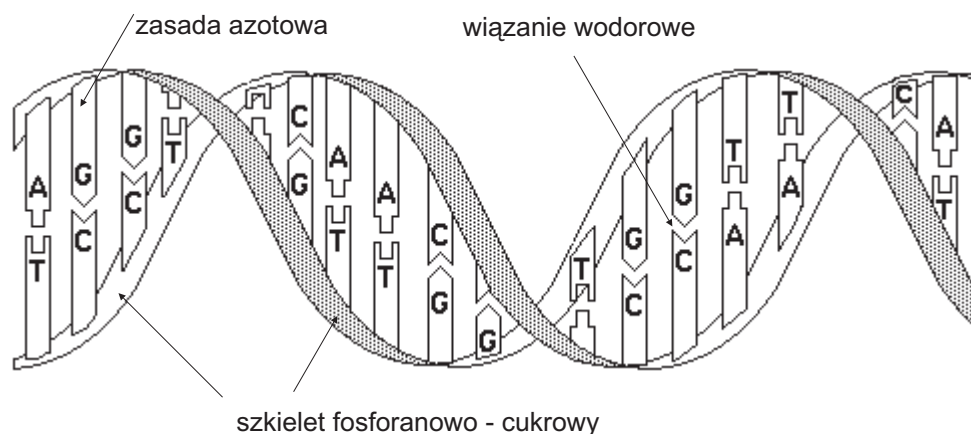


Dodatek A

Metody inżynierii genetycznej wykorzystywane w obliczeniach molekularnych

A.1 Budowa DNA

DNA (kwas dezoksyrybonukleinowy) jest nośnikiem informacji genetycznej w organizmach żyjących na Ziemi. Cząsteczka DNA składa się z dwóch oplatających się, oddalonych od siebie o 2 nm nici złożonych z nukleotydów. Taką strukturę nazywa się podwójną helisą. **Nukleotyd** składa się z nukleozydu i reszty kwasu fosforowego. W skład nukleozydu DNA wchodzi z kolei cząsteczka dezoksyrybozy i zasada azotowa. Dezoksyryboza jest to cząsteczka cukru prostego z grupy pentoz, czyli cukrów pięciowęglowych. Do środka helisy skierowane są połączone z cząsteczkami cukru zasady azotowe. Skok helisy wynosi 3,4 nm. Budowę podwójnej helisy DNA przedstawiono na rys. A.1. W naturze występują 4 różne nukleotydy DNA, które różnią się jedynie zasadami azotowymi. Istnieją następujące zasady azotowe: Adenina, Guanina, Cytosyna i Tymina, więc istnieją następujące nukleotydy DNA: kwas dezoksyadenylowy, kwas dezoksyguanylowy, kwas dezoksytydylowy, kwas dezoksytymidylowy. W skrócie nukleotydy oznaczają się literami



Rysunek A.1: Budowa podwójnej helisy DNA.

A,G,C,T. Chemicznie otrzymano jeszcze wiele innych typów nukleotydów charakteryzujących się ciekawymi własnościami. Zgodnie z regułą Chargaffa (ustaloną, gdy nie znano jeszcze struktury DNA) ilość moli zasad purynowych (Adenina, Guanina) jest równa ilości moli zasad pirymidynowych (Cytosyna, Tymina). Reguła ta znalazła potwierdzenie w istnieniu **komplementarności zasad**, co znaczy, że adenina z jednej nici łączy się zawsze z tyminą z drugiej nici, a guanina z cytozyną. To selektywne oddziaływanie jest często wykorzystywanym mechanizmem w obliczeniach molekularnych. Wiązania łączące zasady należą do grupy wodorowych, czyli słabych, łatwych do zerwania oddziaływań między przeciwnie naładowanymi atomami. Między adeniną a tyminą występują dwa takie wiązania, a między guaniną a cytozyną - trzy.

