

BIOFIZYKA DOŚWIADCZALNA
(60h III semestr specjalności studiów II stopnia: ZFBM „Biofizyka molekularna” i FIZYKA „Biofizyka”)

1. Biofizyka i jej miejsce w badaniach z zakresu fizyki i nauk biomedycznych; kluczowe zagadnienia współczesnej BIOFIZYKI MOLEKULARNEJ (*1 wykład*)
2. **Struktura chemiczna**, metabolizm i biologiczna rola biopolimerów: kwasy nukleinowe DNA i RNA, białka, lipidy i polisacharydy (przegląd i uzupełnienie wiadomości; *3 wykłady*)
3. **Oddziaływania elektrostatyczne i hydrofobowe** stabilizujące struktury makrocząsteczek i kompleksów molekularnych; rola rozpuszczalnika (*1 wykład*)
4. Przegląd zaawansowanych TECHNIK badania struktur i dynamiki **biopolimerów** i ich kompleksów z **rozdzielczością molekularną**; techniki **mikroskopowe**; manipulacja **pojedynczymi** cząsteczkami (*1 wykład*)
5. GENOMIKA i **sekwencjonowanie DNA**; budowa chromosomu, **mutageneza, kancerogeneza** i systemy naprawcze, **nanostruktury DNA** (*3 wykłady*)
6. PROTEOMIKA: zastosowania spektrometrii mas (MS); **przestrzenne struktury** białek: dyfrakcja rentgenowska i magnetyczny rezonans jądrowy NMR, **dynamika** ruchów molekularnych, interaktom (*6 wykładów*)
7. ZWIJANIE BIAŁEK (*PROTEIN FOLDING*): **mechanizmy zwijania** *in vitro* i *in vivo*: badania doświadczalne i teoretyczne; „błędne” zwijania i agregacja (formy amyloidalne) (*3 wykłady*)
8. **Struktury przestrzenne** i zwijanie **RNA**; TRANSKRYPTOMIKA: **mikromacierze** i profile ekspresji, biosensory, sekwencjonowanie (*2 wykłady*)
9. **Kompleksy** biomolekularne: kinetyka i dynamika asocjacji (*1 wykład*)
10. SPECYFICZNE ODDZIAŁYWANIA biopolimerów w kompleksach białko-ligand białko-DNA i białko-RNA; projektowanie leków (*DRUG DESIGN*); struktury subkomórkowe - **motory molekularne** (*2 wykłady*)
11. BŁONY BIOLOGICZNE i PROCESY BIOENERGETYCZNE:: transfer elektronów i protonów (*2 wykłady*)
12. Biopolimery w „żywej komórce” (*in vivo* NMR, *in cell* NMR) (*1 wykład*)
13. **Obrazowanie NMR** w DIAGNOSTYCE MEDYCZNEJ (*MRI* i *fMRI*); porównanie z tomografią komputerową (*X-ray CT*) i pozytonową (*PET*) (*1 wykład*)
14. „Niekonwencjonalne” wykorzystanie polimerów biologicznych (*1 wykład*)

LITERATURA

1. W. Saenger „Principles of nucleic acid structures”
2. T.E. Creighton „Proteins. Structures and molecular properties”
3. I.N. Serdyuk, N.R. Zaccai i J. Zaccai „Methods in Molecular Biophysics”
4. Aktualne odnośniki literaturowe (podawane na wykładzie)

Biofizyka i jej miejsce w badaniach z zakresu fizyki i nauk biomedycznych; kluczowe zagadnienia współczesnej BIOFIZYKI MOLEKULARNEJ

Wykład 1

Biologiczny charakter obiektów zainteresowania BIOFIZYKI (lub FIZYKI BIOLOGICZNEJ) jako dyscypliny przyrodniczej wymaga wszechstronnego zastosowania **pojęć, metodologii i bazy eksperymentalnej** z zakresu szeroko rozumianej fizyki (fizykochemii). Żywiolowy rozwój biologii w ostatnim pięćdziesięcioleciu nie byłby możliwy bez kilku przełomowych spostrzeżeń i eksperymentów, z których chyba (selekcja mniej lub bardziej arbitralna) najważniejsze to:

(a) **Rozważania Schrödingera** na temat wielkości genów niewątpliwie wpłynęły na nurt badań biologicznych związany z poszukiwaniem systemów naprawczych materiału genetycznego, gdy okazało się, że stopień złożoności molekularnej DNA w chromosomie wyklucza jego stabilność termodynamiczną.

(b) Semikonserwatywny model replikacji DNA, tzn. że dwie dwuniciowe cząsteczka DNA powstają w wyniku syntezy nici potomnych na matrycy każdej z nici pierwotnych po ich rozwinięciu, został udowodniony w **doświadczeniu Meselona–Stahla z ultrawirowaniem DNA**, które izolowano z bakterii hodowanych na pożywkach wzbogacanych N-15.

(c) Model **podwójnej helisy DNA** został opracowany przez Watsona i Cricka na podstawie obrazów dyfrakcji promieniowania rentgenowskiego (na włóknach).

Wiodące kierunki badań biofizyki molekularnej, które będą bardziej szczegółowo przedstawione dalej, są ściśle związane z fizyką, chociaż **badane obiekty** mają charakter biologiczny lub chemiczny. Kilka wybranych przykładów.

1. **Proteomika strukturalna** ma na celu scharakteryzowanie struktur przestrzennych wszystkich białek kodowanych przez genomy, w tym człowieka, poznane z dokładnością do jednego nukleotydu składowego (zsekwencjonowane). Ujmując w skrócie, proteomika strukturalna to techniki fizyczne, rentgenografia i magnetyczny rezonans jądrowy (wyznaczanie struktur) plus bioinformatyka (bazy danych zebranego materiału).

2. **Termodynamiczna charakterystyka złożonych obiektów**, białek, kwasów nukleinowych i kompleksów molekularnych tych cząsteczek (największe, dla których struktury określono rentgenograficznie sięgają 3.5 mln jedn. masy atom.) stanowi wyzwanie dla fizyki statystycznej. **Kwantowe rozszerzenia klasycznej dynamiki**

molekularnej (komputerowe, numeryczne rozwiązanie równań ruchu atomów polimerów w roztworze) wymagają analizy układu częściowo klasycznego a częściowo kwantowego.

Z drugiej strony, badania z zakresu fizyki w biologii wpłynęły zapładniająco na szereg dziedzin fizyki, czego klasycznym przykładem jest żywiolowy rozwój badań obiektów typu „*nano*”, przewidywane przez Richarda Feynmana w końcu lat 50-tych ([review: Wong et al. Genes & Develop.27, 2397, 2013](#)). W badaniach tego typu **nie ma podziałów na dziedziny** - trzeba łącznie stosować fizykę, biologię i chemię w opisie zjawisk i budowie modeli oraz planowaniu eksperymentów. Manipulacja pojedynczymi cząsteczkami wymaga **narzędzi laserowych** typu „szczypiec optycznych” (A. Ashkin nagroda Nobla z fizyki w 2018 r.), a **mikromacierze** to przykład drogi od technik epitaksjalnych ciała stałego do określania profili ekspresji genów w komórce. Mikroskopia sił atomowych **AFM** znajduje zastosowanie od biologii molekularnej do fizyki ciała stałego. Zainteresowania analizą żywych i funkcjonujących układów biologicznych zaowocowały rozwojem **termodynamiki stanów nierównowagowych**. W chwili obecnej na konferencjach czysto „fizycznych” (np. w USA) **ok. 30% czasu wystąpień jest poświęconych obiektom biologicznym**. Wielu fizyków „przekwalifikowało” się i podjęło tematykę biologiczną uzyskując sukcesy w wyniku **nowatorskiego spojrzenia** wyniesionego z terenu „czystej” fizyki. Przykłady z „własnego podwórka” to:

(a) **Władysław Minor** przeszedł od fizyki jądrowej do badań rentgenograficznych białek, opracowując najpowszechniej stosowany na świecie *software* do uzyskiwania struktur przestrzennych na podstawie obrazów dyfrakcyjnych promieniowania rentgenowskiego: "Processing of X-ray diffraction data collected in oscillation mode" [Otwinowski & Minor Methods in Enzymology 276, 3017, 1997](#); **ponad 36 000 cytowań w/g Web of Science**.

