějšími úlohami, jakými je např. protein folding. Prakticky pro možnost ověření by se část nasbíraných dat využila pro učení programu, zbytek pak pro ověření platnosti predikce. Součástí jazyka R je i několik balíčků vhodných pro tyto účely.

Celkem bylo v této práci analyzováno 27 oligonukleotidů a 9 duálně značených sond na možnou korelaci mezi plochou a koncentrací a vlivem sekvence a modifikace. Byla prokázána lineární závislost plochy a absorbance čistého produktu, na základě které se poté zkoumal vliv sekvence a modifikace na směrnici této lineární závislosti. Z dat vyplývá

korelace mezi plochou pod píkem a koncentrací oligonukleotidu, která byla následně využita k tvorbě predikčních modelů, jež byly porovnány s reálně naměřenými daty.

POUŽITÁ LITERATURA

- [1] DEVLIN, Thomas M., ed. *Textbook of biochemistry: with clinical correlations*. 7th ed. Hoboken, NJ: John Wiley & Sons, 2011. ISBN 978-0-470-28173-4.
- [2] COBB, Matthew. 60 years ago, Francis Crick changed the logic of biology. *PLOS Biology* [online]. 2017, **15**(9), e2003243. ISSN 1545-7885. Dostupné z: doi:10.1371/journal.pbio.2003243
- [3] Nucleic Acid Bases. In: George P. RÉDEI *Encyclopedia of Genetics, Genomics, Proteomics and Informatics* [online]. Dordrecht: Springer Netherlands, 2008 [vid. 2020-10-30], s. 1371–1371. ISBN 978-1-4020-6753-2. Dostupné z: doi:10.1007/978-1-4020-6754-9_11634
- [4] BERG, Jeremy M., John L. TYMOCZKO, Lubert STRYER a Lubert STRYER. *Biochemistry*. 5th ed. New York: W.H. Freeman, 2002. ISBN 978-0-7167-3051-4.
- [5] ZOCCHI, Giovanni. *Molecular machines: a materials science approach*. 2018. ISBN 978-1-4008-9006-4.
- [6] CREIGHTON, Thomas E. *The biophysical chemistry of nucleic acids*. Eastbourne, East Sussex: Helvetian Press, 2011. ISBN 978-0-9564781-1-5.
- [7] LODISH, Harvey F., ed. *Molecular cell biology*. 4th ed. New York: W.H. Freeman, 2000. ISBN 978-0-7167-3136-8.
- [8] SURAM, Anitha, Jagannatha K. S. RAO, Latha K. S. a Viswamitra M. A. First Evidence to Show the Topological Change of DNA from B-DNA to Z-DNA Conformation in the Hippocampus of Alzheimer's Brain. *NeuroMolecular Medicine* [online]. 2002, **2**(3), 289–298. ISSN 1535-1084. Dostupné z: doi:10.1385/NMM:2:3:289
- [9] LAFER, E. M., R. P. VALLE, A. MÖLLER, A. NORDHEIM, P. H. SCHUR, A. RICH a B. D. STOLLAR. Z-DNA-specific antibodies in human systemic lupus erythematosus. *The Journal of Clinical Investigation* [online]. 1983, **71**(2), 314–321. ISSN 0021-9738. Dostupné z: doi:10.1172/jci110771
- [10] HERBERT, Alan. Mendelian disease caused by variants affecting recognition of Z-DNA and Z-RNA by the Z α domain of the double-stranded RNA editing enzyme ADAR. *European journal of human genetics: EJHG* [online]. 2020, **28**(1), 114–117. ISSN 1476-5438. Dostupné z: doi:10.1038/s41431-019-0458-6
- [11] GARDNER, Paul P. a Hisham ELDAI. Annotating RNA motifs in sequences and alignments. *Nucleic Acids Research* [online]. 2015, **43**(2), 691–698. ISSN 1362-4962, 0305-1048. Dostupné z: doi:10.1093/nar/gku1327
- [12] PARLEA, Lorena G., Blake A. SWEENEY, Maryam HOSSEINI-ASANJAN, Craig L. ZIRBEL a Neocles B. LEONTIS. The RNA 3D Motif Atlas: Computational methods for extraction, organization and evaluation of RNA motifs. *Methods* [online]. 2016, **103**, 99–119. ISSN 10462023. Dostupné z: doi:10.1016/j.ymeth.2016.04.025