Nukleotyd jest cząsteczką niesymetryczną - wykazuje polarność, ponieważ cząsteczka cukru nie jest symetryczna. Na jednym końcu nukleotydu, oznaczanym 5', występuje grupa fosforanowa (która jest przyłączona do piątego atomu węgla dezoksyrybozy), a na drugim, oznaczanym 3', znajduje się grupa hydroksylowa związana z trzecim atomem węgla dezoksyrybozy. Kolejne nukleotydy tworzą łańcuch polinukleotydowy łącząc się szeregowo, 5' koniec z 3' końcem. Nukleotydy połączone są ze sobą wiązaniami fosfodie-

strowymi, które są silne i tworzą tak zwany szkielet fosforanowo - cukrowy. W podwójnej helisie DNA nici ułożone są naprzemianlegle, to znaczy naprzeciwko 5' końca jednej nici jest 3' koniec drugiej nici. Oznacza to, że łańcuchy mają przeciwną polarność - jeden z nich biegnie w jedną stronę, a drugi w przeciwną.

Heliks DNA można porównać do skręconej spiralnie drabiny, której szczeble tworzą oddziałujące ze sobą zasady, natomiast pionowe listwy stanowią połączone wiązaniami fosfodiestrowymi cząsteczki dezoksyrybozy. Konieczność spiralnego skręcenia całej struktury wynika z rozmiarów poszczególnych elementów nukleotydów. Na jeden pełny obrót heliksu przypada 11 par zasad. Heliksy DNA są prawoskrętne, łańcuchy polinukleotydowe wydłużają się, okręcają się wokół siebie w kierunku zgodnym z ruchem wskazówek zegara.

Cząsteczki DNA występujące w komórkach organizmów owijają się wokół białek tworząc złożone struktury i ulegając wielokrotnej spiralizacji (w jądrze komórki człowieka znajdują się cząsteczki DNA o łącznej długości około 2 metrów). Przy opisie reakcji z dziedziny obliczeń molekularnych stosuje się uproszczony sposób zapisu cząsteczek DNA. Podaje się jedynie sekwencje nukleotydów oraz kierunki nici DNA. Tak więc pewien fragment podwójnej helisy można zapisać:



Stosując taki zapis traci się informację o kształcie cząstki, który może mieć duży wpływ na wydajność reakcji. Gdy opisuje się sekwencję pojedynczej nici zazwyczaj koniec 5' znajduje się po lewej stronie. Jeżeli obrazuje się dwuniciową cząsteczkę DNA, to nić górna ma 5' koniec po lewej stronie. W zapisie czasami pojawia się litera N oznaczająca dowolny nukleotyd (z A, T, G, C).

A.2 Hybrydyzacja i denaturacja

Ponieważ wiązania wodorowe utrzymujące dwie nici podwójnej helisy ze sobą są relatywnie słabe, nici te można rozpleść i rozdzielić przy niewielkim nakładzie energii. Żeby rozpleść helisę, wystarczy doprowadzić do temperatury bliskiej temperaturze wrzenia wodny roztwór DNA – wysoka temperatura powoduje rozerwanie wiązań wodorowych nie uszkadzając silniejszych wiązań, utrzymujących razem nukleotydy w poszczególnych niciach. Zjawisko to nosi nazwę **denaturacji DNA**. Denaturacja zachodzi również pod wpływem wysokiego pH (powyżej 13). Temperatura równowagi termodynamicznej dla danej cząsteczki zależy od tworzących ją nukleotydów i można obliczyć ją dla krótkich (do 20 par zasad) łańcuchów DNA z zależności A.1, gdzie a jest liczbą par AT, zaś b jest liczbą par GC.

$$t[^\circ\text{C}] = 2 * a + 4 * b \tag{A.1}$$

Komplementarne, pojedyncze nici odtwarzają strukturę dwuniciową w procesie zwanym **renaturacją**. Renaturację zachodzącą pomiędzy łańcuchami kwasów nukleinowych o różnym pochodzeniu (np. DNA i RNA, lub DNA z dwóch różnych organizmów) nazywa się **hybrydyzacją**. Hybrydyzacja jest oddziaływaniem selektywnym - hybrydują do siebie tylko cząsteczki o komplementarnych sekwencjach. Hybrydyzacja jest podstawowym zjawiskiem wykorzystywanym do przeprowadzania obliczeń molekularnych, ma także duże znaczenie w inżynierii genetycznej (służy do wyszukiwania komplementarnych odcinków DNA w banku genów, stosuje się ją do badań ewolucyjnych oraz badania stopnia pokrewieństwa między poszczególnymi osobnikami).

A.3 Polimeryzacja DNA, PCR

Enzymy są to białka, wyodrębnione z żywych organizmów, najczęściej z bakterii. Niektóre enzymy pozwalają przeprowadzać pewne operacje na cząstecz-

kach DNA i z tego powodu są powszechnie wykorzystywane przez inżynierię genetyczną.

