

STRUKTURA I FUNKCJA MAKROZĄSTECZEK BIOLOGICZNYCH

IV semestr „Zastosowania fizyki w biologii i medycynie” 30h W + 30h Ćw

1. Fizyczne podstawy struktur makrocząsteczek oraz oddziaływań wewnątrz- i międzycząsteczkowych (2 wykłady)
 - **energia** cząsteczki; przybliżenie **pola siłowego** dla dużych biomolekuł
 - parametry termodynamiczne układu cząsteczek w roztworze (**energia swobodna Gibbsa**)
 - struktura przestrzenna (**konformacja**) cząsteczki i oddziaływania stabilizujące: niewiążące **oddziaływania elektrostatyczne** (wiązania wodorowe, mostki solne, nieukierunkowane oddziaływania van der Waalsa), **oddziaływania hydrofobowe; oddziaływania warstwowe** (stacking) $\pi - \pi$ i kation- π , rola **rozpuszczalnika**
 - **specyficzność** wzajemnego rozpoznawania biomolekuł w **kompleksach molekularnych: komplementarność zasad**, wzajemne dopasowanie konformacyjne - modele „induced fit” i „fly-casting”.
2. Ogólny podział metod badania struktur i dynamiki biopolimerów: wykorzystanie różnic mas cząsteczkowych; oddziaływanie promieniowania elektromagnetycznego z cząsteczkami, mikroskopia, manipulacje pojedynczymi biomolekułami; modelowanie i teoretyczne przewidywanie struktur (1 wykład)
3. Kwasy nukleinowe DNA i RNA (4 wykłady)
 - replikacja DNA; transkrypcja; degradacja
 - biologiczna rola kwasów nukleinowych
 - struktura pierwszorzędowa - **sekwencjonowanie**
 - struktury wyższorzędowe: drugorzędowe struktury **helikalne** i **niehelikalne**, motywy struktury trzeciorzędowej RNA: pseudowzrost motyw A-minor, zamek rybozowy struktura tRNA, budowa chromatyny, **superskrętność**.
4. Białka globularne, włókniste i błonowe (4 wykłady)
 - biosynteza białek (translacja mRNA), topogeneza, degradacja
 - różnorodność funkcji białek
 - struktura pierwszorzędowa - **sekwencjonowanie**
 - struktury wyższorzędowe: wykres **Ramachandrana**, **domeny** i ich klasyfikacja, **podjednostki**
 - zwijanie białek (**protein folding**): komputerowe, *in vitro*, *in vivo*
 - enzymy białkowe i RNA; fizyczne podstawy kinetyki reakcji enzymatycznych
5. Lipidy; budowa i dynamika błon biologicznych (1 wykład)
6. Polisacharydy jako materiał zapasowy i budulcowy (1 wykład)
7. Kompleksy biomolekularne białek i kwasów nukleinowych (1 wykład)
8. Biofizyka w biotechnologii i medycynie (wybrane zagadnienia) (1 wykład)



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI



UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



LITERATURA

G. Ślósarek „Biofizyka molekularna. Zjawiska, Instrumenty, Modelowanie” PWN 2011

L. Stryer „Biochemia”

plansze, tekst i rozwiązania zadań: <http://www.biogeo.uw.edu.pl/sfmb>

WYKŁAD 1.

Zagadnienia wstępne

Wykład **STRUKTURA I FUNKCJA MAKROCZĄSTECZEK BIOLOGICZNYCH** [**PLANSZA 1: tytuł wykładu i logo**] będzie obejmował zagadnienia struktury, dynamiki i oddziaływań polimerów biologicznych [**PLANSZA 2: plan i literatura**], które warunkują wykonywanie przez te biopolimery określonych funkcji w komórce i organizmie wielokomórkowym. Szczególna uwaga będzie zwrócona na biopolimery o kluczowej roli w procesie **ekspresji genu**, czyli przekazywaniu informacji genetycznej z genów do funkcjonalnych cząsteczek białek i RNA [**PLANSZA 3: biopolimery - kwasy nukleinowe i białka**]:

- **kwasy nukleinowe DNA (genom)**: nośniki informacji genetycznej o budowie, funkcjonowaniu i podziale komórkowym;
- **kwasy nukleinowe RNA (transkryptom)**: bezpośrednie nośniki „roboczych” kopii informacji genetycznej, składniki struktur komórkowych oraz cząsteczki o różnorodnych funkcjach regulatorowych;
- **peptydy i białka (proteom)**: wielofunkcyjne polimery bezpośrednio decydujące o „materiałowych” podstawach zachodzenia procesów biologicznych oraz ich regulacji na różnych poziomach organizmu.

Kompleksowe podejście do wymienionych zagadnień jest rozwijane w ramach nowopowstałych dziedzin na styku biologii i biofizyki molekularnej: **genomiki**, **transkryptomiki** i **proteomiki**. Ponadto, omówione zostaną dwie klasy biopolimerów o bardziej wyspecjalizowanych funkcjach:

- **lipidy** jako składniki błon biologicznych;
- **polisacharydy** stanowiące materiał zapasowy i budulcowy organizmów.

Metody badań struktur i dynamiki biopolimerów [**PLANSZA 2**] będą jedynie ogólnie sklasyfikowane, bez szczegółowego omawiania (równoległy wykład "Metody

biofizyki molekularnej") na bazie podstaw fizykochemicznych, np. klasa metod opartych na procesach oddziaływania promieniowania elektromagnetycznego z cząsteczkami. Wyniki konkretnych metod będą przywoływane w trakcie wykładu, przy okazji omawiania poszczególnych zagadnień strukturalno - dynamicznych.

Założenie metodologiczne (paradygmat), które warunkuje ujęcie zagadnienia **roli biopolimerów [PLANSZA 3]**, polega na przypisaniu ich szeroko rozumianej strukturze na poziomie molekularnym zasadniczego znaczenia w określaniu mechanizmów ich biologicznego funkcjonowania w żywej komórce, co określa się terminem **SAR** (*structure activity relationship*). **Struktury chemiczne oraz przestrzenne** biomolekuł i **dynamika** ich ruchów molekularnych oraz **przekształceń strukturalnych** warunkują zdolność do **tworzenia kompleksów** molekularnych o różnym stopniu złożoności **[PLANSZA 4: kompleks białka i helikalnego kwasu dezoksyrybonukleinowego DNA]**. Wzajemne **specyficzne rozpoznawanie** biomolekuł jest niezbędne dla zachodzenia **skoordynowanych, podległych regulacji** wspólnych działań biomolekuł w procesach biochemicznych. Współczesne metody strukturalne, przede wszystkim rentgenografia, pozwalają określać z rozdzielczością atomową struktury przestrzenne kompleksów i całych fragmentów subkomórkowych o łącznej masie cząsteczkowej składników rzędu 2,7 mln jedn. masy atom. (Da), takich jak np. rybosom bakteryjny czy 3,3 mln Da jak rybosom eukariotyczny drożdży (*Ben-Shem et al. Science 330, 1203, 2010*). Przez **strukturę przestrzenną** rozumieć należy (najogólniej) zestaw **WSPÓŁRZĘDNYCH** wszystkich „atomów” cząsteczki (termin określony dalej) $r_i = (x_i \ y_i \ z_i) \ i = 1, 2, \dots, N$ atomów, w dowolnie wybranym, na ogół kartezjańskim układzie współrzędnych.

Funkcjonalna analiza biopolimerów w ujęciu SAR wykroczyła w przypadku białek i kwasów nukleinowych poza ramy badań ściśle podstawowych w kierunku masowej („przemysłowej”) **analizy wszystkich białek** kodowanych przez **zsekwencjonowane genomy**. W 2000 r. określono sekwencję **genomu ludzkiego 6×10^9 par zasad (bp)** i genomów szeregu innych organizmów, co stanowi jedno z podstawowych zadań genomiki. **[PLANSZA 3]**. Z kolei proteomika jako zasadniczy cel stawia **pełne scharakteryzowanie proteomu**, czyli w przypadku człowieka charakterystyka ok. 300 tys białek kodowanych przez ok. 30 tys. genów. W ramach proteomiki strukturalnej działalność bardzo dużych, często międzynarodowych grup badawczych LSF (*large scale facilities*) koncentruje się na kompleksowym i wysoceprzepustowym (*high throughput*) wyznaczeniu struktur dużych

klas białek (**docelowo**: wszystkich białek proteomu), współdziałających w ramach określonych odcinków szlaków metabolicznych. Pozwala to konstruować molekularne mechanizmy procesów biochemicznych w ramach **metabolomiki**, a w konsekwencji dokonywać wyboru tzw. „tarcz” czyli receptorów białkowych, w tym enzymów, dla których znana struktura na poziomie molekularnym umożliwia **racjonalne projektowanie specyficznym wiązanych przez nie ligandów jako środków farmakologicznych** (*rational drug design*): antynowotworowych, antywirusowych, antybakteryjnych. Badania w zakresie **transkryptomiki** wymagają z kolei zastosowania metod pozwalających na wyznaczenie **profilu ekspresji genów**.

Fizyczne podstawy struktur makrocząsteczek oraz oddziaływań wewnątrz- i między-cząsteczkowych

Utworzenie przez biomolekułę niezbędnej do właściwego działania **natywnej** struktury w **roztworze wodnym** lub **środowisku hydrofobowym błony biologicznej** i tworzenie przez nią **funkcjonalnych kompleksów** z innymi molekułami jest **zasadniczo uwarunkowane budową chemiczną**. Z fizycznego punktu widzenia, niezbędnego w wyjaśnianiu stabilizacji struktury molekulej i kompleksów molekulej, cząsteczka organiczna jest zbiorem tzw. **zrębów atomowych**: jąder i elektronów powłok wewnętrznych [**PLANSZA 5: struktura chemiczna, konformacja i powierzchnia molekularna cząsteczki**]. Zręby atomowe są połączone między sobą wiązaniami kowalencyjnymi, utworzonymi przez **pary elektronów** powłok zewewnętrznych, tzw. elektronami walencyjnymi, które są odpowiedzialne za wiązania chemiczne. Ten stabilny układ dodatnich i ujemnych ładunków jest opisany przez prawa **mechaniki kwantowej**.

Podstawowym **postulatem mechaniki kwantowej** jest przypisanie cząstkom tworzącym molekułę, dodatnio naładowanym jądom i ujemnie naładowanym elektronom **własności falowych (dualizm korpuskularno-falowy)**. Długość fali:

$$\lambda = h/mv, m - \text{masa}, v - \text{prędkość ruchu obiektu},$$

$$\text{stała Plancka } h = 6.626 \cdot 10^{-34} \text{ J}\cdot\text{s}$$

Własności te będą się przejawiać przede wszystkim w przypadku obiektów o małej masie czyli o dużych długościach fali: elektronów i protonów, a w mniejszym stopniu cięższych jąder składowych biomolekuł: węgla, azotu, fosforu i siarki (najważniejsze pierwiastki składowe cząsteczek biologicznych) oraz jonów metali, często związanych z białkami i kwasami nukleinowymi, lub tworzących dla nich środowisko buforowe: Na^+ , K^+ , Mg^{+2} ,

Zn^{+2} , Ca^{+2} . Ponieważ cząsteczki składają się zarówno z cięższych jąder jak i lżejszych elektronów **klasyczne podejście zawsze będzie tylko przybliżeniem, możliwym do zastosowania jeśli chodzi o pewne aspekty zachowania dużych cząsteczek.**

W ścisłym **opisie kwantowomechanicznym wszystkie** parametry ruchu cząsteczki jako obiektu fizycznego i jej elementów składowych uzyskuje się przez rozwiązanie **równania Schrödingera**. Ma to dwie zasadnicze konsekwencje, które będą podstawą dalszych rozważań:

1. **Energia cząsteczki [PLANSZA 6: diagram energetyczny cząsteczki]** nie ma charakteru ciągłego, ale może przyjmować wartości skwantowane, tworzące układ poziomów energetycznych charakterystyczny dla tej cząsteczki i zależny od jej budowy chemicznej i przestrzennej, a także, w znacznie mniejszym stopniu, od jej oddziaływań z innymi cząsteczkami;
2. Obowiązuje tzw. **zasada nieoznaczoności Heisenberga**: jednoczesne uzyskanie pewnych informacji o układzie kwantowym nie jest możliwe z dowolną precyzją, np. możemy **JEDNOCZEŚNIE** wyznaczyć energię danego poziomu i czas życia τ tego poziomu **TYLKO** ze skończoną dokładnością $\delta E \cdot \tau \geq \hbar/2$, podobnie jak nie można jednocześnie wyznaczyć z dowolną dokładnością położenia i pędu.

Na **energię pojedynczej cząsteczki** składa się w **przybliżeniu zaniedbania wpływu jąder jako obiektów ciężkich na ruch lekkich elektronów** (stosunek mas proton/elektron = 1836): energia elektronowa E_{el} (energia ruchu elektronów w polu jąder cząsteczki i energia oddziaływania między elektronami), energia oscylacyjna E_{osc} (drgań jąder), energia rotacyjna E_{rot} (obroty całej cząsteczki i obroty wewnętrzne), energia translacyjna E_{tr} (ruch postępowy cząsteczek, jedyna o charakterze ciągłym, nieskwantowanym), energia magnetycznego oddziaływania E_{magn} (między orbitalnymi i spinowymi momentami magnetycznymi jąder i elektronów) i, w przypadku przyłożenia pól zewnętrznych np. magnetycznego o indukcji B lub elektrycznego o natężeniu E : E_{ext} (energia oddziaływania elektronów i jąder z tym polem). Tak wprowadzony podział energii cząsteczki pozwala wyznaczyć schematycznie układ poziomów energetycznych, który zmodyfikowany na użytek spektroskopii nazywa się **diagramem Jabłońskiego**. Układ poziomów pokazuje wzajemne relacje między wartościami energii pochodzącymi od poszczególnych rodzajów ruchów:

$$\Delta E_{el} > \Delta E_{osc} > \Delta E_{rot} > \Delta E_{magn}, \Delta E_{ext}$$

Energia translacyjna E_{tr} jest nieskwantowana dla cząsteczek swobodnych w gazie i roztworze i może być bardzo różna (proporcjonalna do temperatury). Odpowiednie różnice energii między sąsiadującymi poziomami są następujące: $\Delta E_{el} \sim 10$ eV, $\Delta E_{osc} \sim 10^{-2}$ eV, $\Delta E_{rot} \sim 10^{-4}$ eV, $E_{magn}, E_{ext} \sim 10^{-6}$ eV, gdzie 1 elektronovolt = $1.6 \cdot 10^{-19}$ J jest energią kinetyczną elektronu przyspieszonego w polu elektrycznym o różnicy potencjałów 1V.

Wszystkie skwantowane poziomy energetyczne **z wyjątkiem poziomu podstawowego**, na którym cząsteczka ma najniższą energię elektronową, tzw. zerowe oscylacje (cząsteczka nie może całkowicie przestać drgać) oraz nie rotuje i nie oddziałuje z polami zewnętrznymi, mają skończony **średni czas życia** τ . Cząsteczka, która się tam znalazła wraca w skali czasowej τ na niższy poziom (podstawowy) **oddając** w różny sposób **energię**. Z zasady nieoznaczoności każdy poziom ma skończoną szerokość: δE .

W przypadku badania większości zagadnień związanych z ruchami molekularnymi i oddziaływaniami **biopolimerów i ich kompleksów** można je traktować w przybliżeniu jako zbiory klasycznych "atomów" **[PLANSZA 5]**. Abstrahując od elektronów (ruchy elektronów nie podlegają wtedy analizie w tym ujęciu) cząsteczka jest **zbiorem klasycznych "atomów" C, O i H, czyli punktów albo ściślej sfer materialnych o ładunku q dodatnim lub ujemnym każdy, na które składają się ładunki jąder wraz z określonym rozkładem ładunku elektronów powłok wewnętrznych (zrąb) plus odpowiednio przypisany ładunek elektronów walencyjnych**. Atomy trzymane są w cząsteczce przez siły wiązań chemicznych i **NIE mają** własności fizykochemicznych atomów swobodnych. Jest to aktualnie powszechnie stosowane podejście dla dużych biomolekuł, których **pełny opis kwantowy, jako znacznie bardziej skomplikowany numerycznie i analitycznie niż opis klasyczny, jest niemożliwy**. Dla nierotującej i spoczywającej cząsteczki o danej strukturze chemicznej **energia minimalna** odpowiada pewnemu przestrzennemu ustawieniu wzajemnemu atomów, które wynika z możliwości obrotu wokół wiązań pojedynczych i częściowo podwójnych. Czyli każdej dopuszczalnej przez strukturę chemiczną **konformacji cząsteczki** odpowiada określona energia i istnieje jedna lub kilka konformacji o energii najniższej, **zwykle odpowiadającej strukturze NATYWNEJ (funkcjonalnej)**.

Klasyczne „atomy” cząsteczki obdarzone są **wypadkowymi ładunkami elektrycznymi** i oddziałują ze sobą i z otaczającymi molekułami rozpuszczalnika jak punkty (sfery) materialne za pośrednictwem sił opisanych przez klasyczny potencjał

oddziaływania na który składają się **[PLANSZA 7: potencjał klasycznego pola siłowego]**:

$$V = (1/2)\sum_{(\text{bonds})}K_b(b - b_0)^2 + (1/2)\sum_{(\text{bond angles})}K_\theta(\theta - \theta_0)^2 + (1/2)\sum_{(\text{dihedral angles})}K_\phi[1 + \cos(\phi - \delta)] + \sum_{(\text{nonbonded pairs})}[A/r^{12} - C/r^6 + (q_1q_2)/(Dr)]$$

- potencjały **harmoniczne** zmian długości wiązań b , kątów płaskich θ i torsyjnych ϕ (stałe siłowe K) *plus ew. wychyleń grupy atomów z płaszczyzny pierścienia do którego są przyłączone*: $K_w w^2$; b_0 , θ_0 - długości wiązań i kąty płaskie w stanie równowagi;
- potencjał **kulombowski** oddziaływania punktowych, niespolaryzowanych ładunków q atomów odległych o r w ośrodku o danej stałej dielektrycznej D ;
- potencjał **van der Waalsa** (stałe Lenarda-Jonesa A i C) oddziaływań **indukcyjnych** i **dyspersyjnych** ($1/r^6$) oraz **odpychanie** powłok walencyjnych elektronów ($1/r^{12}$).

Siły działające na punkt i -ty są pochodnymi potencjału po położeniu w tym punkcie \vec{r}_i :

$$\mathbf{F}_i = -\nabla_{\vec{r}_i} V \quad i = 1, 2, \dots, N \text{ „atomów”} \quad \nabla_{\vec{r}_i} V = \left(\frac{\partial V}{\partial x} + \frac{\partial V}{\partial y} + \frac{\partial V}{\partial z} \right) \Big|_{\vec{r}_i}$$

Ogólnie: atomy związane chemicznie przez jedno, dwa i trzy wiązania trzymane są w polu wytwarzanym przez elektrony wiązań - oddziaływania „**wiążące**” czyli kowalencyjne, a odległe o więcej wiązań w strukturze chemicznej poprzez siły Coulomba i van der Waalsa, czyli oddziaływania „**niewiążące**”.

Oddziaływanie z otaczającym **rozpuszczalnikiem wodnym i jonami** **[PLANSZA 5]** w nim zawartymi można uwzględnić przez wprowadzenie stałej dielektrycznej (ekranowanie ładunków) lub traktując każdą cząsteczkę otaczającej wody (innego rozpuszczalnika) jako dodatkowe trzy punkty (H_2O) oddziałujące z biomolekułą i między sobą. To przybliżenie nosi nazwę **pola siłowego** FF (ang. *force field*) i jest stosowane np. przy komputerowych symulacjach ruchu w ramach tzw. **klasycznej dynamiki molekularnej** (MD, *molecular dynamics*). „Kwantowość” jest tu ukryta w **[PLANSZA 7]** wartościach ładunków przypisywanych zębom atomowym, wartościach stałych siłowych i stałych Lenarda-Jonesa, które można wyznaczyć tylko przez rozwiązanie równania Schrödingera. Jeśli potraktujemy **białko jako obiekt złożony z 20 (22 jak się przyjmuje obecnie) typów aminokwasów, a kwas nukleinowy z 4 nukleotydów** potrzebne wielkości liczymy ściśle kwantowo dla małych elementów składowych i przypisujemy

odpowiednim cegiełkom składowym w polimerze. Możliwość **klasycznego podejścia** zapewnia **twierdzenie Hellmanna-Feynmana**:

"jeśli rozkład gęstości ładunku w cząsteczce jest wyznaczony metodami kwantowymi to ruch tych ładunków odbywa się pod wpływem klasycznych sił elektrostatycznych"

Wszystkie oddziaływania warunkujące **kształt przestrzenny cząsteczki** (konformacja) i stabilizujące **kompleksy molekularne** mają więc charakter **elektrostatyczny**.

Oddziaływanie odpychające elektronów walencyjnych (Ar^{-12}) wyznacza **powierzchnię van der Waalsa albo powierzchnię molekularną cząsteczki** **[PLANSZA 5], [PLANSZA 4]**, która stanowi barierę dostępu innych fragmentów tej samej cząsteczki, i innych molekuł, w tym cząsteczek rozpuszczalnika. Powierzchnia ta powstaje przez połączenie sfer van der Waalsa poszczególnych zrębów atomowych. Dla rozpuszczalnika wodnego można ją również określić jako **powierzchnię kontaktu**, tocząc wokół danej cząsteczki cząsteczkę wody w postaci sfery o promieniu 1.4 Å. Tak zdefiniowana powierzchnia odgranicza wnętrze molekuly, którego zręby atomowe nie wchodzą w kontakt z otoczeniem. Cząsteczki wody i jony przy powierzchni cząsteczki tworzą tzw. **sferę hydratacyjną - ogólnie sferę solwatacyjną**, w której cząsteczki rozpuszczalnika są sztywniej związane (mniej ruchliwe) i mają inne własności niż w jego dalszej części, np. inną gęstość. Między sferą hydratacyjną i pozostałym rozpuszczalnikiem zachodzi oczywiście wymiana jonów i cząsteczek wody.

WYKŁAD 2

Klasyczne metody badań molekularnych zajmują się **makroskopowymi układami** cząsteczek w ciele stałym, cieczy czy roztworze lub w fazie gazowej, gdzie obowiązują prawa **termodynamiki** i **mechaniki statystycznej** [**PLANSZA 8: termodynamika**]. Należy zaznaczyć, że pojawiły się ostatnio **metody badania i manipulacji pojedynczych cząsteczek** o których jeszcze wspomnę. Jednak komórka czy organizm funkcjonują jako układy makroskopowe i **stosują się do praw termodynamiki**. Procesy fizyczne i chemiczne w skali makro są regulowane nie tylko **relacjami energetycznymi** ale również stopniem uporządkowania (lub nieuporządkowania) struktur molekularnych, który jest opisany przez pojęcie **entropii**. Jest to funkcja termodynamiczna, która opisuje stopień nieuporządkowania układu, np. interesującego nas zbioru cząsteczek. Im więcej możliwych **stanów mikroskopowych P** (możliwych struktur, konformacji - stopni swobody) może przyjąć układ cząsteczkowy w danych warunkach ciśnienia p, temperatury T, objętości V, masy M i innych parametrów **makroskopowych**, które go opisują, tym większa jest entropia:

$$S = k \cdot \ln P(p, T, \dots) \quad k = 1.38045 \cdot 10^{-23} \text{ J} \cdot \text{K}^{-1} \text{ jest stałą Boltzmanna}$$

a wartość **maksymalną entropii** układ izolowany **adiabaticznie** przyjmuje w **warunkach równowagi**. **Doprowadzenie ciepła Q do układu** powoduje wzrost entropii:

$$Q = T \Delta S \quad \text{gdzie } T \text{ jest temperaturą wyrażoną w stopniach Kelvina [K]}$$

Wszelkie procesy zachodzące w makroskopowym, **otwartym** układzie o **energii U** (oznaczenie inne dla odróżnienia od energii E pojedynczej cząsteczki), który wymienia energię na sposób **ciepła Q**, wykonuje **pracę W** i wymienia **masę M** z otoczeniem opisane są przez **pierwszą i drugą zasadę termodynamiki**, które łącznie można zapisać jako bilans energetyczny:

$$\Delta U = W + Q + \Delta M \quad \text{gdzie: } W = -p\Delta V + W_{\text{nieobjętościowa}}, \quad \Delta M = \sum_i \mu_i \Delta n_i$$

gdzie μ_i oznaczają **potencjały chemiczne** poszczególnych składników układu w ilości n_i moli każdego z nich. Znaki „+” i „-” określają czy wymiana zachodzi do układu czy z układu do otoczenia, zgodnie z przyjętą konwencją.

Każdy układ dąży do stanu równowagi, który to stan musi **jednocześnie spełniać dwa kryteria:**

(1) **minimum energii wewnętrznej U**

(2) **maksimum entropii S (nieuporządkowania)**

Układy typu **żywa komórka** lub cały organizm, w których zachodzą procesy biochemiczne, stosują się do praw termodynamiki, ALE znajdują się daleko od położenia równowagi. Ten stan utrzymywany jest przez stały dopływ energii z zewnątrz w postaci substancji odżywczych (ΔM), tlenu, światła (rośliny). W końcu jednak każdy żywy organizm znajdzie się w stanie równowagi termodynamicznej z otoczeniem, która oznacza jego śmierć.

Procesy molekularne, które są przedmiotem badań biofizyki molekularnej (a więc i naszego wykładu), zachodzą wtedy, gdy pewne, opisujące je potencjały termodynamiczne, dążą do minimum. Rozpatrywane na wykładzie przypadki procesów zmian struktury chemicznej i konformacji biomolekuł będą się z reguły odbywały w warunkach stałej temperatury i ciśnienia. Procesy takie zachodzą w kierunku **osiągania minimum** przez jeden z takich potencjałów, **energię swobodną Gibbsa G**, inaczej **entalpią swobodną**:

$$G = U + pV - TS$$

$$H = U + pV - \text{entalpia}$$

Minimalizacja G obejmuje jak widać **oba kryteria**: osiągnięcie minimalnej energii (małe U) i maksymalnej entropii (duże S - znak minus). Przykładowo:

(a) jeśli mamy do czynienia z tworzeniem kompleksu molekularnego złożonego z kilku cząsteczek to:

$$\Delta G = G_{\text{kompleks}} - (G_{\text{składnik}_1} + G_{\text{składnik}_2} + \dots) < 0$$

(b) jeśli w danych warunkach otoczenia białko globularne tworzy funkcjonalną formę natywną to:

$$G_{\text{natywna}} < G_{\text{zdenaturowana}} \quad \text{inaczej} \quad \Delta G_{\text{przejścia}} = G_{\text{natywna}} - G_{\text{zdenaturowana}} < 0$$

W przypadku podziałania czynnikiem denaturującym, np. mocznikiem relacja między wartościami G dla obu form białka ulegnie odwróceniu.

Dotychczas wprowadziliśmy pojęcie energii pojedynczej cząsteczki E i energię makroskopowego układu cząsteczek U . Energia U jest sumą energii E_i (energii wewnątrzcząsteczkowych) i energii oddziaływań między cząsteczkami E_{ij} . W przybliżeniu słabych oddziaływań międzycząsteczkowych w porównaniu z wewnątrzcząsteczkowymi, U jest w sumą energii E_i wszystkich cząsteczek:

$$U = \sum_k E_k + \sum_{j < k} E_{kj} \approx \sum_k E_k = \sum_i N_i E_i$$

gdzie poszczególne cząsteczki zajmują skwantowane poziomy energii **[PLANSZA 6]** zgodnie z **rozkładem Boltzmanna** **[PLANSZA 8]**. Ilości cząsteczek N_i na poziomie E_i (wzór nie uwzględnia degeneracji poziomu) w temperaturze T układu wynosi:

$$N_i = N_0 Z^{-1} \exp(-E_i / kT) \quad N_0 - \text{liczba wszystkich cząsteczek}; \quad Z = \sum_i \exp(-E_i / kT)$$

Energia makroskopowego układu molekularnego jest więc proporcjonalna do ilości cząsteczek. Wyraża się ją zwykle w [**J/mol**] lub częściej [**kJ/mol**]. Ponadto, im wyższa temperatura tym większa energia układu ponieważ więcej cząsteczek zajmuje wyżej położone poziomy energetyczne.

Struktura przestrzenna cząsteczek (spoczywających i nierotujących dla uproszczenia) o określonej budowie chemicznej (skład atomów i układ wiązań) i strukturze przestrzennej czyli **konformacja jest uwarunkowana** przez oddziaływania między poszczególnymi „atomami” i oddziaływania z otaczającymi cząsteczkami rozpuszczalnika, w przypadku białek i kwasów nukleinowych cząsteczkami wody i jonami buforu lub cząsteczkami kwasów tłuszczowych w przypadku białek zakotwiczonych w błonach biologicznych. W sumie oddziaływania te składają się na stabilizację określonej struktury poprzez czynniki **energetyczne i entropowe**.

Ilość dopuszczalnych konformacji w cząsteczkach białek i kwasów nukleinowych jest olbrzymia, np. białko o 100 aminokwasach może przyjąć ponad 2^{100} (10^{30}) różnych konformacji. **Czynniki entropowe** sprzyjają więc strukturze w której w równowadze dynamicznej istnieje astronomiczna ilość różnych form przestrzennych biomolekuły. Jest to struktura tzw. **kłębka statystycznego** jaką mają białka zdenaturowane i pojedyncze nici kwasów nukleinowych. Aby powstała **funkcjonalna struktura (konformacja) natywna**

lub niewielka ilość funkcjonalnych konformacji ten niekorzystny czynnik entropii konformacyjnej musi być przewyższony przez **wewnątrzcząsteczkowe oddziaływania stabilizujące w biocząsteczce** z udziałem **oddziaływań makromolekuła-cząsteczki rozpuszczalnika** i oddziaływań **wewnątrz samego rozpuszczalnika**. Każde z nich charakteryzuje się relatywnie niewielką energią, co zapewnia odwracalność procesów.

Wyróżnia się następujące oddziaływania stabilizujące konformację natywną biomolekuły **[PLANSZA 9: oddziaływania stabilizujące]**:

1. **Wiązania wodorowe (ww)**. Polegają na uwspólnieniu protonu między dwoma atomami elektroujemnymi, donorem i akceptorem, np. O lub N. Niekiedy donorem może być **węgiel**, $C_{\alpha}-H\dots O=C$ w peptydach. Tworzy się w ten sposób liniowy układ trzech atomów: donor - wodór - akceptor **z wolną parą elektronową**, o nadmiarowych ładunkach minus - plus - minus. Przy silnych wiązaniach wodorowych donor i akceptor są odległe od ok. 2.7 Å do 3.1 Å, a układ trzech atomów jest zbliżony do liniowego (odstępstwa od 180° dla kąta D-H-A nie przekraczają 25°). **Energia wiązania wodorowego** jest znacznie mniejsza niż energia typowego wiązania kowalencyjnego, rzędu kilka razy **energia związana z ruchami cieplnymi atomów**, których miarą jest $RT = 2.5$ kJ/mol w temperaturze pokojowej 300 K; $R = 8.3143$ kJ/(mol·K) - stała gazowa). **Ukierunkowany charakter ww** jak w „prawdziwym” wiązaniu uzyskało podstawę w ostatnich pracach teoretycznych, których wyniki wskazują na pewien (niewielki) **charakter kowalencyjny** wiązania wodorowego.

2. **Mostki solne** powstają z **grup atomów silnie naładowanych** o przeciwnych znakach w wyniku dysocjacji protonu H^+ z jednej grupy (ładunek ujemny) i przyłączeniu protonu przez drugą (ładunek dodatni). Siła Coulomba stabilizuje parę np. dodatnio naładowana grupa boczna lizyny $CH_2CH_2CH_2CH_2NH_3^+$ i ujemnie naładowana kwasu glutaminowego $CH_2CH_2COO^-$ w białku lub lizyna i grupa fosforanowa $-O-PO_2^- -O-$ kwasu nukleinowego. **Energia mostka** jest rzędu energii wiązania wodorowego.

3. **Kontakty van der Waalsa** (nieukierunkowane). Atomy i grupy atomów w cząsteczce, nawet jeśli cała cząsteczka jest elektrycznie obojętna lub zjonizowane/uprotonowane grupy równoważą się, mają nadmiarowe ładunki dodatnie lub ujemne. Fragment lekko naładowany dodatnio indukuje zwiększony ładunek ujemny w sąsiednim fragmencie i następuje przyciąganie (słabe) przez oddziaływanie indukcyjne, a korelacja ruchu elektronów daje przyciąganie dyspersyjne van der Waalsa, wyrażane przez potencjał pola siłowego **[PLANSZA 7]**.

4. **Oddziaływania hydrofobowe.** Fragmenty cząsteczkowe źle oddziałujące z wodą (źle rozpuszczalne) wolą agregować ze sobą np. grupy alifatyczne CH₃ czy pierścienie benzenowe. Dobrym wizualnym modelem takiego oddziaływania jest agregacja kropli tłuszczu w wodzie. Mówimy wtedy o oddziaływaniu hydrofobowym, w którym decydujące znaczenie mają **efekty entropowe, związane z reorganizacją wzajemną cząsteczek wody** w obecności hydrofobowych powierzchni cząsteczek rozpuszczonych. Układ dąży do minimalizacji powierzchni kontaktu z wodą i ponieważ pojawia się więcej cząsteczek „swobodnych” wody (poza sferą hydratacyjną) w konsekwencji wzrasta entropia. Typowym przykładem znacznego wkładu oddziaływania hydrofobowego jest:

5. **oddziaływanie warstwowe (stacking) dwóch (lub więcej) pierścieni aromatycznych.** Ustawiają się one równolegle w odległości ok. 3.4 Å, czyli dwóch promieni van der Waalsa atomów w oddziałujących pierścieniach, a woda jest wyparta spomiędzy nich.