- [13] CHOJNOWSKI, Grzegorz, Tomasz WALEŃ a Janusz M. BUJNICKI. RNA Bricks—a database of RNA 3D motifs and their interactions. *Nucleic Acids Research* [online]. 2014, **42**(D1), D123–D131. ISSN 0305-1048, 1362-4962. Dostupné z: doi:10.1093/nar/gkt1084
- [14] PAN, Xiaoyong, Peter RIJNBEEK, Junchi YAN a Hong-Bin SHEN. Prediction of RNA-protein sequence and structure binding preferences using deep convolutional and recurrent neural networks. *BMC Genomics* [online]. 2018, **19**(1), 511. ISSN 1471-2164. Dostupné z: doi:10.1186/s12864-018-4889-1
- [15] ROBERTS, Thomas C., Robert LANGER a Matthew J. A. WOOD. Advances in oligonucleotide drug delivery. *Nature Reviews Drug Discovery* [online]. 2020, **19**(10), 673–694. ISSN 1474-1776, 1474-1784. Dostupné z: doi:10.1038/s41573-020-0075-7
- [16] ROY, Subhadeep a Marvin CARUTHERS. Synthesis of DNA/RNA and Their Analogs via Phosphoramidite and H-Phosphonate Chemistries. *Molecules* [online]. 2013, **18**(11), 14268–14284. ISSN 1420-3049. Dostupné z: doi:10.3390/molecules181114268
- [17] REESE, C. The H-phosphonate approach to the solution phase synthesis of linear and cyclic oligoribonucleotides. *Nucleic Acids Research* [online]. 1999, **27**(4), 963–971. ISSN 13624962. Dostupné z: doi:10.1093/nar/27.4.963
- [18] STAWINSKI, Jacek a Roger STRÖMBERG. Di- and Oligonucleotide Synthesis Using H-Phosphonate Chemistry. In: Piet HERDEWIJN *Oligonucleotide Synthesis* [online]. New Jersey: Humana Press, 2004 [vid. 2020-11-21], s. 081–100. ISBN 978-1-59259-823-6. Dostupné z: doi:10.1385/1-59259-823-4:081
- [19] GILHAM, P. T. a H. G. KHORANA. Studies on Polynucleotides. I. A New and General Method for the Chemical Synthesis of the C₅'-C₃' Internucleotidic Linkage. Syntheses of Deoxyribo-dinucleotides¹. *Journal of the American Chemical Society* [online]. 1958, **80**(23), 6212–6222. ISSN 0002-7863, 1520-5126. Dostupné z: doi:10.1021/ja01556a016
- [20] KHORANA, H. Total synthesis of a gene. *Science* [online]. 1979, **203**(4381), 614–625. ISSN 0036-8075, 1095-9203. Dostupné z: doi:10.1126/science.366749
- [21] REESE, Colin B. The chemical synthesis of oligo- and poly-nucleotides by the phosphotriester approach. *Tetrahedron* [online]. 1978, **34**(21), 3143–3179. ISSN 00404020. Dostupné z: doi:10.1016/0040-4020(78)87013-6
- [22] REESE, Colin B. Oligo- and poly-nucleotides: 50 years of chemical synthesis. *Organic & Biomolecular Chemistry* [online]. 2005, **3**(21), 3851. ISSN 1477-0520, 1477-0539. Dostupné z: doi:10.1039/b510458k
- [23] LETSINGER, R.L., E.P. GROODY, N. LANDER a T. TANAKA. Some developments in the phosphotriester method for synthesis of oligonucleotides. *Tetrahedron* [online]. 1984, **40**(1), 137–143. ISSN 00404020. Dostupné z: doi:10.1016/0040-4020(84)85112-1
- [24] BEAUCAGE, S.L. a M.H. CARUTHERS. Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Letters* [online]. 1981, **22**(20), 1859–1862. ISSN 00404039. Dostupné z: doi:10.1016/S0040-4039(01)90461-7

- [25] CAMMACK, Richard. *Oxford dictionary of biochemistry and molecular biology*. New York: Oxford University Press, 2006. ISBN 978-1-61344-113-8.