Cząsteczka DNA jest dwuniciowa i nici są do siebie wzajemnie komplementarne, więc występuje redundancja informacji. **Replikacja DNA** jest procesem, w którym główną rolę odgrywają enzymy zwane **polimerazami DNA**. Podczas replikacji nici helisy są rozplatane w określonym rejonie i każda z nich służy jako matryca dla nowej, komplementarnej nici. Tak więc w każdej z nowo powstających dwuniciowych cząsteczek DNA jedna z nici jest odziedziczona po cząsteczce macierzystej, a druga - nowo syntezowana. Enzym polimerazy DNA syntezuje cząsteczkę komplementarną do matrycowej nici DNA. Polimeraza katalizuje reakcję wydłużania syntetyzowanego łańcucha DNA przez kolejne dołączanie nukleotydów do wolnej grupy 3'-OH. Kierunek syntezy jest zawsze taki sam - od 5' do 3' końca. Do syntezy nici DNA polimerazy wymagają komplementarnego, krótkiego jednoniciowego odcinka DNA zwanego **starterem** (ponieważ potrafią jedynie wydłużać nić w oparciu o nić matrycową).

Często potrzebne jest zwiększenie ilości cząsteczek DNA o danej sekwencji. Można to osiągnąć umieszczając tę sekwencję w łańcuchu DNA organizmu żywego. Mechanizmy powielania komórek sprawiają, że interesująca nas sekwencja będzie powielana. Taki proces nazywa się **klonowaniem**. Jest on najbardziej wydajny, gdy klonujemy w komórkach bakteryjnych, ponieważ ich podział następuje co około 20 minut. Proces taki jest stosowany w biologii molekularnej, jednak jest on bardzo czasochłonny, dlatego w dziedzinie obliczeń molekularnych nie jest spotykany.

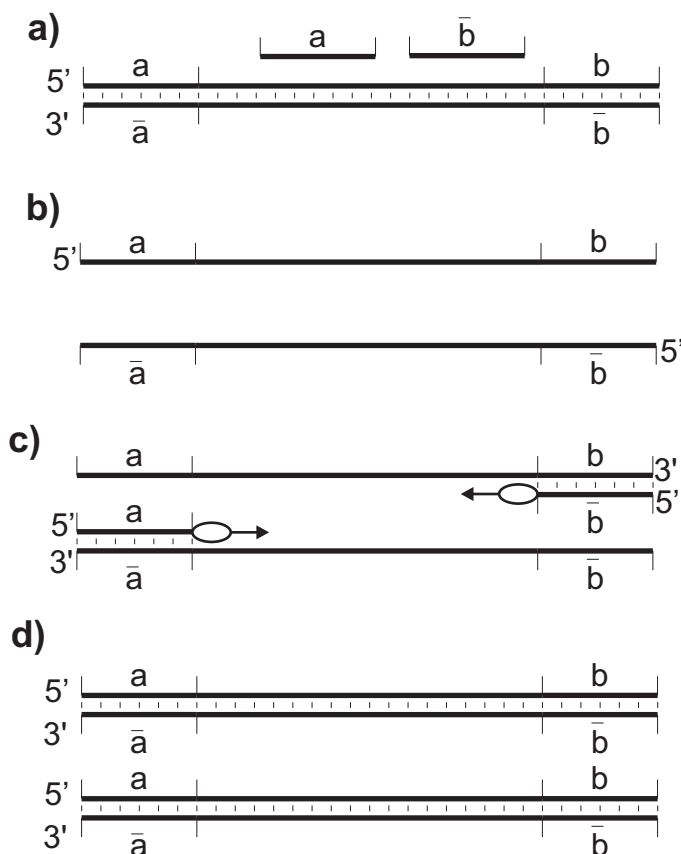
Polimerazowa reakcja łańcuchowa (PCR) jest techniką powielania *in vitro* fragmentów DNA. Polega ona na powielaniu fragmentu DNA na matrycy kwasu nukleinowego z użyciem starterów i termostabilnej polimerazy. PCR pozwala na namnożenie określonych odcinków DNA praktycznie w dowolnych ilościach. W metodzie PCR chcąc namnożyć określony odcinek DNA, dodaje się startery - krótkie (z reguły do 20 nukleotydów) jednoniciowe cząsteczki DNA, komplementarne do dwóch nici żądanego odcinka DNA na jego