W stackingu jest również **udział elektrostatyki** w postaci sił van der Waalsa. Jeśli z pierścieniem aromatycznym oddziałuje fragment naładowany dodatnio, pierścieniowy lub nie, jak np. reszta lizyny, mówimy o tzw. **oddziaływaniu lub stackingu kation - π**.

Specyficzne, silne wiązanie biomolekuł jest podstawą współdziałania różnych cząsteczek w skomplikowanych procesach biologicznych, takich jak replikacja, transkrypcja czy translacja. Aby współdziałać, makromolekuły muszą się rozpoznać i utworzyć **funkcjonalne kompleksy molekularne** **PLANSZA 10: specyficzne kompleksy molekularne**. **Miarą siły i specyficzności wiązania** jest równowagowa **stała asocjacji kompleksu** wyrażona przez stężenia składników [składnik_i] i kompleksu [kompleks]:

$$K_{as} = \frac{[\text{kompleks}]}{([\text{składnik}_1] \cdot [\text{składnik}_2] \dots [\text{składnik}_n])}$$

Im wyższa stała asocjacji tym stabilniejszy kompleks. Niekiedy używa się odwrotności stałej asocjacji czyli stałej dysocjacji: $K_d = 1/K_{as}$. Stała asocjacji wiąże się bezpośrednio z energią swobodną Gibbsa utworzenia kompleksu:

$$\Delta G^\circ = G^\circ_{\text{kompleks}} - (G^\circ_{\text{składnik}_1} + G^\circ_{\text{składnik}_2} + \dots) = -RT \ln K_{as}$$

gdzie R jest stałą gazową; „^o” oznacza G **standardowe** odniesione do stężeń jedno-molowych, co daje jednocześnie usunięcie problemu miana pod logarytmem.

Elektrycznie obojętne makromolekuły, ze względu na niesymetryczny rozkład ładunku elektrycznego zwykle przyciągają się (siły van der Waalsa) ale tego typu oddziaływanie nie jest silne i ma charakter niespecyficzny. Stałe asocjacji nie przekraczają wtedy 10⁴ M⁻¹ (dla kompleksów dwuskładnikowych). Oddziaływanie **specyficzne**

charakteryzuje się **dopasowaniem powierzchni van der Waalsa biomolekuł i dopasowaniem ich rozkładów ładunków [PLANSZA 4]**, aby maksymalizować oddziaływania elektrostatyczne: wiązania wodorowe, mostki solne, oddziaływania van der Waalsa i oddziaływania hydrofobowe. Specyficzne kompleksy białko - kwas nukleinowy charakteryzują się stałymi asocjacji (kompleks dwuskładnikowy) dochodzącymi do $K_{as} \approx 10^{13} \text{ M}^{-1}$ co odpowiada $\Delta G \approx 80 \text{ kJ/mol}$ w 300K. Najsilnie związanym (ze znanych układów) jest kompleks (strept)awidyna - biotyna $K_{as} \approx 10^{15} \text{ M}^{-1}$ [PLANSZA 10].

Ważnym przykładem specyficznego rozpoznawania jest powstawanie struktury podwójnej helisy, oparte na tworzeniu **komplementarnych par zasad Watsona- Cricka: G : C i A : T lub A : U**. Pierwsza stabilizowana jest przez trzy, a druga przez dwa wiązania wodorowe. Pary niekomplementarne mogą się tworzyć tylko z udziałem mniejszych ilości wiązań wodorowych. W parach WC donory i akceptory protonu zasad są odpowiednio zlokalizowane przestrzennie względem siebie. Kształty zasad pozwalają na dostateczne zbliżenie donor-akceptor tak, że w efekcie powstaje stabilny układ, trzymany przez **siły elektrostatyczne, które zależą od odległości ładunków**.

Tworzeniu kompleksu towarzyszą efekty związane z rozpuszczalnikiem o charakterze entropowym (korzystne lub nie), takie jak:

- (a) **efekt polielektrolityczny:** wymiana jonów, w tym protonów między sferami hydratacyjnymi biomolekuł (lub wnętrzem molekuly) i resztą rozpuszczalnika wodnego;
- (b) **hydratacja/dehydratacja:** wymiana cząsteczek wody między sferami hydratacyjnymi oddziałujących cząsteczek i rozpuszczalnikiem.

Tworzenie kompleksu nie ma jednak charakteru statycznego ale dynamiczny. Oddziałujące biomolekuły tworzą przejściowo kompleks spotkaniowy (*encounter complex*) a następnie zmieniają swoje konformacje, często w znacznym zakresie, aby dopasować do siebie powierzchnie molekularne - **wzajemne dopasowanie konformacyjne** (*induced fit*), co zapewnia krótsze odległości między ładunkami, a konsekwencji większe siły. Warto zaznaczyć, że ok. 30% funkcjonalnych białek w organizmach eukaryotycznych w stanie swobodnym występuje w formie **całkowicie lub częściowo nieuporządkowanej** (kłębek statystyczny). Te, tzw. białka natywnie rozwinięte (IDP; *intrinsically disorderd proteins*) dopiero przy tworzeniu kompleksu molekularnego zwijają się do formy natywnej „na tarczy” partnera (*fly-casting model*) w sprzężonym procesie tworzenia kompleksu i zwijania (omawiane w części o kompleksach biomolekularnych, pkt 7).

WYKŁAD 3

Ogólny podział metod badania struktur i dynamiki biopolimerów

Ponieważ duże i skomplikowane biopolimery badane są z punktu widzenia różnych informacji, jakie chcemy uzyskać o ich strukturze i funkcjonowaniu w komórce, rozwinięto w tym celu znaczną ilość różnych metod. Na wykładzie krótko przedstawię ogólny podział metod fizycznych, tych metod które pozwalają uzyskać przede wszystkim dane o strukturach molekularnych (chemicznej, przestrzennej i dynamicznej), w tym **strukturach pierwszorzędowych** czyli sekwencji i **strukturach przestrzennych, drugo- i ew. czwartorzędowych**, z rozdzielczością do położenia poszczególnych zębów atomowych. Pod względem zakresu informacji strukturalnej, metody można zasadniczo podzielić na kilka typów **[PLANSZA 11: metody badania struktur biomolekuł]:**

(A) Metody wyznaczania **mas cząsteczkowych i separacji** cząsteczek na podstawie różnic mas:

ultrawiórowanie

elektroforeza żelowa (GE) i kapilarna (CE): **sekwencjonowanie DNA**

spektrometria masowa (MS): **sekwencjonowanie białek**

(B) Metody wyznaczania **struktur przestrzennych z rozdzielczością atomową**:

dyfrakcja promieniowania rentgenowskiego (rentgenografia)

magnetyczny rezonans jądrowy (NMR)

(C) Metody określania **kształtu** makrocząsteczek:

mikroskopia sił atomowych (AFM) i mikroskopia elektronowa (cryoEM)

(D) Metody badania **dynamiki ruchów molekularnych**:

magnetyczny rezonans jądrowy (NMR) i inne metody spektroskopowe (fluorescencja czasowo-rozdzielcza, wielowymiarowe metody IR i ES)

komputerowe symulacje dynamiki molekularnej (MD i rMD)

Ultrawiórowanie **[PLANSZA 12: ultrawiórowanie]** jest historycznie jedną z pierwszych metod określania masy cząsteczkowej w dużym przedziale wartości i rozdziału dużych biomolekuł. Makrocząsteczka w roztworze poddana szybkiemu wirowaniu z częstością ω kilkudziesięciu tysięcy obrotów na minutę (rpm) porusza się ze stałą prędkością dr/dt od osi rotora ultrawiórowki w wyniku zrównoważenia sił: siła odśrodkowa minus siła wyporu i siła tarcia. Dla cząsteczek w odległości r od osi rotora (granica: zawiesina/klarowna ciecz) można wyznaczyć prędkość ich ruchu, czyli **szybkość**

sedymencacji przez pomiar zmian stężenia c w czasie, w funkcji r (równanie Lamma), metodami optycznymi np. pomiar **absorpcji** w zakresie UV lub **fluorescencji** (prawa Lamberta-Beera). Stąd wyznacza się tzw. **profil sedymencacji**:

$$g(S^*) = (dc/dS^*) = (dc/dt)(dt/dS^*) \quad S^* = \ln(r/r_m)/\omega^2 t \quad r_m - \text{promień menisku}$$

Z położenia maksimum profilu uzyskuje się **współczynnik sedymencacji**

$$S = (dr/dt)/(\omega^2 r) \text{ wyrażony w swedbergach, } 1 S = 10^{-13} s$$

używany przez biologów jak masę cząsteczkową do określania wielkości cząsteczek i kompleksów molekularnych. **Masa cząsteczkowa** jest proporcjonalna do S :

$$M = N_A m = RTS/[D(1 - \rho v)]$$

$N_A = 6.02252 \cdot 10^{23}$ - liczba Avogadro, m - masa cząsteczki, v - objętość właściwa cząsteczki, ρ - gęstość rozpuszczalnika, D - współczynnik dyfuzji,

Technika pozwala **rozdzielać cząsteczki** różniące się masą cząsteczkową (większe poruszają się szybciej). Lepsze rozdziały uzyskuje się przez wirowania w **gradiencie gęstości** np. cukru lub chlorku cezu. Gradient zapobiega ruchom konwektywnym i daje wyraźne przesuwające się pasma dla poszczególnych cząsteczek. Do wyznaczenia masy cząsteczkowej z podanego wzoru trzeba wyznaczyć niezależnie współczynnik dyfuzji D , np. z **szerokości profilu** sedymencacji lub z niezależnego pomiaru. **Alternatywnie, można zastosować eksperyment stacjonarny**: długie wirowanie (kilkadziesiąt godzin) przy niewielkich częstościach obrotu ω (ok. 5 000 rpm) prowadzi do ustalenia gradientu stężenia makromolekuł w wyniku równowagi dyfuzji i migracji w polu siły odśrodkowej.

Zastosowanie metody ultrawirowania w „eleganckim” **eksperymentcie Meselsona-Stahla** pozwoliło udowodnić semikonserwatywny model replikacji.

Elektroforeza **[PLANSZA 13: elektroforeza]** podobnie jak ultrawirowanie pozwala oszacowywać masy cząsteczkowe i rozdzielać mieszaniny makromolekuł na podstawie różnic mas. ksperyment prowadzi się w dwóch wariantach:

- elektroforeza **żelowa** GE (*gel electrophoresis*), na warstwach żelu np. poliakrylamidu lub sefarozy, o różnym stopniu usieciowania (rozmiarach porów);
- elektroforeza **kapilarna** CE (*capillary electophoresis*) w roztworze przepływającym przez cienkie kapilary kwarcowe o średnicy 20 do 100 μm .

Cząsteczki o ładunku q poruszają się w na **żelu** pod wpływem pola elektrycznego o natężeniu E (stałe napięcia ok. 200 V) z prędkością $v = (Z \cdot E)/f$ proporcjonalną do natężenia pola i ładunku Z , a odwrotnie proporcjonalną do współczynnika tarcia f .

Białka migrują w postaci zdenaturowanej, tzn. po zniszczeniu struktury natywnej i powierzchniowym związaniu anionu siarczanu dodecyłu $\text{CH}_3(\text{CH}_2)_{11}\text{SO}_4^{-2}\text{Na}^+$, który nadaje białku duży ładunek elektryczny (**SDS PAGE**: *sodium dodecyl sulphate polyacrylamide gel electrophoresis*). Położenie białek na żelu uzyskuje się w postaci prążków przez wybarwienie lub znakowaniu przeciwciałami z grupami fluoryzującymi. Czulość metody sięga **10 nanogramów** (10^{-8} g) białka, a dokładność oznaczenia masy cząsteczkowej 2% na podstawie równoległej elektroforezy markerów białkowych o znanych masach cząsteczkowych. **Kwasy nukleinowe** mają własny duży ładunek ujemny grup fosforanowych więc migrują na żelu bez SDS. Na żelach poliakrylamidowych **fragmenty jednoniciowego DNA i RNA do ok. 400 nukleotydów dzielą się pod względem długości z dokładnością do jednego nukleotydu**, a czulość jest podobna jak w przypadku białek. W **kapilarze** rozdzielanie cząsteczek naładowanych następuje w wyniku ich migracji w buforze poruszającym się w cienkiej kapilarze kwarcowej długości kiludziesięciu cm między dwoma naczyniami z buforem, pod wpływem pola E o różnicy potencjałów do **30 kV**. **Różnicowanie przepływu cząsteczek** (odpowiednik usieciowania żelu) następuje w wyniku jonizacji grup $=\text{SiOH} \rightarrow =\text{SiO}^-$ na ściankach kapilary (pH buforu > 3) i utworzeniu dyfuzyjnej warstwy podwójna kationów, która daje **napięcie ścinające** między stałą i ruchomą warstwą jonów. **Czulość** zależy od typu detektora: spektrometr masowy (najczulszy), fluorescencja, absorpcja.

Sekwencjonowanie fragmentów DNA do 400n wykonuje się standardowo **metodą Saengera** z wykorzystaniem **GE** lub **CE** w automatycznych sekwenatorach (patrz dalej).

W **spektrometrii masowej (MS)** [**PLANSZA 14: spektrometria masowa**] cząsteczki poddaje się **desorpcji** do zjonizowanego gazu o ładunku Z, przyspiesza w polu elektrycznym o potencjale V i po zogniskowaniu wiązki poddaje detekcji. W detektorze magnetycznym (już rzadko stosowanym) siła Lorentza zakrzywia tor ruchu jonów, a **promień krzywizny r** zależy od stosunku masy jonu m do ładunku Z: $m/Z = (H^2/2V) r$. W technice detekcji TOF (*time-of-flight*) mierzy się **czas przelotu jonu t** na drodze **d** (e - ładunek elementarny): $m/Z = (2V/ed^2)t^2$. W detektorze **kwadrupolowym QIT** (*quadruple ion trap*) pułapkuje się jony w odpowiednio ukształtowanym (zmiennym) polu elektrycznym. Wykres **intensywności sygnału jonów mierzonych w funkcji m/Z daje widmo masowe**, co przy znanym ładunku pozwala określić **masę jonu**. Technika jest bardzo czuła: wymaga jedynie nanogramów substancji i dla mniejszych cząsteczek osiąga

dokładność wyznaczania masy poniżej 5 ppm (0.0005%). Obserwuje się różne piki od cząsteczek lub ich fragmentów o różnym składzie izotopowym.

Przy jonizacji cząsteczka ulega fragmentacji i dopiero techniki desorpcji MALDI (*matrix-assisted laser desorption ionization*) i ESI (*electrospray ionization*) dają tzw. piki molekularne. W pierwszej z nich desorpcja zachodzi pod wpływem promieniowania lasera w silnie absorbującym roztworze wodno/organicznym małych, cząsteczek organicznych (np. *3,5-dimethoxy-4-hydroxycinnamic acid*), tzw. matrycy, rekrystalizujących z utkwionymi badanymi cząsteczkami. W drugiej zachodzi rozpylanie roztworu na drobnokropelkowy spray w polu elektrycznym pod wpływem dużego ładunku powierzchniowego na kroplach i odparowanie w przeciwbieżnym strumieniu gazu. Widmo ESI-MS składa się z serii jonów molekularnym (bez fragmentacji), różniących się o ładunek jednostkowy. Masę cząsteczkową m określa się z dużą dokładnością rzędu kilku Da dla białek ok. 100 kDa.

Kontrolowana fragmentacja krótkich peptydów ok. 20 AA długości w systemie tandemu MS/MS jest podstawą wydajnego **sekwencjonowania białek** (patrz dalej)

Magnetyczny rezonans jądrowy (NMR - *nuclear magnetic resonance*) **[PLANSZA 15: magnetyczny rezonans jądrowy]** jest metodą spektroskopową, w której obserwuje się przejścia pomiędzy poziomami energii jąder danego typu w cząsteczce np. ^1H , ^{13}C , ^{15}N , ^{31}P w silnym, zewnętrznym, stałym polu magnetycznym o indukcji \mathbf{B} od 2.5 do 22 T [tesli], co odpowiada promieniowaniu fal radiowych, ν od ok. 10 do 900 MHz. Postać widma zależy w decydującym stopniu od **stanu skupienia** układu cząsteczkowego. Widmo w stanie stałym to jedna szeroka linia rzędu kilka - kilkadziesiąt tysięcy Hz (kHz). W cieczech i roztworach (a także gazach) obserwuje się szereg wąskich linii o szerokościach od dziesiątych części Hz dla małych cząsteczek organicznych do kilkudziesięciu Hz dla biopolimerów. **Widma wysokiej zdolności rozdzielczej (*high resolution*) w ciele stałym** uzyskuje się przez wirowanie próbki pod „kątem magicznym” $\theta = 54,7$ deg do kierunku indukcji \mathbf{B} : $\langle 3\cos^2\theta - 1 \rangle = 0$.

(a) Częstości rezonansowe poszczególnych jąder w cząsteczce różnią się (kilku do kilkudziesięciu kHz) wskutek przesłaniania (ekranowania) zewnętrznego pola magnetycznego przez elektrony **[PLANSZA 16: eksperyment 1D NMR]** - **przesunięcie chemiczne δ** w jednostkach ppm (part per million) względem wybranego wzorca.

(b) Sygnały rezonansowe mogą się rozszczepiać wskutek tzw. **sprzężeń skalarnych** $J(k,l)$ między jądrami, takimi samymi i innego typu. poprzez elektrony wiązań.

Na podstawie widm z **wielowymiarowych eksperymentów NMR** można zidentyfikować poszczególne jądra w makrocząsteczce czyli wykonać tzw. **przyporządkowanie sygnałów widma NMR jądrom w cząsteczce przy znanej strukturze chemicznej** (strukturze pierwszorzędowej):

- widma 2D ^1H NMR [**PLANSZA 17: eksperyment 2D ^1H NMR**] dla cząsteczek **do ok. 10 kDa**;

- widma 3D i 4D NMR jąder ^1H , ^{13}C , ^{15}N NMR cząsteczek wzbogaconych podwójnie (^{13}C , ^{15}N) i/lub potrójnie (^{13}C , ^{15}N , ^2H) [**PLANSZA 18: eksperyment 3D NMR**] dla cząsteczek **do ok. 40 kDa** - wykorzystanie separacji sygnałów wzdłuż osi heterojądra (^{13}C lub ^{15}N) [**PLANSZA 19: rozdzielenie sygnałów cząsteczek znakowanych**].

- widma 3D i 4D NMR jąder ^1H , ^{13}C , ^{15}N NMR cząsteczek wzbogaconych podwójnie/potrójnie z implementacją sekwencji TROSY (*Transverse Relaxation Optimized Spectroscopy*) i CRINEPT (*Cross-correlated Relaxation-enhanced Polarization Transfer*) w spektrometrach o indukcji B odpowiadającej częstości rezonansowej $\omega(^1\text{H}) \geq 900$ MHz, „teoretycznie” **bez ograniczenia co do wielkości cząsteczek**.

Wyznaczanie struktur przestrzennych białek i kwasów nukleinowych w roztworze metodami NMR przebiega według następującego schematu [**PLANSZA 20: wyznaczenie parametrów strukturalnych i dynamicznych biopolimerów**]:

(A) Wykonanie **przyporządkowania sygnałów NMR** poszczególnym jądrom cząsteczki o **znanej sekwencji aminokwasów lub nukleotydów**: przypisanie sygnałów poszczególnym typom monomerów (aminokwasy, nukleotydy) i przyporządkowanie sekwencyjne wzdłuż łańcucha, na podstawie znanej sekwencji, oddziaływaniom skalarnym i ew. dipolowym jąder na sąsiednich monomerach (nie będę omawiał szczegółów).

(B) Wyznaczenie **parametrów opisujących strukturę przestrzenną cząsteczki**:

- **odległości jądro - jądro** z efektu Overhausera; intensywność oddziaływania dipolowo - dipolowego NOE_{ij} zależy jak r_{ij}^{-6} od odległości między tymi jądrami;

- **kąty dwuścienne** θ_{ij} wokół wiązań z wartości stałych sprzężenia skalarnego dwóch jąder odległych o trzy wiązania (relacja Karplusa): $^3J(i,j) = A\cos^2\theta_{ij} + B\cos\theta_{ij} + C$, A, B, C - stałe zależne od typu fragmentu chemicznego i rodzaju sprzęgających się jąder .

- **orientacja fragmentów w strukturze globalnej** cząsteczek w roztworze częściowo uporządkowanych przez lipidowe micelle, na podstawie np. resztkowych sprzężeń dipolarnych (RDC, *residual dipolar coupling*)

(C) Przejście od odległości, kątów i innych informacji przestrzennych na podstawie struktury chemicznej (długości wiązań, kąty płaskie) do położenia zębów atomowych cząsteczki czyli **współrzędnych atomowych** (x_i, y_i, z_i) $i = 1, 2, \dots, N$ atomów:

- metody geometryczne tzw. geometria odległościowa (*distance geometry*);

- dynamika molekularna z wiązaniami doświadczalnymi NMR **rMD**.

(D) Weryfikacja struktury.

Przebieg łańcucha białka BPTI uzyskany metodą geometrii odległości i białka eIF4E metoda rMD, w postaci **zbioru nałożonych struktur** zgodnych z parametrami (wiązaniami) NMR prezentuje **[PLANSZA 21: wyznaczenie struktur przestrzennych w roztworze metodą wielowymiarowego NMR]**. Przedstawienie w formie nałożonych struktur wynika z **niepewności wartości** doświadczalnych parametrów (długości i kątów), co odpowiada **dynamicznemu charakterowi** cząsteczki.

Rentgenografia pozwala uzyskać struktury przestrzenne cząsteczek w monokryształach przez rozpraszanie promieniowania X. Jest najszerszej stosowaną metodą dla białek, kwasów nukleinowych i ich kompleksów. Kryształ molekularny **[PLANSZA 22: wyznaczenie struktury metodą dyfrakcji rtg]** tworzy regularną strukturę makroskopową wynikającą z symetrii mikroskopowej jednej lub kilku cząsteczek w komórce elementarnej. Wszystkie elementy symetrii można zebrać w tzw. przestrzenną grupę symetrii i jednoznacznie scharakteryzować przez podanie symbolu grupy, jednej z 230 możliwych. Promieniowanie rentgenowskie o **długości fali λ rzędu odległości międzyatomowych** (1 Å) ulega **dyfrakcji w wyniku rozproszenia na elektronach zębów atomowych lub inaczej odbiciu od płaszczyzn sieciowych (hkl)**. W pewnych kierunkach, określonych przez symetrię mikroskopową następuje interferencja wzmacniająca i pojawia się prążek dyfrakcyjny zgodnie z prawem Bragga lub równoważnymi warunkami Lauego:

$$2d\sin\theta = n\lambda, \text{ gdzie } n \text{ jest liczba całkowita } 1, 2, 3 \text{ itd., a } d \text{ odstępem płaszczyzn sieciowych w których leżą centra rozpraszające}$$

(A) Układ prążków dyfrakcyjnych jest uwarunkowany przez grupę symetrii kryształu i wymiary komórki elementarnej (wartości $a, b, c, \alpha, \beta, \gamma$);

(B) Intensywności prążków dyfrakcyjnych są uwarunkowane przez rodzaje atomów i ich położenie w komórkach elementarnych.

Mierząc intensywności poszczególnych prążków w dyfraktogramie można wyznaczyć mapy gęstości elektronowej $\rho(x, y, z)$ [PLANSZA 23: struktura RNA z mapy gęstości elektronowej] w komórce elementarnej, a stąd wstawić w odpowiednich miejscach poszczególne atomy cząsteczki. Podstawowym problemem w uzyskaniu mapy gęstości elektronowej z dyfraktogramu jest tzw. **problem fazowy** [PLANSZA 24: **problem fazowy**]. Wynika on stąd, że czynnik strukturalny, zależny od położenia atomów i ich siły rozpraszania, i który umożliwia wyliczenie intensywności prążków dyfrakcyjnych jest liczbą zespoloną $F = a + bi = \sqrt{a^2 + b^2} (\cos\varphi + i\sin\varphi)$, $i = \sqrt{-1}$. Mając strukturę można łatwo policzyć intensywności jako pierwiastek z wartości bezwzględnej $I = |F| = \sqrt{a^2 + b^2}$, ale aby wyznaczyć czynnik strukturalny z intensywności, a stąd $\rho(x, y, z)$, trzeba niezależnie wyznaczyć fazę $\varphi = \arctg(a/b)$ dla każdego prążka (hkl). **Problem fazowy, kluczowy dla rentgenografii** rozwiązuje się za pomocą jednej z trzech metod::

(1) **podstawienie molekularne** (*molecular replacement*): czyli zastosowanie znanych faz podobnych cząsteczek, które podobnie wykryły, a których struktura przestrzenna w kryształach została już wyznaczona;

(2) **podstawienie izomorficzne MIR** (*multiple isomorphous replacement*): rejestracja i analiza dyfraktogramów od jednakowych (izomorficznych) kryształów: jeden zwykły, drugi z ciężkim atomem lub jonem w każdej komórce elementarnej, który bardzo silnie rozprasza promienie X i stąd można określić jego położenie w komórce jako punkt odniesienia;

(3) **rozpraszanie anomalne** (MAD *multiwave anomalous diffraction*) w kilku długościach fali λ przy użyciu promieniowania synchrotronowego; zastosowanie w białkach, które zamiast metioniny mają **selenometioninę** (ciężki selen silnie rozprasza niż siarka).

Procedura wyznaczania struktury w kryształach jest następująca:

(A) **wyhodowanie kryształu**, wstawienie go do dyfraktometru lub na wyjściu wiązki z akceleratora, zwykle szybkie zamrożenie (witryfikacja) w strumieniu ciekłego azotu;

(B) uzyskanie **dyfraktogramu** w detektorze np. *imaging plate*, w której pod wpływem promieniowania X powstają barwne centra;

(C) wyznaczenie **struktury wstępnej** na podstawie intensywności prążków i rozwiązanego problemu fazowego;

(D) iteracyjne **uzgadnianie dyfraktogramu doświadczalnego i liczonego teoretycznie** aż do uzyskania wystarczającej zgodności opisanej przez tzw. **czynnik R** - powinien być mniejszy od 25% (procentowe odstępstwo intensywności mierzonych z dyfraktogramu i liczonych teoretycznie dla danej struktury). Czynnik ten zależy oczywiście od jakości kryształu (regularność, brak defektów), opisanej przez tzw. **zdolność rozdzielczą** w Å: zdolność rozdzielcza poniżej 2 Å zapewnia określenie położenia atomów **wszystkich ciężkich atomów** cząsteczki (oprócz wodorów) z dokładnością do tysięcznych części Å.

Kluczowym zagadnieniem dla rentgenografii jest **relacja między strukturą przestrzenną biomolekuły w kryształach i jej strukturą natywną w roztworze wodnym**, gdzie funkcjonuje. Kryształy białek i kwasów nukleinowych zawierają na szczęście znaczne ilości nieustrukturyowanej wody, nawet do 60% masy kryształu.

Bezpośrednie określanie struktur cząsteczkowych z rozdzielczością uzyskiwaną metodami rentgenografii czy NMR (poniżej 1 Å) nie jest możliwe ani w mikroskopii optycznej ani elektronowej. W przypadku najlepszych mikroskopów elektronowych skanujących pojedyncze cząsteczki w niskich temperaturach cryoEM rozdzielczość jest rzędu 1 nm (5 ÷ 10 Å). Podobne efekty jeśli chodzi o wyznaczanie kształtu makromolekuł uzyskano za pomocą **mikroskopów siły atomowej AFM** (*atomic force microscopy*) **[PLANSZA 25: mikroskopia sił atomowych AFM]**. Pojedynczą makrocząsteczkę w kropli roztworu na płaskiej powierzchni skanuje się bardzo ostrym rylcem (ostrzem). Siły molekularne van der Waalsa między ostrzem i podłożem rzędu pikonewtonów (1 pN = 10^{-12} N) utrzymuje się na stałym poziomie podczas skanowania dzięki ujemnemu sprzężeniu zwrotnemu: wychylenie wspornika po natrafieniu na zakrzywienie powierzchni jest rejestrowane jako zmiana siły i kompensowane odpowiednim przesunięciem rylca tak aby odległość ostrze - powierzchnia skanowania pozostawała stała. Służy do tego celu najczęściej wiązka laserowa odbita od wspornika. W ten sposób rejestrowany jest **obraz powierzchni a więc kształt umieszczonej na niej cząsteczki** białka, kwasu nukleinowego czy kompleksu molekularnego **[PLANSZA 26: struktura chromatyny AFM]**.

Rejestracja obrazów w czasie pozwala śledzić zmiany zachodzące w cząsteczkach w skali długoczasowej. Zaczepienie rylcem cząsteczki przytwierdzonej do podłoża **[PLANSZA 25]** pozwala poruszać ją i rozciągać oraz mierzyć związane z tym siły molekularne od kilku do kilkuset pN przy odpowiednim wyskalowaniu wychyleń

wspornika jak wagi sprężynowej - **manipulacje pojedynczą cząsteczką**. Podobnie w **szczypcach optycznych** lub **magnetycznych** przyczepia się do makrocząsteczki paciorka dielektryczny, np. polistyrenowy (pierwsza metoda) lub ferromagnetyczny (druga metoda) średnicy ok. 1 - 3 μm . Wiązka laserowa powoduje utrzymywanie paciorka w środku wiązki i przesuwanie (przekaz pędu fotonu - „pułapka optyczna”). Pole magnetyczne umożliwia przesuwanie i dodatkowo rotację magnetyka wraz z przyczepioną cząsteczką.

Dynamikę ruchów molekularnych cząsteczek biologicznych opisuje się na gruncie **modelu dyfuzyjnego**; cząsteczki wykonują **przypadkowe ruchy**:

(a) globalne translacje i obroty całej cząsteczki,

(b) wewnętrzne ruchy lokalne: obroty wokół wiązań (zmiany konformacyjne); na gruncie modelu Lipari-Szabo **[PLANSZA 27: wyznaczenie parametrów dynamicznych biopolimerów]**, gdzie „giętkość” opisana przez parametr $S^2 = 1$ dla fragmentów sztywnych, $S^2 = 0$ dla całkowicie swobodnych ruchów wewnętrznych.

Mierzonymi doświadczalnie parametrami, które opisują takie ruchy są **średnie czasy życia** τ_c dla ruchów globalnych i τ_{int} dla danej formy strukturalnej (konformacji), gdzie $\tau = 1/k$ [s], k - średnie szybkości obu procesów dynamicznych. Drugim parametrem są **średnie amplitudy ruchów**, wyrażone w Å (zmiany odległości), stopniach (zmiany kątowe) lub procentach zmian. Metody pomiarowe to:

- wyznaczania **czasów relaksacji** T_1 i T_2 i efekt NOE $^1\text{H} - ^{15}\text{N}$ technikami NMR;
- pomiary czasowe zmian **anizotropii fluorescencji** po wzbudzeniu elektronowym cząsteczki impulsem (laserowym) światła spolaryzowanego liniowo;
- **wielowymiarowe techniki IR** (*infra red*) i ES (*electron spectroscopy*) – bardzo szybkie procesy.

Pełną informację o ruchach molekularnych, nie tylko w postaci podanych parametrów, ale rejestracji zmian położenia poszczególnych atomów uzyskuje się za pomocą komputerowych symulacji **dynamiki molekularnej MD** (*molecular dynamics*). Cząsteczkę przybliża się jako zbiór **klasycznych punktów (sfer) materialnych** na które działają siły ze strony innych atomów wyznaczone z pola siłowego FF **[PLANSZA 7]**. Ruch atomów wynika z rozwiązywania numerycznego zwykłych równań Newtona w krótkich skokach czasowych ok. 1 femtosekundy ($1 \text{ fs} = 10^{-15}\text{s}$) przez okres taki na jaki pozwalają nowoczesne superkomputery tzn. nanosekund ($1 \text{ ns} = 10^{-9}\text{s}$), a w zmodyfikowanych wersjach, np. brownowska MD do mikrosekund ($1 \mu\text{s} = 10^{-6}\text{s}$)

Dynamikę molekularną stosuje się często do uzyskiwania struktur przestrzennych cząsteczek na podstawie mierzonych parametrów NMR [\[PLANSZA 21\]](#). ponieważ układ dąży do konformacji o **minimalnej energii (swobodnej)**. Aby to zrobić w czasie krótszym niż dostępne czasy symulacji wprowadza się dodatkowe pseudosiły "ściąające" cząsteczkę do konformacji opisanej znanymi z NMR parametrami przestrzennymi: odległościami jądro-jądro (NOE) i kątami dwuściennymi (sprzężenia skalarne) - wersja dynamiki molekularnej z więzami **rMD** (*restrained molecular dynamics*).

WYKŁAD 4

Kwasy nukleinowe DNA i RNA (1)

Kwasy nukleinowe są liniowymi biopolimerami, zbudowanymi z zasadniczo czterech monomerów, **nukleotydów**. „Zasadniczo” oznacza że mogą zajść określone modyfikacje chemiczne nukleotydów. Dwa podstawowe typy kwasów nukleinowych, różniące się budową chemiczną nukleotydów i funkcjonowaniem w komórkach to kwas deoksyrybonukleinowy DNA i kwas rybonukleinowy RNA.