- [26] SANDAHL, Alexander F., Thuy J. D. NGUYEN, Rikke A. HANSEN, Martin B. JOHANSEN, Troels SKRYDSTRUP a Kurt V. GOTHELF. On-demand synthesis of phosphoramidites. *Nature Communications* [online]. 2021, **12**(1), 2760. ISSN 2041-1723. Dostupné z: doi:10.1038/s41467-021-22945-z
- [27] DEVIVAR, Rodrigo V., Steve L. KOONTZ, William J. PELTIER, James E. PEARSON, Todd A. GUILLORY a Jill D. FABRICANT. A new solid-phase support for oligonucleotide synthesis. *Bioorganic & Medicinal Chemistry Letters* [online]. 1999, **9**(9), 1239–1242. ISSN 0960894X. Dostupné z: doi:10.1016/S0960-894X(99)00177-8
- [28] GUZAEV, Andrei P. Solid-Phase Supports for Oligonucleotide Synthesis. *Current Protocols in Nucleic Acid Chemistry* [online]. 2013, **53**(1) [vid. 2020-12-15]. ISSN 1934-9270, 1934-9289. Dostupné z: doi:10.1002/0471142700.nc0301s53
- [29] *Universal Oligo Supports | Glen Research* [online]. [vid. 2021-10-22]. Dostupné z: <https://www.glenresearch.com/products/universal-supports-and-synthesis-supplies/universal-supports.html>
- [30] *ATDBio - Solid-phase oligonucleotide synthesis* [online]. [vid. 2021-10-22]. Dostupné z: <https://atdbio.com/nucleic-acids-book/Solid-phase-oligonucleotide-synthesis#Oligonucleotide-synthesis-with-universal-supports>
- [31] RUSSELL, Mark A., Andrew P. LAWS, John H. ATHERTON a Michael I. PAGE. The mechanism of the phosphoramidite synthesis of polynucleotides. *Organic & Biomolecular Chemistry* [online]. 2008, **6**(18), 3270. ISSN 1477-0520, 1477-0539. Dostupné z: doi:10.1039/b808999j
- [32] WEI, Xia. Coupling activators for the oligonucleotide synthesis via phosphoramidite approach. *Tetrahedron* [online]. 2013, **69**(18), 3615–3637. ISSN 00404020. Dostupné z: doi:10.1016/j.tet.2013.03.001
- [33] CARUTHERS, Marvin H. Chemical synthesis of DNA. *Journal of Chemical Education* [online]. 1989, **66**(7), 577. ISSN 0021-9584, 1938-1328. Dostupné z: doi:10.1021/ed066p577
- [34] EISENSTEIN, Michael. Enzymatic DNA synthesis enters new phase. *Nature Biotechnology* [online]. 2020, **38**(10), 1113–1115. ISSN 1087-0156, 1546-1696. Dostupné z: doi:10.1038/s41587-020-0695-9
- [35] PERKEL, Jeffrey M. The race for enzymatic DNA synthesis heats up. *Nature* [online]. 2019, **566**(7745), 565–565. ISSN 0028-0836, 1476-4687. Dostupné z: doi:10.1038/d41586-019-00682-0
- [36] HOFF, Kendall, Michelle HALPAIN, Giancarlo GARBAGNATI, Jeremy S. EDWARDS a Wei ZHOU. Enzymatic Synthesis of Designer DNA Using Cyclic Reversible Termination and a Universal Template. *ACS Synthetic Biology* [online]. 2020, **9**(2), 283–293. ISSN 2161-5063, 2161-5063. Dostupné z: doi:10.1021/acssynbio.9b00315

- [37] HERMANSON, Greg T. Nucleic Acid and Oligonucleotide Modification and Conjugation. In: *Bioconjugate Techniques* [online]. B.m.: Elsevier, 2013 [vid. 2021-01-26], s. 959–987. ISBN 978-0-12-382239-0. Dostupné z: doi:10.1016/B978-0-12-382239-0.00023-6
- [38] POOLE, Colin, Salvatore FANALI, P. R HADDAD, C. F POOLE a Marja RIEKKOLA. *Liquid Chromatography (Second Edition)*. Amsterdam: Elsevier., 2017. ISBN 978-0-12-805392-8.