(b) **Marek Cieplak** prowadzi liczące się na świecie badania w zakresie teoretycznego zwijania i rozwijania białek „*protein (un)foldng*” w założonej i kierowanej przez siebie grupie fizyki biologicznej w IF PAN.

Wykład BIOFIZYKA DOŚWIADCZALNA **[PLANSZA 1]** będzie obejmował **[PLANSZA 2]** współczesne, kluczowe (tzw. „gorące”) zagadnienia biofizyki molekularnej, związane ze **strukturą, dynamiką i oddziaływaniami kwasów nukleinowych, białek i lipidów**, czyli biopolimerów o fundamentalnym znaczeniu w funkcjonowaniu, budowie i podziale komórki, badane przy użyciu wybranych, zaawansowanych, **biofizycznych technik** eksperymentalnych. Będą także omówione wybrane aspekty zastosowań biofizyki w biotechnologii, niektórych obszarach nauk

biomedycznych i biochemii procesów *in vivo* (omawiane zagadnienia są ujęte w planie wykładu). Podstawowa **literatura** jest dostępna w formie anglojęzycznych podręczników*, m. in. wymienione dwa podstawowe jeśli chodzi o kwasy nukleinowe i białka, oraz wyczerpujący podręcznik metod badawczych, ale wykład jest w dużej mierze oparty na oryginalnych pracach, opublikowanych w ostatnim okresie 10-ciu lat w czasopismach naukowych, których odnośniki będą podawane na wykładzie w trakcie omawiania poszczególnych zagadnień. Nacisk będzie położony na **zagadnienia biofizyczne** i nie będzie omawiana szczegółowo strona biologiczna i biochemiczna tych zagadnień. **PLANSZE** są w języku angielskim w celu bliższego zaznajomienia się z terminologią angielską, wymaganą do studiowania literatury przedmiotu (oraz zakładając, że docelowo wykład będzie prowadzony w języku angielskim).

W **interdyscyplinarnym podejściu [PLANSZA 3]** BIOFIZYKA wykorzystuje szeroko rozumianą metodologię nauk matematyczno-przyrodniczych, przede wszystkim fizyki, ale także chemii, oraz matematyki wraz z informatyką, do badania **objektów biologicznych** *in vitro* oraz *in vivo* o bardzo różnym stopniu złożoności. W szczególności BIOFIZYKA MOLEKULARNA **[PLANSZA 4]** zajmuje się **biopolimerami** warunkującymi zachodzenie wszelkich procesów metabolicznych w komórce i organizmie oraz procesów które prowadzą do podziału i różnicowania komórek:

- **kwasy nukleinowe: DNA:** zapis informacji genetycznej, i **RNA:** zapis informacji genetycznej, budowa struktur komórkowych (np. rybosom) oraz funkcje regulacyjne;
- **białka:** podstawa procesów zapewniających budowę i funkcjonowanie komórki;
- **lipidy:** błony biologiczne;
- **oligo- i polisacharydy:** materiał zapasowy oraz budulcowy struktur komórkowych.

Biopolimery tworzą **kompleksy molekularne** warunkujące ich funkcjonowanie w komórkach. Podczas **ekspresji genów [PLANSZA 5]** zachodzą kolejno dwa procesy:

- **transkrypcja**, czyli przepisanie informacji genetycznej z DNA na RNA,
- **translacja mRNA**, czyli biosynteza funkcjonalnych białek.

Dodatkowo, komórka dzieli się na komórki potomne w wyniku procesu **replikacji DNA**.

Białka, które pełnią zasadniczą (obok RNA) rolę "wykonawczą" w organizmach mają strukturę chemiczną i przestrzenną, zakodowaną w kwasie DNA. Dlatego po osiągnięciu

*Interentowa Księgarnia Wydawnictw Importowanych GAMBIT prezentuje szeroki zestaw podręczników biofizyki w języku angielskim.

odpowiedniego poziomu metod inżynierii genetycznej naczelnym zadaniem **genomiki** stało się ustalenie pełnej sekwencji nukleotydowej DNA człowieka i innych organizmów (**struktury pierwszorzędowej**). W 2000 r. połączone przedsięwzięcia Human Genom Project i prywatnej firmy Celera Genomics osiągnęły wykonanie sekwencjonowania wszystkich chromosomów ludzkich, 6·10⁹ par zasad (bp) na 23 parach chromosomów (długość genomu 3·10⁹ bp). Po zsekwencjonowaniu genomu ludzkiego i szeregu innych organizmów biologia i biofizyka molekularna skoncentrowała się na ustaleniu funkcjonalnych struktur wszystkich białek kodowanych przez genomy, których liczbę szacuje się na kilkaset tysięcy (ok. 300 tys. w genomie ludzkim, kodowanych przez ok. 30 tys. genów). **Proteomika** stara się do tego problemu podejść w sposób kompleksowy, przez wysoce przepustowe (tzw. *high throughput*) ustalanie struktur, w tym struktur przestrzennych, całych klas białek współdziałających w ramach określonych odcinków szlaków metabolicznych tak, aby można było konstruować **molekularne mechanizmy** współdziałania całych kompleksów białkowych. Jednocześnie **celem praktycznym** badań strukturalnych i dynamicznych jest wybór właściwych "tarcz" czyli receptorów, głównie białkowych dla różnego typu inhibitorów czy aktywatorów jako potencjalnych środków leczniczych: antywirusowych, antynowotworowych, cytostatycznych czy też sprzyjających proliferacji określonych typów komórek. Zespoły naukowe krystalografów i spektroskopistów NMR powołane do realizacji badań strukturalnych w ramach tzw. Large Scale Facilities, w USA (finansowanie przez NIH), w Unii Europejskiej (finansowane w ramach Programów Ramowych), w Japonii i w Kanadzie, ściśle współpracują z firmami farmaceutycznymi. Wysoce wydajne metody badania złożonych problemów z zakresu genetyki i struktury białek zostały opracowane w kolejnych działach typu **-omics**. **Transkryptomika** zajmuje się całościowym podejściem do puli RNA w funkcjonującej komórce, a **metabolomika** poddaje analizie całość procesów biochemicznych w funkcjonującej komórce.

Konstrukcja wykładu oparta jest na wyeksponowaniu aktualnie **wiodących tematów i projektów badawczych w biofizyce światowej [PLANSZA 6]:**

- wyznaczanie **struktur przestrzennych** biopolimerów, kompleksów i struktur subkomórkowych **z rozdzielczością atomową** metodami dyfrakcji rentgenowskiej, magnetycznego rezonansu jądrowego i molekularnego modelowania komputerowego, w tym białek, w ramach **PROTEOMIKI STRUKTURALNEJ**, oraz **PROJEKTOWANIE**

LEKÓW na bazie rozwiązanych struktur; do tych technik stopniowo zbliża się pod względem rozdzielczości kriomikroskopia elektronowa pojedynczej cząsteczki (cryoEM);

- analiza problemu ZWIJANIA BIAŁEK i KWASÓW RNA do funkcjonalnych form natywnych, w tym przewidywanie struktur przestrzennych na podstawie sekwencji i określanie dróg zwijania, prawidłowego i błędnego;

- ENERGETYKA komórki związana z transferem elektronów i protonów;

- obrazowanie na **średnim poziomie rozdzielczości**, manipulowanie i spektroskopia POJEDYNCZYCH BIOPOLIMERÓW i kompleksów;

- wyjście ponad poziom struktur molekularnych do struktur o wymiarach „NANO”;

- obrazowanie strukturalne i obrazowanie metabolizmu komórek, tkanek i narządów w zastosowaniach w DIAGNOSTYCE MEDYCZNEJ.