Duże cząsteczki **kwasu DNA** [**PLANSZA 28: kwas deoksyrybonukleinowy DNA**] pełnią rolę **nośnika informacji genetycznej**. **Podwójna helisa DNA** w chromosomach liczy $\sim 10^8$ bp (par zasad) i ma masę cząsteczkową $\sim 10^{10} \div 10^{11}$ Da (długość do kilku centymetrów po rozwinięciu). Komplementarna para zasad, ściślej komplementarna para nukleotydów A:T i G:C w helisie ma średnią masę cząsteczkową 660 Da. DNA składa się z **2'-deoksyrybonukleotydów**: dwóch zawierających zasady purynowe dGMP, dAMP, i dwóch pirymidynowych, dCMP i TMP, w których cukrem jest β -D-2'-deoxyryboza [**PLANSZA 29: tautomeria oxo-amino i protonacja/dysocjacja zasad**]. Numeracja w pierścieniach zasad i w części cukrowej (cyfry primowane) jest jednakowa dla deoxy- i rybonukleotydów [**PLANSZA 28**]. Konfiguracje wokół atomów węgla C(1') i C(4'), podstawionych czterema różnymi podstawnikami, opisują odpowiednio **symbole β i D** [**PLANSZA 29**]. Takie cząsteczki nie pokrywają się ze swoim odbiciem zwierciadlanym i występują w postaci izomerów optycznych, skręcających płaszczyznę światła spolaryzowanego liniowo w przeciwnych kierunkach. **Zasada azotowa**: guanina, adenina, cytozyna i tymina, plus **cukier** czyli **nukleozyd**: 2'-deoxyguanozyna, 2'-deoxyadenozyna, 2'-deoxycytydina i tymidyna, łączą się ze sobą poprzez grupy fosforanowe w **nić polinukleotydową** o wyróżnionym kierunku od 5'-końca do 3'-końca [**PLANSZA 28**]. **Sekwencja** nukleotydów (zasad) w kierunku 5' \rightarrow 3' określa **strukturę pierwszorzędową** kwasu DNA. W warunkach pH komórkowego **grupy fosforanowe** w nici polinukleotydowej są zjonizowane i obdarzone ładunkiem ujemnym (pK ok.1,5), zneutralizowane przez kationy roztworu, np. K^+ , Na^+ czy Mg^{2+} . Swobodne nukleotydy mogą w tych warunkach oddysocjować dwa protony; przy czym pK drugiej jonizacji jest ok. 6.7. Zasady azotowe są **chromoforami** tzn. absorbują promieniowanie elektromagnetyczne dając trzy pasma typu $\pi^* \leftarrow \pi$ i $\pi^* \leftarrow n$ w bliskim nadfiolecie [**PLANSZA 30: absorpcja uv nukleotydów**] o współczynnikach ekstynkcji molowej

$\epsilon \sim 10^4$ [$\text{l}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$]. Nałożenie pasm wszystkich zasad prowadzi do absorpcji DNA w zakresie 180 – 220 nm i 240 – 280 nm z **charakterystycznym maksimum przy 260 nm**.

Kwasy dezoksyrybonukleinowe są syntetyzowane w fazie S cyklu komórkowego w procesie **replikacji** [**PLANSZA 31: replikacja DNA**], w którym na dwóch niciach pierwotnych syntetyzowane są przez polimerazy DNA nici potomne na zasadzie komplementarności (potwierdzone przez eksperyment Meselsona-Stahla ultrawiwrowania DNA bakterii hodowanych na pożywce wzbogaconej N-15). Naprzeciw dA w matrycy włączane jest T, naprzeciw dG włączane jest dC, naprzeciw T włączane jest dA a naprzeciw dC włączane jest dG [**PLANSZA 28**]. Niektóre wirusy syntetyzują DNA na matrycy RNA w procesie **odwrotnej transkrypcji**.

Kwas rybonukleinowy RNA [**PLANSZA 32: kwas rybonukleinowy RNA**] pełni szereg funkcji zależnie od struktury:

(a) **mRNA** - informacyjny (*messenger*) RNA jest roboczą matrycą do syntezy łańcucha białkowego w procesie translacji; **długość $\sim 1000\text{n} \div 2000\text{n}$** (dla średniego białka globularnego), zależnie od kodowanego łańcucha peptydowego i obszarów niekodujących.

(b) **rRNA** - rybosomalny RNA jest wraz z białkami składnikiem struktury rybosomu; **długości $\sim 100\text{n} \div 4000\text{n}$** (kilka rodzajów)

(c) **tRNA** - transferowy RNA łączy się z właściwym aminokwasem i wprowadza go na rybosom, gdzie zachodzi włączanie do powstającej nici peptydowej; **długość $75\text{n} \div 90\text{n}$** ;

(d) **snRNA** - mały jądrowy (*small nuclear*) i **snoRNA** - mały jąderkowy (*small nucleolar*) RNA jest składnikiem splicesome'u w procesie składania mRNA w jądrze komórkowym; **długości od $\sim 100\text{n} \div 200\text{n}$** (kilka rodzajów)

(e) **miRNA** - mikro RNA pełni różnorodne funkcje regulacyjne w procesach komórkowych, **długość $\sim 20\text{n}$** ;

(f) **RNA wirusów** - materiał genetyczny niektórych wirusów (HIV); **długość $\sim 10^4\text{n}$**
[**PLANSZA 33: struktura wirusa HIV**]

RNA składa się z rybonukleotydów: dwóch purynowych GMP, AMP, i dwóch pirymidynowych, CMP i UMP - **brak grupy metylowej w pozycji 5 uracylu** w porównaniu z tyminą; cukrem jest **β -D-ryboza** - zawiera **2'-OH** w przeciwieństwie do 2'-deoxyrybozy [**PLANSZA 29**]. Rybonukleozydy: guanozyna, adenozyzna, cytydyna i urydyna łączą się w nić polinukleotydową poprzez grupy fosforanowe. Sekwencja nukleotydów (zasad) w kierunku $5' \rightarrow 3'$ wyznacza strukturę pierwszorzędową RNA. Jonizacja grup fosforanowych jest podobna jak w przypadku DNA.

Kwasy rybonukleinowe powstają w procesie **transkrypcji** [PLANSZA 34: **transkrypcja**]: polimerazy RNA syntetyzują nić polinukleotydową na matrycy jednej nici DNA, na zasadzie komplementarności zasad: dA:U, T:A, dG:C i dC:G. „*Primary transcripts*” RNA są zwykle modyfikowane do funkcjonalnych RNA, w szczególności **eukariotyczne mRNA** ulega **procesowi składania** czyli *splicingu* [PLANSZA 35: **składanie mRNA w eukariotach**] do właściwej matrycy do syntezy białek oraz **dojrzewiania**, czyli dołączania struktury kapu na 5' końcu i łańcucha poli(A) na 3' końcu [PLANSZA 36: **biosynteza białek w rybosomach**]. RNA wirusów jest syntetyzowane przez RNA-*directed* polimerazy RNA (RNA replikaza lub syntetaza). W przeciwieństwie do DNA cząsteczki RNA ulegają **degradacji** przez rybonukleazy (po wykonaniu „zadań” w komórce) hydrolizujące wiązanie fosfodwuestrowe lub kompleksy enzymatyczne, o ile nie są odpowiednio zabezpieczane, np. przez kap na 5' końcu eukariotycznego mRNA.

Badania krystalograficzne i spektroskopowe pokazują, że **plaskie** zasady azotowe w kryształach i roztworze występują w **dominujących formach tautomerycznych amino i oxo (keto)** [PLANSZA 29]; czułość bezpośrednich metod doświadczalnych sięga 99% czyli 10^{-2} . Zmiana ustawienia protonów i reorganizacja wiązań podwójnych [PLANSZA 37: **tautomeria enolo-imino i mutagenesa**] stwarza możliwość powstawania tzw. **mniejszościowych tautomerów imino i enolo (hydroxy)**. Rzadkie tautomery były postulowane przez Watsona i Cricka (teoria rozwinięta przez *Topal i Fresco: Nature 263, 285, 1976*) jako czynnik inicjujący mutacje punktowe, tranzycje i transwersje [PLANSZA 38: **typy mutacji punktowych**], wskutek niekomplementarnego parowania zasad. Wiele nukleozydów z **modyfikowanymi chemicznie zasadami** np. izoguanozyna (zamiana grup NH_2 i $=\text{O}$ przy C2/C6) ma charakter promutagenny wskutek równowagi tautomerycznej dwóch form o porównywalnej zawartości (populacji), z których każda może parować komplementarnie z inną zasadą. Potwierdzenie hipotezy W.C. wymaga aby rzadkie formy tautomerycznie występowały z częstością większą lub równą częstości mutacji punktowych, tranzycji i transwersji, nieskorygowanych przez mechanizmy naprawcze, czyli ok. 10^{-4} na nowotworzoną parę. Częstości mutacji spadają do 10^{-8} - 10^{-11} w wyniku działania **systemów naprawy DNA**. Zagadnienie tautomerii zasad zwykłych i modyfikowanych stało się jednym z centralnych problemów badań doświadczalnych i obliczeń kwantowomechanicznych w biofizyce. Problem okazał się **złożony** m. in. ze względu na silny wpływ rozpuszczalnika na równowagę tautomeryczną (zmiana stałej

równowagi o 2 rzędy wielkości) oraz przesuwanie tej równowagi w kompleksach zasad. Mimo, że odkryto ok. 50% udział rzadkiej formy enolowej w guanozynie i ok. 10% formy imino w cytydynie [PLANSZA 37] techniką spektroskopii IR w niskotemperaturowych matrycach gazowych, **hipoteza inicjacji mutacji punktowych przez rzadkie formy tautomeryczne pozostaje nierozstrzygnięta**: dla adenozyiny i urydyny (tymidyny) szacuje się udział rzadkich tautomerów poniżej 10^{-8} .

Konkurencyjną hipotezą (*Sowers et al. Mutation Research 177, 201, 1987*) jest inicjacja mutacji punktowych przez błędne parowanie z udziałem **zasad w formach uprotonowanych i zdysocjowanych** [PLANSZA 29]. Stwierdzono, że adenina naprzeciw cytozyny w syntetycznym, helikalnym 12-merze DNA (*Hunter et al. Nature 320, 552, 1986*) [PLANSZA 39: **parowanie C:A⁺ helisie w 12-bp DNA**] w kryształach jest uprotonowana. Wartości pK naturalnych zasad różnią się o ok. 3 jednostki od pH komórkowego co pozwala szacować prawdopodobieństwo wystąpienia tych form na 10^{-3} .

Eksperymenty z włączaniem przez polimerazę I DNA trójfosforanu deoxyrybozydu z **difluorotoluenem** jako zasadą **zamiast tyminy** [PLANSZA 40: **parowanie adenina : difluorotoluen**] naprzeciw adeniny (*Kool et al. Biopolymers 48, 3, 1998*) stały się podstawą stwierdzenia, że błędne parowanie nie wymaga szczególnych form a **komplementarność kształtu** i steryczne dopasowanie zasad odgrywa nie mniejszą rolę niż komplementarność wynikająca z wiązań wodorowych. Parowanie o mniejszej liczbie wiązań wodorowych typu **wobble G:T** [PLANSZA 41: **parowanie G:T typu „wobble”**], błędne w DNA, jest "normalne" jako parowanie G:U kodon mRNA - antykodon tRNA.

Nukleotyd jest cząsteczką **konformacyjnie "giętką"** tzn. ze względu na możliwość rotacji wokół wiązań pojedynczych występuje w roztworze w **równowadze dynamicznej** szeregu konformacji [PLANSZA 42: **równowaga konformacyjna w nukleotydach**], które w krótkim **czasie rzędu nanosekund** (ns) przechodzą jedna w drugą. W makroskopowym obrazie każda z konformacji jest obsadzona przez pewien ułamek wszystkich cząsteczek czyli tzw. **populację**. Każda konformacja jest scharakteryzowana przez wartość kąta dwuściennego między wybranymi podstawnikami, przyłączonymi do wiązania wokół którego zachodzi obrót. Według nomenklatury oznaczeń komisji IUPAC-IUB [PLANSZA 43: **nomenklatura IUPAC-IUB konformerów**] poszczególnym zakresom kąta dwuściennego odpowiada jednoznaczna nazwa. Podstawniki mogą być ustawione transoidalnie względem siebie: **anti** lub **antiperiplanar** (ap), nakładać się cisoidalnie w rzucie Newmana: **syn** lub **synperiplanar**

(sp) lub ustawiać w pozycji naprzemiennej: **synclinal**, z dodatnią/ujemną wartością kąta ($\pm sc$) - dawne określenie ustawienia **gauche**. **Swobodne nukleozydy** w roztworze preferują określone typy konformacji **[PLANSZA 42]**.

(a) **Zasada** może wykonywać obrót wokół wiązania glikozydowego z cukrem i przyjmować dwa typy ustawień: **syn** i **anti**, opisane przez wartości kąta χ O(4')-C(1')-N(1)-C(2) (pirymidyny) i O(4')-C(1')-N(9)-C(4) (puryny) **[PLANSZA 45]**. Zakres syn odpowiada $-90^\circ < \chi < +90^\circ$ a zakres anti $+90^\circ < \chi < 270^\circ$;

(b) **Pierścień cukru** nie jest płaski i przyjmuje konformacje, opisane przez 5 kątów dwuściennych wewnątrz pierścienia; Cukier może przyjmować dwa podstawowe formy "skrajne" **[PLANSZA 44: pseudorotacja pierścieni pięciocłonowych]**: kopertową **E**, w której jeden atom pierścienia x jest wychylony w górę (endo) ${}^x E$ lub w dół (exo) ${}_x E$ z płaszczyzny czterech pozostałych i **skręconą T**, w której dwa sąsiednie atomy x i y są maksymalnie wychylone z płaszczyzny w przeciwne strony ${}_x y T$ oraz wszystkie formy pośrednie między E i T (mniejsze wychylenia sąsiednich atomów). W **modelu pseudorotacji** każda z tych konformacji może być opisana przez dwa parametry: **kąt pseudorotacji P** od 0 do 360° i **maksymalne pofałdowanie pierścienia** ν_{max} , który zwykle zawiera się w granicach 35° do 45° (kątowne wychylenie maksymalne w kopercie). **Tylko dwa parametry jednoznacznie wyznaczają pięć kątów dwuściennych w pierścieniu (!)**. Konformacje przechodzą jedna w drugą w sposób ciągły i cały zakres P obejmuje wszystkie możliwe przypadki. W cyklopentanie pseudorotacja jest rzeczywistym ruchem wewnętrznym cząsteczki, nie wymagającym przechodzenia przez bariery energetyczne. W pierścieniach podstawionych takich jak cukrowe dopuszczalne są tylko dwie konformacje, z **zakresu N** (*north* na kole pseudorotacji) z P ok. 0° do 36° (${}^3_2 T$ do ${}^3_4 T$, czyli wokół C(3')endo i z **zakresu S** (*south*) z P ok. 135° do 190° (${}^2_1 T$ do ${}^2_3 T$, wokół C(2')endo) **[PLANSZA 42]**.

(c) **Grupa egzocykliczna** -C(5')H₂O obraca się wokół wiązania C(4') - C(5') i może przyjmować trzy klasyczne formy naprzemianległe +sc, -sc i ap, z kątami dwuściennymi γ bliskimi odpowiednio $+60^\circ$, -60° , i 180° .

(d) **Grupa fosforanowa** może przyjmować w zasadzie trzy klasyczne formy wokół wiązania C(5') - O podobnie jak egzocykliczna ale silnie preferowana jest jedna konformacja ap z transoidalnym ustawieniem P i C(4'), kąt β ok. 180° . Ponieważ dla

nukleotydu w łańcuchu polinukleotydowym nie można jednoznacznie przypisać fosforanu pozycji 5' trzeba również rozpatrywać **3'-nukleotydy**. Fosforan preferuje silnie konformację zbliżoną do transoidalnego ustawienia P i C(4') z kątem $\epsilon \sim 220^\circ$ (40° odchylenia od ap).

Energie wymienionych konformerów są zbliżone i nukleotydy w roztworze przechodzą z jednej konformacji w drugą pokonując **niewielkie bariery** między nimi, nie przekraczające 25 kJ/mol czyli poniżej 10RT w temperaturze pokojowej (RT jest miarą energii ruchów cieplnych).

Po włączeniu nukleotydu do **łańcucha polinukleotydowego** DNA lub RNA pojawia się możliwość obrotu wokół wiązań fosfoestrowych i dodatkowe konformacje, opisane **kątami** α i ζ [**PLANSZA 45: parametry konformacyjne łańcucha polinukleotydowego**]. W sumie strukturę przestrzenną łańcucha polinukleotydowego opisuje 6 wartości kątów dwuściennych $\alpha - \zeta$ dla każdego nukleotydu, a pełny opis cząsteczki kwasu nukleinowego wymaga dodatkowo wartości kąta glikozydowego χ oraz konformacji pierścienia cukru, parametry P i ν_{max} . Kąt δ jest wspólny dla łańcucha i pierścienia (nie definiuje jednoznacznie konformacji tego ostatniego). Struktura przestrzenna cząsteczki kwasu nukleinowego może tworzyć różne formy (typy) lokalne i globalne, które określa się terminem **struktur wyższorzędowych**: drugorzędowych i trzeciorzędowych.

WYKŁAD 5

Kwasy nukleinowe DNA i RNA (2)

Pierwszym krokiem w ustaleniu pełnej struktury cząsteczki kwasu nukleinowego jest określenie **sekwencji nukleotydów od 5'- do 3'-końca** **[PLANSZA 45]** czyli **STRUKTURY PIERWSZORZĘDOWEJ**. Naczelnym zadaniem **genomiki** jest ustalenie pełnej sekwencji nukleotydowej DNA człowieka i innych organizmów. W 2000 r. badacze w ramach Human Genom Project i prywatna firma Celera Genomics ogłosiły (wspólnie) wykonanie sekwencjonowania DNA wszystkich chromosomów ludzkich: **6·10⁹ par zasad (bp) na 46 chromosomach**. Niezależnie od zastosowanej w obu przypadkach techniki cząsteczkę DNA trzeba:

- **pociąć przy pomocy endonukleaz restrykcyjnych na zestaw krótkich fragmentów jednoniciowych;**
- **sklonować fragmenty w *E. coli*** (przypadkowe biblioteki);
- **ustalić sekwencję każdego fragmentu** przy pomocy jednej z dostępnych technik;
- **złożyć sklonowane fragmenty we właściwej kolejności**, np. „na zakładki” jednakowych sekwencji’.

Ustalanie sekwencji w krótkich fragmentach do ok. 400 n można było wykonać za pomocą jednej z metod:

- (1) **chemiczne rozszczepianie** Maxam-Gilbert'a (starsza, rzadko obecnie stosowana);
- (2) **kontrolowane przerywanie replikacji** Sangera. **[PLANSZA 46: metoda Sangera sekwencjonowania DNA]**.

Metoda Sangera polega na enzymatycznej syntezie (polimeraza I) komplementarnej kopii do nici DNA o poszukiwanej sekwencji. Materiał DNA czyli sklonowany fragment jest rozdzielany na 4 porcje i dla każdej z nich jest prowadzona synteza, przerywana na ściśle określonej zasadzie, komplementarnej do zasady **analogu 2',3'-dideoxy-ATP, 2',3'-dideoxy-GTP, 2',3'-dideoxy-CTP lub 2',3'-dideoxy-TTP**, który po włączeniu do nici **terminuje syntezę wskutek braku 3'OH**. Odpowiednia ilość 2',3'-dideoxy-NTP w każdej z 4 reakcji zapewnia jego włączenie do nowo-syntetyzowanego DNA średnio 1:100 w stosunku do zwykłych NTP. To zapewnia powstanie porównywalnych ilości fragmentów różnej długości. Primery wymagane do zapoczątkowania działania polimerazy w każdym z układów, lub same analogi 2',3'-dideoxy-NTP posiadają inaczej fluoryzującą etykietę

(znacznik fluorescencyjny). **Elektroforeza** pod detektorami czułymi na 4 różne długości fali pozwala identyfikować **fragmenty różniące się długością o jeden nukleotyd i zakończone na 3'-końcu określonymi zasadami** (identyfikacja po długości fali emisji), dając sekwencję nici komplementarnej, a stąd wyjściowej po uwzględnieniu primera.

Dominującą techniką składania zsekwencjonowanych fragmentów we właściwej kolejności było **whole-genome shotgun** cloning (WGS) opracowany przez J. Craig Ventera (Celera Genomics). Wykorzystuje się w tym celu komputery dużej mocy obliczeniowej, które dopasowują „na zakładkę” **miliony** fragmentów.

Oba etapy: sekwencjonowanie i składanie fragmentów można w pełni **zautomatyzować**. Np. w sekwenatorze ABI Prism3730XL stosuje się do rozdziału elektroforetycznego CE 96 kapilar, co daje wydajność ok. $1.6 \cdot 10^5$ nukleotydów/doba. Łatwo policzyć że sekwencjonowanie genomu ludzkiego $1.2 \cdot 10^{10}$ n przy pomocy 100 maszyn trwa ok. 2,5 roku. **Celem praktycznym genomiki jest opracowanie wysoce „zrównoleglowanej” (*massively parallel*) techniki umożliwiającej zsekwencjonowanie genomu (indywidualnego) człowieka w cenie ok. \$1 000 w przeciągu ok. tygodnia**. Opracowuje się szereg metod, tzw. **platform sekwencyjnych** dysponujących komercyjnie dostępnymi sekwenatorami (review: Gundry & Vijg Mutation Research 729, 1, 2012). Aktualnie wiodące systemy to:

- Roche/454, sekwenator *GS FLX Titanium*;
- Life Technologies, sekwenator *SOLID*;
- Illumina, sekwenator *Hi-Seq 2000* (poprawiona wersja *SOLEXA Genome Analyser*);
- Pacific Biosciences, sekwenator *PacBio RS* „trzeciej” generacji.

Platformy różnią się od metody Sangera, ale wykazują wspólne cechy, np. zastosowanie ogólny model sekwencjonowania jako **skoordynowana seria powtarzalnych, zautomatyzowanych kroków** dla kilkuset tysięcy do kilku milionów cząsteczek DNA jednocześnie. Ze względu na krótkość fragmentów ich składanie odbywa się w odniesieniu do sekwencji znanych. To ukierunkowuje zastosowania na wyszukiwanie różnic, np. SNP (*single nucleotide polymorphism*), a nie sekwencjonowanie *de novo* całych genomów.

Kluczowym, a jednocześnie trudnym zagadnieniem (przede wszystkim w eukariotach) dla którego sekwencje stanowią „materiał” wyjściowy jest **annotacja regionów kodujących** czyli „określenie genów”.

Globalna, funkcjonalna struktura przestrzenna DNA i RNA różni się od siebie w znaczący sposób w wyniku odmiennych funkcji biologicznych obu rodzajów kwasów

nukleinowych. Cząsteczki **DNA są zintegrowane z białkami w nici chromatynowej** jako nośnik informacji genetycznej. Różne rodzaje **RNA przyjmują zróżnicowaną strukturę globularną**, raczej przypominającą białka globularne. **Lokalne** zwinięcie nici polinukleotydowej i związane z tym różne formy struktury drugorzędowej są **podobne w DNA i RNA**, i uwarunkowane przyjmowaniem określonych konformacji **[PLANSZA 42]** przez pojedyncze nukleotydy.

Struktury drugorzędowe DNA i RNA:

- (1) **helisy podwójne** (dupleksy);
- (2) **helisy złożone z większej ilości nici**: potrójne (tripleksy) i poczwórne (tetrapleksy lub quadrupleksy);
- (3) **pętle**: szpilki do włosów **hairpin** i pętle **wewnętrzne** w duplexach, w tym **wybrzuszenia** (bulldges).

Klasyczne duplexy złożone z dwóch przeciwnie skierowanych i komplementarnych nici polinukleotydowych zostały wymodelowane na podstawie dyfrakcji we włóknach. Wyróżnia się trzy główne **rodziny helis podwójnych: B, A i Z [PLANSZA 47: klasyczne helisy podwójne]**. DNA może występować w z każdej form, natomiast RNA **NIE WYSTĘPUJE** w formach z rodziny B ze względu na **steryczne przeszkody między fosforanami i grupami 2'-OH**. Każda rodzina zawiera dodatkowe typy helis, np. A', Z', a rodzina B: helisy C, D, E. Mogą także powstawać **hybrydy złożone z nici DNA i nici RNA**, podczas transkrypcji. Niektóre formy helikalne, np. D, E obserwuje się w syntetycznych oligomerach, np. poli(dA-T), poli(dG-dC).

W każdej z tzw. **klasycznych, regularnych helis** wszystkie nukleotydy składowe opisane są przez jeden zestaw wartości **kątów dwuściennych [PLANSZA 48: klasyczne helisy podwójne: parametry], [PLANSZA 45]**.

KĄTY DWUŚCIENNE

	α	β	γ	δ	ϵ	ζ	χ
B-DNA	314	214	36 (+sc)	156 (S)	155	264	262(anti)
A-RNA	293	186	50 (+sc)	82 (N)	207	294	202(anti)
Z-DNA							
dG	52	207	187 (ap)	76 (N)	288	102	89 (syn)
dC	250	192	54 (+sc)	147 (S)	257	296	201(anti)

PARAMETRY HELIKALNE

	symetria X_k	skok P (Å)	skok/ bp $h=P/X$	Skręt helisy $\Omega=360/X$ (°)	nachylenie bp η (°)	środek bp (Å)	duża bruzda S G (Å)	mała bruzda S G(Å)
B-DNA	10_1	33.8	3.38	36	-6	0.5	11.7 8.5	5.7 7.5
A-RNA	11_1	30.0	2.73	32.7	18	4.4	2.7 13.5	11.0 2.8
Z-DNA	6_5	43.5	7.25	-60 (-45 -15)	-5		brak	3.7 8.8

Podane parametry różnią się w niewielkim stopniu w przypadku A-DNA i A-RNA oraz Z-DNA i Z-RNA. Wartości kątów dwuściennych w pełni opisują konformację helisy, ale nie oddają „poglądowo” kształtu poszczególnych helis. Dlatego wprowadza się **parametry helikalne**. Helisa ma **symetrię śrubową** oznaczaną symbolem X_k . Obrót helisy o kąt $(k/X) \cdot 360^\circ$ i jednocześnie przesunięcie wzdłuż osi o $(k/X)P$, gdzie **P** jest długością jednego skrętu helisy, $k \leq X$ **[PLANSZA 49: symetria osi śrubowej]**, zachowuje te same co poprzednio położenia poszczególnych nukleotydów. **Symetria nie jest pełna** (pseudosymetria), gdyż taka operacja nie przeprowadza atomów w identyczne atomy, ale całe nukleotydy w inne, niekoniecznie identyczne nukleotydy. Helisa może być **prawoskrętna** jak B lub A gdy obrót zachodzi zgodnie z ruchem śruby prawoskrętnej lub **lewoskrętna** jak Z gdy ruch jest przeciwny: obrót o 300° w prawo i przesunięcie o $5/6$ skoku jest równoważny obrotowi o -60° czyli w lewo i przesunięciu o $1/6$ skoku. W przypadku helis A i B każdy nukleotyd przechodzi w sąsiedni wzdłuż nici lub para zasad (nukleotydów) komplementarnych w sąsiednią parę; każdy nukleotyd jest równoważny.

W helisie Z, którą zaobserwowano dla **naprzemiennych** sekwencji poli(dG-dC) i poli (G-C), cytozyna nie jest równoważna konformacyjnie guanozynie **[PLANSZA 48]** i dwunukleotydowy fragment przechodzi w dwunukleotydowy fragment położony obok w sekwencji (dwie pary zasad komplementarnych). Na jedną parę zasad przypada skręcenie Ω kolejno o 45° i 15° . Z parametrów helikalnych widać **charakterystyczne cechy kształtu helis** **[PLANSZA 47]**: nachylenie par zasad do osi opisane kątem η : małe dla B i Z, duże dla A, odstęp geometrycznego środka par zasad od osi helisy i wielkość dwóch wcięć w walcowym kształcie helisy, **dużej i małej bruzdy**, o szerokości S i głębokości G.

Dupleksy są stabilizowane przez dwa rodzaje oddziaływań wewnątrzcząsteczkowych i oddziaływanie z rozpuszczalnikiem:

- (1) **wiązania wodorowe** w parach zasad komplementarnych, trzy w parze GC i dwa w parze AT;
- (2) **oddziaływania warstwowe (stacking)** między sąsiednimi parami zasad;
- (3) **oddziaływanie z jonami rozpuszczalnika**: odpychanie elektrostatyczne ujemnie naładowanych nici jest ekranowane przez dodatnie kationy związane z ujemnie naładowanymi resztami fosforanowymi.

Pierwsze dwa typy oddziaływań są silniejsze w przypadku par GC i fragmenty helisy DNA i RNA bogate w te pary są stabilniejsze niż obszary bogate w pary AT lub AU. Obliczenia kwantowe DLA PRÓŻNI dają **energię (entalpię) wiązania** pary G:C mostkami wodorowymi ok. **100 kJ/mol**, a pary A:T ok. **50 kJ/mol**; dla stackingu wartości zawierają się w przedziale od **60 do 15 kJ/mol**, zależnie od typu pary i sąsiadów.

Ostateczne potwierdzenie struktury helikalnej kwasów nukleinowych nastąpiło dopiero po badaniach rentgenograficznych na **monokryształach** (a nie włóknach) i NMR w roztworze krótkich, syntetycznych fragmentów *self*-komplementarnych. Miało to szczególne znaczenie dla DNA, gdzie cały **model przekazu informacji genetycznej**: kod genetyczny, replikacja i transkrypcja jest oparty na strukturze podwójnej helisy. Badania te pokazały, że struktura helisy najczęściej odbiega od regularnej, klasycznej i każdy nukleotyd jest opisany przez indywidualny, nieco różny, zestaw kątów dwuściennych. W dalszym ciągu jednak **nieregularne helisy** należą do określonej rodziny A, B lub Z. Trzeba więc wprowadzić więcej parametrów helikalnych opisujących te nieregularności [**PLANSZA 50: parametry helisy nieregularnych**], głównie chodzi o nieregularności we wzajemnym ustawieniu zasad, co przenosi się na nieregularności łańcucha fosfocukrowego. **Parametry te mają zerowe wartości w helisach regularnych**. Podane już parametry: Ω , h (na **PLANSZY** Dz), η oraz D (na **PLANSZY** Dx i Dy) dalej obowiązują, ale mają **inne wartości** w każdym fragmencie helisy [**PLANSZA 51: zmiany parametrów w d(GGAAT^FUCC)₂**]. Przykład nieregularnej helisy A-RNA w monokryształach pokazuje [**PLANSZA 24**] m. in. załamania samej osi helisy, która w helisach regularnych jest linią prostą.

Przyjęcie przez DNA i RNA określonego typu struktury helikalnej jest uwarunkowane przez dwa czynniki:

(1) **sekwencja zasad:**

(2) **warunki środowiska:** temperatura, stężenie soli, pH, skład rozpuszczalnika, warunki krystalizacji, wilgotność włókien, itp.

W kryształach obserwuje się krótkie DNA we wszystkich trzech formach: A, B i Z. Forma Z wymaga naprzemiennych sekwencji dG-dC. RNA występuje w kryształach w formie A, a naprzemienne sekwencje G-C mogą wystąpić w formie Z. W roztworze wodnym fragmenty DNA występują w formie B. Dodanie **etanolu** wymusza formę A. Naprzemienne sekwencje dG-dC DNA i G-C RNA przechodzą z formy B w Z przy znacznych stężeniach soli, przy czym formie Z znacząco sprzyja **metylacja cytozyny w pozycji 5**. **Natywne, funkcjonalne DNA** przyjmuje najprawdopodobniej wszystkie trzy typy form helikalnych. Istnieją dowody biologiczne na obecność formy Z - przeciwciała anty-Z.

WYKŁAD 6

Kwasy nukleinowe DNA i RNA (3)

Struktura helikalna jest stabilna w roztworze w określonych warunkach środowiska. Wzrost temperatury, obniżenie stężenia soli - szczególnie w przypadku krótkich helis, skrajne pH, powodują **denaturację helisy** i przejście do formy pojedynczych nici. Proces jest odwracalny i helisa odtwarza się przy powrocie do warunków sprzyjających. Takie **odwracalne topienie helis** np. wywołane zmianami temperatury można śledzić **[PLANSZA 52: profile topnienia helis podwójnych]:**

(a) spektrofotometrycznie, wykorzystując efekt **hypochromazji**: absorpcja helisy jest mniejsza niż absorpcja dwu pojedynczych nici;

(b) kalorymetrycznie metodą DSC (różnicowa kalorymetria skaningowa): zmiana pojemności cieplnej (ciepła właściwego) układu w stałym ciśnieniu w funkcji temperatury. Topienie i powstawanie helisy zachodzi w wąskim przedziale temperatur. Przy formowaniu się helisy ok. 3 pary zasad tworzą **stabilne jądro**, a dalej proces ma charakter **kooperatywny**: już utworzona struktura sprzyja łączeniu się następnych par zasad i proces przypomina zamykanie zamka błyskawicznego. Całkowita stała równowagi w przybliżeniu modelu „*all-or-none*”

$$K = \beta s^N \quad \beta - \text{parametr nukleacji } \sim 10^{-3} \text{ M}^{-1}; s - \text{parametr wzrostu łańcucha } \sim 10 \\ \text{dla } T \ll T_m, s = 1 \text{ dla } T = T_m; N - \text{ilość par zasad}$$

W temperaturze **T_m** , tzw. **temperaturze topnienia helisy**, współlistnieje 50% formy helikalnej i 50% pojedynczonicowej. Im wyższa wartość temperatury topnienia tym stabilniejsza helisa. Wzrost T_m następuje przy:

- (a) wydłużaniu łańcucha polinukleotydowego;
- (b) wzroście stężenia soli w roztworze;
- (c) wzroście zawartości par zasad GC.

"Spłaszczanie" profilu topnienia i jego "pofałdowanie" świadczy o **zmniejszeniu kooperatywności** przejścia helisa - pojedyncze nici. Topienie i odtwarzanie helisy może wtedy następować niezależnie w różnych fragmentach i charakteryzować się nie jedną wartością T_m ale ich rozkładem. W zależności od eksperymentu otrzymuje się różne entalpie procesu:

- spektroskopia absorpcyjna daje z **równania van't Hoffa** na zależność $K(T)$ tzw. **entalpię van't Hoffa** procesu topienia/tworzenia helisy ΔH_{vH} ;

- DSC daje „bezpośrednią” **entalpię kalometryczną** ΔH_{cal} .