- [39] OLSON, Nathan D a Jayne B MORROW. DNA extract characterization process for microbial detection methods development and validation. *BMC Research Notes* [online]. 2012, **5**(1), 668. ISSN 1756-0500. Dostupné z: doi:10.1186/1756-0500-5-668
- [40] KEER, Jacquie T. a Lyndsey BIRCH, ed. *Essentials of nucleic acid analysis: a robust approach*. Cambridge, UK: Royal Society of Chemistry, 2008. ISBN 978-0-85404-367-5.
- [41] GERSTEIN, Alan S. Nucleotides, Oligonucleotides, and Polynucleotides. In: Alan S. GERSTEIN, ed. *Molecular Biology Problem Solver* [online]. New York, USA: John Wiley & Sons, Inc., 2001 [vid. 2020-10-28], s. 267–289. ISBN 978-0-471-37972-0. Dostupné z: doi:10.1002/0471223905.ch10
- [42] NWOKEOJI, Alison O., Peter M. KILBY, David E. PORTWOOD a Mark J. DICKMAN. Accurate Quantification of Nucleic Acids Using Hypochromicity Measurements in Conjunction with UV Spectrophotometry. *Analytical Chemistry* [online]. 2017, **89**(24), 13567–13574. ISSN 0003-2700, 1520-6882. Dostupné z: doi:10.1021/acs.analchem.7b04000
- [43] *ATDBio - Ultraviolet absorbance of oligonucleotides* [online]. [vid. 2021-11-03]. Dostupné z: <https://atdbio.com/nucleic-acids-book/Ultraviolet-absorbance-of-oligonucleotides>
- [44] *Oligonucleotide Quantification* [online]. [vid. 2021-07-14]. Dostupné z: <https://www.sigmaaldrich.com/CZ/en/technical-documents/technical-article/genomics/pcr/quantitation-of-oligos>
- [45] KIBBE, W. A. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Research* [online]. 2007, **35**(Web Server), W43–W46. ISSN 0305-1048, 1362-4962. Dostupné z: doi:10.1093/nar/gkm234
- [46] BANIHASHEMIAN, Seyedeh, Vengadesh PERIASAMY, Seyed MOHAMMADI, Richard RITIKOS a Saadah RAHMAN. Optical Characterization of Oligonucleotide DNA Influenced by Magnetic Fields. *Molecules* [online]. 2013, **18**(10), 11797–11808. ISSN 1420-3049. Dostupné z: doi:10.3390/molecules181011797
- [47] COOPER, Alan. *Biophysical chemistry*. Cambridge: Royal Society of Chemistry, 2004. Tutorial chemistry texts, 16. ISBN 978-0-85404-480-1.
- [48] O'CONNOR, Gavin, Carol DAWSON, Alison WOOLFORD, Kenneth S. WEBB a Tim CATTERICK. Quantitation of Oligonucleotides by Phosphodiesterase Digestion Followed by Isotope Dilution Mass Spectrometry: Proof of Concept. *Analytical*

- Chemistry* [online]. 2002, **74**(15), 3670–3676. ISSN 0003-2700, 1520-6882. Dostupné z: doi:10.1021/ac0255375
- [49] JAIN, Aakanchha, Richa JAIN a Sourabh JAIN. Estimation of DNA by Diphenylamine Reaction. In: Aakanchha JAIN, Richa JAIN a Sourabh JAIN *Basic Techniques in Biochemistry, Microbiology and Molecular Biology* [online]. New York, NY: Springer US, 2020 [vid. 2021-03-23], Springer Protocols Handbooks, s. 69–71. ISBN 978-1-4939-9860-9. Dostupné z: doi:10.1007/978-1-4939-9861-6_22
- [50] ENGELHARDT, Heinz. *Practice of High Performance Liquid Chromatography: Applications, Equipment and Quantitative Analysis*. ISBN 9783642692253
- [51] VÉKEY, Károly, András TELEKES a Akos VERTES, ed. *Medical applications of mass spectrometry*. 1st ed. Amsterdam ; Boston: Elsevier, 2008. ISBN 978-0-444-51980-1.