Wybór tematyki ma oczywiście charakter częściowo arbitralny, ale w przypadku tego wykładu jest nakierowany na zagadnienia pozostające w sferze mniej lub bardziej bezpośredniego zainteresowania prowadzących badania w Zakładzie Biofizyki IFD.

Fundamentalne założenie metodologiczne, które warunkuje sposób ujęcia problematyki polega na przypisaniu szeroko rozumianej **strukturze** na poziomie molekularnym:

- struktura **chemiczna**,

- struktura **przestrzenna**: położenia zębów atomowych (atomów), czyli trzy współrzędne przestrzenne (x_i y_i z_i) dla każdego i-tego atomu w wybranym układzie współrzędnych,

- **dynamika** ruchów molekularnych,

zasadniczego znaczenia w określaniu mechanizmów **biologicznego funkcjonowania** cząsteczek (biopolimerów) w żywej komórce. Założenie to ujmuje się często w formie terminu **SAR (structure-activity relationship)** oraz **QSAR (quantitative structure-activity relationship)**. W przypadku QSAR wyznacza się korelacje aktywności biologicznej związków z ich dobrze mierzalnymi parametrami fizykochemicznymi.

Struktury biomolekuł i **dynamika** ich ruchów cząsteczkowych oraz przekształceń strukturalnych warunkuje m. in. zdolność do **tworzenia kompleksów** molekularnych o różnym stopniu złożoności. Wzajemne **specyficzne rozpoznawanie** biomolekuł jest niezbędne dla zachodzenia skoordynowanych, podległych regulacji wspólnych działań biomolekuł w procesach biochemicznych. Należy zaznaczyć, że rola dynamiki ruchów molekularnych jest niezwykle istotna i funkcjonalnej struktury nie należy traktować w sposób "sztywny". Przykładowo, duża grupa białek, oceniana na 30% u eukariotów, jest całkowicie lub częściowo nieustrukturowana i pełni określone role biologiczne jako tzw.

białka natywnie rozwinięte (*natively unfolded proteins, intrinsically disordered proteins* IDPs: [review: Uversky & Dunker: Biochim. Biophys. Acta 1804, 1231, 2010](#)). Białka te uzyskują formę ustrukturowaną (lub częściowo ustrukturowaną) dopiero w momencie tworzenia funkcjonalnych kompleksów.

Wymienione na wstępie podstawowe podręczniki [\[PLANSZA 2\]](#) są skoncentrowane na biopolimerach, białkach i kwasach nukleinowych, oraz na metodach ich badania i stanowią bezpośredni materiał pomocniczy do wykładu. Nie wyklucza to możliwości korzystania ze znacznej ilości dostępnych na rynku, anglojęzycznych podręczników, poświęconych biofizyce molekularnej, ujętej pod względem różnych aspektów (niekoniecznie pod kątem obiektów czy metodologii), zależnie od założonych celów informacyjnych i dydaktycznych ([oferta tytułów](#) Internetowej Księgarni Wydawnictw Importowanych GAMBIT).

Struktura chemiczna, metabolizm i biologiczna rola biopolimerów: kwasy nukleinowe DNA i RNA, białka, lipidy, polisacharydy (przegląd i uzupełnienie wiadomości)

Wykład 2

Kwasy nukleinowe DNA i RNA są liniowymi biopolimerami, zbudowanymi z zasadniczo czterech monomerów, **nukleotydów**. „Zasadniczo” oznacza że mogą zajść ściśle 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 **kwasy deoksyrybonukleinowego DNA** [PLANSZA 7] pełnią rolę **nośnika informacji genetycznej**. **Podwójna helisa DNA** liczy od 10^4 bp (par zasad) w mitochondiach, wirusach DNA, poprzez genomy bakteryjne (genom *E. coli* długości $2,5 \times 10^6$ bp), do genomów organizmów wielokomórkowych $\sim 10^9 \div 10^{10}$ bp, gdzie poszczególne chromosomy zawierają podwójne helisy DNA o masie cząsteczkowej $\sim 10^{10} \div 10^{11}$ Da i długość do kilku centymetrów po rozwinięciu. Komplementarna para zasad, ściślej komplementarna para nukleotydów dA:T i dG:dC w helisie ma średnią masę cząsteczkową ok. 660 Da.

Kwasy dezoksyrybonukleinowe są syntetyzowane w fazie S cyklu komórkowego w procesie **replikacji** [PLANSZA 8], w którym na matrycach dwóch nici pierwotnych syntetyzowane są przez polimerazy DNA nici potomne w kierunku $5' \rightarrow 3'$ na zasadzie komplementarności; **semikonserwatywny** charakter replikacji potwierdzony przez eksperyment Meselsona-Stahla. Naprzeciw dA w matrycy włączane jest T, naprzeciw dG włączane jest dC, naprzeciw T włączane jest dA a naprzeciw dC włączane jest dG, dostarczane przez odpowiednie trifosforany NTP. Niektóre wirusy syntetyzują DNA na matrycy RNA w procesie **odwrotnej transkrypcji**. Błędna, nieskorygowana inkorporacja nukleotydu niekomplementarnego może prowadzić do powstania **mutacji punktowej** [PLANSZA 9] (nie są to wszystkie rodzaje mutacji) z **częstotliwością** na nowosyntetyzowaną parę zasad: od $10^{-3} \div 10^{-8}$ dla wirusów pozbawionych możliwości naprawy DNA, do $10^{-8} \div 10^{-11}$ dla organizmów obdarzonych **systemami naprawy DNA**.

Kwas DNA jako nośnik informacji genetycznej funkcjonuje w chromosomach:
- bakteryjnym, tzw. **genofor** u prokariota;

- eukariotycznym, w połączeniu z białkami histonowymi w jądrze komórkowym.

DNA składa się **[PLANSZA 10]** zasadniczo z czterech **2'-deoksyrybonukleotydów**: dwóch zawierających zasady purynowe dGMP, dAMP, i dwóch pirymidynowych, dCMP i TMP, w których cukrem jest β -D-2'-deoxyryboza. Numeracja w pierścieniach zasad i w części cukrowej (cyfry primowane) jest jednakowa dla deoxy- i rybonukleotydów (powinna być znana). Konfiguracje wokół atomów węgla C1' i C4', podstawionych czterema różnymi podstawnikami, opisują odpowiednio **symbole β i D**. Takie cząsteczki nie pokrywają się ze swoim odbiciem zwierciadlanym i występują w postaci izomerów optycznych, skręcających światło spolaryzowane liniowo w przeciwnych kierunkach.