Dla całkowicie kooperatywnego procesu wartości są jednakowe; dla procesów niekooperatywnych stosunek $\Delta H_{vH}^0/\Delta H_{cal}^0 < 1$ i stanowi miarę (nie)kooperatywności.

Zmiany warunków środowiska mogą prowadzić również do przekształcenia jednej formy helikalnej w inną, np. do przejść $A \leftrightarrow B$ czy $B \leftrightarrow Z$. Przykładowo, przy dużej zawartości naprzemiennych sekwencji dGdC (GC), znaczne stężenie soli (chlorek litu) wymusza przejście od B-DNA do Z-DNA [**PLANSZA 53: przejście konformacyjne B \leftrightarrow Z DNA**]. Zanik formy B i pojawianie się formy Z widać na widmach ^{31}P NMR: zanik pojedynczego sygnału od równoważnych fosforów formy B i powstawanie dwóch sygnałów od nierównoważnych fosforów formy Z wskutek "załamania" łańcucha fosfocukrowego i nierównoważności chemicznego otoczenia fosforów w dGpdC i dCpdG. Podejmowane są próby znalezienia mechanizmów tego typu przejść, np. mechanizm solitonowy (*Biopolymers* **23**, 2211, 1984). Prowadzi się w tym celu m. in. symulacje TMD (*targeted molecular dynamics*; *Biophysical Journal* **91**, 2976, 2006)

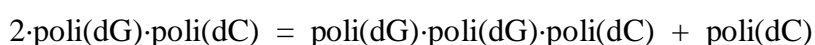
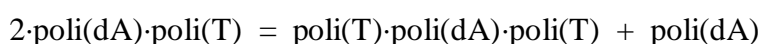
W przeciwieństwie do skoordynowanego, kooperatywnego przekształcenia helisa - kłębek statystyczny, w helisach podwójnych obserwuje się zjawisko **przypadkowego (niekooperatywnego), lokalnego otwierania helisy** na odcinku jednej pary zasad komplementarnych, tzw. „**oddychanie**” helisy (*breathing*). Wiązania wodorowe stabilizujące parę ulegają przejściowemu zerwaniu, a uzewnętrznienie zasad do rozpuszczalnika powoduje wymianę protonów NH z protonami cząsteczek buforu wodnego, jedna wymiana na kilka aktów otwarcia pary zasad. Proces jest katalizowany zasadowo i zachodzi w czasie: **od poniżej 1 ms** dla krańcowych par zasad (*frying*) do **ok. 600 ms** dla silnie stabilizowanych par zasad w środku helisy w nieobecności katalizatora. Średnio, na ok. 10^5 do 10^6 bp jedna para jest „otwarta”.

W końcu lat 90-tych XX w. pojawiły się prace dotyczące kilku nowych, ciekawych struktur helisy podwójnych. **M-DNA** jest kompleksem DNA z dwuwartościowymi jonami metali Zn^{+2} , Co^{+2} lub Ni^{+2} [**PLANSZA 54: helisa podwójna M-DNA**]. W warunkach lekko alkalicznych ($\text{pH} > 8$) badania NMR pokazały, że protony iminowe NH wiązań wodorowych par komplementarnych wymieniają się na jony metali, co wprowadzie w niewielkim stopniu modyfikuje strukturę typu B-DNA, ale nadaje helisie własności **przewodzącego drutu molekularnego**.

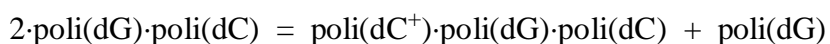
Praca Allemand *et al.* w PNAS [**PLANSZA 55: helisa podwójna P-DNA**] relacjonuje uzyskanie nowej struktury **P-DNA** w 17 kDa plazmidowym B-DNA o dodatniej superskrętności (bufor fosforanowy pH 8), w wyniku ciągnięcia ze skręceniem z siłą 3 pN magnesem za paciorek magnetyczny przyczepiony do jednego końca helisy łączem streptawidyna – biotyna; drugi koniec unieruchomiony na płaskiej powierzchni szklanej łączem digoxigenin (DIG) – przeciwciało (antyDIG). Struktura ta, dłuższa o 75%, z łańcuchami fosfocukrowymi splecionymi w środku, zasadami na zewnątrz, 2.62 nukleotydu na skret helisy, była postulowana przez Corey'a i Paulinga w 1953 r. jako konkurencyjny model w stosunku do modelu Watsona i Cricka. Pojawiły się doniesienia o możliwości tworzenia tego typu struktury *in vivo* w fagu Pfl.

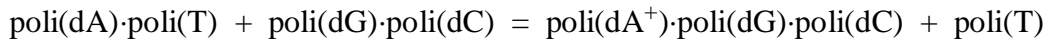
xDNA [**PLANSZA 56: helisa podwójna xDNA**] (nazwa wzięta z nomenklatury oznaczeń rozmiarów garderoby *X large*) jest eleganckim przykładem modelowania molekularnego. Rozmiary zasad azotowych zwiększono przez wprowadzenie dodatkowych skondensowanych pierścieni benzenu $T \rightarrow xT$ i $dA \rightarrow dxA$ bez naruszenia układu potencjalnych wiązań wodorowych. Pojedyncze duplekсы z udziałem powiększonych zasad: $xT:dA$ i $T:dxA$ destabilizują wprawdzie helisę, ale **DNA zbudowane z naprzemiennych sekwencji tego typu par lub helisy z jedną nicią xT a druga dA jest stabilniejsze niż analogiczne, zwykłe DNA**, ze względu na wzrost stackingu.

Podobnie jak w przypadku struktur duplexów, **zmiany warunków środowiska** przy odpowiedniej sekwencji nukleotydów oraz **oddziaływania wewnątrzcząsteczkowe** mogą prowadzić do przejścia cząsteczki kwasu nukleinowego w **różne od helisy podwójnej struktury drugorzędowe**. W helisach podwójnych, złożonych z jednej nici homopurynowej i jednej homopirymidynowej, $poli(dA) \cdot poli(T)$, $poli(A) \cdot poli(U)$, $poli(dA) \cdot poli(U)$, $poli(A) \cdot poli(T)$, oraz w podobnych z udziałem G i C: $poli(dG) \cdot poli(dC)$ itd., w obecności jonów magnezu Mg^{+2} następuje rozseparowanie nici jednej z helis i powstanie **helisy potrójnej czyli „triplexu”** i jednej nici pojedynczej:



lub w obniżonym pH (5-6):





i podobnie dla nici RNA i nici mieszanych DNA/RNA. Trzecia nić wchodzi w dużą bruzdę helisy podwójnej i wiąże z jedną z nici dupleksu poprzez **parowanie typu Hoogsteena** [**PLANSZA 57: struktura triplexu DNA**]. Długości dwóch wiązań wodorowych Hoogsteena są takie jak W.C., 2.8 - 2.9 Å a kąty D-H...A nie odbiegają więcej niż 20° od liniowości. Podobne są także energie stabilizacji w obu typach parowania. **PLANSZA** obrazuje tryplex DNA zbudowany z pojedynczej nici DNA o odpowiednio dobranej sekwencji. Trypleks jest biologicznie ważną strukturą przejściową w procesie **rekombinacji DNA** tzw. **general recombination** [**PLANSZA 58: rekombinacja; złącze Hollidaya**], który prowadzi do wymiany części nici w dwóch homologicznych podwójnych helisach DNA (*crossing over*, mejoza).

Homopirymidynowe nici oligo(dC) w obniżonym pH ok. 4.5 mogą tworzyć ciekawą strukturę **tetrapleksu**, tzw. **I-motif DNA** [**PLANSZA 59: struktura tetraplexu oligoC**], złożoną z dwóch równoległych dupleksów (a nie antyrównoległych jak w zwykłych duplexach). Równoległe duplexy są stabilizowane przez 3 wiązania wodorowe **dC⁺:dC** i ustawione antyrównoległe do siebie z **naprzemiennym stackingiem par cytozynowych**, tzn. para od jednego duplexu interkaluje między dwie pary drugiego. Cała helisa jest prawoskrętna, jej wszystkie cztery nici są konformacyjnie równoważne i należą do rodziny B-DNA. Motyw występuje w wielu przypadkach DNA zawierających trakty oligoC/oligoG (nici bogate w cytozyny) w *Tetrahymena thermophila*, telomerach ludzkich chromosomów (3-"zwisające", niezwiązane komplementarnie końce chromosomów eukaryotycznych), sekwencjach centromerycznych i promotorach onkogenów związanych ze wzrostem i proliferacją komórek.

Inną strukturę tetrapleksową (*G-quadruplex*) mogą wykazywać fragmenty DNA bogate w guanozyny (komplementarne do fragmentów bogatych w cytozyny), zawierające sekwencje **tandem repeat** $d(\text{A}\underline{\text{lub}}\text{T})_{(1-4)}d\text{G}_{(1-8)}$. Badania NMR w roztworze pokazały przechodzenie w obecności jonów K^+ dupleksów $d(\text{G}_3\text{T}_4\text{G}_3)/d(\text{C}_3\text{A}_4\text{C}_3)$ [**PLANSZA 60: struktura tetraplexu $d(\text{G}_3\text{T}_4\text{G}_3)_2$**] w dwuniciowe tetrapleksy $d(\text{G}_3\text{T}_4\text{G}_3)_2$ i dwie nici pojedyncze $d(\text{C}_3\text{A}_4\text{C}_3)$. Stabilizacja 4xdG zachodzi poprzez parowanie Hoogseen'a.

DNA i RNA mają potencjalną możliwość tworzenia **helis równoległych** stabilizowanych przez wiązania wodorowe par A:T typu odwrócona para W.C. (*reverse*

W.C.; do samodzielnego narysowania). Postuluje się, że **psDNA** (*parallel stranded DNA*) może mieć istotne znaczenie w regulacji replikacji, transkrypcji, splicingu itd. Helisę równoległą zaobserwowano we fragmencie 12-merowym:

5'-d(CCATAATTTACC-3'

5'-d(CCTATTAAATCC-3'

o pełnej komplementarności środkowego oktameru w ustawieniu równoległym i braku komplementarności w ustawieniu antyrównoległym (*Parvathy et al. Nucleic Acids Res. 30, 1500, 2002*). Helisa jest dodatkowo spinana przez pary C⁺:C na końcach.

Efekty wpływu **niekomplementarnego parowania**, tzw. *mismatch* na strukturę helis kwasów nukleinowych badano szeroko w kryształach i roztworze, praktycznie dla wszystkich możliwych kombinacji zasad. W **DNA** zagadnienie jest ważne ze względu na problem powstawania **mutacji punktowych** bez udziału rzadkich tautomerów czy form protonowanych/zjonizowanych, a w **RNA** ze względu na parowanie kodon mRNA - antykodon tRNA, np.: kodony 5'-UUU i 5'-UUC dla fenyloalaniny parują z jednym antykodonom tRNA^{Phe} 3'-AAGm (Gm ozn. 2'-OMe-G) **[PLANSZA 41]**. Różnice energii par niekomplementarnych DNA w stosunku do komplementarnych (destabilizacja helisy podwójnej) wynoszą:

$$\Delta G^\circ = RT \ln K = +17 \text{ do } 28 \text{ kJ/mol}$$

największe dla C:C i C:A, najmniejsze dla G:T

To odpowiada stałej równowagi $K \sim 10^{-3} \div 10^{-5}$, czyli **częstości mutacji *in vitro*** (bez udziału systemów naprawczych), i stanowi argument na rzecz hipotezy że błędne parowanie nie wymaga żadnych szczególnych form zasad, a wynika jedynie ze spełniania przez pary zasad praw termodynamiki: komplementarność kształtu i steryczne dopasowanie zasad *implicite* **[PLANSZA 40]**

W przypadkach osłabienia stabilizacji struktury helikalnej:

(1) **brak komplementarności** na dwóch niciach podwójnej helisy;

(2) **struktura helikalna w oligo(T)/oligo(dA)**, np. przy niższym stężeniu soli,

mogą tworzyć się różnego rodzaju **pętle**. Dobrze scharakteryzowane strukturalnie są tzw. **wybrzuszenia** (*bulges*) **[PLANSZA 61: struktura wybrzuszenia]**, przykładowo dla *extra* dA₅ lub dAATAA obserwuje się ostre załamania (*kink*), których pozycja zależy od sekwencji warunkującej stacking w wybrzuszeniu i na styku wybrzuszenie-helisa. Struktura **wybrzuszenia DNA** jest potencjalnym miejscem wiążącym środki farmaceutyczne oraz białko p53 (*tumor suppression protein*). W RNA wybrzuszenia są ważnym elementem strukturalnym w self-splicingu i zwijaniu RNA (*tertiary folding*).

Kolejną strukturą drugorzędową jest **pętla hairpin** [**PLANSZA 62: szpilki do włosów hairpin DNA i RNA**], występująca w RNA i w krzyżowych formach DNA. Szpilka do włosów jest zbudowana z trzonu helikalnego i jednoniciowej pętelki, w której **zasady są zestackingowane**. Najbardziej stabilne są pętle DNA złożone z 4 do 5 nukleotydów i pętle RNA 6 do 7 nukleotydów. Jest to związane z najczęściej występującą budową pętli [**PLANSZA 63: zasada budowy pętli hairpin**]. Jeżeli w helisie B-DNA urwiemy jedną nić z zachowaniem końcowej grupy fosforanowej a drugą kontynuujemy w kierunku 3' to minimum odległości między fosforanami wystąpi dla **trzeciego** z kolei dobudowywanego nukleotydu. Tę lukę ok. 12 Å można zamknąć jednym lub dwoma nukleotydami o charakterze "ostrego zwrotu". W helisie A-RNA, tu hairpin antykodonu tRNA^{Phe}, minimum odległości P–P ok. 13 Å następuje przy rozbudowie nici w kierunku 5' dla **piątego** nukleotydu i lukę znów można zamknąć jednym lub częściej dwoma nukleotydami. Pętle te mogą być dodatkowo stabilizowane przez wiązania wodorowe, np. w hairpin antykodonu, między U33N(1)H i tlenem grupy fosforanowej A35p(+2)A36.

Omówione zasady nie wykluczają innych struktur typu pętli: **czteropętla** (*tetraloop*) sekwencji 5'-GNRA (N - dowolny nukleotyd, R = G lub A) w intronach grupy I, rRNA, retrowirusach, oraz **trzyplet** UNR.

Struktury trzeciorzędowe DNA, czyli organizacja przestrzenna struktur drugorzędowych - ogólny fold cząsteczek DNA, różnią się zasadniczo od struktur trzeciorzędowych RNA. Funkcjonalne DNA tworzy, z wyjątkami takimi jak niektóre wirusy DNA czy plasmidy, strukturę trzeciorzędową **chromosomu** w powiązaniu z białkami [**PLANSZA 64: etapy upakowywania DNA w chromosomie**], [**PLANSZA 26**]. Następuje silne zwinięcie helikalnego DNA w następujących po sobie etapach. W pierwszym etapie podwójna helisa nawija się na białka histonowe w których występują częściowo specyficzne wobec DNA fragmenty SPKK (seryna, prolina, lizyna) tworząc **nukleosomy** ("paciorki na nici"). Dalej tworzy się 30 nm (średnica) **nić chromatynowa**, która dalej fałduje się w jądrze eukaryotów o wymiarze ok. 2 µm. aż do **upakowania** sięgającego $10^5 = \text{długość nici DNA/wymiar jądra} (0.1\text{m} / 10^{-6}\text{m})$

Zdolność do takiego upakowania wynika z giętkości długich podwójnych helis DNA, które bada się również w zrelaksowanym, pozbawionym białek DNA np. plasmidów. Ta zdolność i elementy struktury trzeciorzędowej DNA które ją warunkują spełniają ważną biologiczną rolę przez umożliwienie **kontaktów fragmentów DNA**

odległych w strukturze pierwszorzędowej. Przykładem takiej lokalnej struktury, która wykracza poza prostą strukturę drugorzędową jest **forma krzyżowa DNA**, obserwowana pod mikroskopem AFM [**PLANSZA 65: forma krzyżowa fragmentu *inverted repeat***]. Forma ta może powstać w sekwencjach tzw. **odwróconego powtórzenia** (*inverted repeat*) przedzielonych fragmentem bogatym w mniej stabilne pary A:T. Tworzą się wtedy dwie, skierowane do siebie trzonami szpilki do włosów.

Inną ważną cechą struktur trzeciorzędowych DNA, także funkcjonującego w chromosomie, jest **superskrętność** [**PLANSZA 66: superskrętność podwójnej helisy DNA**]. Przejawia się w DNA kolistym jak w **plasmidach**, lub w DNA którego fragment jest unieruchomiony na bliskopolożonych końcach jak w **nukleosomie** [**PLANSZA 67: struktura krystalograficzna nukleosomu**]. DNA nie może wtedy zmieniać topologicznej liczby skrętów **L** czyli tego ile razy jedna nić jest opleciona wokół drugiej (ilość przecięć krawędzi „wstażki”). Jeżeli jest ona mniejsza niż ilość skrętów **T** narzucanych przez typ helisy (10 nukleotydów na jeden skok helisy w klasycznym B-DNA) wtedy pojawiają się ujemne superskręty w liczbie **W**, tak aby:

$$L = T + W.$$

Odwijanie helisy bez rozcinania w jednym miejscu powoduje "super"zapętlenie w innym. **DNA w nukleosomach** można traktować jako „zamkniętą” pętlę ze względu na unieruchomienie końców fragmentu 147 bp przez białka histonowe. Widać prawie dwa superskręty nawinięte na histony w postaci spirali, która topologicznie odpowiada lewoskrętnym (znak „-”) superskrętom. Wynika stąd, że funkcjonalne DNA w chromatynie ma stopień superskrętności $\lambda = (L - L_0)/L_0 \sim W/L_0$ od **-0.03** do **-0.09**, gdzie L_0 jest topologiczną liczbą skrętów w zrelaksowanym DNA. Enzymami regulującymi topologiczną ilość skrętów są **topoisomerazy**, które odwracalnie rozcinają jedną (typ I) lub dwie (typ II) nici i zwijają lub odwijają nici przy udziale ATP ze zmianą superskrętności, a potem łączą powtórnie nici.

Kluczowym intermediatem w prawie wszystkich procesach rekombinacji jest trzeciorzędowa struktura DNA **złącza Hollidaya** (*Holliday junction* lub *four-way junction*) [**PLANSZA 58**], stabilizowana *in vivo* przez system białkowy RuvABC. Cztery helikalne ramiona A (niebieski), B (zielony), C (czerwony) i D (żółty) z dwukrotną osią symetrii

(prostopadła do płaszczyzny rysunku) wykazują w obszarze złącza odstępstwa od konformacji B-DNA dla czterech nukleotydów, ozn. C(7A), C(7C), A(6A) i A(6C) **[PLANSZA 48]**: kąt skręcenia helisy $\Omega = 25^\circ$ (zamiast 36°), kąt glikozydowy $\chi = -152^\circ$ (zamiast 262°) i $\varepsilon = -73,6^\circ$ dla adenozyiny 6A i adenozyiny 6C (zamiast 155°). W tej konformacji stabilizacja struktury złącza zachodzi przez:

- (1) silne nakrywanie zasad nici na złączu (stacking): guanina-5 i cytozyna-7;
- (2) wiązania wodorowe $N^4-H \dots O=P$ między cytozynami C(8C), C(8A) i fosforanami łańcucha A(6C)pC(7A), A(6A)pC(7A);
- (3) parowanie puryna-puryna A:G w dwóch ramionach.

WYKŁAD 7

Kwasy nukleinowe DNA i RNA (4)

Struktura trzeciorzędowa cząsteczek RNA jest bardzo bogata i przypomina struktury przestrzenne białek globularnych [PLANSZA 32]. W stosunkowo niewielkich (75 do 90n) cząsteczkach tRNA [PLANSZA 68: struktura tRNA] elementy struktury drugorzędowej (głównie szpilki do włosów) tworzą charakterystyczny motyw liścia koniczyny i przez wewnątrzcząsteczkowe oddziaływania organizują się w strukturę trzeciorzędową, przypominającą literę L. Dwa A-helikalne trzony: T i aminokwasowy tworzą jedną **pseudociągłą helisę (złącze dwóch helis)** a dwa pozostałe: D i antykodonowy drugą, w sumie dając **dwa ramiona L na giętkiej zawiasie**. Dwa podstawowe typy oddziaływań stabilizują strukturę trzeciorzędową:

- (1) **interkalacyjny stacking**, w którym zasada z jednego fragmentu nici wchodzi pomiędzy dwie stackingujące zasady drugiego fragmentu np. **C13** między U8/A9, zwykle na styku trzech nici, tu ściąganej przez wiązanie wodorowe m⁷G46:G22:C13 (**h**); w sumie w inerkalacyjny stacking zawiasy zaangażowanych jest 9 nukleotydów. Pozostałe trójki interkalacji: **G18**→m¹A58/G57, **G57**→G18G19 oraz **A9**→G45/m⁷G46, gdzie zasada **A9** zestakingowana jako układ do interkalacji (U8/A9) sama może interkalować inną parę.
- (2) **nietypowe wiązania wodorowe**: wobble (**a**), Hoogsteen (**b**), odwrócone W.C. (**e**), puryna-puryna (**d**), trójki zasad (**f**, **g**, **h**) zwykle z udziałem **nukleotydów modyfikowanych** (pseudourydyna Ψ, 7-metyloG), których jest ok. 10% w tRNA.

Powszechnie spotykanym w RNA elementem trzeciorzędowym jest **pseudowęzeł (pseudoknot)** [PLANSZA 69: motyw pseudowęzła RNA]. W wersji **H** (od: *hairpin*) następuje parowanie W.C. nukleotydów w pętli hairpin z jednoniciowym fragmentem, co prowadzi do struktury zbudowanej z dwóch obszarów helikalnych i trzech pętli. W niektórych wersjach tego pseudowęzła jedna z pętli może być zredukowana do zera. Pseudowęzły zapewniają ściągnięcie odległych fragmentów jednoniciowego RNA w zwartą strukturę przestrzenną, np. w 5S rRNA rybosomów. Spełniają różne funkcje:

- regulacja ekspresji genu na poziomie translacji wirusów, np. **ribosomal frameshifting** onkogennego wirusa sutka myszy (*mouse mammary tumor virus*); dwa białka z jednego RNA w określonym stosunku ilościowym przez przesunięcie ramki odczytu;
- regulacja procesu translacji: pseudowęzły na **5'UTR w regionie IRES** (*internal ribosome entry side*) mRNA, do rekrutacji rybosomu w kap-niezależnej translacji;

- **mimikra tRNA** (TLS) ramienia akceptorowego 3'-końca genomu wirusów roślinnych dla zapewnienia reaktywności katalizowanej przez enzym rozpoznający tRNA.

Inne typy pseudowęzłów: B-type (wybrzuszenie), I-type (pętla wewnętrzna), M-type (pętla wielogałęziowa), H-H *kissing loop* (dwie pętle hairpin), podwójny (jeden w drugim; *nested*) w *hepatitis delta virus* z pięcioma fragmentami helikalnymi (przegląd struktur: *Bierley et al. Nature Rev. 5, 598, 2007*).

Motyw A-minor [**PLANSZA 70: motyw A-minor RNA**] jest powszechnie występującym elementem upakowania dużych cząsteczek RNA, w którym helikalne fragmenty wiążą się przez oddziaływanie niesparowanej adeniny jednego elementu drugorzędowego z komplementarną parą G:C drugiego elementu drugorzędowego, od strony małej bruzdy. Klasyfikacja: typ I lub typ II oparta jest na położeniu 2'OH adenozyiny względem grup 2'OH cukrów pary G:C. Stabilizacja motywu zachodzi poprzez wiązania wodorowe między grupą N²H₂ guaniny par G:C i N3 adeniny oraz N1 adeniny, plus wiązania wodorowe z udziałem grup 2'-OH cukrów.

Zamek rybozowy [**PLANSZA 71: zamek rybozowy RNA**] między helisami stabilizowany jest przez wiązania wodorowe grup 2'OH obu helis oraz między hydroksylami 2' a N3 (puryny) i C2=O (pirymidyny).

Motyw Z-kotwica (*Z-anchor*) w spliceosomie, wiążący helisy (IC i I(ii)) w ustawieniu „*side-by-side*” w układzie architektonicznym „*inside-out*” przypominającej P-DNA (*N. Toor et al. Curr. Opin. Struct. Biol 19, 260, 2009*), z łańcuchem fosforanowym w środku i zasadami na zewnątrz, które oddziałują z dwoma sąsiednimi fragmentami łańcucha [**PLANSZA 55**].

Wymienione **MOTYWY STRUKTURALNE**: pseudociągłe helisy i złącza większej ilości helis, pseudowęzły, motywy A minor, zamki rybozowe. **Z-kotwice** plus **ODDZIAŁYWANIA STRUKTUR DRUGORZĘDOWYCH**: interkalacyjny **stacking**, nietypowe parowanie zasad, wiązania wodorowe z udziałem grup hydroksylowych 2'OH zapewniają zwinięcie RNA w funkcjonalne struktury natywne [**PLANSZA 32**].

Duże cząsteczki kwasów nukleinowych podlegają stałym, niekontrolowanym zmianom i modyfikacjom jako złożone układy termodynamiczne. Jest to szczególnie istotne w przypadku olbrzymich i "nieodnawialnych" cząsteczek DNA, które powinny być możliwie trwałe w czasie ze względu na niesioną informację genetyczną. Systemy naprawcze **obniżają częstości mutacji** z 10⁻⁴ w systemach replikacyjnych *in vitro* do 10⁻⁸

do 10^{-11} w żywych komórkach, od bakterii do eukaryotów. Latentne DNA w komórce podlega szeregu procesom **[PLANSZA 72: uszkodzenia DNA]**.

(a) **spontanycznym: np. dezaminacja** cytozyny do uracylu, który paruje z A zamiast z G, depurynacja (odłączenie zasady purynowej);

(b) **indukowanym przez czynniki chemiczne kancero- i mutagenne**, które reagują m. in. z zasadami (czynniki alkilujące, metoksyamina) i modyfikują je do form o innym lub niejednoznacznym (tautomeria) typie parowania; przykładowo metoksyamina modyfikuje adeninę do N⁶-metoksyadeniny, która wykazuje **tautomerię**: forma aminowa paruje z T a iminowa z C (tranzycja); związki wielopierścieniowe (piren) oddziałują warstwowo z parami zasad DNA - **interkalacja** bez rozrywania helisy, zaburzając strukturę helisy;

(c) **indukowanym przez czynniki fizyczne np. promieniowanie ultrafioletowe**, prowadzące do reakcji fotochemicznych zasad w elektronowym stanie wzbudzonym, np. dimeryzacja tymin lub oksydacja zasad purynowych do pochodnych 8-oxo pod działaniem rodników tlenowych.

Zaburzenia chemiczne i strukturalne, które mogą być przyczyną mutacji są ustawicznie usuwane przez systemy **naprawy DNA**. W szczególności rozpoznawane są błędne pary zasad **po replikacji** i usuwana jest zasada niekomplementarna w nowej nici a zachowana "właściwa" na nici starej. Dwa zasadnicze typy mechanizmów leżą u podstaw szeregu złożonych procesów usuwania różnych uszkodzeń/zmian DNA.

(A) **Bezpośrednie** przywrócenie nieuszkodzonego DNA (*direct reversal*);

(B) Naprawa przez **wycięcie** (*excision pathway*).

Pierwszy typ naprawy funkcjonuje przy usuwaniu niektórych modyfikacji chemicznych i fotodimeryzacji pirymidyn **[PLANSZA 73: systemy naprawcze DNA]**. Enzym fotoreaktywujący rozpoznaje uszkodzenie nici DNA i przy pomocy absorbowanego światła z zakresu widzialnego (a nie UV jak przy uszkodzeniu) **katalizuje reakcję fotochemiczną** rozcinania dimeru do dwóch niepołączonych zasad. **Drugi mechanizm naprawczy** jest kompleksem enzymatycznym. Uszkodzenie na nici, np. błędną parę czy fotodimer, jest specyficznie rozpoznawane przez odpowiednią endonukleazę, która wycina uszkodzoną zasadę lub cały fragment nici, która ją zawiera. Polimeraza DNA syntetyzuje brakujący kawałek na matrycy drugiej, nieuszkodzonej nici a ligaza zszywa nowopowstały fragment na końcu z resztą nici.

Białka globularne, włókniste i błonowe (1)

Białka **[PLANSZA 21]** są biopolimerami liniowymi, zbudowanymi zasadniczo z 20 typów aminokwasów włączanych w procesie **translacji mRNA**, a w niektórych *archaea* i *eubacteria* z 22 typów aminokwasów. Te tzw. **α -aminokwasy** **[PLANSZA 74: struktura chemiczna aminokwasu]** mają przy węglu oznaczonym α dwie grupy funkcyjne: **aminową NH₂** i **karboksylową COOH**, a poszczególne aminokwasy różnią się typem **reszty bocznej R**. Każdy aminokwas ma więc co najmniej dwie wartości pK, odpowiadające protonacji do NH₃⁺ i dysocjacji do COO⁻. W pH zbliżonym do obojętnego dominują więc formy jonów obojętnych (ze względu na wymienione grupy). Wartość pH w której wszystkie cząsteczki są obojętne nosi nazwę **punktu izoelektrycznego pHi**. Wylicza się ten punkt jako średnią wartość z dwóch pK: $1/2(pK_1 + pK_2)$ ($pK_1 \sim 1,71 \div 2,38$, $pK_2 \sim 8,80 \div 10,78$), z wyjątkiem aminokwasów zasadowych z dodatkową grupą zasadową w R: lizyny ($pK_3 = 10,53$), argininy ($pK_3 = 12,84$) i histydyny ($pK_3 = 9,17$), gdzie $pH_i = 1/2(pK_2 + pK_3)$. Węgiel C α wszystkich aminokwasów z wyjątkiem glicyny jest asymetryczny i aminokwasy mogą występować w jednej z form izomerycznych L lub D. Wszystkie aminokwasy włączane do łańcucha białkowego w procesie translacji mRNA są **typu L**: po ustawieniu trzech węgli C-karbonylowy, C α i C β w płaszczyźnie, grupa aminowa jest skierowana "do przodu" a proton "do tyłu". Można to również przedstawić w określonej konwencji na płaskim rysunku. Jeśli w aminokwasach występują dodatkowe izomerie optyczne ze względu na inne węgle podstawione czterema różnymi podstawnikami, to tylko jeden typ izomeru jest wbudowywany do białka.

Aminokwasy w biochemii oznacza się standardowymi trzyliterowymi skrótami **[PLANSZA 75: typy reszt bocznych aminokwasów]**, a na potrzeby komputerowych baz danych sekwencji białkowych skrótami jednoliterowymi. W momencie utworzenia przez aminokwasy łańcucha peptydowego o **ich własnościach w białku decydują reszty boczne**:

(A) Aminokwasy **alifatyczne, hydrofobowe**, których łańcuchy boczne preferują kontakty między sobą lub ze środowiskiem niepolarnym a unikają wody: glicyna, alanina, walina, leucyna, izoleucyna i prolina. **Prolina jest iminokwasem** z drugorzędową grupą aminową, po połączeniu między C δ reszty bocznej i azotu amidowego. Stosunkowo często spotyka się w białkach także **γ -hydroksyprolinę** (modyfikacja post-translacyjna; patrz dalej).

(B) Aminokwasy **zasadowe**: lizyna, arginina i histydyna, w dwóch pierwszych reszty boczne w białkach są zwykle **naładowane dodatnio**, w histydynie pK jest bliskie pH fizjologicznego i może mieć ładunek dodatni lub nie, zależnie od otoczenia białkowego.

(C) Aminokwasy **kwaśne, ujemnie naładowane** z grupą karboksylową w reszcie bocznej: kwas asparaginowy i kwas glutaminowy.

(D) Aminokwasy **amidowe, silnie hydrofilowe**, ale niezjonizowane, preferujące środowisko wodne: asparagina i glutamina; ich grupy boczne CONH₂ są jednocześnie donorami i akceptorami protonów w wiązaniach wodorowych.

(E) Aminokwasy zawierające **siarkę**, słabo hydrofobowe; cysteina i metionina, oraz zawierające hydroksyl, słabo hydrofilowe: seryna i treonina; grupa SH cysteina jest bardzo reaktywna i połączenie (utlenienie) dwóch grup cysteinowych daje **mostek dwusiarczkowy –S–S–**, wiązanie kowalencyjne, które stabilizuje struktury białkowe (aminokwas taki nazywa się cystyną).

(F) Aminokwasy zawierające **pierścienie aromatyczne**: fenyloalanina, tyrozyna i tryptofan; dwie ważne ich cechy to:

(1) **absorpcja i fluorescencja w zakresie UV [PLANSZA 76: widma absorpcji UV-VIS i fluorescencji aminokwasów]**: białka absorbują promieniowanie z max. 280 nm (kwasy nukleinowe 260 nm) i wydajnie emitują promieniowanie; zwykle emisja tryptofanowa wskutek FRET: Tyr → **Trp** (emisja tyrozynowa zachodzi przy braku Trp);

(2) wchodzenie w oddziaływania warstwowe (stacking) π - π , lub w tzw. oddziaływania kation - π z grupami dodatnio naładowanymi.

(G) Aminokwasy 21 i 22: **pyrolizyna (ozn. Pyl, O)** z podstawnikiem pyrrolowym zamiast protonu na grupie aminowej reszty bocznej lizyny i **selenocysteina (Sec, U)**, z atomem Se w miejscu siarki **[PLANSZA 77: "nietypowe aminokwasy" w archea i eubacteria]**.