końcach (rys. A.2a). Aby startery mogły się przyłączyć, matrycowa cząsteczka DNA musi najpierw ulec denaturacji (jak pokazano na rys. A.2b), co osiąga się podnosząc temperaturę do około 94 °C. Następnie temperatura zostaje obniżona do około 40–50 °C (zależnie od sekwencji starterów). W tej temperaturze startery hybrydują w odpowiednich miejscach, co zilustrowano na rys. A.2c. Startery są dodane w stężeniu tak wysokim, że prawdopodobieństwo hybrydyzacji startera jest o wiele wyższe niż renaturacja dwóch nici matrycy. W końcu podnosi się temperaturę do 72 °C i w tej temperaturze przeprowadza się polimeryzację (polimeraza dobudowuje nowe nici zaczynając od starterów). Tym samym uzyskuje się z jednej cząsteczki DNA dwie takie same, co pokazano na rys. A.2d. Po tym kroku cykl można powtórzyć.

W każdym cyklu ilość cząsteczek DNA ulega podwojeniu, więc po niewielu cyklach uzyskuje się żadaną liczbę cząstek. Metoda jest możliwa do zastosowania dzięki temu, że enzym termostabilnej polimerazy nie jest niszczonej w temperaturze 95 °C. Używa się polimerazy izolowanej z bakterii żyjących w gorących źródłach, zwanej Taq polimerazą. Metoda PCR ma wiele zastosowań i jest podstawową techniką wykorzystywaną do budowy urządzeń obliczeniowych przedstawionych w niniejszej rozprawie.

A.4 Cięcie enzymami restrykcyjnymi

Enzymy restrykcyjne (zwane także **restrykatazami**) są białkami zdolnymi do rozpoznawania specyficznych sekwencji w DNA i do przecinania dwuniciowej cząsteczki DNA. Cięcie polega na zerwaniu silnego wiązania fosfodiesterowego, utrzymującego razem nukleotydy w poszczególnych niciach. Największe znaczenie praktyczne mają enzymy klasy II, które rozpoznają sekwencje składające się z 4 – 12 nukleotydów, zaś miejsce przecięcia znajduje się w obrębie rozpoznawanej sekwencji lub w ściśle określonej odległości od niej. Znanych jest obecnie kilka tysięcy restryktaz, z których kilkaset jest sprzedawanych przez firmy zajmujące się handlem odczynnikami chemicznymi. Tab. A.1 zawiera sekwencje rozpoznawane i miejsca cięcia niektórych



Rysunek A.2: Polimerazowa reakcja łańcuchowa, namnażanie łańcucha o końcach oznaczanych a, b; a) cząsteczka wejściowa oraz startery; b) denaturacja nici matrycowej; c) przyłączenie starterów oraz polimeryzacja; d) wynikowe cząsteczki.

enzymów restrykcyjnych.

W wyniku trawienia enzymem restrykcyjnym mogą powstać cząsteczki z jednoniciowymi końcami. Mówimy wtedy, że enzym zostawia **lepkie końce**. Dzieje się tak dlatego, że enzym tnie obie nici DNA w różnych miejscach. Przykładem takiego enzymu jest *EcoRI*. Jeżeli miejsce przecięcia obu nici jest takie samo, to cząsteczki po przecięciu mają **tepe końce**. Tak tnie *HaeIII*. Lepkie końce mają duże znaczenie, ponieważ sekwencje lepkich końców powstałych w wyniku działania tego samego enzymu są komplementarne i mogą one łatwo hybrydyzować ze sobą. Można też czasami dobrać dwa enzymy roz-

Nazwa	Rozpoznawana sekwencja	Cząsteczki po przecięciu	
<i>Hinf</i> I	GANTC CTNAG	G CTNA	ANTC G
<i>Eco</i> RI	GAATTC CTTAAG	G CTTAA	AATTC G
<i>Fok</i> I	GGATGNNNNNNNNNNNN CCTACNNNNNNNNNNNN	GGATGNNNNNNNNNN CCTACNNNNNNNNNN	NNNN N
<i>Hae</i> III	GGCC CCGG	GG CC	CC GG
<i>Hpa</i> II	CCGG GGCC	C GCC	CGG C
<i>Pst</i> I	CTGCAG GACGTC	CTGCA G	G ACGTC

Tablica A.1: Niektóre enzymy restrykcyjne, sekwencje rozpoznawane i miejsca cięcia.

poznające różne sekwencje i zostawiające takie same lepkie końce.