Zasada azotowa: guanina, adenina, cytozyna i tymina jest strukturą **płaską**. Po połączeniu wiązaniem glikozydowym z **cukrem** powstaje **nukleozyd**: 2'-deoxyguanozyna, 2'-deoxyadenozyna, 2'-deoxycytydina i tymidyna. Nukleozydy łączą się ze sobą poprzez grupy fosforanowe w **nić polinukleotydową** o wyróżnionym kierunku od 5'-końca do 3'-końca **[PLANSZA 7]**. Sekwencja czyli kolejność nukleotydów (zasad) w kierunku 5' \rightarrow 3' określa **strukturę pierwszorzędową kwasu DNA**, w której zawarta jest **informacja genetyczna**: kolejne trójki zasad DNA kodują jeden z aminokwasów oraz zakończenie procesu translacji mRNA (kodon stop). W warunkach pH komórkowego **grupy fosforanowe w nici polinukleotydowej** są zjonizowane i obdarzone ładunkiem ujemnym ($pK \sim 1,5$), zneutralizowane przez kationy roztworu, np. K^+ , Na^+ czy Mg^{2+} . W DNA występują nukleotydy z zasadami **modyfikowanymi**, np. **reszty adeninowe metylowane na grupie aminowej** (N^6CH_3 -dA) lub **reszty cytozynowe podstawione metylem w pozycji C5** ($5-CH_3$ -dC) (epigenetyka). Zasady azotowe są **chromoforami** i absorbują promieniowanie elektromagnetyczne, dając trzy pasma typu $\pi^* \leftarrow \pi$ i $\pi^* \leftarrow n$ w bliskim nadfiolecie **[PLANSZA 11]** o współczynnikach ekstynkcji molowej $\epsilon \sim 10^4$ [$l \cdot mol^{-1} \cdot cm^{-1}$]. Nałożenie pasm wszystkich zasad prowadzi do absorpcji DNA w zakresie $180 \div 220$ nm i $240 \div 280$ nm, z **charakterystycznym maksimum przy 260 nm**. **Fluorescencja zasad** jest bardzo słaba i charakteryzuje się wydajnościami kwantowymi ok. $3 \cdot 10^{-4}$. Umożliwia jednak pomiary fluorescencji naturalnego DNA w zakresie $300 \div 400$ nm, z maksimum w 327 nm ([Vaya et al. J. Am. Chem. Soc. 132, 11834, 2010](#)). W niskich i wysokich pH ok. 2,3 do 3 jednostek poniżej i powyżej pH fizjologicznego (pH 7,2) zasady ulegają odpowiednio **protonacji** i **dysocjacji** **[PLANSZA 10]**. Badania krystalograficzne i spektroskopowe pokazują, że **płaskie** zasady azotowe w standardowych warunkach w kryształach i roztworze występują w **dominujących formach tautomerycznych amino i oxo**

(keto); czułość bezpośrednich metod doświadczalnych sięga 99% czyli 10^{-2} . Zmiana ustawienia protonów i reorganizacja wiązań podwójnych stwarza możliwość powstawania **mniejszościowych tautomerów imino i enolo (hydroxy)** (patrz dalej).

Obok struktury chemicznej opisanej składem atomowym, układem i wartościami długości wiązań oraz kątów płaskich, o strukturze przestrzennej decyduje **konformacja** czyli przestrzenne ustawienie zrębów atomowych uwarunkowane obrotami wokół wiązań pojedynczych. Konformację określa się przez podanie wartości odpowiednich **kątów dwuściennych** (torsyjnych) lub bardziej jakościowo ich zakresów według nomenklatury oznaczeń komisji IUPAC-IUB. Dwa wybrane podstawniki w danym fragmencie molekularnym mogą mieć ustawienie transoidalne czyli *antiperiplanar* (ap), ustawienie gauche czyli *synclinal* (\pm sc) i cisoidalne czyli *synperiplanar* (sp). Kąty są obrazowane i definiowane co do wartości i znaku w **rzucie Newmana**: Konformację nici polinukleotydowej wyznaczają wartości kątów α , β , γ , δ , ε , ζ w łańcuchu **[PLANSZA 12]**:



oraz kąta χ zasady wokół wiązania glikozydowego: formy syn (synperiplanar) i anti (antiperiplanar). Pięć kątów ν_j $j = 0, 1, 2, 3, 4$, w pentafuranozie określa konformację **niepłaskiego** pierścienia cukrowego na gruncie tzw. **modelu pseudorotacji [PLANSZA 13]** z dwoma parametrami: kąt pseudorotacji P od 0 do 360° i maksymalne pofałdowanie pierścienia ν_{\max} , w granicach 35° do 45° (kątowne wychylenie maksymalne w kopercie). Pierścień pięciocłonowy może przyjmować dwa podstawowe formy: 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 skrzyżowaną T, w której dwa sąsiednie atomy x i y są wychylone z płaszczyzny w przeciwne strony $^x T$.

Nukleotyd jest cząsteczką **konformacyjnie "giętką"** tzn. ze względu na możliwość rotacji wokół wiązań pojedynczych występuje w roztworze w **równowadze dynamicznej** szeregu konformacji **[PLANSZA 14]**, 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**. W łańcuchu polinukleotydowym swoboda przejść konformacyjnych jest znacznie bardziej **ograniczona** ale **przejścia takie zachodzą**. Ponadto, **konformacja monomeru wyznacza dopuszczalne konformacje łańcucha polinukleotydowego**. Nukleotyd może przyjmować następujące konformacje.

(1) **Zasada** może wykonywać obrót wokół wiązania glikozydowego z cukrem i przyjmować dwa typy ustawień: **syn** i **anti**, opisane przez wartości kąta χ **[PLANSZA 12]**. Zakres syn odpowiada $-90^\circ < \chi < +90^\circ$ a zakres anti $+90^\circ < \chi < 270^\circ$;

(2) **Pierścień cukrowy** występuje w dwóch konformacjach: z zakresu **N** (*north*) na kole pseudorotacji, z P ok. 0° do 36° (3_2T do 3_4T), czyli wokół C(3')endo i z **zakresu S** (*south*) na kole pseudorotacji, z P ok. 135° do 190° (2_1T do 2_3T), wokół C(2')endo **[PLANSZA 14]**.

(3) **Grupa egzocykliczna C(5')H₂-O** 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° .

(4) **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° . W 3'-nukleotydach 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). Odpowiada to średnim czasom życia konformerów ~nanosekund.

Lokalne zwiniecie łańcucha polinukleotydowego w struktury drugorzędowe jest w **DNA** podobne jak w **RNA** i daje w wyniku:

- **helisy**: podwójne, potrójne i poczwórne;

- **pętle**: *hairpin*, wewnętrzne (w helisach podwójnych) i wybrzuszenia (*bulges*).

Struktura trzeciorzędowa funkcjonalnego DNA w chromatynie jest tworzona poprzez kompleks z białkami histonowymi (patrz dalej), dając upakowanie od 10^3 razy w komórce bakteryjnej do 10^5 razy w jądrze komórki eukariotycznej.

Kwas rybonukleinowy RNA **[PLANSZA 15]** pełni funkcje zależnie od struktury:

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

(b) **rRNA**: rybosomalny RNA jest wraz z białkami składnikiem struktury rybosomu; **długości od** $\sim 100n$ do $4000n$ (kilka rodzajów);

- (c) **tRNA**: transferowy RNA łączy się i wprowadza aminokwasy w miejsce na rybosomie, gdzie zachodzi włączanie do powstającej nici peptydowej; **długość od 75n do 90n**;
- (d) **snRNA**: mały jądrowy (*small nuclear*) RNA jest składnikiem splicesome'u w procesie składania mRNA w jądrze komórkowym; **długości od ~100n do 200n** (kilka rodzajów) i **snoRNA** mały jądrowy (*small nucleolar*) RNA modyfikujący inne RNA;
- (e) **miRNA** i **siRNA**: mikro i *small interfering* RNA pełni różnorodne funkcje regulacyjne w procesach komórkowych (udział w regulacji 30% genów u człowieka), **długości ~20n** ([review: C. Arenz: Angew. Chem. Int. Ed. 45, 5048, 2006](#)).
- (f) **RNA wirusów**: materiał genetyczny niektórych wirusów (HIV); długość ~10⁴n lub bp.

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 10](#)]. Rybonukleozydy: guanozyna, adenozyzna, cytydyna i urydyna łączą się w nić polinukleotydową poprzez grupy fosforanowe. Sekwencja nukleotydów (zasad) w kierunku 5' → 3' wyznacza **strukturę pierwszorzędową RNA**. **Jonizacja** grup fosforanowych, **jonizacja/protonacja** zasad oraz **tautomeria zasad** jest podobna jak w przypadku DNA. W przeciwieństwie do DNA **niekomplementarne parowanie w RNA** nie ma tak „restrykcyjnego” charakteru i jest dopuszczalne na trzecim miejscu kodonu mRNA: G:U, U:G, A:I, I - inozyna, modyfikowany nukleotyd („guanozyna” bez grupy aminowej w pozycji 2). Modyfikowane nukleozydy, zarówno w części zasadowej jak i na cukrze mogą stanowić znaczny udział w RNA, do 10% w tRNA.