W łańcuchach bocznych aminokwasów mogą zachodzić **obroty wokół wiązań pojedynczych**, które prowadzą do różnych konformacji reszt bocznych; przykładowe konformacje wokół wiązania C α -C β pokazuje **[PLANSZA 78: konformacje aminokwasów]**. W swobodnych aminokwasach w roztworze występuje równowaga dynamiczna między klasycznymi konformerami, a odpowiednie wartości populacji można wyznaczyć ze stałych sprzężenia proton - proton. Konformację **pierścienia proliny** można opisać na gruncie modelu pseudorotacji analogicznie jak konformację pierścienia cukrowego **[PLANSZA 44]** zamieniając tylko tlen O(4') na azot. W roztworze obserwuje się równowagę dynamiczną konformacji N (bliska C γ endo) i S (bliska C γ exo) o

populacjach po ok. 50%. Jedynym **problemem tautomerycznym** jest migracja protonu $N(1)H \leftrightarrow N(3)H$ w (płaskim) pierścieniu histydy w nieuprotowanej [PLANSZA 75]. W formie zwitterjonowej stosunek populacji obu tautomerów jest 1 do 3.

Aminokwasy łączą się w łańcuch peptydowy poprzez utworzenie **wiązania peptydowego** [PLANSZA 79: **wiązanie peptydowe**] między grupą karboksylową pierwszego i grupą aminową następnego w procesie **translacji mRNA na rybosomie** [PLANSZA 36]. Powstaje w ten sposób liniowy polimer [PLANSZA 80: **łańcuch peptydowy**] o określonej **kolejności czyli sekwencji aminokwasów w kierunku od aminokwasu z wolną grupą aminową (N-koniec) do aminokwasu z wolną grupą karboksylową (C-koniec)**. Kolejność aminokwasów w łańcuchu peptydowym określa informacja zapisana w nici kodującej DNA jako **kod genetyczny** [PLANSZA 81: **kod genetyczny**]. Piryrolizyna jest kodowana przez kodon UAG, a selenocysteina przez UGA [PLANSZA 77], które są **kodonami „stopu” translacji** i ulegają „przeprogramowaniu” w przypadkach kodowania aminokwasów. Rozwinięty **łańcuch peptydowy zwija się** (*protein folding*) do funkcjonalnej formy natywnej stosownie do informacji zapisanej w sekwencji aminokwasów już w trakcie translacji (zwijanie *co-translacyjne*) oraz po translacji (*post-translacyjne*), często przy pomocy różnych mechanizmów komórkowych podnoszących **efektywność zwijania** (*szczegóły: zwijanie in vivo*). Jednym z nich jest pomoc ze strony **białek opiekuńczych** czyli chaperonów, np. kompleks chaperoninowy **GroEL/GroES *E. coli*** [PLANSZA 82: **zwijanie i degradacja białek: chaperonin GroEL/ES i proteasom**]. W trakcie zwijania tworzą się **mostki dwusiarczkowe** stabilizujące strukturę trzeciorzędową pomiędzy ściśle określonymi (jeśli jest więcej niż dwie) resztami cysteinowymi, katalizowane (*shuffling* par cystein i reakcja łączenia) przez izomerazę *protein-disulfide isomerase*, przy udziale tripeptydu glutationu jako koenzymu.

Białka „niepotrzebne” są **degradowane do aminokwasów w strukturalnie podobnym do chaperoninowego kompleksie proteasomu** po wyznakowaniu białkiem ubikwityną.

Białka są *co-* i *post-translacyjnie* modyfikowane do ostatecznej formy w procesach:
(A) **Usunięcie pierwszej metioniny** przez Met-aminopeptydazę w eukariotach i kolejne usuwanie grupy formylowej z metioniny (deformylaza) i samej metioniny (Met-aminopeptydaza) w prokariotach.

(B) Proteoliza **sekwencji sygnałowych** (*signal* lub *transit peptide (sequence)*) czyli usunięcie jednego lub większej ilości fragmentów kilkunastu/kilkudziesięciu aminokwasów na N-końcu. Sekwencje sygnałowe decydują o miejscu przeznaczenia

(funkcjonowania) białka i proteoliza jest powiązana z przechodzeniem białka przez błony w procesie **topogenezy** [**PLANSZA 83: topogeneza**].

(C) **Proteolityczne wycinanie fragmentów** wewnętrznych przez proteazy (*PRO*protein → protein), np. zymogenów enzymów proteolitycznych przewodu pokarmowego do enzymów aktywnych, proinsuliny do insuliny [**PLANSZA 84: fragmenty wycinane w proinsulinie**], **rozcinanie poliprotein wirusowych do odrębnych białek** (HIV [**PLANSZA 32**])

(D) **Składanie** (*splicing*): po translacji może zachodzić wewnątrzcząsteczkowa reorganizacja poprzez samousuwanie sekwencji inteinowych i łączenie ekstein (podobnie jak w RNA).

(E) W nielicznych przypadkach **posttranslacyjne dodanie** leucyny lub fenyloalaniny na N-końcu przekazywanej z tRNA (*E. coli*).

(F) **Fosforylacja** przez kinazy białkowe grup OH seryny, tyrozyny, treoniny, rzadziej grup aminowych asparaginy, histydyny lub lizyny.

(G) **Modyfikacja łańcuchów bocznych**: wykryto w białkach ponad 200 różnych typów aminokwasów, np. γ -hydroksypolina [**PLANSZA 76**].

(H) **Dodanie** części różnych od aminokwasów w białkach **złożonych**: **glikoproteidy**, **nukleoproteidy**, **lioproteidy**, **chromoproteidy**, itp.

Peptydy o nietypowej strukturze, np. zawierające D-aminokwasy, wiązania kowalencyjne między aminokwasami różne od wiązań peptydowych są syntetyzowane nierybosomalnie (antybiotyki).

WYKŁAD 8

Białka globularne, włókniste i błonowe (2)

Ilość aminokwasów w białku może być bardzo różna. W komórce funkcjonują już krótkie peptydy o długości kilku aminokwasów: dwupeptyd karnozyna (β Ala-His) w mięśniach, tripeptyd glutation: Glu-Cys-Gly, 9-peptyd oksytocyna (hormon). „Klasyczne” białko BPTI (*basic pancreatic trypsin inhibitor*) wszechstronnie badane pod względem strukturalnym rozwijanymi metodami 2D NMR **[PLANSZA 21]** ma 58 aminokwasów, ale znana jest rodzina białek CMTI (*cucurbita maxima trypsin inhibitor*) ok. **30 aminokwasów**; w przypadku obu typów białek zachodzi stabilizacja przez 3 mostki dwusiarczkowe, podobnie jak w 50 AA insulinie **[PLANSZA 84]**. Średnia długości białek globularnych to 300 ÷ 400 aminokwasów, ale spotykane są bardzo długie białka rzędu 3 000 aminokwasów np. kolagen (długość 28 000Å; MW ok. 300 000 Da).

Białka pełnią w komórce bardzo różne funkcje:

- (1) **materiał budulcowy** struktur komórkowych;
- (2) **enzymy**, katalizatory reakcji biochemicznych;
- (3) **przeciwciała**, funkcje odpornościowe organizmu;
- (4) **transport** aktywny przez błony cząsteczek organicznych i jonów;
- (5) **energetyka**, udział w procesach oddychania i fotosyntezy;
- (6) **praca mechaniczna** kosztem energii chemicznej;
- (7) **regulacja** działania genów, enzymów i procesów hormonalnych.

Ze względu na olbrzymią liczbę różnych białek nie można podać prostej **klasyfikacji** w oparciu o strukturę czy funkcje. Najwcześniejsza klasyfikacja dzieliła białka na **globularne**, rozpuszczalne w wodzie i **strukturalne** (włókniste, fibrylarne), nierozpuszczalne w wodzie i odporne na działanie czynników denaturujących oraz proteaz. Ze względu na strukturę chemiczną można podzielić białka na **proste**, które zawierają tylko aminokwasy oraz **złożone**, w których obok aminokwasów występują dodatkowe składniki. Do pierwszej grupy należą m. in. albuminy, globuliny, histony, skleroproteidy (kolagen, elastyna). Wśród białek złożonych można wyróżnić nukleoproteidy (zawierające fragmenty kwasów nukleinowych), glikoproteidy (zawierające oligosacharydy), lipoproteidy (zawierające tłuszcze), chromoproteidy (hemoglobina, cytochrom) itd.

Na wykładzie będą omówione trzy klasy białek usystematyzowane pod względem strukturalnym:

(A) globularne

(B) strukturalne

(C) błonowe

Określenie kolejności aminokwasów od N- do C-końca dokonuje się metodami **sekwencjonowania** (podobnie jak w kwasach nukleinowych). W starszej i obecnie rzadziej stosowanej metodzie **degradacji Edmana** (nie będzie omawiana) odszczepia się kolejne aminokwasy od łańcucha peptydowego i identyfikuje chromatograficznie. Na przełomie lat osiemdziesiątych i dziewięćdziesiątych opracowano nową metodę sekwencjonowania peptydów, rozwiniętą następnie do sekwencjonowania całych białek i mieszanin (kompleksów) wielu białek w bardzo efektywny sposób **metodą spektrometrii mas** [**PLANSZA 14**];. Przykładowo, grupa M. Mann'a z EMBL (European Molecular Biology Laboratory) scharakteryzowała 19 nowych białek kompleksu spliceosom'u komórek HeLa (*Neubauer et al. Nature Genetics* 20, 46, 1998). Białka poddaje się:

(a) elektroforezie na żelu,

(b) trawieniu peptydazami na żelu np. **trypsyną**,

(c) ekstrakcji buforem i skierowaniu mieszaniny peptydów do spektrometru masowego, pracującego w systemie **tandemu MS/MS** [**PLANSZA 85: sekwencjonowanie białek metodą MS**]. Pierwszy spektrometr masowy jest wykorzystywany jako **filtr masowy** do selekcji kolejnych peptydów, najczęściej obdarzonych ładunkiem +2. Peptyd przepuszczany do komory kolizyjnej ulega **kontrolowanej fragmentacji** przy zderzeniu z cząsteczkami gazu (CID). W jej wyniku powstają **fragmenty różniące się kolejno o jeden aminokwas** (Y_n), w wyniku pęknięcia wiązań peptydowych od N-końca z jednakowym, najczęściej pojedynczym ładunkiem dodatnim. W drugim spektrometrze następuje rozdział i powstaje widmo fragmentów, a z różnicy mas identyfikuje się kolejne aminokwasy łańcucha, z wyjątkiem pary leucyna/izoleucyna. Peptydy o znanych sekwencjach składa się następnie w całe białko na podstawie np. przeszukiwania programem **PeptideSearch bazy danych sekwencji nukleotydowych genomu EST** (*expressed sequence tags*) pod kątem „unikalnych” sekwencji peptydowych. Niezwykle wydajna (*high throughput*) technika sekwencjonowania MS jest podstawą w badaniach w zakresie **proteomiki**.

Konformacja łańcucha peptydowego jest uwarunkowana charakterystycznymi cechami wiązania peptydowego C–N [**PLANSZA 79**]. Ma ono częściowo **podwójny**

charakter co widać z jego długości 1.33 Å, pomiędzy 1.45 Å dla typowo pojedynczego i 1.25 Å dla typowo podwójnego. Często strukturę chemiczną fragmentu przedstawia się w postaci równowagi dwóch form mezomerycznych o różnym rozkładzie ładunków i wiązań. Układ złożony z dwóch atomów wiązania i czterech atomów podstawników jest płaski **[PLANSZA 80]**. Konformację wokół wiązania peptydowego opisuje kąt ω , który może przyjąć wartość **bliską 180°**, najczęstsze transoidalne ustawienie węgli C α , lub wartość bliską 0° z cisoidalnym ustawieniem obu węgli **[PLANSZA 79]**. Struktura każdego z układów jest sztywna, a przejście konformacyjne między nimi poprzez **rotację jest zahamowane**. Wysokość bariery szacuje się na ok. **60 ÷ 120 kJ/mol**. Forma trans jest energetycznie uprzywilejowana ok. **8.5 kJ/mol** ze względu na niekorzystne sterycznie oddziaływania w formie cis atomów przy C β danego aminokwasu i C α aminokwasu poprzedzającego w sekwencji **[PLANSZA 86: konformacja cis – trans wokół wiązania peptydowego]**. W przypadku wiązania peptydowego **poprzedzającego prolinę** takie oddziaływanie sterycznie niekorzystne występuje w obu formach i ich energia jest zbliżona. Wiązania peptydowe przed proliną mogą mieć $\omega = 0^\circ$; a w bardzo rzadkich przypadkach także w innych aminokwasach.

Konformację łańcucha peptydowego opisują wartości dwóch kątów dwuściennych dla każdego i-tego aminokwasu **[PLANSZA 87: konformacja łańcucha peptydowego]**:

(1) **kąt Φ** między węglami karbonyłowymi C' we fragmencie C' $_i$ – C α_i – N $_i$ – C' $_{i-1}$, gdzie wartości dodatnie są przy obrocie zgodnym z ruchem wskazówek zegara patrząc wzdłuż C $\alpha_i \rightarrow N_i$ czyli przeciwnie do kierunku łańcucha;

(2) **kąt Ψ** między azotami amidowymi we fragmencie N $_i$ – C α_i – C' $_i$ – N $_{i+1}$, gdzie wartości dodatnie są przy obrocie zgodnym z ruchem wskazówek zegara patrząc wzdłuż C $\alpha_i \rightarrow C'_i$ czyli zgodnie z kierunkiem łańcucha.

Dopuszczalne wartości kątów Φ i Ψ wynikają głównie z **oddziaływań sterycznych atomów wokół tych wiązań i reszt bocznych aminokwasów typu 1 – 4** („trzeci sąsiad”) i ich zakresy silnie redukują się w porównaniu z zakresami dostępnymi dla pojedynczych aminokwasów lub bardzo krótkich peptydów. Można to zobrazować na tzw. **wykresach Ramachandrana** **[PLANSZA 88: diagram Ramachandrana]**, które prezentują mapy poziomicowe dozwolonych obszarów kątów Φ i Ψ lub dopuszczalne zakresy kątów Φ i Ψ z energią na trzeciej osi. Z map Ramachandra wynika, że łańcuch przyjmuje zasadniczo **dwa typy konformacji: helikalną, zwiniętą prawoskrętnie (kąty Φ i Ψ po ok. -60°) i**

rozciągniętą (kąty Φ i Ψ po ok. 120 - 140°) **[PLANSZA 89: parametry struktur drugorzędowych białek]**. Helisy zwinięte lewoskrętne są teoretycznie możliwe, ale nie występują ze względu na sterycznie niekorzystne oddziaływania łańcuchów bocznych aminokwasów w konfiguracji L. Poszczególne fragmenty łańcucha peptydowego o różnej długości zwijają się więc **lokalnie w różne struktury drugorzędowe**. Wyróżnić tu można kilka nieco różnych konformacji helikalnych, z których główne to **[PLANSZA 90: struktury helikalne białek]**:

- (a) prawoskrętna, klasyczna **helisa α** , oznaczana często jako 3.6₁₃;
- (b) prawoskrętna, klasyczna helisa 3₁₀;
- (c) prawoskrętna, klasyczna **helisa π** , oznaczana często jako 4.4₁₆.

Helisy te określane są jako klasyczne gdyż każdy aminokwas w takiej strukturze drugorzędowej ma jednakowe wartości kątów Φ i Ψ **[PLANSZA 89]**. Poszczególne helisy różnią się stopniem zwinięcia: najbardziej zwinięta jest helisa 3₁₀ z 3 aminokwasami na jeden skręt, dalej helisa α z 3.6 aminokwasami na skręt i helisa π z 4.4 aminokwasami na skręt (symbole rodzaju helisy). Helisa jest stabilizowana przez wiązania wodorowe, w których donorem protonu jest azot amidowy wiązania peptydowego **[PLANSZA 79]**, a akceptorem tlen karbonylowy tego wiązania. W helisie 3₁₀ takie wiązanie występuje między każdym **i-tym** i **i+3-cim** aminokwasem (1-4, 2-5, 3-6 itd), w helisie α między **i-tym** i **i+4-tym**, a w helisie π między **i-tym** i **i+5-tym**. Wiązanie wodorowe "zamyka" jak gdyby fragment łańcucha **[PLANSZA 90]**, a ilość atomów wynosi odpowiednio 10, 13 i 16 dla trzech typów helis, co wyjaśnia znaczenie subskryptu w symbolu każdej z helis. Helisy w białkach mogą być różnej długości; najkrótsza musi umożliwić powstanie chociaż jednego stabilizującego wiązania wodorowego (odpowiednio 4, 5 i 6 aminokwasów). W białkach globularnych helisy mają długości średnio 10 ÷ 15 aminokwasów. ale spotyka się niekiedy nawet do 50 aminokwasów długości.

Dwa lub więcej fragmentów rozciągniętego łańcucha peptydowego, tzw. **β -nici** może ustawiać równoległe (kierunki N→C takie same) lub antyrównoległe (kierunki N→C przeciwne) do siebie, tworząc strukturę płaską **β -kartki (ang. β -sheet) [PLANSZA 91: struktury β - karteł białek]**. Taka struktura drugorzędowa jest **stabilizowana przez wiązania wodorowe między niciami** z udziałem grup NH i C=O wiązań peptydowych. Kąty dwuścienne **[PLANSZA 89]** nie odpowiadają niciom maksymalnie rozciągniętym Φ , $\Psi = 180^\circ$, ale odbiegają od nich, przez co β -kartka jest **pofołdowana**. Najczęściej

występują β -kartki antyrównoległe, złożone z 2 ÷ 3 nici, rzadziej mieszane, częściowo równoległe, częściowo antyrównoległe od 3 do 15 nici, a najrzadziej równoległe, zwykle złożone z ponad 4 nici. Obserwuje się również jedno- i trzyniciowe **β -helisy** i **β -spirale** (*Schultz & Ficner Curr. Op, Struct. Biol. 21, 232, 2011*)

Trzecim typem struktur drugorzędowych są **ciasne pętle**, tzw. **β -skręty** (β -szpilki) i **γ -skręty** [**PLANSZA 92: struktury ciasnych skrętów białek**] na końcach struktur antyrównoległej β -kartki, gdy łańcuch peptydowy wykonuje ciasny skręt. W skręcie γ jeden aminokwas wypada z układu wiązań wodorowych, a w przypadku skrętu β wypadają dwa aminokwasy. Struktury takich skrętów są dobrze zdefiniowane poprzez wartości kątów Φ i Ψ [**PLANSZA 89**]. Istnieje ok. 8 różnych rodzajów β -skrętów różniących się konformacją; dane na kąty dwuścienne przytoczone są dla jednego typu tej struktury.

Statystycznie, w białkach globularnych [**PLANSZA 93: cząsteczka białka globularnego eIF4E**] ok. 31% aminokwasów przyjmuje struktury helikalne, 28% przyjmuje struktury β -kartek, a pozostałe występują w pętlach, z których tylko ciasne skręty β i γ uważa się za dobrze zdefiniowane struktury drugorzędowe. Dłuższe pętle są **obszarami nieustrukturowanymi drugorzędowo**. Helisy **oznacza się** spiralami lub walcami natomiast β -nici tworzące β -kartki w postaci strzałek wskazujących kierunek N→C.

Prolina ze względu na brak protonu amidowego nie może tworzyć wiązań wodorowych stabilizujących helisy i β -kartki. Poliprolina tworzy dwa odrębne typy regularnej struktury drugorzędowej: poliprolina-I ze wszystkimi wiązaniami peptydowymi w formie cis i poliprolina-II ze wszystkimi wiązaniami peptydowymi trans [**PLANSZA 89**]. Forma I jest prawoskrętną helisą z 3,3 AA na skręt a forma II lewoskrętną helisą z 3,0 AA na skręt. **Helisę II zaobserwowano w funkcjonalnych białkach globularnych, m. in. w inhibitorze trypsyny trzustkowej i cytochromie c₅₅₁**. **Poliglicyna** może także przyjmować nietypowe formy drugorzędowe ze względu na brak łańcucha bocznego w glicynie (giętkość konformacyjna), np. **formę poliglicyny-II podobną do poliproliny-II**.

Struktury drugorzędowe organizują się przestrzennie w tzw. **globalny fold pojedynczego łańcucha peptydowego** czyli **strukturę trzeciorzędową**. Struktury trzeciorzędowe białek (a dokładniej domen, patrz dalej) próbuje się od dawna

klasyfikować i w najprostszej klasyfikacji ze względu na struktury drugorzędowe wyróżnia się **[PLANSZA 94: klasyfikacja struktur trzeciorzędowych białek]**:

- białka **klasy α** z dominacją form helikalnych;
- białka **klasy β** z dominacją form β -karkowych;
- białka **klasy $\alpha+\beta$** oraz **α/β** z różnymi rozkładami obu podstawowych form.

Znacznie bardziej „zaawansowana” **klasyfikacja typów foldów trzeciorzędowych pod względem topologicznym** stara się uchwycić podobieństwa strukturalne i powtarzalne formy w bardzo różnych białkach i zebrać je w **bazy danych**. Bazy **SCOP** i **CATH** opierają się na **podejściu hierarchicznym** z kolejnymi poziomami uogólnienia podobieństwa topologicznego. Jest to szczególnie istotne w **zawodach CASP** (*critical assesment of structure prediction*) polegających na teoretycznym przewidywaniu (modelowaniu) struktur przestrzennych białek o znanych sekwencjach i niepublikowanych strukturach wyznaczonych doświadczalnie (szerzej omówione w temacie „zwijanie białek”). W chwili obecnej rozpoznanych jest **ok. 700 różnych foldów, ale ich ilość wzrasta** w miarę odkrywania nowych struktur; kilka przykładowych dla białek enzymatycznych podaje **[PLANSZA 95: przykładowe typy foldów trzeciorzędowych białek enzymatycznych]** (szerzej podczas ćwiczeń).

Ważną cechą topologiczną większości białek (globularnych) jest brak „węzłków” w przebiegu łańcucha peptydowego – pociągnięcie za N-koniec i C-koniec daje niezapętłony, rozciągnięty łańcuch. W miarę rozwiązywania kolejnych struktur białkowych rośnie liczba wyjątków od tej prawidłowości. W chwili obecnej znanych jest **ponad 100** struktur trzeciorzędowych zawierających węzły. Przykładowa struktura krystalograficzna syntetazy adenosylometioniny (MAT) pokazuje węzeł utworzony przez β -nić B9 przez pętelkę B1-[central domain]-B5-H3-B6 **[PLANSZA 96: węzeł na łańcuchu peptydowym]**, najprawdopodobniej wskutek utworzenia β -kartki przez nici B1 i B6 przy elongacji B9 podczas procesu translacji.

WYKŁAD 9

Białka globularne, włókniste i błonowe (3)

Dla białek zbudowanych z jednej nici peptydowej struktura trzeciorzędowa jest ostateczną, globalną strukturą. Funkcjonalne białka mogą być zbudowane z kilku, najczęściej jednakowych nici peptydowych czyli **podjednostek** [PLANSZA 97: **struktura czwartorzędowa białek**], każda o własnej strukturze trzeciorzędowej, połączonych specyficznymi oddziaływaniami. Takie białka mają określoną **strukturę czwartorzędową**. Ilość podjednostek może być różna: dwie, trzy, cztery, pięć, sześć (heksamer insuliny), osiem i bardzo wiele w otoczkach wirusów [PLANSZA 33].

W zależności od roli i środowiska funkcjonowania następuje wyraźny podział białek pod względem charakterystyki strukturalnej na poziomie struktury trzeciorzędowej. **BIAŁKA GLOBULARNE** przyjmują fold trzeciorzędowy w wyniku określonych, głównych oddziaływań stabilizujących plus pozostałe - van der Waalsa [PLANSZA 9].

(A) Oddziaływania hydrofobowe reszt bocznych aminokwasów alifatycznych, które prowadzą do powstania tzw. **jądra hydrofobowego** [PLANSZA 98: **struktura NMR interleukiny-4**] białek globularnych. Charakter hydrofobowy ma także **oddziaływanie warstwowe** (stacking) aromatycznych pierścieni takich aminokwasów jak tryptofan, tyrozyna i fenyloalanina oraz oddziaływania typu kation- π dodatnio naładowanych łańcuchów bocznych (lizyna, arginina) i pierścieni aromatycznych.

(B) Aminokwasy polarne, a przede wszystkim naładowane występują w przewodzie na powierzchni białka, zapewniając elektrostatycznie korzystne **oddziaływanie z wodą i jonami rozpuszczalnika**. Woda jest w znacznym stopniu wyparta z wnętrza białek globularnych i może tam występować jedynie w formie klastrów niewielkiej ilości cząsteczek.

(C) Pary aminokwasów naładowanych o przeciwnych znakach wewnątrz białka, np. dodatnio naładowana lizyna czy arginina i ujemnie naładowany kwas asparaginowy czy glutaminowy, wiążą się oddziaływaniem coulombowskim w tzw. **mostki solne**.

(D) Aminokwasy hydrofilowe wewnątrz białka tworzą **wiązania wodorowe**.

(E) Kowalencyjne wiązania między grupami $-SH$ cystein **S-S**, szczególnie w przypadku mniejszych białek, np. najkrótsze (najprawdopodobniej) białka o regularnej, natywnej strukturze trzeciorzędowej: CMTI ok. 30 AA, insulina 50 AA [PLANSZA 84] oraz inhibitor trypsyny BPTI 58 AA, mają po 3 mostki dwusiarczkowe.

Struktura trzeciorzędowa mniejszych białek jest najczęściej "zwarta" w pofałdowaną formę mniej lub bardziej sferyczną. W przypadku białek większych, ok. 200 aminokwasów, może pojawić się struktura podzielona na **domeny** **[PLANSZA 99: struktura wielodomenowa białka globularnego]**. Każda domena tworzy zwartą strukturę z własnym jądrem hydrofobowym, a połączenia między domenami mają luźniejszy charakter i umożliwiają większe ruchy molekularne domen względem siebie (trzeba zwrócić uwagę na różnice pojęć domena - podjednostka). **Powierzchnia** cząsteczki białkowej jest dobrze zdefiniowana geometrycznie; wymiar:

$$A_s = 6.3(M)^{0.73} [\text{Å}^2] \quad M - \text{masa cząsteczkowa}$$

co stanowi 25 ÷ 45% powierzchni rozwiniętego łańcucha peptydowego i odgranicza od rozpuszczalnika i innych cząsteczek **wnętrze białka** o objętości V, gęsto upakowane (75%) wypełniającymi atomami:

$$V = 1.27M [\text{Å}^3]$$

Definiuje się dwa typy powierzchni. Wypadkowa sfer van der Waalsa **[PLANSZA 5, PLANSZA 7]** połączonych kowalencyjnie atomów, tzw. powierzchnia kontaktu (*contact surface*) tworzy **powierzchnię molekularną** (*molecular surface*). **Powierzchnie dostępu** (*accessible surface*) definiuje się jako położenia środka (miejsce geometryczne) sondy wodnej w formie sfery o promieniu 1,4 Å toczonej po powierzchni białka. Wnętrze białka globularnego jest. **Pojedyncze cząsteczki** lub niewielkie **klastry wody** we wnętrzu białka są unieruchomione w wydrążeniach. Stanowią tzw. ustrukturyzowaną wodę, która może tworzyć wiązania wodorowe z polarnymi resztami wewnętrznych aminokwasów.

Niekiedy pomiędzy strukturą drugorzędową i trzeciorzędową wyróżnia się **struktury pośrednie** tzw. **superdrugorzędowe** **[PLANSZA 100: struktury superdrugorzędowe]**. Są to pewne charakterystyczne i często spotykane motywy białkowe, zbudowane z kilku elementów drugorzędowych blisko siebie położonych w sekwencji, np.:

- **motyw β-α-β**: helisa łącząca dwie nici równoległej β-kartki,
- **klucz grecki**: dwu- i trzyniciowa β-kartka, ustawione równoległe do siebie.

Szereg **motywów rozpoznających kwasy nukleinowe** (dalsza część wykładu) można klasyfikować jako struktury superdrugorzędowe, np. motyw helisa-skრეტ-helisa.

BIAŁKA STRUKTURALNE (włókniste, fibrylarne), w przeciwieństwie do globularnych, charakteryzują się nierozpuszczalnością w wodzie, odpornością na czynniki

denaturujące i enzymy proteolityczne. Struktury drugorzędowe są w białkach strukturalnych takie jak w globularnych, ale różnią się zdecydowanie strukturą trzeciorzędową. Są to przeważnie bardzo wydłużone struktury w jednym kierunku, gdzie jedna lub kilka typów struktur drugorzędowych powtarza się. **Struktura trzeciorzędowa i czwartorzędowa są jakby przedłużeniem struktur(y) drugorzędowych w jednym kierunku** (kierunku włókna).

(A) **Kolagen** [**PLANSZA 101: struktura kolagenu**], składnik szeregu tkanek, zawiera **powtarzalne sekwencje aminokwasowe typu (-Gly-Xaa-Yaa-)_n** ze znaczną zawartością proliny w miejscach Xaa i Yaa i lizyny w Yaa. Prolina jest często hydroksylowana w pozycji γ do hydroksyproliny. Kolagen tworzy strukturę **potrójnej helisy**: trzy łańcuchy peptydowe, każdy w formie helisy poliproliny-II z 3 AA na skręt [**PLANSZA 89**], są zwinięte prawoskrętnie względem siebie, a skok wynosi ok. 30 Å, tzn. 10 aminokwasów. Struktura ta jest stabilizowana przez wiązania wodorowe między łańcuchami w helisie (NH glicyny, C=O aminokwasu Xaa), ze znacznym udziałem hydroksyli hydroksyprolin jako donorów protonu. Obserwowano cząsteczki kolagenu o długości do 28 000 Å przy szerokości helisy 14 Å.

(B) Szereg **cytoplazmatycznych białek nadających kształt komórkom** tworzy struktury **superhelikalne**, tzw. helisa helis (coiled coils), w których [**PLANSZA 102: amfipatyczność helis białkowych**] dwie prawoskrętne α -helisy zwinięte są równolegle i lewoskrętne względem siebie. Stabilizacja układu superhelikalnego zachodzi dzięki **amfipatyczności** helis. W strukturze pierwszorzędowej występują powtarzające się kolejno sekwencje siedmiu aminokwasów (ok. 2 razy skok helisy α , 3,6 AA/skręt): **a-b-c-d-e-f-g-**, gdzie dwa aminokwasy **a** i **d** są hydrofobowe (Leu, Ile, Ala), **e** i **g** naładowane przeciwnie (Glu, Arg), a pozostałe dowolne. Jedna strona helis jest więc hydrofobowa co umożliwia ich łączenie, wspomagane przez przyciąganie elektrostatyczne aminokwasów naładowanych. W ten sposób mogą powstawać także superhelisy z trzech i czterech helis.

(C) **Jedwab** jest zorganizowany w segmenty β -karkowe o regularnej strukturze pierwszorzędowej, złożonej głównie z Gly, Ala i Ser (sekwencje długości ok. 60 AA), poprzedzielane obszarami o słabo zdefiniowanej strukturze (przyjmowano, że jest nieregularna). Badania NMR w ciele stałym jedwabiu pajaków [**PLANSZA 103: struktura NMR jedwabiu pajaków**] potwierdzają regiony β -karkowe. Dla pozostałych obszarów, również bogatych w glicynę, została wymodelowana nowa **struktura helisy 3₁**

(nie mylić z helisą 3_{10}), zgodnie ze ścisłym oznaczeniem symetrii śruby prawoskrętnej **[PLANSZA 49]**; wartości kątów $\Phi = -60^\circ$ i $\Psi = 135^\circ$.

BIAŁKA BŁONOWE są strukturalnie podobne do białek globularnych, ale naturalne środowisko przynajmniej dla fragmentu cząsteczki ma charakter hydrofobowy: podwójna warstwa lipidowa, złożona z reszt kwasów tłuszczowych (szczegółowe omówienie w części poświęconej lipidom). Fragmenty białek błonowych mogą wystawać do środowiska wodnego i stosować się do reguł budowy charakterystycznych dla białek globularnych. Jednak jako całość białka te agregują w wodzie i są w niej nierozpuszczalne, ze względu na **hydrofobowość powierzchni kontaktu z błoną**. Dopiero ostatnio udało się opracować metody krystalizacji białek tego typu w środowisku buforowym z udziałem detergentów i ilość znanych struktur jest stosunkowo niewielka (> 150). Obserwuje się dwa podstawowe typy foldów, które zapewniają możliwość ich funkcjonowania jako kanałów i transporterów przez błonę **[PLANSZA 104: struktury białek błonowych]**:

- wiązka helis
- β -baryłka

Białka te są zwykle złożone z **dużej ilości aminokwasów**. Przykładem może być 492 aminokwasowy transporter glukozy przez błonę erytrocytów **[PLANSZA 105: transporter glukozy]**. Helisy struktury drugorzędowej białka tworzą transportery do aktywnego transportu przez błonę: wyłapywanie cząsteczki cukru na zewnątrz i wypuszczenie w środku komórki. W białku można zaobserwować charakterystyczny rozkład aminokwasów hydrofobowych, polarnych oraz dodatnio i ujemnie naładowanych, w rodzaju amfipatyczności. Naładowane i polarne znajdują się głównie we fragmentach kontaktujących się z wodą, ale i tam występują hydrofobowe dla utworzenia jąder hydrofobowych odpowiednich, wystających z błony domen. Z kolei znaczna przewaga aminokwasów hydrofobowych występuje we fragmentach zaczepionych w błonie, ale i tu występują w odpowiednich miejscach aminokwasy polarne, stabilizujące strukturę kanałów i ułatwiające wnikanie polarnych cząsteczek cukru.

Oddziaływania molekularne stabilizujące białka błonowe są stosunkowo słabo poznane i są ciągle przedmiotem kontrowersji. Zakotwiczone w błonie części mają wysoce stabilne struktury odporne na termiczną denaturację. Postuluje się m. in., że **ograniczenie efektu hydrofobowego** wewnątrz struktury (w porównaniu z białkami globularnymi) jest kompensowane przez lepsze upakowanie reszt bocznych oraz większą stabilizację

wiązaniami wodorowymi aminokwasów polarnych w środku, w tym **slabymi wiązaniami wodorowymi typu** $C_{\alpha}-H\dots O=C$.