- [52] ARMAREGO, Wilfred L. F. *Purification of laboratory chemicals*. Eighth edition. Amsterdam Boston Heidelberg: BH, Butterworth-Heinemann, an imprint of Elsevier, 2017. ISBN 978-0-12-805456-7.
- [53] *Analytical Separation Science*. Place of publication not identified: John Wiley & Sons, Inc. ISBN 978-1-5231-1058-2.
- [54] FANALI, Salvatore. *Liquid Chromatography*. Place of publication not identified: Elsevier., 2017. ISBN 978-0-12-809345-0.
- [55] DONG, M. W. *HPLC and UHPLC for practicing scientists*. ISBN 978-1-5231-2826-6.
- [56] CREIGHTON, Thomas E. *The physical and chemical basis of molecular biology*. Eastbourne: Helvetian Press, 2010. ISBN 978-0-9564781-0-8.
- [57] HARRISON, Roger G, Paul TODD, Scott R RUDGE a Demetri P PETRIDES. *Bioseparations science and engineering*. 2020. ISBN 978-0-19-756276-5.
- [58] CHEN, Yuxin, Colin T. MANT a Robert S. HODGES. Preparative reversed-phase high-performance liquid chromatography collection efficiency for an antimicrobial peptide on columns of varying diameters (1mm to 9.4mm I.D.). *Journal of Chromatography A* [online]. 2007, **1140**(1–2), 112–120. ISSN 00219673. Dostupné z: doi:10.1016/j.chroma.2006.11.052
- [59] SCHELLINGER, Adam P. a Peter W. CARR. Isocratic and gradient elution chromatography: A comparison in terms of speed, retention reproducibility and quantitation. *Journal of Chromatography A* [online]. 2006, **1109**(2), 253–266. ISSN 00219673. Dostupné z: doi:10.1016/j.chroma.2006.01.047
- [60] Chapter 6 Desalting and concentrating oligonucleotide solutions. In: *Laboratory Techniques in Biochemistry and Molecular Biology* [online]. B.m.: Elsevier, 1976 [vid. 2021-04-17], s. 295–301. ISBN 978-0-7204-4215-1. Dostupné z: doi:10.1016/S0075-7535(08)70279-1
- [61] SEMENYUK, Andrey, Matilda AHN FELT, Camilla ESTMER NILSSON, Xiao YONG HAO, Andras FÖLDESI, Yu-Shu KAO, Hong-Huei CHEN, Wei-Chen KAO, Konan PECK a Marek KWIATKOWSKI. Cartridge-based high-throughput purification of

- oligonucleotides for reliable oligonucleotide arrays. *Analytical Biochemistry* [online]. 2006, **356**(1), 132–141. ISSN 00032697. Dostupné z: doi:10.1016/j.ab.2006.05.008
- [62] GILAR, Martin. Analysis and Purification of Synthetic Oligonucleotides by Reversed-Phase High-Performance Liquid Chromatography with Photodiode Array and Mass Spectrometry Detection. *Analytical Biochemistry* [online]. 2001, **298**(2), 196–206. ISSN 00032697. Dostupné z: doi:10.1006/abio.2001.5386
- [63] GILAR, Martin a Edouard S.P BOUVIER. Purification of crude DNA oligonucleotides by solid-phase extraction and reversed-phase high-performance liquid chromatography. *Journal of Chromatography A* [online]. 2000, **890**(1), 167–177. ISSN 00219673. Dostupné z: doi:10.1016/S0021-9673(00)00521-5
- [64] ZHANG, Qiulong, Huanhuan LV, Lili WANG, Man CHEN, Fangfei LI, Chao LIANG, Yuanyuan YU, Feng JIANG, Aiping LU a Ge ZHANG. Recent Methods for Purification and Structure Determination of Oligonucleotides. *International Journal of Molecular Sciences* [online]. 2016, **17**(12), 2134. ISSN 1422-0067. Dostupné z: doi:10.3390/ijms17122134