Trawienie enzymem restrykcyjnym (oraz inne reakcje wykorzystujące enzymy) można zatrzymać przez grzanie przez 10-15min w temperaturze 65-70 °C lub ekstrakcją fenolem (dodanie nasyconego fenolu sprawia, że DNA przechodzi do fazy wodnej, którą się zbiera). Istnieją termostabilne enzymy restrykcyjne, które nie ulegają termicznej inaktywacji.

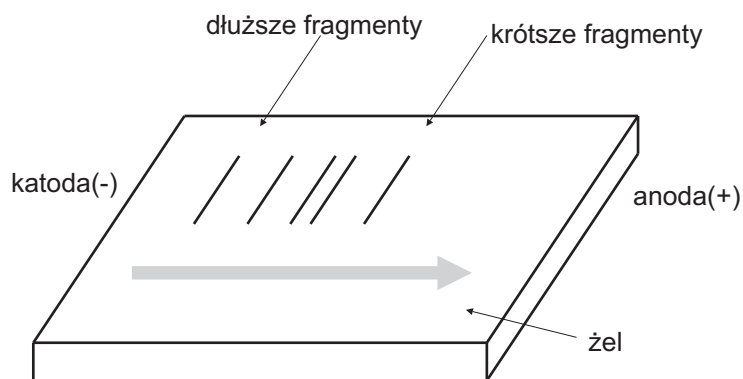
A.5 Ligacja DNA

Powstanie długiego łańcucha DNA wymaga łączenia wielu krótkich fragmentów w jedną nić. Odpowiedzialny za to jest enzym zwany **ligazą DNA**, który katalizuje powstawanie wiązań między fragmentami DNA. **Ligacja** polega na tworzeniu kowalencyjnych wiązań fosfodiesterowych pomiędzy grupą fosforanową na 5' końcu jednego łańcucha a grupą hydroksylową z 3' końca

drugiego łańcucha. Jest to reakcja odwrotna do reakcji cięcia enzymem restrykcyjnym. Ligazy wymagają obecności ATP, który jest źródłem energii. Reakcję ligacji przeprowadza się w niskiej temperaturze 14 – 16 °C, ponieważ taka temperatura sprzyja powstawaniu wiązań wodorowych między lepkiemi końcami fragmentów DNA. W przypadku ligowania cząsteczek z tępymi końcami konieczne jest dodanie glikolu polietylenowego do buforu ligacyjnego, aby zwiększyć prawdopodobieństwo połączenia tępych końców ze sobą.

A.6 Odczytywanie długości łańcuchów DNA

Elektroforeza w żelach jest techniką stosowaną standardowo przy rozdzielaniu, analizie i oczyszczaniu kwasów nukleinowych. Pozwala ona badać długość DNA z dokładnością do 1 nukleotydu, eliminując w wielu przypadkach użycie drogich mikroskopów elektronowych. Po umieszczeniu w polu elektrycznym cząsteczki DNA migrują w żelu, zgodnie ze swym ujemnym ładunkiem. Szybkość migracji zależy od wielkości cząsteczki. Schematycznie proces ten pokazano na rys. A.3.



Rysunek A.3: Elektroforeza - rozdzielanie cząsteczek DNA ze względu na ich długość.

Elektroforezę można prowadzić w warunkach natywnych bądź denaturujących. Żel posiada warunki **denaturujące**, gdy następuje w nim rozdział

dwuniciowej cząsteczki DNA na jednoniciowe. Jeśli takie zjawisko nie występuje mówimy, że elektroforeza jest prowadzona w warunkach **natywnych**.

Szybkość migracji w żelu jednoniciowych cząsteczek DNA, zależy nie tylko od ich wielkości i ładunku, ale również w dużym stopniu od kształtu, jaki przybiera cząsteczka. Aby dokładnie określić wielkość takich cząsteczek eliminuje się różnice w szybkości migracji wywołane kształtem stosując żele denaturujące. Są to najczęściej żele akrylamidowe, w których czynnikiem denaturującym jest mocznik.