Wyznaczanie struktur białek z rozdzielczością atomową metodami rentgenografii, NMR i przez zwijanie komputerowe jest naczelnym zadaniem proteomiki strukturalnej. Struktury te w postaci **położeń poszczególnych atomów** zapisywane są w bazach danych w standardowych formatach tzw. **PDB-files**. W chwili obecnej znajduje się tam ponad 100 tys. struktur białkowych (ilość szybko wzrasta), z czego ponad 80% to struktury uzyskane metodą dyfrakcji promieniowania rentgenowskiego na monokryształach.

Ustalona struktura trzeciorzędowa cząsteczki **swobodnego** białka **NIE** jest warunkiem koniecznym jego funkcjonowania biologicznego. Wzrasta liczba doniesień o **białkach strukturalnie nieuporządkowanych IDPs** (*intrinsically disordered; natively unfolded*). Przyjmuje się, że ok. 30% białek eukariotycznych wykazuje obszary częściowo nieuporządkowane na długości ponad 50 AA lub całkowity brak nieuporządkowania. W związku z tym postuluje się rozszerzenie paradygmatu SAR (*structure - activity relationship*) dla uwzględnienia nowych form strukturalnych, oprócz formy natywnej i kłębka statystycznego (*random coil*), także na formy pośrednie między nimi, **stopionej globuły** (*pre*)molten globule. W białkach IDPs przejście do uporządkowanej formy natywnej zachodzi pod wpływem tworzenia funkcjonalnych kompleksów molekularnych. Proces określa się terminem „*folding upon binding*”. Połączenie wiązania ze zwijaniem może zapewnić:

- zdolność specyficznego wiązania kilku różnych strukturalnie partnerów bez konieczności tworzenia kilku centrów wiążących (oszczędność na wielkości cząsteczki);
 - wysoka specyficzność wiązania przy względnie niskim powinowactwie.
- Oba efekty zapewniają możliwość funkcjonowania białek natywnie rozwiniętych jako elementów przekąźnikowych (sygnałowych) typu „*one- to-many signaling*”.

Stabilizacja struktur natywnych cząsteczek białkowych nie jest wysoka, rzędu kilkadziesiąt kilojouli na mol [**PLANSZA 106: czynniki stabilizacji/destabilizacji formy natywnej**] i wynika z (omówionych) **oddziaływań stabilizujących**: hydrofobowe, mostki solne, wiązania wodorowe, **przeważające nad wkładem destabilizującym entropii konformacyjnej**. Udział mają również oddziaływania z rozpuszczalnikiem, a w przypadku wody także wymiana jonów i cząsteczek wody między sferą hydratacyjną białka i „*bulk solvent*”. Obliczenia poszczególnych wkładów obdarzone są niestety

błędami, które kumulują się w ostatecznej wartości teoretycznej (zsumowanie daje 170 kJ/mol). Dla białek globularnych i błonowych zmiany warunków środowiska mogą spowodować **odwrócenie relacji stabilizacja - destabilizacja** struktury natywnej i przejście białka w formę zdenaturowaną kłębka statystycznego [**PLANSZA 107: denaturacja białek**]. Czynnikiem denaturującymi dla białek globularnych są:

(a) **podwyższona temperatura**: ruchy cieplne rozbijają oddziaływania elektrostatyczne o energiach rzędu RT;

(b) chemiczne **denaturanty** np. mocznik, chlorek guanidyny: wiążą się z powierzchnią białka dążąc do jej maksymalizacji, a tak jest w strukturze rozwiniętej;

(c) **pH roztworu**: jonizacja i protonacja grup bocznych niszczy oddziaływania stabilizujące struktur natywnych.

Należy wspomnieć, że białka niektórych oraganizmów wykazują znaczną stabilność w skrajnych warunkach np. białka **bakterii termofilnych** nawet do 100 ÷ 120°C. Jak widać z wykresów zależności stosunku zawartości forma zdenaturowana/forma natywna

$$K = [native]/[unfolded] = (1 - \alpha)/\alpha \quad \alpha - \text{frakcja (\%)} \text{ białka rozwiniętego}$$

przejście zachodzi dla wielu białek ostro jak topnienie i ma **charakter odwracalny**: usunięcie czynnika denaturującego powoduje powtórne **zwinięcie się** białka zdenaturowanego do formy natywnej, tzw. **refolding**. Wynika to stąd, że **informacja o strukturze przestrzennej formy natywnej białka jest określona przez sekwencję aminokwasów**. Stąd, znając sekwencję aminokwasów istnieje możliwość teoretycznego zwinięcia białka do jego formy natywnej, która według tzw. hipotezy Anfinsena jest stanem o **minimum energii swobodnej G** (zdarzają się wyjątki od tej reguły).

Podobnie jak dla procesu tworzenie/denaturacja helis DNA i RNA, z zależności stałej równowagi od temperatury K(T) czyli z równania van't Hoffa otrzymuje się wartość entalpii van't Hoffa przejścia ΔH_{vH}

$$\partial \ln K / \partial T = \Delta H_{vH} / RT^2$$

a rejestracja zmiany ciepła właściwego przy stałym ciśnieniu w funkcji temperatury $c_p(T)$ daje (DSC) entalpię kalorymetryczną ΔH_{cal} . Stosunek $\Delta H_{cal} / \Delta H_{vH}$ jest miarą kooperatywności; np. dla białek dwudomenowych może przyjmować wartość 1.8 ÷ 1.9.

Ujemna zmiana $\Delta c_p = c_p[native] - c_p[unfolded] < 0$ powoduje nieliniowy przebieg zależności van't Hoffa ($\ln K$ w funkcji $1/T$) **[PLANSZA 108: termodynamika przejścia *native-unfolded*]** i w związku z tym entalpia swobodna stabilizacji w przejściu *native-unfolded* $\Delta G^\circ = -RT \ln K$ jest rezultatem **kompensacji** dużych, nieliniowych i porównywalnych co do wielkości członów entropowych i entalpowych (błąd wyznaczania ΔG°) **[PLANSZA 106]**, oraz może dawać niestabilność białka nie tylko w wysokich, ale także w niskich temperaturach – „zimna” denaturacja **[PLANSZA 108]**.

Struktura białek ma charakter **dynamiczny**. W szerokim zakresie skali czasowej obserwuje się:

- rotacje reszt bocznych wokół wiązań pojedynczych, w tym grup metylowych: 10^{-11} s do 10^{-8} s i *flip-flop* pierścieni fenyloalaniny i tyrozyny: 10^{-9} s do 1 s;
- nanosekundowe, lokalne ruchy aminokwasów w łańcuchu ($\sim 10^{-9}$ s);
- przemieszczanie domen: 10^{-7} s do 10^{-8} s;
- odwracalne procesy rozwijanie/zwijanie fragmentów lub całych cząsteczek w długiej skali czasowej 10^4 s do 10^{12} s.

Ten ostatni proces można śledzić poprzez wymianę proton-deuter grup NH z rozpuszczalnikiem **[PLANSZA 109: wymiana proton-deuter; mechanizm EX₂]** rozpuszczając białko w ciężkiej wodzie. Wymiana ma charakter podobny jak w przypadku „oddychania” helis kwasów nukleinowych, ale jest katalizowana zarówno przez kwasy jak i przez zasady. Mechanizm jest typu EX₂ z kilkoma aktami otwarcia/zamknięcia na jeden akt wymiany.

WYKŁAD 10

Zwijanie białek *in vitro*

Temat **zwijanie białek** (*protein folding*) jest centralnym problemem biofizyki molekularnej i obejmuje:

- (a) poszukiwanie algorytmów komputerowych **wyznaczenia struktury** natywnej na podstawie sekwencji aminokwasów;
- (b) poszukiwanie molekularnych **mechanizmów (dróg) zwijania** różnymi metodami, zarówno doświadczalnymi jak i teoretycznymi.

Problem leży w znalezieniu **jednej funkcjonalnej konformacji o minimum entalpii swobodnej (hipoteza Anfinsena)** z astronomicznej ilości możliwych, ponad 10^{30} dla białka o 100 aminokwasach. Ponieważ **czasy zwijania różnych białek wynoszą od mikrosekund (10^{-5} s; cytochrom c drożdży: 0.1 ms) do minut a nawet godzin (10^2 s; mioglobina: 10 s)**, więc są za długie do szukania właściwej konformacji metodami **dynamiki molekularnej MD** w reprezentacji pełnoatomowej.

W standardowej (pełnoatomowej) wersji MD rozwiązuje się komputerowo klasyczny układ równań Newtona dla atomów makrocząsteczki w polu sił wszystkich pozostałych atomów **[PLANSZA 7]**. Rozpuszczalnik można uwzględnić poprzez wprowadzenie stałych dielektrycznych roztworu wodnego ($D \sim 80$) i wnętrza makromolekuły ($D \sim 2 \div 20$) lub, co jest znacznie lepszym rozwiązaniem jeśli chodzi o jakość wyników, uwzględniać *explicite* cząsteczki rozpuszczalnika w równaniach ruchu. Numeryczne rozwiązanie wykonuje się w krótkich skokach czasowych ~ 1 fs (10^{-15} s) uzyskując **trajektorię układu w przestrzeni fazowej położenia i pędów**, czyli obraz dynamiki molekule w czasie **[PLANSZA 110: symulacje rMD fragmentu peptydowego]**. Ze względu na dużą ilość stopni swobody układu takiego jak białko w roztworze najdłuższe symulacje MD przy użyciu superkomputerów wyznaczają dynamikę w czasie poniżej mikrosekund ($\square 10^{-6}$ s). Przy symulacjach dłuższych procesów „ściąga się” molekule do konformacji znanej z doświadczenia (rMD) i/lub **redukuje ilość stopni swobody** wprowadzając uproszczone modele cząsteczki z fragmentami wieloatomowymi jako sferami, materialnymi w ciągłym ośrodku; np. białko w formie sfer centrowanych na C_α , wiązaniu peptydowym i C_β (reszta boczna). Oddziaływanie w postaci „niefizycznego” potencjału konstruuje się zależnie od modelu. W **brownowskiej dynamice molekularnej** rozwiązuje się równania Langevina (np. algorytm Ermaka i McCammona) z siłami

oddziaływań między centrami makromolekuły, siłami tarcia z rozpuszczalnikiem i fluktuującymi siłami stochastycznymi Browna. Nawet przy najlepszych obecnie superkomputerach czasy symulacji dla średniej wielkości białek nie przekraczają znacząco mikrosekund ($\square 10^{-5}$ s), można więc uchwycić zaledwie efekty początkowe, zaczątki struktury natywnej białka np. formowanie się helisy czy β -spinki.

Komputerowe techniki zwijania głównie białek globularnych wymagają opracowania algorytmu pozwalającego przełożyć „**astronomiczną**” ilość sekwencji **różnej długości** złożonych z 20 (22) typów aminokwasów na **trzy** zasadnicze typy **foldów drugorzędowych**: helisa, β -kartka, pętla (skręt) oraz ponad **700 foldów trzeciorzędowych**, trudniejszych do opisanie, których ilość wzrasta. Problem pozostaje nierozwiązany (Święty Graal biofizyki molekularnej), ale różnorodne podejścia z zakresu **projektowania molekularnego** stosuje się z sukcesem w trzech typach zagadnień:

- (A) Wyznaczanie struktur **drugorzędowych**, a więc lokalnych;
- (B) Wyznaczanie struktur całkowitych w białkach o dużej **homologii**, czyli podobieństwie (identyczności) sekwencji do białek o strukturze znanej;
- (C) Wyznaczanie struktur trzeciorzędowych (i superdrugorzędowych) *de novo* niezbyt dużych białek, metodami **statystycznego próbkowania** konformacji dla uproszczonych (gruboziarnistych) modeli cząsteczek białkowych, np. zwijanie na dyskretnych siatkach.

Wyznaczanie struktur drugorzędowych opiera się na preferencjach czy prawdopodobieństwie aminokwasów do występowania w określonych typach tych struktur [**PLANSZA 111: preferencje aminokwasów do struktur drugorzędowych**]. Dla każdego i-tego z 20 aminokwasów można określić liczbowo takie **preferencje** $P_{x,i}$ przebywania w x: helisie, β -kartce i skręcie (pętli), na podstawie **analizy statystycznej baz danych sekwencji i struktur białek**: np. kwas glutaminowy (Glu) bardzo chętnie przebywa w helisie ($P_{\alpha} > 1$), niechętnie w β -kartce ($P_{\beta} < 1$) i nie ma preferencji ani braku preferencji jeśli chodzi o skręty ($P_t \sim 1$). Odwrotnie zachowuje się cysteina, a prolina „unika” helisy i β -kartek (brak protonu amidowego). Na tej podstawie tworzy się **algorytm przekładający sekwencje poszczególnych fragmentów białka na jego struktury drugorzędowe**, np. Chou i Fasmana (jeden z pierwszych):

- określenie centrum nukleacji struktury drugorzędowej (helisa, kartka) długości 6 aminokwasów;

- elongacja struktury w obu kierunkach wzdłuż łańcucha aż do natrafienia na prolinę lub 4 kolejne aminokwasy terminujące aktualnie budowaną strukturę drugorzędową.

Po rozszerzeniu algorytmu o dodatkowe informacje, np. z danych doświadczalnych, homologii sekwencji i poprawie jego metodologii (techniki *neural network*) **dokładność przewidywań sięgająca 80%** oddaje **lokalne uwarunkowania** struktur drugorzędowych, które mogą być dodatkowo uwarunkowane przez globalny fold białka.

Zwijanie na podstawie homologii sekwencji wykorzystuje znaną zależność, że **30% identyczności sekwencji zbliżonych pod względem funkcji białek daje ponad 70% podobieństwa ich struktury trzeciorzędowej** (bazy CATH, SCOP) **[PLANSZA 112: etapy modelowania homologicznego]**. Punktem wyjścia jest właściwe **uszeregowanie sekwencji** (*sequence alignment*) białka o nieznannej strukturze w stosunku do homologicznego białka lub białek matrycowych (wzorców) o strukturach znanych **[PLANSZA 113: uszeregowanie sekwencji]**. Problem nie jest trywialny. Uzyskanie identyczności w jednym miejscu wymaga jej poświęcenia w innym i trzeba w tym celu skonstruować algorytm punktowania: punkty dodatnie za identyczność/podobieństwo aminokwasów, ujemne za zamiany aminokwasów różniących się własnościami oraz za wprowadzanie luk. **Schemat punktowania** (*scoring scheme*) jest ujęty w macierzy podstawień 20x20 (BLOSUM) oraz modelu uwzględniania luk Δ ; w przykładzie na **PLANSZY 113** system punktowania (najprostszy): +1, 0, -1. Właściwe uszeregowanie (programy FASTA, BLAST, CLUSTAL W) daje **obszary strukturalnie zachowane SCR** (*structurally conserved regions*), najczęściej **drugorzędowe [PLANSZA 114: obszary strukturalnie zachowane]**, które trzeba połączyć pętlami, np. biorąc „gotowe” pętle z baz danych o podobnej sekwencji i długości. Program, np. MODELLER, **buduje strukturę** nieznannej molekule na szkieletcie molekule homologicznej poprzez wprowadzenie więzów przestrzennych na strukturę startową i optymalizację funkcji wyrażającej te więzy (gęstość prawdopodobieństwa) na podstawie mechaniki statystycznej oraz analizy baz danych struktur; „ściągnięcie” struktury startowej do struktur(y) tarcz(y) (wzorca). Procedurę szukania modelu kończy minimalizacji energii i/lub krótka symulacja MD dla **usunięcia naprężeń [PLANSZA 112]**. **Weryfikacja modelu** odbywa się w oparciu o szereg kryteriów: optymalne kontakty hydrofobowe, optymalna powierzchnia dostępu rozpuszczalnika, „przeciągnięcie” (*threading*) sekwencji przez strukturę, sprawdzaniu lokalizacji aminokwasów na wykresie Ramachandrana. Przy większej homologii sekwencji procedura jest standardowa (komputerowe pakiety komercyjne), ale przy małej homologii wymaga „twórczego” wkładu metodologicznego (patrz dalej CASP).

W trzeciej grupie technik przybliża się cząsteczkę białka przez model uproszczony, np. łańcuch reprezentowany przez sfery (pseudoatomy) centrowane na C_α , $C(O)-N$ (wiązanie peptydowe) i C_β (łańcuch boczny). Oddziaływania przyciągające (wiązania wodorowe i hydrofobowe) oraz odpychające (*hard core* van der Waalsa) powodują zmiany położenia sfer w sposób nieciągły na siatce. Próbkowanie dyskretnych stanów konformacyjnych odbywa się oszczędnie czasowo, w porównaniu z MD, **metodami Monte Carlo**, aż do znalezienia **GLOBALNEGO minimum energetycznego**. Kolejna, przypadkowo wybrana konformacja jest akceptowana lub nie stosownie do określonego kryterium. W **kryterium Metropolisa** akceptacja konformacji $i + 1$ zachodzi jeśli:

(1) $E_{i+1} < E_i$ - przejście w kierunku mniejszej energii układu;

(2) $E_{i+1} > E_i$ - jeśli liczba x wybrana losowo z przedziału $[0, 1]$ spełnia relację:

$$x < \exp[-(E_{i+1}-E_i)/RT] \text{ co daje przeskoki przez bariery lokalne.}$$

Markowski łańcuch konformacji Monte Carlo jest równoważny trajektorii MD. Ulepszenia metodologii i wzrastająca sprawność komputerów pozwala na symulacje zwijania mniejszych białek metodą Monte Carlo w reprezentacji pełnoatomowej.

W modelowaniu komputerowym struktur rozgrywa się "zawody sportowe" w **ramach CASP** (*Critical Assessment of Structure Prediction*): opracowane doświadczalnie struktury deponuje się bez publikowania i różne grupy badawcze starają się je wysymulować swoimi metodami. Aktualnie rozgrywane konkurencje:

- modelowanie homologiczne (CM, *comparative modeling*)
- rozpoznawanie foldów (FR, *fold recognition*)
- nowe foldy (NF, *new fold*)
- przewidywanie struktur drugorzędowych

Szukanie przez cząsteczkę białka natywnej konformacji nie może zachodzić metodą prób i błędów. Zagadnienie formułuje tzw. **paradoks Levinthala**: dla łańcucha peptydowego 100 aminokwasów takie próbkowanie 10^{30} konformacji z czasem 10^{-11} s dla każdej z nich wymagałoby 10^{19} s czyli 10^{11} lat, dłużej niż wiek wszechświata. Model zwijania dobrze oddaje **lejek zdarzeń równoległych** (*folding funnel*) **[PLANSZA 115: lejek zwijania]**. Każdy punkt na powierzchni lejka reprezentuje **entalpię swobodną G** (oś pionowa) określonej **konformacji** (osie poziome Φ i Ψ). Zwijanie ma **charakter dyfuzyjnych**, przypadkowych ruchów termicznych i ruchów Browna łańcucha, a poszczególne cząsteczki „wędrują” niezależnie od siebie od startowych form kłęбка

statystycznego (obrzeże lejka) jak spływające krople wody. Zysk energetyczny (entalpowy) kompensuje spadek entropii związany z zacieśnianiem ilości dostępnych struktur w dół lejka. Proces ma charakter **KOOPERATYWNY: wcześniejsze etapy zwinięcia - utworzenie natywnych kontaktów stabilizujących sprzyjają następnym**. Przeszukiwaniu przypadkowemu konformacji odpowiadałoby błądzenie po płaskiej powierzchni w poszukiwaniu małego otworu w środku. Powolne (wielokrotność) foldingi są reprezentowane przez **lejki pofaldowane** gdzie łańcuchy mogą trafiać na minima lokalne czyli formy pośrednie (pułapki). Koncepcja lejka umożliwia również opis zjawiska **nieprawidłowego zwijania** (wielokrotne lejki zwijania) do form agregujących (formy amyloidalne, priony chorobotwórcze) [**PLANSZA 116: złogi amyloidalne**].

Teoretyczna analiza procesu zwijania wymaga konstrukcji lejka czyli „krajobrazu energetycznego” (*energy landscape*) dla konkretnego białka o znanej strukturze przestrzennej. Odpowiednie symulacje mają charakter dynamiki molekularnej lub innego typu próbkowania statystycznego mikrostanów przejściowych z wykorzystaniem uproszczonych struktur białkowych na gruncie tzw. **modeli Go** (*Go: Annual Review of Biophysics and Bioengineering 2, 183, 1983*), które dopuszczają tylko oddziaływania (kontakty) natywne. **Dynamika brownowska** pozwala uchwycić jedynie zaczątki tworzenia struktur drugorzędowych: α -helisa, β -hairpin w czasie rzędu mikrosekundy.

Dobrym parametrem, który stanowi punkt styyczny między teoretykami i doświadczalnikami jest **Φ -value** wprowadzona przez Fershta (*Itasaki et al. Journal of Molecular Biology 254, 260, 1995*) i charakteryzująca zespół mikrostanów przejściowych przy foldingu:

$$\Phi = (\Delta\Delta G^\ddagger)/(\Delta\Delta G^\circ) \quad \Delta\Delta G^\ddagger = \Delta G^\ddagger_{\text{mut}} - \Delta G^\ddagger_{\text{wt}} = -RT\ln(k_{\text{mut}}/k_{\text{wt}})$$

$$\Delta\Delta G^\circ = \Delta G^\circ_{\text{mut}} - \Delta G^\circ_{\text{wt}}$$

gdzie k_{mut} , k_{wt} są szybkościami zwijania mutantu i białka „dzikiego”, $\Delta\Delta G^\ddagger$ są różnicami entalpii swobodnych barier, a $\Delta\Delta G^\circ$ różnicami entalpii swobodnych stanów natywnych.

Φ -value jest wygodną miarą ułamka struktury natywnej w zbiorze mikrostanów przejściowych wokół miejsca mutacji. $\Phi \sim 1$ oznacza, że zmiana entalpii swobodnej w wyniku mutacji jest podobna w stanie natywnym i przejściowym, czyli natywne kontakty z udziałem mutowanego aminokwasu są już utworzone w stanie

przejściowym. $\Phi \sim 0$ oznacza, że lokalne otoczenie miejsca mutacji ma charakter rozwinęty.

Stosuje się dużą ilość metod doświadczalnych w badaniu procesów zwijania, także jako **testy teoretycznych modeli** „*energy landscape*”: parametry mierzone potwierdzają model lub stanowią podstawę jego odrzucenia. Znajdują tu zastosowania wszystkie techniki badania **dynamiki/kinetyki** układów molekularnych: fluorescencja czasoworozdzielacza, „*stopped flow*” NMR, wielowymiarowe techniki NMR [**PLANSZA 117: eksperymenty NMR w badaniu zwijania**]. W widmie zatrzymanego przepływu ^{19}F NMR **reduktazy kwasu dwuhydrofaliowego** (będzie omawiana jako przykład oddziaływania enzym - substrat) znakowanego fluorem w pozycji C6 tryptofanów widać zanikanie w czasie sygnałów od białka zdenaturowanego i pojawianie się sygnałów formy natywnej – możliwość wyznaczania **czasów zwijania**. Widmo $^1\text{H},^{15}\text{N}$ -HSQC (*fingerprint*) α -laktalbuminy pokazuje przejście od słabo rozdzielonych sygnałów struktury kłębka do znacząco większej ich dyspersji w strukturze natywnej poprzez **etapy pośrednie**, np. **stopiona globuła** (*molten globule*) na drodze pomiędzy formą natywną a kłębką statystycznym, czyli formy o foldzie globalnym takim jak struktura natywna ale bez właściwego i stabilnego ustawienia reszt bocznych aminokwasów.

Jedną z metod o charakterze strukturalnym jest wyłapywanie **intermediatów z przypadkowo utworzonymi mostkami dwusiarczkowymi** jak w białku BPTI [**PLANSZA 118: określanie struktur na drodze zwijania**]. Trzy mostki S-S muszą się utworzyć w ściśle określonej kolejności aby doprowadzić białko do formy natywnej. Można przerywać tworzenie struktury natywnej w określonych momentach czasu (dla BPTI przez zakwaszenie) i określać struktury form przejściowych. W przypadku BPTI widać, że wcześniejsze utworzenie właściwego mostka 14-38, niż wymaga tego prawidłowa droga zwijania, prowadzi do niewłaściwej formy. Forma ta musi ulec rozwinięciu, żeby białko weszło na właściwą drogę zwijania. **Prawidłowa droga zwijania prowadzi przez przejściowe utworzenie niewłaściwych mostków 5-38 i 5-14.**

Mimo olbrzymiej ilości prac doświadczalnych i teoretycznych szereg zagadnień w mechanizmie zwijania (białek globulanych) pozostaje przedmiotem spornym, np.:

- znaczna ilość różnych modeli początków zwijania, czyli **struktur inicjujących folding** na obrzeżu lejka zwijania;

- zagadnienie **hierarchiczności foldingu**: czy zwijanie zachodzi kolejno poprzez struktury drugorzędowe które oddziałując ze sobą dają globalny fold ?

czy

najpierw zachodzi kolaps hydrofobowy do foldu trzeciorzędowego a następnie „wymuszanie” struktur drugorzędowych przez już ukształtowany fold trzeciorzędowy ?

Natomiast istnieje duża zgodność jeśli chodzi o akceptację formy **stopionej globuły**, chociaż i tu mnoży się etapy pośrednie, np. *premolten globule*.

WYKŁAD 11

Zwijanie białek *in vivo*; zwijanie RNA

Zdolność samoistnego zwijanie białek *in vitro* jest warunkiem koniecznym dla zachodzenia zwijania *in vivo*. W komórce wykształciły się mechanizmy, które podnoszą efektywność zwijania.

(A) **Maszynaria translacyjna** rybosomu sprzyja przynajmniej częściowemu zwijaniu białek w trakcie ich powstawania na rybosomie. Szybkość syntezy łańcucha peptydowego $50 \div 300$ AA/min jest czasem dostatecznie długim dla zwijania **ko-translacyjnego**.

(B) **Wewnętrzna kataliza foldingu** przez sekwencje mogące działać jak foldazy stabilizujące stany przejściowe foldingu (np. w α -lytic protease), a następnie wycinane proteolitycznie.

(C) **Enzymatyczna kataliza dwóch powolnych procesów: tworzenie mostków dwusiarczkowych** między właściwymi parami cystein (*shuffling*) przez białkową izomerazę dwusiarczkową (*protein disulfide isomerase*) oraz **izomeryzacja trans \rightarrow cis** w przypadku niektórych wiązań peptydowych, zwykle poprzedzających prolinę, przez izomerazę peptydylo-prolylową (*peptydyl-prolyl isomerase*).

(D) Przynajmniej dla części białek wymagany jest udział w zwijaniu pomocniczego systemu **białkowych kompleksów opiekuńczych** tzw. **chaperonów** [**PLANSZA 119: systemy białek opiekuńczych**], często związanych z białkami szoku termicznego HSP. Najlepiej poznany jest system bakteryjny **GroEL/GroES 800+70 kDa**, którego obie części zostały wykrywane [**PLANSZA 120: krystalografia systemu GroEL/GroES**]. System pracuje przy udziale cząsteczek ATP. Po związaniu GroES z jednej strony cylindra GroEL, do środka wchodzi białko źle zwinięte. Pomimo znanej struktury kompleksu mechanizm jego działania nie jest do końca potwierdzony w sposób jednoznaczny. Postuluje się dwie możliwości:

- białko wewnątrz kompleksu GroEL/GroES jest rozwijane tak, aby po wyjściu mogło się zwinąć właściwie; model pasywny IAM (*iterative annealing model*)
- wewnątrz kompleksu GroEL/GroES zachodzi właściwe zwijanie w „ochronnym otoczeniu” białkowym; model aktywny klatki Anfinsena ACM (*Anfinsen cage model*)

Niewłaściwe zwinięte białka mogą być przyczyną szeregu **chorób typu neurodegeneracyjnego**:

Złogi β -amyloidu [PLANSZA 121: struktura złoży amyloidnego] odkładane w tkance mózgowej są uznawane za czynnik wywołujący chorobę Alzheimera. Struktura ma formę 4-warstwy złożonej z dwóch β -krotek (*cross- β*), gdzie nici tworzące kartkę są połączone wygiętym segmentem, a obie kartki **oddziałują hydrofobowo**.

Priony o zmienionej strukturze [PLANSZA 122: działanie prionu chorobotwórczego] wywołują choroby: kuru, Creutzfelda-Jakoba, wściekłych krów i in., stanowiąc czynnik zakaźny poprzez wiązanie ze zwykłymi białkami prionowymi i indukowanie w nich przekształcanie w formy chorobotwórcze, odkładane w postaci złoży powodujących ubytki w tkance nerwowej („dziury” w mózgu).

Zwijanie RNA obejmuje te same zagadnienia co zwijanie białek:

- przewidywanie struktury drugo- i trzeciorzędowej na podstawie sekwencji nukleotydów;
- określanie molekularnego mechanizmu przejścia od formy rozwiniętej do funkcjonalnej formy natywnej (drogi zwijania),
- udział mechanizmów komórkowych w zwijaniu RNA *in vivo*.

Skala czasowa zwijania *in vitro* jest podobna jak w przypadku białek ($10^{-5} \div 10^2$ s). W komórce RNA zwija się średnio ok. 10 razy szybciej niż *in vitro* w wyniku pomocy ze strony układu transkrypcyjnego (zwijanie ko-transkrypcyjne), udziału specyficznie oraz niespecyficznie wiążących się białek, najprawdopodobniej poprzez likwidację pułapek kinetycznych na drodze zwijania, oraz zapewnienie optymalnego stężenia jonów mono- i dwuwartościowych. Pejzaż energetyczny (lejek zwijania) RNA jest bardziej pofałdowany niż w przypadku białek, z licznymi pułapkami kinetycznymi i strukturalnymi.

Jak na razie, znaczące osiągnięcia uzyskano w przewidywaniu jedynie struktur drugorzędowych na podstawie sekwencji nukleotydów, natomiast brak jest podobnie efektywnych jak dla białek algorytmów komputerowych do przewidywania struktur trzeciorzędowych. Proces zwijania ma wyraźnie hierarchiczny charakter: tworzenie struktur drugorzędowych, ich organizacja poprzez motywy trzeciorzędowe w globalny fold cząsteczki.

Kompleksy biomolekularne: enzymy białkowe (1)

Tworzenie **specyficznych kompleksów** molekularnych wymaga w opisie **wzajemnego dopasowywania struktur przestrzennych i rozkładów ładunków oddziałujących cząsteczek** uwzględnienia dynamiki ruchów molekularnych i przekształceń strukturalnych. Najprostszy (najwcześniejszy) mechanizm polegający na dopasowaniu sztywnych struktur jak „klucz do zamka” (*rigid body docking*) okazał się niewystarczający. Koshland zaproponował (w 1958 r.) model **wzajemnego, konformacyjnego dopasowania enzymu i substratu (*induced fit*): $E \cdot S \rightarrow ES$** **[PLANSZA 123: dynamiczne modele tworzenia kompleksów]**. Model potwierdzono w szeregu przypadków (przegląd: *Williamson Nature Structural Biology* 7, 834, 2000). Wzajemne konformacyjne dopasowanie wymaga dodatkowych nakładów energetycznych (spadek ΔG°) ALE obniża barierę energii tworzenia kompleksu ΔG^\ddagger i daje efektywniejsze kontakty ładunków (wzrost ΔG°). Konkurencyjny model to rozpoznanie poprzez **wstępnie selekcjonowane konformery** optymalne dla wiązania (*conformational selection mechanism*; *Ma et al. Protein Engineering* 12, 713, 1999). Oba modele nie uwzględniają procesów dynamicznych niezbędnych do tworzenia kompleksów z udziałem białek natywnie rozwiniętych. Model *fly-casting* („wędkowanie na muchę”) ujmuje wiązanie w kompleks i zwijanie jako **sprzężone procesy kooperatywne** (*Shoemaker et al. Proceeding of the national Academy of Sciences USA* 97, 8868, 2000). Rozwinięte białko jest zdolne do tworzenia wstępnych kontaktów z partnerem w warunkach gdy białko w pełni zwinięte pozostaje poza sferą oddziaływań z powodu znacznie mniejszych fluktuacji strukturalnych. Umożliwia to „nawinięcie” na drugą molekułę po wstępnym „zahaczeniu”.

ENZYMY, cząsteczki **białek** lub **RNA (rybozomy)**, katalizują wszystkie kontrolowane reakcje biochemiczne. Enzym może składać się z samej cząsteczki białkowej (monomer lub oligomer o strukturze czwartorzędowej) lub zawierać **koenzym**, czyli związaną (niekowalencyjnie lub kowalencyjnie) cząsteczkę (niebiałkową), niezbędną do zachodzenia katalizowanej reakcji:



Enzymy połączone w układy wieloenzymatyczne np. związane z błoną mogą katalizować sprzężone reakcje, w których produkty z jednego enzymu są przekazywane jako substraty na drugi enzym. **Zasadnicze cechy procesu katalizy enzymatycznej** to:

- **swoistość substratowa** w stosunku do określonego związku chemicznego lub klasy związków danego typu;
- **swoistość katalizowanej reakcji** jednego typu, zwykle określonej reakcji i reakcji odwrotnej;
- **optymalne warunki** aktywności enzymu: temperatura, pH, siła jonowa, potencjał oksydoredukcyjny ośrodka, itp. (warunki stabilnej struktury natywnej);
- możliwość **regulacji aktywności** przez cząsteczki inhibitorów, aktywatorów, ew. same substraty lub produkty.