Struktura przestrzenna RNA jest opisana przez kąty dwuścienne zdefiniowane analogicznie jak dla DNA [[PLANSZA 12](#)]: Kwasy rybonukleinowe RNA charakteryzują się znacznie większym polimorfizmem strukturalnym niż DNA. Lokalne fragmenty **struktury drugorzędowej** (helisy i pętle) wiążą się poprzez **charakterystyczne motywy: złącze helis, pseudowęzeł** [[PLANSZA 16](#)], **motyw A-minor, zamek rybozowy, Z-anchor**, w globalną, złożoną **strukturę trzeciorzędową, która bardziej przypomina białka globularne** niż kwasy DNA [[PLANSZA 15](#)].

Kwasy rybonukleinowe powstają w procesie **transkrypcji** [[PLANSZA 17](#)]: 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 RNA ulega **procesowi składania** czyli *splicingu* lub *self-splicingu* (autokataliza) do funkcjonalnych cząsteczek, np. **eukariotyczne mRNA** uzyskuje właściwą formę matrycy do syntezy

białek poprzez wycinanie niekodujących *introns* i łączenie kodujących *exons*, oraz **dojrzewania**, czyli dołączania struktury kapu na 5' końcu i łańcucha poli(A) na 3' końcu **[PLANSZA 18]**. RNA wirusów jest syntetyzowane przez RNA-*directed* polimerazy RNA (RNA replikaza lub syntetaza). Funkcjonalne Kwasy RNA ulegają **procesowi zwijania** (*RNA folding*) podobnie jak białka (zagadnienie będzie szczegółowo omówione przy omawianiu zwijania białek)

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 zabezpieczone, np. przez kap na 5' końcu eukariotycznego mRNA. (reviews: Shyu et al. EMBO J. 27,471, 2008; Collier & Parker Annu. Rev. Biochem. 73, 861, 2004).

Wykład 3

Białka pełnią w komórce różnorodne funkcje:

- (a) materiał budulcowy struktur komórkowych;
- (b) enzymy, katalizatory reakcji biochemicznych;
- (c) przeciwciała, funkcje odpornościowe organizmu;
- (d) transport aktywny przez błony cząsteczek organicznych i jonów;
- (e) energetyka, udział w procesach oddychania i fotosyntezy;
- (f) praca mechaniczna kosztem energii chemicznej;
- (g) regulacja ekspresji genów, hormony.

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 δ -Cys-Gly. Wszechstronnie badane „klasyczne” białko BPTI (*basic pancreatic trypsin inhibitor*) ma 58 aminokwasów, ale znana jest rodzina białek CMTI (*cucurbita maxima trypsin inhibitor*) o długościach łańcuchów ok. **30 aminokwasów**. W przypadku obu typów białek zachodzi stabilizacja przez 3 mostki dwusiarczkowe (patrz dalej). Średnia długości białek globularnych to 300 ÷ 400 aminokwasów, a spotykane są bardzo długie białka rzędu 3 000 AA, np. kolagen (długość 28 000Å; masa cząsteczkowa MW ~300 000 Da).

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 czynniki denaturujące oraz proteazy. 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, fosfoproteidy, chromoproteidy (hemoglobina, cytochrom) i szereg innych. Ze strukturalnego punktu widzenia można podzielić białka na trzy klasy różniące się **strukturami trzeciorzędowymi globalnego ukształtowania** czyli **foldu** łańcucha peptydowego przy zachowaniu bardzo **dużego podobieństwa struktur drugorzędowych lokalnego ukształtowania łańcuchów peptydowych**:

- (1) białka globularne **[PLANSZA 19]**

(2) białka **strukturalne** (włókniste) [\[PLANSZA 20\]](#)

(3) białka **blonowe** [\[PLANSZA 21\]](#)

Białka większości organizmów są zbudowane zasadniczo z 20 typów α -aminokwasów o konfiguracji przestrzennej (stereoizomerii) **typu L** ze względu na asymetryczny węgiel C α podstawiony czterema różnymi podstawnikami [\[PLANSZA 22\]](#), z wyjątkiem glicyny. Poszczególne aminokwasy różnią się strukturą chemiczną i własnościami fizykochemicznymi reszt bocznych [\[PLANSZA 23\]](#), które można podzielić na:

- **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 aminowego;

- **silnie hydrofilowe**, preferujące środowisko wodne: asparagina i glutamina; ich grupy boczne CONH₂ są jednocześnie donorami i akceptorami protonów w wiązaniach wodorowych;

- **naładowane ujemnie** (kwaśne), hydrofilowe, ze zjonizowaną grupą karboksylową w reszcie bocznej: kwas asparaginowy i kwas glutaminowy.

- **naładowane dodatnio** (zasadowe), hydrofilowe: lizyna, arginina i histydyna, w histydynie pK jest bliskie pH fizjologicznego i reszta boczna może mieć ładunek dodatni lub nie, zależnie od otoczenia białkowego;

- **zawierające siarkę**, słabo hydrofobowe: cysteina i metionina, grupa SH cysteina jest bardzo reaktywna i połączenie (utlenienie) dwóch grup cysteinowych daje wiązanie kowalencyjne czyli mostek dwusiarczkowy –S–S–, który stabilizuje struktury białkowe;

- **zawierające tlen**, słabo hydrofilowe: seryna i treonina

- **zawierające pierścienie aromatyczne**: fenyloalanina, tyrozyna i tryptofan.

Ważną cechą aminokwasów z resztami aromatycznymi jest absorpcja i fluorescencja w zakresie ultrafioletu [\[PLANSZA 24\]](#). Białka absorbują promieniowanie UV z max. ok. 280 nm (kwasy nukleinowe 260 nm) i emitują promieniowanie zasadniczo w formie emisji tryptofanowej ze względu na rezonansowy transfer energii wzbudzenia FRET Tyr → Trp. Emisja tyrozynowa jest obserwowana w białkach pozbawionych tryptofanów;

Aminokwasy łączą się w łańcuch peptydowy poprzez utworzenie **wiązania peptydowego** [\[PLANSZA 25\]](#) między grupą karboksylową pierwszego i grupą aminową następnego, w procesie **translacji mRNA na rybosomie** [\[PLANSZA 18\]](#). Powstaje w ten sposób liniowy polimer o określonej **strukturze pierwszorzędowej**, czyli kolejności lub

inaczej 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 26]. Kolejne aminokwasy w sekwencji są kodowanych przez trójki zasad genomu. W niektórych organizmach, np. *Archea* i *Eubacteria* w procesie translacji mRNA włączane są do łańcucha dwa dodatkowe rodzaje aminokwasów [PLANSZA 27]: **selenocysteina**; redefiniowany kodon stopu UGA i **pirolizyna**; redefiniowany kodon stopu UAG (*Atkins & Gesteland Science 296, 1409, 2002*). Pozostałe rodzaje aminokwasów w białkach, tzw. **niestandardowe** powstają w wyniku posttranslacyjnych modyfikacji aminokwasów standardowych.