Elektroforeza w **żelach agarozowych**, jest metodą do rozdzielania cząsteczek o długościach 200 – 10000 par zasad. Elektroforezę przeprowadza się w odpowiednim buforze. Najczęściej stosowanym buforem jest TAE (Tris-octan, EDTA) lub TBE (Tris-boran, EDTA). Przed naniesieniem na żel próbki zawieszona się w buforze zawierającym „obciążacze” (glicerol, Ficoll lub sacharoza) oraz barwnik umożliwiający śledzenie przebiegu elektroforezy.

Elektroforeza w **żelach poliakrylamidowych** stosowana jest do rozdzielania małych cząsteczek DNA. Żel taki powstaje poprzez polimeryzację monomerów akrylamidu w długie łańcuchy z wytworzeniem wiązań poprzecznych przez bisakrylamid. Wielkość porów w żelu zależy od stężenia akrylamidu do bisakrylamidu. Żele poliakrylamidowe stosuje się przy rozdziale fragmentów o wielkości od 6 par zasad (20% poliakrylamid) do 1000 par zasad (3% poliakrylamid) oraz przy rozdziale jednoniciowych cząsteczek DNA. Do elektroforezy stosuje się zazwyczaj bufor TEB (tris-boran, EDTA). W żelach akrylamidowych stosuje się najczęściej barwniki, aby kontrolować przebieg reakcji. DNA w żelu uwidacznia się poprzez znakowanie omówione w dalszej części pracy.

Linijowe cząsteczki DNA migrują w żelu z prędkością odwrotnie proporcjonalną do logarytmu dziesiętnego ich masy cząsteczkowej (wyrażonej w Daltonach lub w liczbie par zasad). Prędkość migracji jest również zależna od stężenia oraz rodzaju żelu, natężenia pola elektrycznego, temperatury. Praktycznie stosuje się kalibrację żelu. Używa się do tego standardy wielkości, czyli cząsteczki DNA o znanej wielkości, poddawane elektroforezie w tym samym

żelu co cząsteczki badane.

A.7 Odczytywanie sekwencji DNA

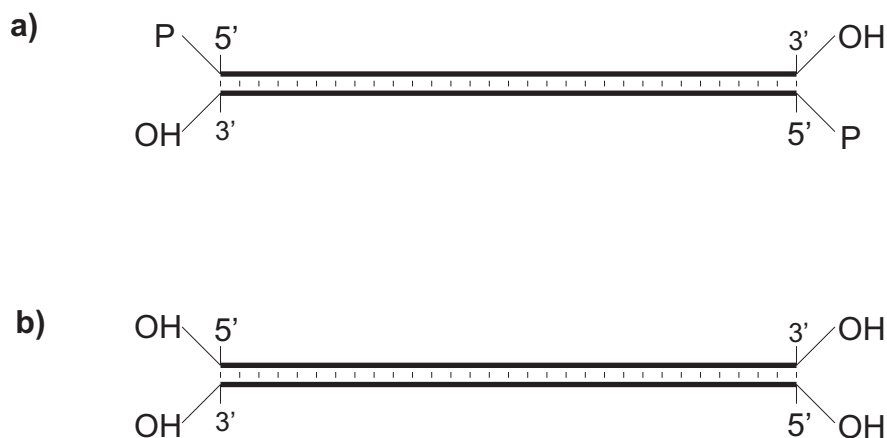
Sekwencjonowanie cząsteczek DNA pozwala poznać kolejność występujących w nich nukleotydów. Obecnie najczęściej stosuje się metodę polegającą na analizowaniu produktów polimeryzacji na jednoniciowej matrycy, počawszy od przygotowanego startera. Prowadzi się równolegle cztery reakcje. Każda z nich zawiera w mieszaninie nukleotydów dideoksytrifosforany nukleotydów, powodujące zakończenie polimeryzacji w pozycji, w której zostaną włączone do syntetyzowanej nici. Powstaje zatem pełen zestaw fragmentów DNA o różnej długości, kończących się odpowiednim nukleotydem. Metoda ta pozwala odczytać sekwencje 400 - 700 nukleotydów w jednym doświadczeniu. Techniki sekwencjonowania DNA są na tyle zautomatyzowane, że stało się możliwe podjęcie projektów poznania sekwencji całych genomów.