Enzymy przyspieszają (ale nie wymuszają) **[PLANSZA 124: szybkości katalizy enzymatycznej]** od 10^5 do 10^{14} (10^{18}) razy zachodzenie reakcji spontanicznych z termodynamicznego punktu widzenia **[PLANSZA 125: hiperboliczna kinetyka reakcji enzymatycznej]**:

$$\Delta G = G(\text{produkty}) - G(\text{substraty}) < 0$$

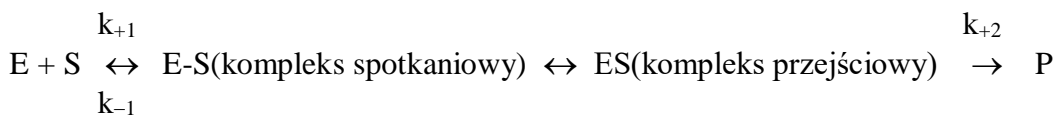
Dla reakcji sprzężonego układu enzymów liczy się całkowity efekt zmiany $\Delta G < 0$. Z „fizycznego” punktu widzenia enzymy **obniżają barierę reakcji czyli entalpię swobodną przejścia $\Delta G^\#$ od substratu(ów) do produktu(ów)** zgodnie ze wzorem Eyringa:

$$\text{stała szybkości reakcji } k_{\text{cat}} = (kT/h)\exp(-\Delta G^\#/RT)$$

$$T - \text{temperatura}; \quad k - \text{stała Boltzmanna}; \quad kT/h = 6.2 \cdot 10^{12} \text{ s}^{-1} \text{ dla } T = 300\text{K}$$

Efekt entalpowy $\Delta G^\#$ jest związany z utworzeniem **stanu przejściowego kompleksu enzym - substrat**, w którym substrat uzyskuje w **centrum aktywnym** enzymu optymalną geometrię (naprężenia), rozkład ładunku i ew. warunki bezwodne. **Efekt entropowy $\Delta G^\#$** wynika ze związania cząsteczek i grup niezbędnych do reakcji w jednym obszarze, a reakcja zachodząca w **centrum katalitycznym** ma charakter wewnątrzcząsteczkowy (jednocząsteczkowy) zamiast przypadkowych zderzeń kilku substratów.

Podstawowe **cechy kinetyki katalizy enzymatycznej** można ująć na przykładzie prostego modelu reakcji nieodwracalnej z przekształcaniem jednego substratu S do jednego produktu P ($k_{\text{cat}} \equiv k_{+2}$):



stała asocjacji $K_{\text{as}} = k_{+1}/k_{-1}$ $k_{+1} \leq k_D \sim 10^9 \div 10^{11} \text{ s}^{-1}\text{M}^{-1}$ dla procesu wiązania kontrolowanego dyfuzyjnie (dolna granica) oraz przy uwzględnieniu oddziaływania elektrostatycznego (górną granicą).

$$d[ES]/dt = k_{+1}[S][E] - k_{-1}[ES] - k_{+2}[ES]$$

$$d[S]/dt = k_{-1}[ES] - k_{+1}[E][S]$$

$$d[E]/dt = k_{-1}[ES] - k_{+2}[ES] - k_{+1}[S][E]$$

$$d[P]/dt = k_{+2}[ES]$$

w stanie stacjonarnym $d[ES]/dt = 0$ i przy założeniu niezmiennego stężenia enzymu

$$[E_0] = [E] + [ES]$$

rozwiązanie (stan ustalony) daje **szybkość początkową reakcji** (duży nadmiar substratu w stosunku do enzymu) w funkcji początkowego stężenia substratu $[S_0]$:

$$v = d[P]/dt = V_{\max}[S_0]/([S_0] + K_m)$$

ilość μmol produktu na sekundę i stężenie enzymu; stężenie enzymu w mg białka na jednostkę objętości lub ilość jednostek enzymu U na jedn. objętości; $1U =$ ilość enzymu katalizująca 1 mmol substratu w ciągu 1 min w warunkach optymalnych.

Parametrami kinetycznymi reakcji do wyznaczenia są:

$$\text{szybkość maksymalna } V_{\max} = k_{+2}[E_0] = k_{\text{cat}}[E_0]$$

stała Michaelisa: $K_m = (k_{-1} + k_{+2})/k_{+1}$; może być równa lub różna od odwrotności stałej substratowej $1/K_{\text{as}}$.

Zależność szybkości v od stężenia początkowego substratu $[S_0]$ jest **hiperbolą - kinetyka Michaelisa-Menten** [**PLANSZA 126: reprezentacje kinetyki Michaelisa-Menten**]: dla $S_0 \ll K_m$ zależność liniowa, dla $v = V_{\max}/2$ zachodzi $K_m = [S_0]$, a dla $[S_0] \gg K_m$ następuje wysycenie wszystkich centrów wiążących enzymu przez cząsteczki substratu. W reprezentacji Lineweavera-Burke'a $1/v = f(1/[S_0])$ i Eadie-Hofstee $v = f(v/[S_0])$ uzyskuje się **zależność liniową**. **Miarą specyficzności enzymu do substratu** jest wartość:

$$k_{\text{cat}}/K_m$$

Uwzględnienie bardziej realistycznej sytuacji reakcji odwracalnej w zasadzie niczego nie zmienia jak długo enzym pracuje w zakresie szybkości początkowych.

Wyróżnia się następujące zasadnicze **mechanizmy hamowania reakcji enzymatycznej** przez **inhibitory** wiążące się niekowalencyjnie z enzymem [**PLANSZA 127: mechanizmy inhibicji**]:

(A) **Hamowanie kompetycyjne:** inhibitor I wiąże się w centrum aktywnym wolnego enzymu (E) konkurencyjnie w stosunku do substratu, ze stałą inhibicji:

$$K_i = [E][I]/[EI]$$

ale nie wiąże się z kompleksem ES .

K_i dla analogów substratu jest rzędu K_m , ale dla **analogów substratu w stanie przejściowym** może osiągać wartości do 10^{-12} M. Zmianie ulega stała Michaelisa a prędkość maksymalna pozostaje nie zmieniona.

(B) Hamowanie niekompetycyjne: inhibitor wiąże się zarówno z wolnym enzymem jak i kompleksem ES, przy czym stała inhibicji jest jednakowa w obu przypadkach. Przy zaniedbaniu stałej szybkości rozpadu EIS do produktu zachodzi modyfikacja prędkości maksymalnej bez zmiany wartości stałej Michaelisa. W przeciwnym wypadku ($k_4 \neq 0$) hamowanie jest częściowo niekompetycyjne i może mieć **kinetykę niehiperboliczną**.

(C) Hamowanie mieszane: inhibitor wiąże się z E i z ES z różnymi stałymi inhibicji:

$$[E][I]/[EI] \neq [ES][I]/[EIS]$$

co prowadzi do zmiany zarówno K_m jak i V_{max} . Jednym z wariantów tego hamowania jest **(D) hamowanie akompetycyjne** (odwrotne do kometetycyjnego) gdy inhibitor wiąże się tylko z kompleksem ES do EIS blokując katalizowaną reakcję.

Większość reakcji enzymatycznych przebiega z udziałem więcej niż jednego substratu (produktu). Opis wraz z terminologią reakcji wielosubstratowych i wieloproduktowych podał W. Cleland w serii prac w *Biochimica et Biophysica Acta* 67, od str. 104, 1963, ujęte w skrócie w *podręczniku „Elementy enzymologii” pod red. J. Witwickiego i W. Ardelta PWN [PLANSZA 128: reakcje wielosubstratowe i wieloproduktowe]*. Liczbę substratów oraz produktów określa się przedrostkami jeden-, dwa-, trzy- itd., np. reakcja dwa-dwa - dwa substraty, dwa produkty (*ang. uni, bi,...*). Wyróżnia się dwa typy kinetycznych mechanizmów reakcji:

(1) sekwencyjne: wszystkie substraty muszą się związać z enzymem w kompleks centralny (zwykle jeden, oznaczony nawiasem) przed uwolnieniem pierwszego produktu. Przyłączenie substratów może mieć charakter **uporządkowany** lub **losowy** (kilka równorzędnych dróg reakcji).

(2) ping-pong: jeden lub więcej produktów uwalnia się przed przyłączeniem wszystkich substratów. Kolejność przyłączania jest „rozbita” na mniejsze sekwencje określone przedrostkami jeden-, dwa- itd.

Wyprowadzenie zależności na szybkość reakcji w funkcji stężeń substratów w stanie stacjonarnym wymaga ułożenia i rozwiązania układu równań różniczkowych na stężenia i jednego równania na zachowanie masy enzymu.

Niektóre mechanizmy reakcji wielosubstratowych-wieloproduktowych nie stosują się do kinetyki hiperbolicznej. **Kinetykę niehiperboliczną** obserwuje się w przypadku większości enzymów. Obok już wymienionych także w białkach enzymatycznych o strukturze czwartorzędowej, w wyniku **kooperacji między centrami katalitycznymi lub centrami katalitycznymi i centrami allosterycznymi, w obu przypadkach zlokalizowanymi na różnych podjednostkach**. Allosteria polega na zmianie powinowactwa enzymu do substratu w wyniku związania czasteczki efektora poza centrum aktywnym (katalitycznym) i zmianie konformacji tego centrum.

Kinetykę niehiperboliczną enzymów złożonych z kilku podjednostek można ująć w ramach dwóch modeli:

- **model sprzężony** (*concerted*) MWC zaproponowany przez Monoda i wsp.;
- model sekwencyjny (*sequential*) AKNF zaproponowany przez Adaira i wsp.

Prostszy model MWC **[PLANSZA 129: model MWC enzymów oligomerycznych]** zakłada, że enzym jest zbudowany z jednakowych protomerów (podjednostek) występujących w dwóch konformacjach: zrelaksowanej R i naprężonej T ze stałą równowagi L. **Sprężone** przejście konformacyjne $R \leftrightarrow T$ zachodzi **jednocześnie we wszystkich protomerach i ulegają mu czasteczki enzymu nie związane z substratem**. W formie niezwiązanej równowaga jest przesunięta w kierunku formy T, co wyraża duże L, o mniejszym powinowactwie do substratu, co wyraża małe $c = K_R/K_T$, niezależnie od ilości przyłączonych czasteczek substratu. Ze schematu kinetycznego reakcji wyznacza się ułamek Yapełnienia miejsc dostępnych i prędkość reakcji wyraża się jako:

$$v = V_{\max} Y$$

Zależność $v = f([A])$, [A] – stężenie substratu, jest **krzywą sigmoidalną**, tzn. najpierw przeważa forma T_0 i dodawany w niewielkiej ilości substrat wiąże się głównie z formą R_0 tworząc R_1 a więc szybkość rośnie powoli ze stężeniem substratu. Przesuwanie równowagi $R \leftrightarrow T$ w lewo pod wpływem przyłączania substratu daje **szybki wzrost (sprężenie) nowych miejsc wiążących czyli przyspieszenie szybkości w wyniku dodatniego efektu kooperatywnego**, aż do nasycenia.

Krzywa sigmoidalna przechodzi w hiperbolę dla:

$$n = 1, \text{ czyli dla enzymu monomerycznego;}$$

L bliskie 1 i/lub relatywnie większe c, czyli przy wzroście powinowactwa substratu do formy T.

Model MWC pozwala jakościowo interpretować również bardziej skomplikowane przypadki mechanizmów działania **efektorów allosterycznych**, np. w 12-merze transkarbamylazy kwasu asparaginowego (6 podjednostek katalitycznych i 6 allosterycznych; przyłączanie karbamyłu $\text{NH}_2\text{-C(O)O}^-$ do grupy aminowej kwasu asparaginowego) **[PLANSZA 130: struktura i kinetyka transkarbamylazy asparaginianu]**. **Aktywator** allosteryczny **ATP** wiąże się z formą R dając spadek L a co za tym idzie zanik sigmoidalności. **Inhibitor** allosteryczny **CTP** wiąże się z formą T dając wzrost L i pogłębienie przebiegu sigmoidalnego.

Model MWC nie opisuje kinetyki z **ujemnym efektem kooperatywnym**: wzrost szybkości reakcji przy niskich stężeniach substratu oraz znacznie wolniejsze dochodzenie do stanu nasycenia enzymu (w porównaniu z przebiegiem hiperbolicznym) **[PLANSZA 131: dodatnia i ujemna kooperacja]**. Regulacja działania enzymu przez dodatnią kooperatywność wzmacnia **czułość sygnału metabolicznego**: mała zmiana stężenia ligandu ma duży efekt przyspieszenia reakcji. Przy ujemnej kooperatywności zachodzi **poszerzenie zakresu odpowiedzi enzymu**: nie zachodzi wyłączenie enzymu nawet przy wysokich stężeniach substratu.

WYKŁAD 12

Kompleksy biomolekularne: enzymy białkowe (2); rybozomy

Zagadnienia molekularnych mechanizmów działania enzymów będą omówione na kilku „klasycznych” już przykładach **enzymów białkowych** oraz na przykładach **rybozymów** czyli enzymów RNA.

Mechanizm działania **proteaz serynowych**, enzymów takich jak trypsyna i chymotrypsyna, katalizujących hydrolizę wiązań peptydowych i zawierających resztę serynową jako główną grupę funkcjonalną w centrum katalitycznym, należy do jednego z najdokładniej poznanych. Kluczową rolę w katalizie pełni **triada katalityczna** (*catalytic triad*) czyli trzy aminokwasy centrum: seryna, histydyna i kwas asparaginowy [**PLANSZA 132: centrum katalityczne proteaz serynowych**]. W wyniku przekazania protonu na sąsiednią histydynę słabo reaktywna grupa –OH seryny staje się silnym nukleofilem $-O^{(-)}$, który atakuje węgiel karbonylowy wiązania peptydowego. Dodatkowo naładowany pierścień imidazolowy histydyny jest stabilizowany przez ujemnie naładowaną resztę kwasu asparaginowego. Proteazy wiążą fragment peptydowy mniej lub bardziej specyficznie w zależności od aminokwasów tworzących kieszeń hydrofobową, np. trypsyna łańcuch po lizynie lub argininie. W **pierwszym etapie** [**PLANSZA 133: mechanizm katatalizy chymotrypsyny**] następuje utworzenie **tetrahedralnego stanu przejściowego**, dodatkowo stabilizowanego przez dwa wiązania wodorowe z grupami NH łańcucha, pęknięcie wiązania peptydowego $R-N(H)-C(O)-R'$ i oddysocjowanie fragmentu H_2N-R . W **drugim etapie** pozostały fragment (acyl-enzym) jest hydrolizowany przez cząsteczkę wody poprzez analogiczny **tetrahedralny stan przejściowy** z uwolnieniem fragmentu $R'-COOH$.

Reduktaza kwasu dwuhydrofoliowego (DHFR) jest obok syntazy tymidylowej kluczowym enzymem - „tarczą” dla projektowania środków przeciwnowotworowych [**PLANSZA 134: synteza TMP z 2'deoxyUMP**] blokujących syntezę DNA. Enzym odtwarza kwas N^5,N^{10} - tetrahydrofoliowy z kwasu dwuhydrofoliowego po przekazaniu grupy CH_3 (N^5CH_2) syntazie tymidylowej, która metyluje pozycję 5 uracylu. **Methotrexate (MTX)**, analog kwasu dwuhydrofoliowego z grupą NH_2 zamiast hydroksylowej w pozycji C(4) i podstawnikiem $N^{10}CH_3$ jest inhibitorem DHFR i blokuje reakcję transferu CH_3 . MTX jest bardzo silnym inhibitorem o stałej $K_I = 5.8 \cdot 10^{-11}$ M co powoduje, że na widmie ^{13}C NMR związku wzbogaconego C-13 w pozycji C(2)

obserwuje się w kompleksie z DHFR dwa rozdzielone sygnały: jeden pochodzący od wolnego związku a drugi (szerszy) od związku w kompleksie z enzymem. Wynika to z powolnej w skali NMR wymiany inhibitora między stanem wolnym i związanym [**PLANSZA 135: MTX w centrum DHFR; ¹³C NMR**]. Przy miareczkowaniu w szerokim zakresie od pH 5 do pH 10 dla związanego MTX **nie obserwuje się** charakterystycznych dla wolnego inhibitora zmian przesunięcia chemicznego w związku z protonacją/dysocjacją (pK 5.7) pozycji N(1); wiązanie MTX przez DHFR zachodzi w formie uprotonowanej i forma ta utrzymuje się do pH prawie 5 jednostek powyżej pK. Pochodna MTXO methotrexatu [**PLANSZA 136: MTXO i NADPH w kompleksie z pcDHFR i hDHFR**] jest dopuszczonym w chemioterapii środkiem antynowotworowym o silnym działaniu ubocznym, podobnie jak 5-fluorouracyl, inhibitor syntazy tymidylowej. Oba inhibitory uniemożliwiają syntezę DNA, ale bardziej uderzają w szybko namnażające się komórki nowotworowe niż w komórki zdrowe. Różnice w strukturach kompleksów MTXO/NADPH z reduktazą bakteryjną *Pneumocystis carinii* i reduktazą ludzką (*Cody et al. Acta Crystallographica D53, 638, 1997*) wyjaśniają różnice w danych kinetycznych i pozwalają na wnioski o **mechanizmie selektywności działania środka leczniczego**. W chemioterapii wykorzystuje się z reguły (choć nie zawsze) **mniej specyficzną specyficzność enzymów kodowanych przez wirusy, bakterie czy komórki nowotworowe** w stosunku do inhibitora, w porównaniu z enzymami kodowanymi przez zdrowe komórki gospodarza. W konsekwencji enzymy o mniejszej specyficzności są silniej blokowane przez podobne do substratu inhibitory.

Ciekawy przykład zmian konformacyjnych enzymu pokazały symulacje Brownowskiej dynamiki molekularnej **esterazy acetylocholiny** (grupa A. McCammona) [**PLANSZA 137: mechanizm katalizy esterazy acetylocholiny**], enzymu hydrolizującego neurotransmitter acetylocholiny na synapsach nerwowych. Krystalograficzna struktura enzymu stwarzała problemy z wyjaśnieniem mechanizmu katalizy: zbyt wąskie „gardło” dla wejścia substratu do centrum aktywnego i wyjścia produktów. 119 ps symulacji i analiza powierzchni kontaktu białka (toczenie kulki wody 1,4 Å) **pokazały przejściowe otwieranie się na ok. 0,3 ps „bocznego kanału” dla substratu i produktów, niewidocznego w strukturze krystalograficznej**. Rozwiązanie równania Poissona-Boltzmana pokazuje linie sił pola elektrostatycznego w miejscu otwartego (*open*) wejścia do bocznego kanału, przyciągającego dodatnio naładowany substrat. Pole to znika w momencie wejścia substratu (*closed*).

Własności katalityczne cząsteczek RNA czyli **rybozymów** wynikają z tworzenia przez nie natywnych struktur przestrzennych, zdolnych do specyficznego wiązania ligandu w formie przejściowej na drodze reakcji. **RNA katalizuje następujące typy reakcji:**

- cięcie RNA czyli **aktywność rybonukleazy**: „samej siebie” w procesach self-splicingu (własności katalityczne intronu) [**PLANSZA 138: self-splicing**] lub innych cząsteczek RNA, rybozomy: hairpin 50n i hammerhead 34n po self-splicingu *Tobacco ringspot virus*, rybozym *Hepatitis delta virus* HDV, *Neurospora varkud satellite* VS;
- dołączanie nukleotydu do łańcucha RNA czyli aktywność **polimerazy**: 395n intron z rRNA pierwotniaka *Tetrahymena* katalizuje (obok cięcia oligoC) dołączanie cytydyny do fragmentu oligonukleotydowego złożonego z cytydyn;
- pełnienie roli katalitycznego **koenzymu** w enzymach białkowych, np. 377n RNA w 20 kDa rybonukleazie P z *E. coli*;
- tworzenie wiązania peptydowego: 23S rRNA jest **transferazą peptydylową**.

Aktywność rybonukleolityczna poszczególnych rybozymów może się różnić co do mechanizmu. Zwykle grupa 2'OH rybozy w formie zdeprotonowanej $-O^{-}$, stabilizowanej przez Mg^{+2} lub zasady RNA w formie kationowej, dokonuje ataku nukleofilowego S_N2 na wiązanie fosfodwuestrowe, utworzenie **dwuanionowego intermediatu oksyfosforanowego** i powstanie dwóch fragmentów, zakończonych cyklicznym 2',3'-cyklofosforanem oraz grupą 5'OH [**PLANSZA 139: mechanizm cięcia RNA przez rybozym**]. Strukturę intermediatu uzyskano na podstawie krystalografii niereaktywnego rybozomu z wanadem (V^{+2}) zamiast fosforu w łańcuchu RNA.

Lipidy; budowa i dynamika błon biologicznych

Błona biologiczna [**PLANSZA 140: błona biologiczna**] jest kompleksem molekularnym zbudowanym z **hydrofobowych** reszt kwasów tłuszczowych (dwa łańcuchy węglowodorowe) połączonych z **fragmentem hydrofilowym**: cukier glicerol połączony z fosfatydylocholiną lub seryną w fosfolipidach, a sfingozyna połączona z pierścieniowym cukrem w glikolipidach. Fragmenty hydrofobowe agregują w **półpłynną warstwę podwójną** z dwiema częściami hydrofilowymi wystawionymi do kontaktu ze środowiskiem wodnym po obu stronach błony. Wewnątrz tkwią białka błonowe oraz **cholesterol usztywniający** fragment, w którym się znajduje. Z zewnętrzną stroną mogą być związane fragmenty komórkowe oraz węglowodanowe (sacharydy). Inne możliwe struktury (fazy) tworzone przez podwójną warstwę lipidowa to [**PLANSZA 141: struktury błon biologicznych**] **pęcherzyk** z wnętrzem hydrofobowym (micella), **liposom** oraz nieco inna forma **heksagonalna** (będzie pokazana dalej). W badaniach często używa się **modelowych systemów** imitujących błony, np. DMPC - fosfolipid dimyristoylphosphatydilcholine. Biosynteza błon (i cholesterolu) jest wieloetapowym procesem, który nie będzie omawiany (patrz Stryer Biochemia).

Dobłą metodą badania błon jest technika magnetycznego rezonansu jądrowego NMR w ciele stałym [**PLANSZA 142: NMR w ciele stałym**]. Punktem wyjścia analizy widm ^{31}P NMR (spin $I = \frac{1}{2}$), ^{14}N NMR i ^2H NMR (spiny $I = 1$, moment kwadrupolowy $eQ \neq 0$) są szerokie linie charakterystyczne dla układów polikrystalicznych i amorficznych (*powder pattern*). Jednak dla błon struktura widm ulega modyfikacji uwarunkowanej przez trzy zasadnicze czynniki:

- dynamiczny charakter półpłynnej warstwy lipidowej: cząsteczki wykonują dyfuzyjne ruchy molekularne;
- obecność innych, tzn. różnych od lipidów, składników błon: jony, białka, itp.
- różne fazy (struktury) warstwy lipidowej: chaotycznie zorientowane fragmenty dwuwarstwowe (system multilamelarny), pęcherzyki, forma heksagonalna.

Doskonałym obiektem badań przy użyciu ^{31}P NMR są zawieszane w wodzie fosfolipidy w systemie multilamelarnym [**PLANSZA 143: ^{31}P NMR fragmentów fosfolipidowych**]. W wyniku szybkiego ruchu dyfuzyjnego wokół osi prostopadłej do powierzchni błony tensor ekranowania σ ulega osiowej symetryzacji do dwóch składowych: σ_{\perp} prostopadłej i σ_{\parallel} równoległej do osi obrotu, dodatkowo modyfikowanych

(σ_{\perp} i σ_{\parallel}) w wyniku ruchu reszty fosforanowej w stożku o ograniczonej amplitudzie. Symulacja z różnymi prędkościami obrotu odtwarza widmo dając w wyniku **czasu korelacji ruchu dyfuzyjnego** $\tau_{c\perp}$ i $\tau_{c\parallel}$ oraz odpowiednie **stałe R dyfuzji rotacyjnej** $\tau_c = 1/(6R)$. Wpływ innych cząsteczek na widmo ^{31}P NMR [**PLANSZA 144: wpływ składników w DMPC na widma NMR**] oraz analogiczne badania wpływu różnych składników (antybiotyki peptydowe, jony, białka) daje NMR jąder o niezerowym momencie kwadrupolowym: ^2H NMR selektywnie deuterowanych fragmentów węglowodorowych i ^{14}N NMR części polarnej lipidu. Z analizy wartości parametrów uporządkowania $S_{CD} = \langle 3\cos^2\vartheta - 1 \rangle$ (^2H NMR; ϑ - kąt między osią główną tensora momentu kwadrupolowego i normalną do powierzchni błony) można charakteryzować model ruchów dyfuzyjnych, a pomiar czasów relaksacji T_1 dostarcza wartości czasów korelacji τ_c w różnych fragmentach błony, a stąd **informacje o jej płynności**. Widma są również uwarunkowane globalną strukturą (fazą) błony [**PLANSZA 145: ^{31}P NMR różnych struktur błony**]: dwuwarstwa, forma heksagonalna, rozmiary pęcherzyków.

NMR pozwala patrzeć na systemy błonowe od strony cząsteczek zintegrowanych z błonami np. białek [**PLANSZA 146: dynamika bakteriorodopsyny w błonie**]. Selektywna deuteracja fenyloalanin bakteriorodopsyny w błonie purpurowej *Halobacterium halobium* dostarcza, na podstawie widm NMR i ich teoretycznej symulacji (uwzględnienie braku symetrii osiowej tensora momentu kwadrupolowego $\eta \neq 0$), informacje o **modelu ruchu dyfuzyjnego**, w tym przypadku mechanizm „flip-flop” przeskoków o 180° i wykluczenie modelu szybkiej dyfuzji rotacyjna o charakterze ciągłym.

WYKŁAD 13

Polisacharydy

Polisacharydy (wielocukry, cukry złożone) to grupa węglowodanów i zarazem biopolimerów, które są złożone z monomerów będących cukrami prostymi o ogólnym **wzorze sumarycznym** $(\text{CH}_2\text{O})_n$, np. glukoza $\text{C}_6\text{H}_{12}\text{O}_6$. Wśród polisacharydów wyróżnia się oligosacharydy zawierające kilka monomerów: disacharydy, trisacharydy, tetrasacharydy itd., oraz wielocukry właściwe, złożone z wielu tysięcy monomerów, do których należą m.in.:

- **skrobia**: materiał zapasowy roślin, gromadzony w owocach, nasionach, korzeniach, liściach, bulwach, rdzeniu łodygi i kłęczach;
- **glikogen**: materiał zapasowy u ludzi i zwierząt, występuje głównie w mięśniach szkieletowych i wątrobie;
- **dekstran**: forma zapasowa dla bakterii i drożdży
- **celuloza i hemiceluloza** (rozgałęziona celuloza) główny składnik budulcowy roślin (ściana komórkowa);
- **pektyna**: składnik (obok hemicelulozy) matrycy ściany komórki roślinnej, w której tkwią mikrofibryle celulozy;
- **chityna**: substancja podporowa budująca pancerze skorupiaków i oskórek owadów;
- **glikozaminoglikany** (żel polisacharydowy): składnik matrycy pozakomórkowej wraz z białkami włóknistymi: kolagen, elastyna;
- **kwasy hialuronowe**: wiążący wodę w skórze człowieka.

Ponieważ każdy monocukier ma kilka grup hydroksylowych, które mogą tworzyć połączenia z innymi monocukrami, potencjalna ilość struktur polisacharydów jest olbrzymia. W szczególności, w przeciwieństwie do białek i kwasów nukleinowych, polisacharydy mogą tworzyć i tworzą **zarówno struktury liniowe jak i rozgałęzione**.

Reakcja syntezy policukrów (polimeryzacja) podpada pod ogólny schemat dla białek i kwasów nukleinowych, tzn. połączenie poprzez grupy funkcyjne (tu dwa hydroksyle) z wydzieleniem cząsteczki wody (szczegóły biochemiczne zależą od typu policukru i nie będą omawiane).

Przykładowa struktura funkcjonalnego cukru **celulozy** zbudowanej z liniowych łańcuchów o długości co najmniej 500 cząsteczek (3 000 do 14 000) glukozy połączonych O-glikozydowymi wiązaniami $\beta 1 \rightarrow 4$ **[PLANSZA 147: struktura celulozy]** i

stabilizowanych **wewnątrzcząsteczkowymi wiązaniami wodorowymi**. Stabilizowane przez **międzycząsteczkowe wiązania wodorowe** warstwy 60÷70 łańcuchów celulozowych tworzą mikrowłókna o średnicy 3,5 nm i do kilku mikrometrów długości.

Wraz z dwoma innymi polocukrami o **strukturze rozgałęzionej**:

- **hemicelulozą**, glukoza $\beta 1 \rightarrow 4$ połączona z łańcuchami bocznymi różnej długości zawierającymi ksyloglukan, galaktozę i fruktozę,

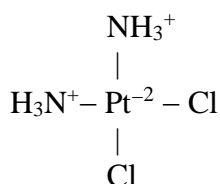
- **pektyną**, kwas D-galaktouronowy $\alpha 1 \rightarrow 4$ z załamaniami (kinks) przez 1,2-połączoną ramnozę,

tworzy **ścianę komórkową**. Układ jest dodatkowo stabilizowany przez związane jony wapnia Ca^{+2} .

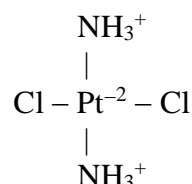
Kompleksy makromolekularne kwasów nukleinowych

Białka i kwasy nukleinowe tworzą ze sobą i z innymi cząsteczkami funkcjonalne kompleksy o różnym stopniu **złożoności** (ilości molekuł w kompleksie) i **różnym stopniu specyficzności** wzajemnego rozpoznawania biomolekuł, wyrażanym przez stałe asocjacji K_{as} . Zagadnienie oddziaływań stabilizujących kompleksy omówię na przykładach struktur i motywów cząsteczkowych biorących udział we wzajemnym rozpoznawaniu.

Najprostszymi cząsteczkami wiązanyymi przez **kwasy nukleinowe** są jony metali w formie „prostej”: Na^+ , K^+ , Mg^{+2} , Ca^{+2} . Ważną klasę stanowią **kompleksy jonowe**, w których jony metali są połączone wiązaniami koordynacyjnymi z podstawnikami, np. cis- i trans-platinum (II):



cis



trans

W wyniku oddziaływań z N(7) i O⁶ guanin oraz grupami keto tyminy w podwójnych helisach DNA zachodzą zmiany konformacyjne i destabilizacja/stabilizacja helis, co może, jak w **przypadku cis-platinum**, wywoływać **transformację nowotworową**.

Liczne związki organiczne o charakterze antybiotyków, mutagenów, kancerogenów lub odwrotnie, środków przeciwnowotworowych oraz środków przeciwwirusowych ma postać **układów wielopierścieniowych** [**PLANSZA 148: interkalacja podwójnej helisy DNA**]. Tego typu związki mogą oddziaływać z helikalnymi formami DNA **interkalując**, czyli wchodząc pomiędzy dwie zestackowane pary zasad i oddziałując z nimi warstwowo (stacking). Struktura helikalna nie jest przy tym niszczone a tylko modyfikowana, co może prowadzić do zaburzeń w ekspresji genów i wymienionych efektów. Interkalatory zwykle wykazują większe powinowactwo do obszarów DNA bogatych w pary GC, czyli **częściową specyficzność**, i mogą przy odpowiednim stężeniu i stałej asocjacji zapełnić do 40% dostępnych miejsc, czyli wejść pomiędzy prawie co drugą parę zasad. Różne interkalatory mogą oddziaływać z różnymi strukturami helikalnymi DNA: duplexy, tripleksy. Silną preferencję do tetrapleksów wykazują cząsteczki z **dwoma inerkalatorami** np. ditercalinium (*Carrasco et al. ChemBiochem. 3, 1235, 2002*)

Miareczkowanie poli(dGm⁵dC)·poli(dGm⁵dC) w formie Z **adriamycyną** (hydroksydaunomycyną) śledzone przy pomocy spektroskopii ³¹P NMR [**PLANSZA 149: wiązanie adriamycyny do podwójnej helisy DNA**] pokazuje przejście do formy B-DNA i stabilizację tej formy przez interkalację. Można więc wiązać działanie tego **cytotoksycznego antybiotyku** z ekspresją genów. Przyjmuje się związek między represją genów i formą Z-DNA *in vivo*.

Aktynomycyna [**PLANSZA 150: oddziaływanie aktynomycyny z helisą DNA**] jest **środkiem przeciwnowotworowym**, który hamuje syntezę mRNA. Analiza NMR interkalacji do krótkich fragmentów helikalnego DNA z dwiema i czterema parami G:C pokazuje struktury kompleksów z jedną i dwiema związanymi cząsteczkami interkalatora; w tym ostatnim przypadku występowanie formy (a) z trzech możliwych. Aktynomycyna stabilizuje fragmenty helisy do których się wiąże. Wykazano również możliwość silnego wiązania 7-aminoaktynomycyny do DNA o strukturze hairpin.

Cząsteczki „białkopodobne” typu **poliamidów** złożonych z trzech pierścieni: pyrołu (Py), 3-hydroksypyrołu (Hp) i imidazolu (Im) połączonych wiązaniami peptydowymi wiążą się z **helisami DNA od strony małej bruzdy** i, w zależności od **kombinacji dwóch pierścieni w pozycjach 3 i 6, selektywnie rozpoznają pary zasad: A:T, T:A, G:C i C:G** [**PLANSZA 151: chemiczne odczytywanie informacji genetycznej**]. Odróżnienie opiera się na niskiej stałej dysocjacji:

$$K_D = [\text{poliamid}][\text{DNA}]/[\text{kompleks}]$$

poniżej 1 nM, w stosunku do wyróżnionej pary. W przypadku 3 pozostałych par zasad i innej kombinacji Py, Hp i Im wartość K_D jest do dwóch rzędów wielkości wyższa. **Specyficzność jest uwarunkowana układem wiązań wodorowych pierścieni poliamidu/pierścieni zasady oraz komplementarności kształtu powierzchni kontaktu (*shape complementarity*)**, co w przypadku rozróżnienia par A:T i T:A wiąże się z rozpoznaniem bardzo małych różnic powierzchni molekularnych.