- [65] CHAVALI, S., A. MAHAJAN, R. TABASSUM, S. MAITI a D. BHARADWAJ. Oligonucleotide properties determination and primer designing: a critical examination of predictions. *Bioinformatics* [online]. 2005, **21**(20), 3918–3925. ISSN 1367-4803, 1460-2059. Dostupné z: doi:10.1093/bioinformatics/bti633
- [66] STEIN, C.A., Chrisanthi SUBASINGHE, Kazuo SHINOZUKA a Jack S. COHEN. Physicochemical properties of phosphorothioate oligodeoxynucleotides. *Nucleic Acids Research* [online]. 1988, **16**(8), 3209–3221. ISSN 0305-1048, 1362-4962. Dostupné z: doi:10.1093/nar/16.8.3209
- [67] JANIK, Borek. *Physicochemical Characteristics of Oligonucleotides and Polynucleotides* [online]. Boston, MA: Springer US, 1971 [vid. 2021-05-20]. ISBN 978-1-4613-4612-8. Dostupné z: https://doi.org/10.1007/978-1-4613-4612-8
- [68] BOWATER, Richard P a Andrew J GATES. Nucleotides: Structure and Properties. In: JOHN WILEY & SONS LTD, ed. *eLS* [online]. Chichester, UK: John Wiley & Sons, Ltd, 2015 [vid. 2021-05-20], s. 1–9. ISBN 978-0-470-01590-2. Dostupné z: doi:10.1002/9780470015902.a0001333.pub3
- [69] KOIKE, Tatsuro a Yasuo INOUE. STRUCTURE AND REACTIVITY OF OLIGONUCLEOTIDES. PART I KINETICS OF THE NON-ENZYMATIC TRANSPHOSPHORYLATION OF ADENYLYL-(3'-5')-ADENOSINE 3'-PHOSPHATE AND OTHER DINUCLEOTIDES. *Chemistry Letters* [online]. 1972, **1**(7), 569–572. ISSN 0366-7022, 1348-0715. Dostupné z: doi:10.1246/cl.1972.569
- [70] LESCRINIER, E. Difference in conformational diversity between nucleic acids with a six-membered „sugar" unit and natural „furanose" nucleic acids. *Nucleic Acids Research* [online]. 2003, **31**(12), 2975–2989. ISSN 1362-4962. Dostupné z: doi:10.1093/nar/gkg407
- [71] BIBA, Mirlinda, Eileen JIANG, Bing MAO, Daniel ZEWGE, Joe P. FOLEY a Christopher J. WELCH. Factors influencing the separation of oligonucleotides using reversed-phase/ion-exchange mixed-mode high performance liquid chromatography

- columns. *Journal of Chromatography A* [online]. 2013, **1304**, 69–77. ISSN 00219673. Dostupné z: doi:10.1016/j.chroma.2013.06.050
- [72] CANTARA, William A., Erik D. OLSON a Karin MUSIER-FORSYTH. Analysis of RNA structure using small-angle X-ray scattering. *Methods* [online]. 2017, **113**, 46–55. ISSN 10462023. Dostupné z: doi:10.1016/j.ymeth.2016.10.008
- [73] BOTTARO, Sandro, Giovanni BUSSI, Scott D. KENNEDY, Douglas H. TURNER a Kresten LINDORFF-LARSEN. Conformational ensembles of RNA oligonucleotides from integrating NMR and molecular simulations. *Science Advances* [online]. 2018, **4**(5), eaar8521. ISSN 2375-2548. Dostupné z: doi:10.1126/sciadv.aar8521
- [74] DJURANOVIC, D. a B. HARTMANN. Conformational Characteristics and Correlations in Crystal Structures of Nucleic Acid Oligonucleotides: Evidence for Sub-states. *Journal of Biomolecular Structure and Dynamics* [online]. 2003, **20**(6), 771–788. ISSN 0739-1102, 1538-0254. Dostupné z: doi:10.1080/07391102.2003.10506894
- [75] REUTER, Jessica S a David H MATHEWS. RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* [online]. 2010, **11**(1), 129. ISSN 1471-2105. Dostupné z: doi:10.1186/1471-2105-11-129