Rozwinięty **łańcuch peptydowy związa 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-translacyjnie*), często przy pomocy różnych mechanizmów komórkowych podnoszących **efektywność związania**, (*szczególuy przy omawianiu związania in vivo*). Jednym z tych mechanizmów jest pomoc ze strony **białek opiekuńczych** czyli chaperonów, np. kompleks chaperoninowy **GroEL/GroES** *E. coli* [PLANSZA 28]. W trakcie związania tworzą się **mostki dwusiarczkowe** [PLANSZA 29] stabilizujące strukturę trzeciorzędową pomiędzy ściśle określonymi (jeśli jest więcej niż dwie) resztami cysteinowymi, katalizowane przez *protein-disulfide isomerase*, przy udziale tripeptydu glutationu jako koenzymu. W procesie **topogenezy** proteolityczne usunięcie jednego lub większej ilości **sekwencji sygnałowych** (*signal sequence*) na N-końcu w powiązaniu z przechodzeniem białka przez błony dostarcza białko do miejsca funkcjonowania w komórce. Białka są *co-* i *post-translacyjnie* modyfikowane do ostatecznej formy przez **usunięcie enzymatyczne pierwszej metioniny**, **proteolityczne wycinanie fragmentów** wewnętrznych przez proteazy, **składanie** (*splicing*) poprzez **samousuwanie** sekwencji inteinowych i łączenie eksteinowych (podobnie jak w RNA), **fosforylacja** przez kinazy białkowe, **modyfikacja łańcuchów bocznych** (wykryto w białkach ponad 200 różnych typów aminokwasów), **dodanie** w białkach fragmentów różnych od aminokwasów, takich jak: cukrowe, nukleinowe, lipidowe, itp. Białka „niepotrzebne” są **degradowane** do aminokwasów w strukturalnie podobnym do chaperoninowego **kompleksie proteasomu** [PLANSZA 28] po wyznakowaniu białkiem ubikwityną.

Struktura przestrzenna białek (drugorzędowa, trzeciorzędowa i czwartorzędowa) jest uwarunkowana przez strukturę pierwszorzędową i wynika z

charakterystycznych cech płaskiego wiązania peptydowego C–N **[PLANSZA 25]**, dopuszczalnych wartości kątów Φ i Ψ opisujących konformację łańcucha głównego oraz konformacjami reszt bocznych, w wyniku obrotów wokół wiązań pojedynczych **[PLANSZA 30]**. Wiązanie peptydowe ma częściowo **podwójny charakter**. Konformację wokół wiązania peptydowego opisuje kąt ω , który może przyjąć wartość bliską 180° **[PLANSZA 25]** przy najczęstszym, transoidalnym ustawieniu węgli $C\alpha$, lub wartość bliską 0° z cisoidalnym ustawieniem obu węgli. Struktura każdego z układów jest sztywna, a przejście konformacyjne między nimi poprzez jest zahamowane (bariera ok. $60 \div 120$ kJ/mol). Forma trans jest energetycznie uprzywilejowana ~ 8.5 kJ/mol ze względu na niekorzystne sterycznie oddziaływania w formie cis węgla $C\alpha$ danego aminokwasu i łańcucha bocznego aminokwasu poprzedzającego w sekwencji W przypadku wiązania peptydowego poprzedzającego prolinę niekorzystne oddziaływanie sterycznie występuje w obu formach i ich energia jest zbliżona, co powoduje że niektóre wiązania peptydowe przed proliną mogą mieć $\omega = 0^\circ$. Konformację łańcucha peptydowego opisują wartości dwóch kątów dwuściennych dla każdego i-tego aminokwasu w sekwencji **[PLANSZA 30]**:

- (1) **kąt Φ** między węglami karbonyłowymi C' we fragmencie $C'_i - C\alpha_i - N_i - C'_{i-1}$, gdzie wartości dodatnie są przy obrocie zgodnym z ruchem wskazówek zegara patrząc wzdłuż $C\alpha_i - N_i$ czyli przeciwnie do kierunku łańcucha;
- (2) **kąt Ψ** między azotami amidowymi we fragmencie $N_i - C\alpha_i - C'_i - N_{i+1}$, gdzie wartości dodatnie są przy obrocie zgodnym z ruchem wskazówek zegara patrząc wzdłuż $C\alpha_i - C'_i$ 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”), a 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 31]**, 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 **trzy typy struktur** **[PLANSZA 19]**:

- **helikalną**, zwiniętą prawoskrętnie, z kątami Φ i Ψ o wartościach ok. -60° : helisa 3_{10} stabilizowana przez wiązania wodorowe $C=O\dots H-N$ między aminokwasami i-tym i i+3-

cim, helisa α z wiązaniami między AA i-tym i i+4-tym i helisa π z wiązaniami między AA i-tym i i+5-tym;

- **rozciągnięta** z kątami Φ i Ψ po ok. 120 - 140°, tworzące pofałdowane β -kartki złożone z nici antyrównoległych, równoległych lub mieszanych, stabilizowana przez międzyniciowe wiązania wodorowe C=O...H-N;
- **ciasne pętle**, tzw. β -skręty (β -szpilki) i γ -skręty na końcach struktur antyrównoległej β -kartki, kiedy łańcuch peptydowy wykonuje ciasny skręt dla uzyskania tej struktury;
- **dłuższe pętle** nie są dobrze określonymi strukturami drugorzędowymi; łączą drugorzędowe fragmenty helikalne i β -kartkowe;
- **poliprolina** (w inhibitorze trypsyny trzustkowej, w cytochromie c_{551}) nie może tworzyć wiązań wodorowych stabilizujących helisy i β -kartki ze względu na brak protonu amidowego i tworzy dwa **odrębne typy regularnej struktury drugorzędowej**: prawoskrętną helisą poliprolina-I ze wszystkimi wiązaniami peptydowymi w formie cis i z 3,3 AA na skręt oraz lewoskrętną helisą poliprolina-II ze wszystkimi wiązaniami peptydowymi trans i z 3,0 AA na skręt [**PLANSZA 20**]. Struktury typu poliproliny występują w szeregu białek ([review: Adzhubel et al. J. Mol. Biol. 425, 2100, 2013](#));
- **poliglicyna** może także przyjmować nietypowe formy drugorzędowe ze względu na brak łańcucha bocznego w glicynie (giętkość konformacyjna): poliglicyna-II podobna 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: białka klasy α z dominacją form helikalnych, białka klasy β z dominacją form kartkowych, białka klasy $\alpha+\beta$ oraz α/β z różnymi rozkładami obu podstawowych form. **Klasyfikacja typów foldów trzeciorzędowych pod względem topologicznym** stara się uchwycić podobieństwa strukturalne i powtarzalne formy w bardzo różnych białkach (szerzej omówione w dalszej części wykładu). 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 ([Moult et al. Proteins 82, 1, 2014](#)). W chwili obecnej rozpoznanych jest **ponad 700 różnych foldów** ale ich ilość wzrasta w miarę odkrywania nowych struktur (także drugorzędowych).

BIĄŁKA GLOBULARNE **[PLANSZA 19]** przyjmują fold trzeciorzędowy w wyniku określonych, głównych oddziaływań stabilizujących (pozostałe - van der Waals):

(A) Oddziaływania hydrofobowe reszt bocznych aminokwasów alifatycznych, które prowadzą do powstania tzw. **jądra hydrofobowego** **[PLANSZA 32]** białek globularnych. Charakter hydrofobowy ma także **oddziaływanie warstwowe** (stacking) $\pi - \pi$ pierścieni aminokwasów aromatycznych oraz stacking 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 korzystne elektrostatycznie **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 wiążą się oddziaływaniem coulombowskim w **mostki solne**, a aminokwasy hydrofilowe wewnątrz białka tworzą **wiązania wodorowe**.