A.8 Modyfikacja końców cząsteczek DNA

Cząsteczki DNA izolowane z organizmów zawierają na 5' końcu **grupę fosforanową**, która jest niezbędna do łączenia dwóch łańcuchów DNA za pomocą ligazy. Takie cząsteczki są pokazane na rys. A.4a. Jeżeli cząsteczki takiej grupy nie mają, to ligacja fragmentów nie jest możliwa.

Usunięcie grupy fosforanowej następuje w wyniku działania **alkalicznej fosfatazy**. W wyniku tej reakcji na końcu 5' zostaje dołączona grupa OH, tak jak pokazano na rys. A.4b.

Działanie odwrotne posiada enzym zwany **kinazą polinukleotydową**. Kinaza taka przenosi grupę fosforanową z nukleotydu ATP na cząsteczkę DNA, zawierającą grupę OH na 5' końcu. Jeżeli cząsteczka z rys. A.4b zostanie poddana reakcji kinazowania, to powstanie cząsteczka z rys. A.4a. Po reakcji kinazowania cząsteczki mogą być łączone za pomocą enzymu ligazy.



Rysunek A.4: Grupy na końcach DNA; a) DNA izolowane z organizmów posiada na 5' końcu grupa fosforanowa, to samo po reakcji kinazowania; b) DNA otrzymywane sztucznie posiada na 5' końcu grupa OH, to samo po działaniu alkalicznej fosfatazy.

Sztucznie syntezowane oligonukleotydy także nie mają grupy fosforanowej na 5' końcu, zatem nie mogą one ligować, jeżeli nie zastosuje się reakcji kinazowania.

A.9 Znakowanie kwasów nukleinowych

Wiele metod biologii molekularnej oraz wiele metod obliczeń molekularnych opiera się na wykorzystaniu znakowanych cząsteczek DNA. Stosuje się techniki znakowania radioaktywnego oraz znakowania fluorescencyjnego.

Metody **znakowania fluorescencyjnego** dzieli się na dwa typy: bezpośrednie i pośrednie. Najbardziej rozpowszechnione są systemy pośrednie, w których dodatkowa cząsteczka wbudowuje się w helisę DNA. Przykładami takich cząsteczek są **bromek etydyny** i biotyna. Najbardziej powszechną metodą znakowania cząsteczek podczas elektroforezy jest dodanie do żelu agarozowego bromku etydyny, który wbudowuje się między obie nici cząsteczki DNA. Bromek etydyny uwidacznia się pod wpływem światła UV.

Systemy bezpośrednio stosują specjalne nukleotydy, które podczas syntezy zostały tak zmodyfikowane, że wykazują własności fluorescencyjne. Takie systemy nie potrzebują dodatkowych związków biorących udział w reakcji.

Mimo dużego postępu technik znakowania fluorescencyjnego, w dalszym ciągu wykorzystuje się **znakowanie radioaktywne**. Używa się nukleotydów, w których jeden z atomów został zastąpiony izotopem radioaktywnym. Najczęściej stosuje się izotopy fosforu ^{32}P lub ^{33}P , oraz izotop siarki ^{35}S . Fosfor ^{32}P emituje promieniowanie o najwyższej energii wśród wykorzystywanych w praktyce biologicznej izotopów (promieniowanie β o energii 1,7 MeV). Uzyskuje się zatem najbardziej aktywne sondy, zapewniana jest wysoka czułość i krótki czas detekcji. Okres półtrwania ^{32}P wynosi 14 dni. Siarka ^{35}S emituje promieniowanie o niskiej energii, nie są wymagane żadne specjalne ekrany ochronne (promieniowanie β o energii 0,17 MeV). Okres rozpadu połowicznego wynosi 87 dni. Fosfor ^{33}P ma własności pośrednie pomiędzy omówionymi powyżej izotopami.

Metody znakowania można także podzielić na metody znakujące równomiernie łańcuch na całej długości, oraz metody znakujące końce nici. Do pierwszej klasy zalicza się metodę zwaną „random priming”, która polega na polimeryzacji nici z użyciem znakowanych nukleotydów. Znakowanie przy użyciu kinazy polinukleotydowej pozwala otrzymać cząsteczkę wyznakowaną na 5' końcu. Znakowanie 3' końca fragmentu DNA jest możliwe przy użyciu **terminalnej transferazy**. Jest to enzym dołączający nukleotydy na 3' końcu, nie wymagając do tego matrycy. Inną metodą znakowania 3' końca jest wypełnianie lepkich końców przy użyciu znakowanych nukleotydów.