- [76] MARKHAM, Nicholas R. a Michael ZUKER. UNAFold. In: Jonathan M. KEITH, ed. *Bioinformatics* [online]. Totowa, NJ: Humana Press, 2008 [vid. 2021-05-20], *Methods in Molecular Biology*TM, s. 3–31. ISBN 978-1-60327-428-9. Dostupné z: doi:10.1007/978-1-60327-429-6_1
- [77] HENDLING, Michaela a Ivan BARIŠIĆ. In-silico Design of DNA Oligonucleotides: Challenges and Approaches. *Computational and Structural Biotechnology Journal* [online]. 2019, **17**, 1056–1065. ISSN 20010370. Dostupné z: doi:10.1016/j.csbj.2019.07.008
- [78] *Factors Affecting Resolution in HPLC* [online]. [vid. 2021-07-14]. Dostupné z: <https://www.sigmaaldrich.com/CZ/en/technical-documents/technical-article/analytical-chemistry/small-molecule-hplc/factors-affecting-resolution-in-hplc>
- [79] GUO, X. Secondary structural characterization of oligonucleotide strands using electrospray ionization mass spectrometry. *Nucleic Acids Research* [online]. 2005, **33**(11), 3659–3666. ISSN 0305-1048, 1362-4962. Dostupné z: doi:10.1093/nar/gki671
- [80] MIRLINDA BIBA, BING MAO, CHRISTOPHER J. WELCH, a JOE P. FOLEY. Liquid Chromatography Methods for the Separation of Short RNA Oligonucleotides. *Lc Gc North America*. 2014, **32**(4), 42–50.
- [81] STURM, Marc, Sascha QUINTEN, Christian G. HUBER a Oliver KOHLBACHER. A statistical learning approach to the modeling of chromatographic retention of oligonucleotides incorporating sequence and secondary structure data. *Nucleic Acids Research* [online]. 2007, **35**(12), 4195–4202. ISSN 1362-4962, 0305-1048. Dostupné z: doi:10.1093/nar/gkm338
- [82] R CORE TEAM. *R: A language and environment for statistical computing*. [online]. Vienna, Austria: R Foundation for Statistical Computin. 2021. R: A language and environment for statistical computing. Dostupné z: <https://www.R-project.org/>.

- [83] WICKHAM, Hadley. *ggplot2: Elegant Graphics for Data Analysis* [online]. 2nd ed. 2016. Cham: Springer International Publishing: Imprint: Springer, 2016. Use R! ISBN 978-3-319-24277-4. Dostupné z: doi:10.1007/978-3-319-24277-4
- [84] WICKHAM, Hadley, Mara AVERICK, Jennifer BRYAN, Winston CHANG, Lucy MCGOWAN, Romain FRANÇOIS, Garrett GROLEMUND, Alex HAYES, Lionel HENRY, Jim HESTER, Max KUHN, Thomas PEDERSEN, Evan MILLER, Stephan BACHE, Kirill MÜLLER, Jeroen OOMS, David ROBINSON, Dana SEIDEL, Vitalie SPINU, Kohnske TAKAHASHI, Davis VAUGHAN, Claus WILKE, Kara WOO a Hiroaki YUTANI. Welcome to the Tidyverse. *Journal of Open Source Software* [online]. 2019, 4(43), 1686. ISSN 2475-9066. Dostupné z: doi:10.21105/joss.01686
- [85] VERSION), José Pinheiro (S, Douglas Bates (up TO 2007), Saikat DebRoy (up TO 2002), Deepayan Sarkar (up TO 2005), EISPACK AUTHORS (SRC/RS.F), Siem Heisterkamp (Author fixed SIGMA), Bert Van Willigen (Programmer fixed SIGMA), Johannes RANKE (VARCONSTPROP()) a R CORE TEAM. *nlme: Linear and Nonlinear Mixed Effects Models* [online]. 25. březen 2022 [vid. 2022-05-12]. Dostupné z: <https://CRAN.R-project.org/package=nlme>