(D) Kowalencyjne wiązania między grupami siarkowymi cystein S-S, szczególnie w przypadku mniejszych białek, np. CMTI ok. 30 AA, insulina 50 AA **[PLANSZA 29]** 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 33]**. Każda domena tworzy własną zwartą strukturę z własnym jądrem hydrofobowym a połączenia między domenami mają luźniejszy charakter i umożliwiają większe ruchy molekularne domen względem siebie **Powierzchnia molekularna** (*molecular surface, contact surface*) cząsteczki białkowej jest dobrze zdefiniowana geometrycznie; wymiar $A_s = 6.3(M)^{0.73}$ [\AA^2], M -masa cząsteczkowa (25 ÷ 45% powierzchni rozwiniętego łańcucha peptydowego) i odgranicza **wnętrze białka** o objętości $V = 1.27M$ [\AA^3] od rozpuszczalnika i innych cząsteczek. Niekiedy pomiędzy strukturą drugorzędową i trzeciorzędową wyróżnia się **struktury pośrednie** tzw. **superdrugorzędowe**, np. motyw β - α - β , klucz grecki, helisa-skret-helisa. Ważną cechą topologiczną większości białek (globularnych) jest brak „węzłków” w przebiegu łańcucha peptydowego. W chwili obecnej znanych jest ponad 100 struktur trzeciorzędowych białek zawierających węzły.

Wykład 4

BIĄŁKA STRUKTURALNE [\[PLANSZA 20\]](#) (włókniste, fibrylarne) 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) jest jakby przedłużeniem struktur(y) drugorzędowych w jednym kierunku - kierunku włókna.**

BIĄŁKA BŁONOWE [\[PLANSZA 21\]](#), zwykle złożone z **dużej ilości aminokwasów**, są strukturalnie podobne do białek globularnych, ale naturalne środowisko przynajmniej dla fragmentu cząsteczki jest hydrofobowe (szczegółowe omówienie w części poświęconej lipidom). Części białek błonowych mogą wystawać do wody i stosować się do reguł budowy białek globularnych, jednak jako całość białka te agregują w wodzie i są w niej nierozpuszczalne, Relatywnie niedawno udało się opracować metody krystalizacji białek błonowych w środowisku buforowym z udziałem detergentów. Prace w tym kierunku są ciągle prowadzone. Przykładem jest zastosowanie systemu **nanocząsteczek** złożonego z lipidów i białka saposina A (*saposin-lipoprotein nanoparticle system*; [Frauenfeld et al. Nature Methods 13, 345, 2016](#)), który zapewnia dobrą rekonstrukcję białek błonowych w naturalnym środowisku. Od uzyskania w 1985 r. struktury krystalograficznej fotosyntetycznego centrum reakcji (RC) rozwiązano metodą dyfrakcji ([review: Moraes et al. Biochim. Biophys. Acta 1838, 78, 2014](#)) i technikami NMR, głównie wysokiej zdolności rozdzielczej w ciele stałym ([perspectives: Cross et al. J. Magn. Reson 238, 100, 2014](#)), **blisko 1 000 struktur**. Obserwuje się dwa podstawowe typy foldów, zapewniających możliwość funkcjonowania białek błonowych jako kanałów transportu przez błonę: **wiązka helis** oraz **β -baryłka**.

Białka **metamorficzne** (*metamorphic*) rodziny CLIC (*chloride intracellular channel proteins*; ok. 240 AA) i ERM (*exrin, radixin, moezin family*) [Jiang et al. Biochim. Biophys. Acta 1838, 643, 2014](#)) są w zależności od warunków albo zdolne do integracji z błoną albo rozpuszczalne w wodzie. Funkcjonują m. in. w procesach związanych z oddziaływaniem błon z cytoskeletonem miozynowym.

Dla białek zbudowanych z **jednej nici peptydowej** struktura trzeciorzędowa jest strukturą funkcjonalną. Białka mogą być zbudowane z kilku, najczęściej jednakowych nici

peptydowych czyli **podjednostek** [\[PLANSZA 34\]](#), każda o własnej strukturze trzeciorzędowej. Takie białka mają **strukturę czwartorzędową** złożoną z podjednostek.

Stabilizacja struktur natywnych cząsteczek białkowych ok. kilkadziesiąt kJ/mol wynika z **oddziaływań stabilizujących, przewyższających wkład destabilizujący, głównie entropii konformacyjnej**. Udział mają również oddziaływania z rozpuszczalnikiem, w tym wymiana jonów i wody między sferą hydratacyjną białka i „*bulk solvent*”. Dla białek globularnych i błonowych zmiany warunków środowiska mogą spowodować **odwrócenie relacji stabilizacja/destabilizacja** struktury natywnej i przejście w formę zdenaturowaną. Czynnikiem denaturującym są **temperatura**, chemiczne **denaturanty** (mocznik, chlorek guanidyny), skrajne **pH**. Usunięcie czynnika denaturującego może spowodować powrót cząsteczek do formy natywnej (*refolding*). Warunkiem refoldingu jest wystarczająca ilość informacji w sekwencji aminokwasowej. Utrata tej informacji podczas modyfikacji posttranslacyjnych uniemożliwia refolding.

Lipidowa 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 lub 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 warstwę lipidową to [\[PLANSZA 35\]](#) **pęcherzyk** z wnętrzem hydrofobowym (*micelle*), **liposom** oraz forma **heksagonalna**. W badaniach często używa się **modelowych systemów** imitujących błony, np. DMPC (fosfolipid *dimyristoylphosphatidylcholine*), ale coraz częściej badania prowadzi się na *stricte* błonach biologicznych ([review: Sturgis et al. Biochemistry 48, 3679, 2009](#)). Mikroskopia AFM jest jedną z częściej używanych technik doświadczalnych w badaniach strukturalnych błon na średnim poziomie rozdzielczości.

Szczególnie dobrą metodą wglądu w strukturę i dynamikę błon jest technika magnetycznego rezonansu jądrowego NMR w ciele stałym (analiza szerokich linii NMR) z zastosowaniem jąder ^{31}P (spin $I = 1/2$), ^{14}N NMR oraz ^2H NMR (spiny $I = 1$, moment kwadrupolowy $eQ \neq 0$), w tym lipidów **selektywnie wzbogacanych deuterami** w różnych miejscach łańcucha ([review: Brown et al. Biochi. Biophys. Acta 1768, 2979, 2007](#)). Szerokie linie charakterystyczne dla układów polikrystalicznych i amorficznych (*powder*

pattern) modyfikowane przez dynamiczny charakter błony i wartości czasów relaksacji T_1 pozwalają analizować:

(a) **dynamikę anizotropowych ruchów dyfuzyjnych, półpłynność** poszczególnych fragmentów błony oraz **modele ruchów dyfuzyjnych**, włącznie z wyznaczaniem **czasów korelacji** i stałych dyfuzji rotacyjnej: $\tau_c = 1/6D$ z pomiarów czasów relaksacji oraz wartości **parametru uporządkowania** S_{CD} (2H NMR):

$$S_{CD} = \langle 3\cos^2\vartheta - 1 \rangle = P_2(\cos\vartheta) \quad \text{- wielomian Legendre'a drugiego stopnia}$$

ϑ - kąt między osią główną tensora jądrowego momentu kwadrupolowego i

normalnej do powierzchni błony; $\langle \rangle$ - uśrednienie po czasie (równowagowe)

(b) obecność i wpływ na własności błony **składników** różnych od lipidów: jony, białka, cholesterol,

(c) różne **fazy** (struktury) warstwy lipidowej: chaotycznie zorientowane fragmenty dwuwarstwowe (system multilamelarny), pęcherzyki, forma heksagonalna.

(d) **modele dynamiki ruchów dyfuzyjnych** cząsteczek w błonie, np. białek.

Polisacharydy (wielocukry, cukry złożone) to grupa węglowodanów i zarazem biopolimerów, które są złożone z monomerów będących cukrami prostymi o ogólnym **wzorze sumarycznym** $(CH_2O)_n$, np. glukoza $C_6H_{12}O_6$. Do wielocukrów złożonych z tysięcy monomerów należą: **skrobia, glikogen i dekstran** (materiał zapasowy roślin, zwierząt, drożdży), **celuloza, hemiceluloza** (rozgałęziona celuloza), **pektyna** (główny składnik ściany komórkowej), **chityna** (składnik pancerzy skorupiaków i owadów), **glikozaminoglikany** (składnik matrycy pozakomórkowej), **kwask hialuronowy** 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 przeciwieństwie do białek i kwasów nukleinowych, polisacharydy mogą tworzyć i **tworzą zarówno struktury liniowe jak i rozgałęzione**.