Szereg oligomerycznych fragmentów RNA i pojedynczonicowych DNA (kilkadziesiąt nukleotydów) wykazuje własności wiązania ligandów niskocząsteczkowych, np. aminokwasy, nukleotydy, ze znacznymi stałymi asocjacji i wysoką specyficznością. Fragmenty te, tzw. **aptamery** pochodzą z selekcji *in vitro*, w których startując z bibliotek sekwencji przypadkowych optymalizuje się te sekwencje w kierunku wysokiego powinowactwa wobec określonych ligandów (SELEX - *systematic evolution of ligands by exponential enrichment*). Przykładem **aptameru DNA** (nie będzie omawiany) jest wiążący białko trombinę kwadrupleks zawierający powtarzalną sekwencję d(G₂T₂G₂TGTG₂T₂G₂).

Własności konformacyjne i sposób wiązania ligandów są podobne w przypadkach wielu **aptamerów RNA**. Przykładem może być 40n aptamer RNA wiążący AMP [**PLANSZA 152: aptamer RNA wiążący AMP**], którego strukturę wyznaczono metodami spektroskopii NMR (widmo odpowiadające protonom wiązań wodorowych). AMP jest wiązane w **kieszeni utworzonej przez giętką pętlę RNA**, która w trakcie wiązania usztywnia się i przybiera właściwą konformację jakby dwóch zintegrowanych szpilek do włosów z jednej nici (induced fit). Obszary helikalne *hairpins* są nachylone do siebie pod kątem 106°. Pętla jest zamknięta przez dwie pary zasad G7:G11 i G17:G34 a AMP jest stabilizowane przez dużą ilość kontaktów jak na małą cząsteczkę AMP: **trzy wiązania wodorowe** (do A12 i dwa do G8) i **stacking typu sandwich** (przy strukturach RNA określałem go jako interkalacyjny) z dwoma pierścieniami, A10 i G11. Ostatnio odkryte tzw. **przełączniki RNA** (*RNA switches; riboswitches*) regulujące ekspresję genów w roślinach, bakteriach i grzybach na poziomie transkrypcji i translacji (np. *adenine riboswitch* w *V. vulnificans* w translacji) zawierają aptamer RNA i platformę ekspresyjną, która przekazuje stan związania ligandu przez aptamer (zmiany konformacyjne RNA). Przekazniki RNA często występują w obszarach 5'UTR na mRNA.

Kompleksy makromolekularne białek z DNA

Centralnym punktem zainteresowania biofizyków i biologów molekularnych są wysoce specyficzne kompleksy **białko - DNA** i **białko - RNA**, które są podstawą zachodzenia i regulacji procesów **ekspresji genu na poziomie transkrypcji, translacji i splicingu**, w tym **reakcji enzymatycznych z udziałem np. endonukleaz restrykcyjnych**. Ze względu na wielkość oddziałujących cząsteczek często badania strukturalne ograniczały się początkowo do fragmentów białkowych i oligonukleotydowych, które bezpośrednio uczestniczą we wzajemnym rozpoznaniu. Obecnie określa się już struktury olbrzymich kompleksów o masie dochodzącej do kilku milionów Da. Według przeglądów znanych jest kilkaset struktur kompleksów białko - DNA i o rząd wielkości mniej struktur białko-RNA (liczby te szybko rosną). **Motywy strukturalne** cząsteczek białkowych (motywy superdrugorzędowe) i **oddziaływania stabilizujące** kompleksy prześledzimy na wybranych, charakterystycznych przykładach.

Kompleks 69 aminokwasów **białka represora** faga 434 z 14 bp **podwójną helisą DNA operatora** zbadano rentgenograficznie [**PLANSZA 153: kompleks represor - operator**]. Dwie cząsteczki (dimer) białka wiążą się z palindromiczną, ściśle określoną sekwencją DNA w postaci charakterystycznego **motywu helisa-skręt helisa** (helix-turn-helix). Jedna z helis motywu stanowi jakby platformę dla drugiej helisy, która bezpośrednio wchodzi w **dużą bruzdę DNA w konformacji B**, przy czym DNA jest nieco wygięte, mocniej zwinięte w środku, o zwężonej małej bruzdzie. Stabilizacja kompleksu zachodzi przez:

- (a) **wiązania wodorowe** (5) warunkujące specyficzność, między resztami bocznymi glutamin 28, 29 i 33 oraz parami AT1, CG2, AT4 i TA5. Dwie glutaminy tworzą wiązania typu podwójnego (*bidentate*), gdzie C=O jest akceptorem a NH₂ donorem protonu do par AT (A-N⁶H/T-O⁴ i A-N⁶H/N7);
- (b) **oddziaływanie kulombowskie** dodatni naładowanych arginin (43, 41 i 16) z ujemnie naładowanymi tlenami grup fosforanowych DNA;
- (c) **oddziaływanie hydrofobowe** grupy metylowej tyminy z pary AT3 z hydrofobową częścią łańcucha Gln29.

Dopasowanie powierzchni obu cząsteczek i właściwe wzajemne rozmieszczenie grup funkcyjnych o odpowiednich ładunkach, zapewniających oddziaływania są niezbędnym warunkiem silnego związania cząsteczek w kompleksie - przy **niespecyficznym**

oddziaływaniu nastąpiło by odsunięcie obu cząsteczek o $2 \div 3 \text{ \AA}$. Z ok. 20 AA tworzących motyw helisa-skręt-helisa **6 kluczowych** jest zakonserwowanych w szeregu innych kompleksów białko-DNA. Motyw ten należy odróżniać od motywu helisa-hairpin-helisa, który występuje w białkach niespecyficznie oddziałujących z DNA: polimeraza β , endonukleaza III.

Omówione charakterystyczne typy oddziaływań stabilizujących kompleksy białko - kwas nukleinowy będą się powtarzać we wszystkich przypadkach, z różnym udziałem każdego z nich. Dodatkowo, częściej w przypadku kompleksów z RNA, stosunkowo rzadko w DNA obserwuje się:

(d) **oddziaływania warstwowe (stacking)** pierścieni aromatycznych aminokwasów i zasad nukleinowych.

Tego typu stabilizacja pojawia się w **potrójnym kompleksie transkrypcyjnym** (struktura krystalograficzna) w eukaryotach: fragment białka TBP, które rozpoznaje TATA-box w DNA, fragment czynnika transkrypcyjnego TFIIB i 14 bp podwójna helisa DNA zawierająca sekwencję TATA [**PLANSZA 154: kompleks TFIIB-TBP-DNA**]. Białko TBP w postaci **dwudomenowej antyrównoległej β -kartki** tworzy strukturę podobną do siodła, która wiąże DNA od strony **małej bruzdy**. DNA ulega częściowemu **rozwinęciu i wygięciu** między dwoma załamaniami (kinks) wywołanymi przez dwie pary fenyloalanin (F165/F148 i F57/F74), spełniające jakby rolę klinów między AT3/AT4 i AT9/GC10. Dwie z fenyloalanin (po jednej w parze) stackingują interkalacyjnie.

Ciekawym motywem białkowym, rozpoznającym sekwencje z dużym udziałem par GC w podwójnie helikalnym DNA jest **palec cynkowy** (*cynk finger*, [**PLANSZA 155: palec cynkowy**]). Jest to struktura superdrugorzędowa ok. **30 AA zbudowana z helisy i dwuniciowej antyrównoległej β -kartki**, utrzymywane przez jon Zn^{+2} skoordynowany przez 4 łańcuchy boczne cystein lub 2 histydyny i 2 cysteiny. Zwykle kilka palców połączonych linkerami o różnej długości oddziałuje z sąsiadującymi sekwencjami po trzy pary. W oddziaływaniu takich trzech palców cynkowych białka Zif263 (*mouse intermediate early protein*) helisa każdego z nich wchodzi w kontakt z **dużą bruzdą DNA** o sekwencji 5'-GCG TGG GCG-3' (jeden palec - 3 bp) poprzez **podwójne wiązania wodorowe między NH_2 reszt argininowych (donory) i $\text{N}7/\text{O}^6$ guaniny**. Dodatkową stabilizację tego układu dają wiązania wodorowe drugiego protonu grupy NH_2 argininy z kwasem asparaginowym w samym białku. Palec cynkowy jest powszechnie występującą strukturą, także w receptorze estrogenowym i czynniku transkrypcyjnym TFIIB. Może

także rozpoznawać **pętlę jednoniciowego RNA**, ale oddziaływanie ma tu inny schemat i występuje zarówno **z łańcuchem fosfocukrowym helisy RNA** jak i **zasadami pętli**.

W przypadku białek dimerycznych, np: aktywator GAL4 drożdży (regulacja katabolizmu galaktozy) dodatkowym czynnikiem stabilizującym kompleks DNA - białko może być **zamek leucynowy** (*leucine zipper*) **[PLANSZA 156: zamek leucynowy]**. Dwie **amfipatyczne helisy** dimeru białkowego oddziałują ze sobą **hydrofobowo** poprzez alifatyczne leucyny, rozmieszczone po jednej stronie każdej z helis.

Charakterystyczny fold dwunukleotydowy złożony z helis i β -kartki (*dinucleotide fold*, *Rossmann fold*) zaobserwowano w strukturze wysoce specyficznej **endonukleazy restrykcyjnej EcoRI** tnącej palindromiczny fragment d(GAA•TTC) zostawiając „lepkie końce 5'-AATT i TTAA-3' **[PLANSZA 157: fold dwunukleotydowy; endonukleaza EcoRI]**. Wysoka specyficzność enzymu w stosunku do sekwencji DNA jest oparta na bardzo rozbudowanej sieci wiązań wodorowych białko-DNA.

WYKŁAD 14

Kompleksy makromolekularne białek z RNA

Fold dwunukleotydowy w formie pięcioniciowej β -kartki i czterech helis występuje również w syntetazie glutaminylo-tRNA^{Gln} i jest odpowiedzialny za wiązanie części akceptorowej tRNA i ATP [PLANSZA 158: fold dwunukleotydowy; syntetaza GlnRS-tRNA]. Leucyna 136 rozbija ostatnią parę helikalnego trzonu akceptora aminokwasu U1:A72, a układ jest stabilizowany poprzez wiązania wodorowe arginin 133 i 192 z fosforanami tRNA i dodatkowo przez wiązanie wodorowe A72–P=O...H₂N–G73. Zasady pętli antykodonu [PLANSZA 158 B] są „wywichnięte” do białka i tworzą bogaty układ wiązań wodorowych z resztami aminokwasów.

Specyficzne kontakty reszt aminokwasowych i łańcucha głównego z regionami jednoniciowymi i pętlami hairpin RNA są typowe dla oddziaływań białko-RNA. Drugą charakterystyczną cechą tych oddziaływań jest stabilizacja kompleksów przez **stacking**. Powszechnie występującą domeną w białkach wiążących RNA jest **motyw RRM** (RNA - recognition motif; RBD - RNA-binding domain) [PLANSZA 159: motyw RRM], pokazywany już poprzednio jako przykład foldu trzeciorzędowego [PLANSZA 95], w postaci β -kartki otoczonej fragmentami helikalnymi i pętlami. Przy wiązaniu RNA międzycząsteczkowy stacking obejmuje 4 układy π -elektronowe: pierścieni i reszty karboksylowej Phe56/A11/C12/Asp92.

Inicjacja procesu translacji w eukaryotach zachodzi poprzez specyficzne rozpoznanie **struktury kapu na 5'-końcu mRNA** [PLANSZA 160: inicjacja translacji; kompleks eIF4E – mRNA 5'kap] przez eukaryotyczny inicjacyjny czynnik białkowy eIF4E (PLANSZA 104). Motyw wiążący kap w formie kieszeni hydrofobowej utworzonej przez wygiętą jak dłoń antyrównoległą β -kartkę, krótką helisę i pętlę S1,2 i S3,4 przypomina motyw RRM, ale nie jest z nim tożsamy. **Kluczowym dla specyficzności oddziaływaniem jest stacking kation- π typu sandwich dodatnio naładowanego** pierścienia 7-metyloguaniny między dwa tryptofany. Podobny stacking obserwuje się w szeregu innych białek specyficznie wiążących kap, przy czym pierścienie tryptofanu mogą być zastępowane przez tyrozyny, Wiązania wodorowe stabilizujące 7-metyloguaninę (trzy: do reszty kwasu glutaminowego E103–COO⁽⁻⁾ i NH w łańcuchu Trp102) **nie** decydują o specyficzności: pierścień guaninowy jest zdolny do utworzenia tych samych wiązań wodorowych ALE nie może zastąpić dodatnio naładowanego 7-metyloguaniny.

Inne motywy zaobserwowane w kompleksach białko-RNA (których nie będę omawiał szczegółowo) to:

- **PWI** w formie wiązki czterech helis w faktory białkowym SRm160 w splicingu;
- motyw **k-TURN** w U4 **snRNA** w splicingu (także w rybosomie): 3-nukleotydowe wybrzuszenie pomiędzy kanoniczną helisą i niekanoniczną helisą z dwu par G:A.

W dziedzinie kompleksów białko-RNA osiągnięto olbrzymi postęp w wyznaczaniu dużych struktur metodami rentgenograficznymi, gdzie wychodzi się poza układy modelowe w kierunku **pełnych układów subkomórkowych**, jak to się określa układów **mezoskopowych** (pomiędzy mikroskopowymi i makroskopowymi). Szereg prób krystalograficznych, np. 70 kDa kompleks fragmentu domeny centralnej 16S rRNA (104n) plus białka S6, S8, S11, S15, S18 [**PLANSZA 161: rybosom: kompleks 16S rRNA z białkami**] doprowadziło w końcu do uzyskania struktur: dużej 50S i małej 30S, podjednostki rybosomu bakterii *Thermus thermophilus*, co umożliwiło skonstruowanie modelu atomowego pełnego rybosomu prokaryotycznego [**PLANSZA 162: struktura krystalograficzna rybosomu**]: mRNA, rRNA, tRNA i kilkadziesiąt białek. Obraz jest złożeniem kilku odrębnych kompleksów krystalograficznych. W pracy *Yusupov et al. Science 292, 883, 2001* opublikowano strukturę rentgenograficzną kompletnego **70S rybosomu, 2,7 mln Da**, aczkolwiek z niską rozdzielczością 5,5 Å. W końcu udało się wykryć całe rybosomy z rozdzielczością 3,7 Å: *Korostelev et al. Cell 126, 1065, 2006* (3 rRNA, 2 tRNA, mRNA, 54 białka). Z rozdzielczością 3,5 Å uzyskano dyfraktogram dla 80S rybosomu drożdżowego, **3,3 mln Da** (*Ben-Shem et al. Science 330, 1203, 2010*; 4 rRNA, 2 tRNA, 79 białek). „Rekord” wielkości układu dla którego pokazano, że można wyznaczyć strukturę NMR to 870 kDa kompleks GroEL/GroES: *Flaux et al. Nature 418, 207, 2002* (grupa K. Wüthricha w ETH w Zurichu).

Dla tak dużych układów molekularnych standardowe podejście w badaniach strukturalnych polega na **połączeniu rentgenografii (~0.01 Å) z technikami o średniej rozdzielczości: cryoEM lub AFM (~5 ÷ 10 Å)**, które „atakują” jakby od dwóch stron: atomowej i globalnej (*hybrid approach*). Technika mikroskopii transmisyjnej cryoEM rozwija się bardzo intensywnie w ostatnich latach w kierunku **uzyskania rozdzielczości atomowej**, aktualnie 1 ÷ 2 Å w mikroskopach 300 kV. Wyznaczenie struktury przestrzennej jest warunkiem koniecznym, ale z reguły niewystarczającym dla pełnej **charakterystyki mechanizmu funkcjonowania** tak złożonych kompleksów i stanowi punkt wyjścia do stosowania technik „dynamicznych”, np. śledzenie translacji w

pojedynczym rybosomie przy zastosowaniu szczypek optycznych: *Wen et al. Nature 452, 598, 2008*. Rozwinięcie podejścia hybrydowego stosuje się do badania dużych struktur subkomórkowych o masach cząsteczkowych kilkudziesięciu mln Da (a nie kilku mln Da jak poprzedni), np. 50 mln Da **nuclear pore complex NPC** złożony z 456 białek w błonie jądrowej (dwie prace *Alber et al. Nature 450, 683, 2007; 695, 2007*). Zastosowane tu „podejście proteomiczne” polega na wykorzystaniu maksymalnej ilości technik doświadczalnych: ultrawirowanie analityczne, cryoEM, immunoEM, chromatografia powinowactwa, rentgenografia składników itp., w celu wygenerowania **zbioru danych, które są przekładane na więzy przestrzenne** wzajemnych relacji między składnikami układu. Po **optymalizacji** zbioru potencjalnych struktur wybiera się te, które są najlepiej zgodne z danymi doświadczalnymi.

Z punktu widzenia termodynamiki specyficznego kompleksowania **[PLANSZA 163: termodynamika kompleksów]** zarówno entalpia jak i entropia asocjacji silnie zależą od temperatury. Jest to związane z niezerową zmianą pojemności cieplnej (ciepła właściwego) przy stałym ciśnieniu $\Delta c_p < 0$ (nieliniowe równanie van't Hoffa). Maksimum obserwowanej stałej równowagi K_{as} przypada w **temperaturze T_H** zerowej entalpii wiązania ΔH° . W temperaturze T_S zeruje się człon entropowy. Poniżej temperatury T_H proces jest napędzany entropowo, poniędzy T_H a T_S entalpowo-entropowo, a powyżej T_S entalpowo. **Kompensacja entropowo-entalpowa zapewnia małą zmienność ΔG° w szerokim zakresie temperatur.** Ujemna zmiana c_p wynika głównie z **usunięcia niepolarnych fragmentów powierzchni cząsteczek ΔA_{np} przy asocjacji z roztworu wodnego**. Podobne zależności termodynamiczne można zaobserwować w przypadku przeniesienia cząsteczek benzenu z wody do czystej cieczy.

Biofizyka w biotechnologii i medycynie (wybrane zagadnienia)

Po zsekwencjonowaniu genomu ludzkiego priorytety biologii i biofizyki molekularnej przesunęły się w kierunku charakterystyki proteomu. Nowa dziedzina, która powstała na gruncie obu wymienionych dyscyplin, **PROTEOMIKA**, koncentruje się na **wysocze przepustowym** (*high throughput*) scharakteryzowaniu proteomu i obejmuje następujące etapy [**PLANSZA 164: proteomika**].

(I) Identyfikacja białek wytwarzanych w określonej komórce danego typu, tkance czy całym organizmie w oparciu o znany genom i bezpośrednią izolację z komórek:

- **dwuwymiarowa elektroforeza żelowa** (separacja według masy cząsteczkowej i punktu izoelektrycznego);
- **spektrometria masowa** do określania struktury pierwszorzędowej: np. electrospray ionization tandem mass spectrometry (ESI-MS/MS)

(II) Określenie współdziałania białek w ramach funkcjonalnych kompleksów molekularnych, np. metoda tandem-affinity-purification (TAP) zastosowana do drożdży (*Gavin et al., Nature 415, 141, 2002*)

(III) Określenie struktur przestrzennych cząsteczek z rozdzielczością atomową (proteomika strukturalna):

- otrzymanie przez **nadekspresję w bakteriach** lub **systemach cell-free** odpowiednich ilości białka:
 - (a) zawierającego **selenometioninę** (**dyfrakcja promieni X** w wiązce synchrotronowej, metoda multiwave anomalous dispersion **MAD** do rozwiązania problemu fazowego), LUB
 - (b) **podwójnie znakowanego** $^{15}\text{N}/^{13}\text{C}$, a dla większych białek **potrójnie znakowanego** $^2\text{H}/^{15}\text{N}/^{13}\text{C}$ (wielowymiarowy magnetyczny rezonans jądrowy **2D, 3D i 4D NMR**),
- znalezienie w bazach danych białek o **znanych strukturach i odpowiednio wysokiej homologii sekwencji** do zwijania komputerowego (*data base search*);
- **przygotowanie próbki białkowej: krystalizacja** (*X-ray crystallography*); uzyskanie **odpowiedniej rozpuszczalności** w wodzie i rozcieńczonych ciekłokrystalicznych roztworach micellarnych (NMR);
- **rozwiązanie struktury** i budowa modelu cząsteczki;
- **weryfikacja** modelu.

PROTEOMIKA STRUKTURALNA, koncentruje się na kompleksowym i wysoce wydajnym (*high throughput*) wyznaczaniu **struktur białek** oraz **związków sekwencja - struktura - funkcja** dla całych grup cząsteczek białkowych, **współdziałających w ramach określonych odcinków szlaków metabolicznych** tak, aby można było konstruować molekularne mechanizmy procesów biochemicznych. **Celem ostatecznym** byłoby funkcjonalne scharakteryzowanie, poprzez strukturę (założenie **SAR - Structure - Activity - Relationship**) wszystkich białek znanych po zsekwencjonowaniu genomu ludzkiego i innych organizmów. **Celem bardziej realnym** wydaje się uzyskanie **reprezentatywnych struktur w każdej rodzinie białek o określonym typie foldu**. Liczbę białek funkcjonujących w samym proteomie człowieka szacuje się na ok. 300 tys., kodowanych przez ok. 30 tys. genów. Wymaga to rozwiązywania **jednej struktury w ok. 2 - 3 tygodnie**. Celem praktycznym badań strukturalnych jest **wybór właściwych "tarcz" czyli receptorów białkowych dla różnego typu inhibitorów czy aktywatorów jako środków farmakologicznych**: antywirusowych, antynowotworowych, cytostatycznych czy sprzyjających proliferacji określonych typów komórek. Następnie **projektuje się takie środki (*rational drug design*)** przez analizę stabilizujących kontaktów chemicznie modyfikowalnych ligandów w centrum aktywnym receptora.

Zespoły naukowe biologów molekularnych, krystalografów, spektroskopistów NMR i informatyków (rozwój software'ów komputerowych) powołane do realizacji w zakresie proteomiki działają w ramach tzw. **Large Scale Facilities (LSF)** w USA, EU, Japonii i Kanadzie. Konsorcja te ściśle **współpracują z firmami farmaceutycznymi, firmami konstruującymi platformy robocze do automatyzacji czynności w ramach proteomiki** (Proteomic Analyser firmy Applied Biosystems [**PLANSZA 165: analizator proteomiczny Applied Biosystem**]) i firmami produkującymi aparaturę typu dyfraktometry, spektrometry NMR, spektrometry MS. Finansowane odbywa się w ramach instytucji takich jak: NIH (USA), Programy Ramowe UE. Znaczny udział w finansowaniu mają też firmy prywatne: farmaceutyczne, Celera Genomics (sekwencjująca genomy), firmy proteomiczne GeneProt (Genewa), MDS Proteomics (Toronto). Szacuje się, że światowy rynek na aparaturę, materiały i usługi związane z proteomiką przekroczył **\$5 mld w 2005**.

Proteomika jest również wyzwaniem dla **bioinformatyki** (hardware i software) określanej w ujęciu „filozoficznym” jako ontologia proteomiki (gromadzenie i dostęp do danych). Jednocześnie chodzi o wypracowywanie metod komputerowych projektowania

molekularnego. Z ponad 100 tys rozwiązanych struktur białkowych, udziały poszczególnych technik wynoszą:

- (A) rentgenografia czyli dyfrakcja promieni X na monokryształach białek ok. 81%;
- (B) magnetyczny rezonans jądrowy NMR w roztworze o.k. 17%;
- (C) modelowanie teoretyczne na podstawie podobieństwa sekwencji aminokwasowej z homologicznymi białkami o znanej strukturze poniżej 2%.

Praktyczne wykorzystanie zdobyczy proteomiki (i biofizyki molekularnej) opiera się na analizie struktury centrum wiążącego ligand w wybranym receptorze białka ‘tarczy’. Do niedawna poszukiwano inhibitorów **metodą „prób i błędów”** modyfikując chemicznie związek wyjściowy o określonym działaniu tak aby zwiększyć efekt. **Racjonalne projektowanie leków** w oparciu o znaną strukturę receptora, np. określonego białka enzymatycznego, jest zdecydowanie bardziej wydajne. Procedura *rational drug design* w połączeniu z testowaniem doświadczalnym tylko dobrze rokujących związków modelowanych **skraca ok. trzykrotnie czas** od badań podstawowych do klinicznych zastosowań leku (średnio z ok. 10 do 3 lat; dane „optymistyczne”).

Jeśli w zainfekowanej komórce pojawia się nowe białko i jest kluczowe dla rozwoju wirusa lub bakterii albo białko tego typu jest zmodyfikowane w stosunku do analogu w komórce zdrowej procedury proteomiczne umożliwiają jego wyłowienie **[PLANSZA 166: proteomika w racjonalnym projektowaniu leków]**, uzyskanie w odpowiedniej ilości i określenie struktury przestrzennej. Zaprojektowana na podstawie tej struktury cząsteczka ligandu (inhibitora) selektywnie blokuje nieporządane białko **[PLANSZA 136]**. Przy słabszym wiązaniu inhibitora przez enzym gospodarza może to uniemożliwić namnażanie i spowodować zniszczenie wirusa lub nowotworu przez system immunologiczny. Przykładem tego typu projektowania może być poszukiwanie inhibitorów **fosforylasy nukleozydów purynowych PNP** (*purine nucleoside phosphorylase*) **[PLANSZA 167: zaprojektowanie inhibitora PNP]**, enzymu występującego od bakterii do ssaków, który rozkłada nukleozydy purynowe do zasady i cukru (przecięcie wiązania glikozydowego) w obecności reszty fosforanowej. Enzym tnie różne nukleozydy, a więc i podawane z zewnątrz środki farmakologiczne które są często pochodnymi nukleozydów purynowych. Podanie silnego inhibitora PNP razem z takim środkiem podniosłoby więc efektywność jego działania. W oparciu o znaną strukturę krystalograficzną białka udało się zaprojektować związki, które oddziałują silnie ze wszystkimi trzema fragmentami centrum wiążącego substraty reakcji - **analogi multisubstratowe**. Przykładowa analiza oddziaływań w części wiążącej zasadę oddaje

charakter postępowania. Zamiana guaniny na 8-aminoguaninę daje jedno więcej wiązanie wodorowe. Z kolei 9-deazaguanina zastępuje jedno słabsze wiązanie silniejszym, a ponadto uniemożliwia enzymowi przecięcie wiązania typu C–C zamiast C–N. Niestety jednoczesne podstawienie 8-NH₂ i zamiana N(9) na C **dają w sumie niekorzystny efekt steryczny** i pogarszenie powinowactwa do enzymu w porównaniu z każdą ze zmian oddzielnie. Jednym z **optymalnych rozwiązań okazuje się 9-deazaguanina jako zasada, chlorobenzen jako odpowiednik cukru a reszta kwasu octowego jako odpowiednik fosforanu**. Uzyskany w ten sposób inhibitor ma wartość stałej inhibicji K_i ponad 100 razy mniejszą niż wszystkie inhibitory znalezione metodą konwencjonalną.

Interesującym przykładem **biotechnologicznego i medycznego** wykorzystania wysokiej specyficzności rozpoznawania wzajemnego biomolekuł jest konstrukcja tzw. **mikromacierzy** (*microarrays*) lub chip'ów. Są one realizowane zasadniczo w dwóch wersjach: DNA chips i protein chips. W obu przypadkach łańcuchy polimerowe jednoniciowego DNA lub polinukleotydu są unieruchomiane jednym końcem na powierzchni (szklanej, krzemowej, złotej, syntetycznego polimeru), przez odpowiednio dobraną cząsteczkę linkera. Każde miejsce próbujące tzw. *probe site* (*spot*) **[PLANSZA 168: mikromacierze DNA]** zawiera jednakowe cząsteczki biopolimeru, czyli o tej samej sekwencji i o określonej **gęstości** (*density*), czyli ilości cząsteczek na jednostkę powierzchni. Cały chip składa się z dużej ilości takich *spots*, każda o innej sekwencji. Ilość różnych sekwencji biopolimeru na chipie warunkuje jego tzw. **gęstość informacyjną** (*complexity*); obie wielkości są parametrami charakteryzującymi chip. Nazwa chip bierze się stąd, że pierwsze tego typu obiekty były wykonywane techniką fotolitograficzną jak w produkcji mikroprocesorów (computer chip).

Chipy DNA są złożone z **syntetycznych oligomerów** długości do 60n lub **cDNA** (complementary DNA) - produktów *polimer chain reaction* (PCR) o długości do 300 nukleotydów, które mogą **wiązać komplementarnie** (hybrydyzacja) pojedyncze nici DNA lub RNA. Mikromacierze DNA wykorzystywane są zasadniczo do dwóch celów:

- (1) **Genotypowanie** (*genotyping*): porównanie sekwencji DNA na chipie i DNA w badanej próbce w celu określenia, jakie geny występują w tej próbce lub jaka jest struktura pierwszorzędowa jeszcze niezsekwencjonowanego DNA.
- (2) **Wyznaczanie profilu ekspresyjnego** (*gene-expression profiling*): które geny w tkance uległy ekspresji.

Przykład tego drugiego zastosowania pokazuje testowanie zmian ekspresji genów np. wątroby jako reakcja na podanie leku. Chip zawiera DNA fragmentów **tysięcy genów**, po jednym na każdy spot. Po pobraniu dwóch próbek komórek wątroby, do jednej wprowadza się testowany lek. Następnie z każdej zbiera się mRNA, przeprowadza odwrotną transkrypcję do cDNA i wprowadza na nie znaczniki fluorescencyjne o różnym kolorze, np. czerwonym dla cDNA z próbki potraktowanej lekiem a zielonym dla próbki bez leku. Po wprowadzeniu znakowanych cDNA do chipu zachodzi **hybrydyzacja chip-DNA/cDNA**, dająca informację które geny uległy ekspresji czyli uległy transkrypcji i powstało mRNA. Za pomocą skanera oblicza się komputerowo stosunek czerwony/zielony w każdym spot i generuje **kolorowy wydruk**. **Czerwony: silny wzrost aktywności genu po podaniu leku; zielony: silny spadek aktywności po podaniu leku; żółty: geny jednakowo aktywne; bez koloru brak aktywności danego genu w obu przypadkach.** Mając dodatkowo różne profile reakcji dla znanych toksyn wątroby można porównać z nimi profil testowanego leku i szybko **wykluczyć lub zaakceptować lek** na podstawie podobieństwa profili (kolorowego schematu). Innym przykładem zastosowania **gene-expression profiling** jest "prognozowanie" rozwoju nowotworów bez konieczności ustalenia mechanizmów. Z podobieństwa profili ekspresji do znanych przypadków rozwoju/hamowania po podaniu środków" przeciwnowotworowych można oceniać prawdopodobieństwo rozwoju choroby i skuteczności terapii. Wydajność hybrydyzacji jest na tyle duża, że pozwala na charakteryzowanie indywidualnych (osobniczych) różnic w sekwencjach genomów czyli **single nucleotide polymorphism** (SNPs). Alternatywą dla konwencjonalnych chips przy detekcji sekwencji są **biosensory** (*Current Opinion Chem. Biol.* 6, 816, 2002). Szczegóły chemiczne i techniczne produkcji DNA microarrays podaje przykład: *Pirrunga'a Angewandte Chemie International Edition* 41, 1276, 2002.

Chipy białkowe [**PLANSZA 169: mikromacierze białkowe**] oparte są na **specyficznym wiązaniu antygen-przeciwciała**. Unieruchomione na chipie cząsteczki przeciwciał w stosunku do charakterystycznych białek kodowanych przez bakterie i wirusy wiążą takie białka z próbki krwi. Dodanie fluorescencyjnie znakowanych przeciwciał w stosunku do innych fragmentów białek ujawnia związanie się antygenu i przeciwciała na chipie przez kompleks typu "sandwich". Zabarwienie określonego miejsca na chipie sugeruje wystąpienie w krwi antygeny pochodzącego od jakiegoś czynnika patogennego. W wersji tzw. **funkcjonalnych chipów białkowych** wprowadza się na chip same białka (a nie przeciwciała). Szczegóły techniczne podaje przegląd *Zhu & Snyder Current Opinion in Chemical Biology* 7, 55, 2003.

Aktualne, „gorące” tematy w biofizyce molekularnej [**PLANSZA 170: kluczowe zagadnienia współczesnej biofizyki molekularnej**] mogą stanowić również dobre **PODSUMOWANIE** kierunków praktycznego wykorzystywania zdobyczy w tej dziedzinie w **biotechnologii i medycynie**.

W wyznaczaniu struktur przestrzennych biomolekuł: szczególnie białka błonowe, włókniste i agregaty amyloidalne stanowią prawdziwe wyzwanie metodologiczne. Olbrzymie bogactwo struktur obserwuje się dla świata cząsteczek RNA, a szczególnie intensywne badania strukturalne DNA dotyczą sposobu organizacji podwójnej helisy w strukturę chromosomu na poziomie zwinięcia nici chromatynowej - domeny chromosomu bakteryjnego i chromosomu eukariotycznego.

Zwijanie białek to charakterystyczny przykład ścisłego współdziałania metod doświadczalnych i komputerowych w poszukiwaniu mechanizmów zwijania i ich wyjaśniania na gruncie fizyki oraz przewidywania struktur przestrzennych na podstawie sekwencji aminokwasowej.

Zastosowanie metod *single molecule* w zakresie spektroskopii fluorescencyjnej, obrazowania i manipulowania stwarza unikalne możliwości śledzenia skomplikowanych procesów biologicznych na poziomie molekularnym.

Badania nanoukładów biologicznych zaczyna mieć kluczowe znaczenie w dostarczaniu środków farmakologicznych do komórek i tkanek oraz konstrukcji wysoce zminiaturyzowanych układów elektrycznych i komputerowych, w tym w systemach hybrydowych z półprzewodnikami.

Nieinwazyjne obrazowanie tkanek i narządów oraz procesów które w nich zachodzą dają możliwość jednoczesnego wglądu w anatomię i morfologię dla celów diagnostyki i terapii.