- [86] WICKHAM, Hadley, Jennifer BRYAN, RStudio (Copyright holder of all R. code and all C/C++ code without explicit copyright ATTRIBUTION), Marcin Kalicinski (Author of included RapidXML CODE), Komarov Valery (Author of included libxls CODE), Christophe Leitienne (Author of included libxls CODE), Bob Colbert (Author of included libxls CODE), David Hoerl (Author of included libxls CODE) a Evan Miller (Author of included libxls CODE). *readxl: Read Excel Files* [online]. 28. březen 2022 [vid. 2022-05-12]. Dostupné z: <https://CRAN.R-project.org/package=readxl>
- [87] WICKHAM, Hadley, Dana SEIDEL a RSTUDIO. *scales: Scale Functions for Visualization* [online]. 13. duben 2022 [vid. 2022-05-12]. Dostupné z: <https://CRAN.R-project.org/package=scales>
- [88] AUGUIE, Baptiste a Anton ANTONOV. *gridExtra: Miscellaneous Functions for „Grid“ Graphics* [online]. 9. září 2017 [vid. 2022-05-12]. Dostupné z: <https://CRAN.R-project.org/package=gridExtra>
- [89] WICKHAM, Hadley, Romain FRANÇOIS, Lionel HENRY, Kirill MÜLLER a RSTUDIO. *dplyr: A Grammar of Data Manipulation* [online]. 28. duben 2022 [vid. 2022-05-12]. Dostupné z: <https://CRAN.R-project.org/package=dplyr>
- [90] PEDERSEN, Thomas Lin a RSTUDIO. *ggforce: Accelerating „ggplot2“* [online]. 5. březen 2021 [vid. 2022-05-12]. Dostupné z: <https://CRAN.R-project.org/package=ggforce>
- [91] 16.7: Critical Values for Grubb's Test. *Chemistry LibreTexts* [online]. 10. leden 2019 [vid. 2022-04-10]. Dostupné z: [https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Analytical_Chemistry_2.1_\(Harvey\)/16%3A_Appendix/16.07%3A_Critical_Values_for_Grubb's_Test](https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Analytical_Chemistry_2.1_(Harvey)/16%3A_Appendix/16.07%3A_Critical_Values_for_Grubb's_Test)
- [92] Appendix 05: Critical Values for the F-Test. *Chemistry LibreTexts* [online]. 2. prosinec 2013 [vid. 2022-04-10]. Dostupné z: [https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Analytical_Chemistry_2.1_\(Harvey\)/05:_Appendix/05.05:_Critical_Values_for_the_F-Test](https://chem.libretexts.org/Bookshelves/Analytical_Chemistry/Analytical_Chemistry_2.1_(Harvey)/05:_Appendix/05.05:_Critical_Values_for_the_F-Test)

z: https://chem.libretexts.org/Ancillary_Materials/Reference/Reference_Tables/Analytic_References/Appendix_05%3A_Critical_Values_for_the_F-Test

- [93] HIGSON, S^{??}amus. *Analytical chemistry* [online]. Oxford: Oxford University Press, 2006 [vid. 2022-04-18]. ISBN 978-0-19-850289-0. Dostupné z: <http://app.knovel.com/web/toc.v/cid:kpAC000001>
- [94] PATIENCE, Gregory S. *Experimental methods and instrumentation for chemical engineers*. Amsterdam: Elsevier, 2013. ISBN 978-0-444-53804-8.
- [95] GILAR, Martin, Kenneth J. FOUNTAIN, Yeva BUDMAN, Uwe D. NEUE, Kurt R. YARDLEY, Paul D. RAINVILLE, Reb J. RUSSELL II a John C. GEBLER. Ion-pair reversed-phase high-performance liquid chromatography analysis of oligonucleotides: *Journal of Chromatography A* [online]. 2002, **958**(1–2), 167–182. ISSN 00219673. Dostupné z: [doi:10.1016/S0021-9673\(02\)00306-0](https://doi.org/10.1016/S0021-9673(02)00306-0)