Przykładem jest 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 36]** i stabilizowanych **wewnątrzcząsteczkowymi wiązaniami wodorowymi**. Stabilizowane przez **międzycząsteczkowe wiązania wodorowe** warstwy $60 \div 70$ łańcuchów celulozowych tworzą mikrowłókna o średnicy 3,5 nm i do kilku mikrometrów długości. Wraz z dwoma innymi policukrami 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 załamaniem (*kinks*) przez 1,2-połączoną ramnozę tworzy **ścianę komórkową**.

Układ jest dodatkowo stabilizowany przez **jony wapnia** Ca^{+2} . Struktura molekularna i dynamika tego typu układów może być badana z rozdzielczością atomową jedynie przy zastosowaniu NMR wysokiej zdolności rozdzielczej w ciele stałym techniką wirowania pod kątem magicznym ([Dick-Perez et al., Biochemistry 50, 989, 2011](#)); technika będzie omówiona w dalszej części wykładu.

Specyficzne, silne wiązanie biomolekuł jest podstawą współdziałania różnych cząsteczek w procesach biologicznych, takich jak replikacja i ekspresja genów. Aby ze sobą współdziałać makromolekuły muszą się rozpoznać i utworzyć **funkcjonalny kompleks molekularny** [\[PLANSZA 37\]](#). Miarą siły a więc i specyficzności wiązania jest wysoka wartość równowagowej **stałej asocjacji** kompleksu K_{as} , wyrażonej przez stężenia poszczególnych składników [składnik_i] oraz stężenie kompleksu [kompleks], z mierzonej wartości której można bezpośrednio wyznaczyć **energią swobodną Gibbsa** ΔG° utworzenia (stabilizacji) kompleksu (termodynamiczna $K_{as} = K_{as}(p, T)$):

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

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

gdzie R stała gazowa; „ $^\circ$ ” standardowe, odniesione do stężeń jednomolowych (jednocześnie usuwa problemu miana pod logarytmem); I siła jonowa.

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 10^4 M^{-1} (dla kompleksów dwuskładnikowych). Oddziaływanie specyficzne charakteryzuje się **dopasowaniem powierzchni van der Waalsa biomolekuł i dopasowaniem ich rozkładów ładunków**, tak aby maksymalizować oddziaływania elektrostatyczne: wiązania wodorowe, mostki solne, słabe oddziaływania van der Waalsa i oddziaływania hydrofobowe. **Silne, specyficzne kompleksy** białko - kwas nukleinowy charakteryzują się stałymi asocjacji (kompleks dwuskładnikowy) dochodzącymi do $K_{as} \approx 10^{13} \text{ M}^{-1}$ co

odpowiada $\Delta G \approx 80$ kJ/mol w 300K. Najsilniej związanym (ze znanych układów) jest tetrameryczny (4 podjednostki białka, 4 ligandy) kompleks (strept)awidyna–biotyna $K_{as} \approx 10^{15}$ * (biotyna, witamina B₇, koenzym karboksylaz) Tworzeniu kompleksów towarzyszą efekty związane z rozpuszczalnikiem o charakterze entropowym: **efekt polielektrolityczny** czyli wymiana jonów między sferami hydratacyjnymi biomolekuł i *bulk solvent* wodny i **hydratacja/dehydratacja**, czyli wymiana cząsteczek wody między sferami hydratacyjnymi oddziałujących cząsteczek i rozpuszczalnikiem wodnym.

Tworzenie kompleksu nie ma jednak charakteru statycznego. Oddziałujące biomolekuły tworzą przejściowo kompleks spotkaniowy (*encounter complex*) a następnie zmieniają konformacje aby dopasować do siebie powierzchnie molekularne, np. **wzajemne dopasowanie konformacyjne** (*induced fit*). Zapewnia to krótsze odległości między ładunkami, a w konsekwencji silniejsze związanie. Tworzenie kompleksów z udziałem białek natywnie rozwiniętych opisuje model *fly-casting* („wędkowanie na muchę”) wiązania w kompleks i zwijania białka jako **sprzężone procesy kooperatywne**.

Powszechnie występujący proces katalizy reakcji biochemicznej zachodzi poprzez tworzenie kompleksów przez **enzymy**, cząsteczki **białek** lub **RNA (rybozomy)** z substratami, kofaktorami i cząsteczkami regulującymi (inhibitory, efektory allosteryczne). Przykładem szczególnie wysokiej specyficzności są reakcje przecinania podwójnej helisy DNA przez **endonukleazy restrykcyjne**. Tworzenie kompleksów kwasów nukleinowych DNA z **interkalatorami** (pierścienie aromatyczne) może być czynnikiem warunkującym procesy mutagenyzy i kancerogenyzy. **Aptamery**, oligomeryczne fragmenty RNA i pojedynczoniciowe DNA wykazują własności wiązania niskocząsteczkowych ligandów (aminokwasy, nukleotydy, małe białka) ze znacznymi stałymi asocjacji i wysoką specyficznością. W połączeniu z platformą ekspresyjną aptamery RNA tworzą **przełączniki RNA** (*RNA switches; riboswitches*) regulujące ekspresję genów w roślinach, bakteriach i grzybach na poziomie transkrypcji i translacji, np. w obszarach 5'UTR na mRNA. Centralnym punktem zainteresowania biofizyków i biologów molekularnych są wysoce specyficzne kompleksy **białko - DNA** i **białko - RNA** jako podstawa zachodzenia oraz regulacji procesów **ekspresji genu** na poziomach **transkrypcji, translacji, splicingu**.

* ([Weber et al. Science243, 85, 1989](#)): (1) *high protein-ligand shape complementarity (antiparallel β -sheets) and water displacement*; (2) *stabilization of two streptavidin flexible loops which close over biotin* (3) *protein binding pocket lined with polar and aromatic amino acids: extensive network of hydrogen bonds to biotin heteroatoms (first and second shell) and hydrophobic interaction with the biotin hydrophobic parts, respectively*.

W kompleksach tego typu charakterystyczne (superdrugorzędowe) motywy białkowe specyficznym rozpoznają i wiążą fragmenty drugorzędowe DNA i RNA: helisy, pętle *hairpin* o odpowiedniej sekwencji, poprzez oddziaływania elektrostatyczne i hydrofobowe, takie jak oddziaływania warunkujące tworzenie struktur natywnych biopolimerów. Udało się wykryć całość rybosomu bakteryjnego 70S, 2,7 mln Da (3 rRNA, 2 tRNA, mRNA, 54 białka) z rozdzielczością 3,7 Å ([Korostelev et al. Cell 126, 1065, 2006](#)). Z podobną rozdzielczością uzyskano dyfraktogram i wyznaczono strukturę dla 80S rybosomu drożdżowego (4 rRNA, 2 tRNA, 79 białek), 3,3 mln Da ([Ben-Shem et al. Science 330, 1203, 2010](#)). „Rekord” wielkości układu dla którego pokazano, że można wyznaczyć strukturę techniką NMR to 870 kDa kompleks GroEL/GroES ([Flaux et al. Nature 418, 207, 2002](#)). 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*). Wyznaczenie struktury przestrzennej jest warunkiem koniecznym, ale z reguły niewystarczającym dla pełnej **charakterystyki mechanizmu funkcjonowania** i stanowi punkt wyjścia do stosowania szeregu innych technik. Rozwinięcie podejścia hybrydowego zaczęto stosować do badania dużych struktur subkomórkowych o masach cząsteczkowych **kilkudziesięciu mln Da**, np. 50 mln Da **nuclear pore complex NPC** złożony z 456 białek w błonie jądrowej (dwie prace: Alber et al. Nature 450, 683 & 695, 2007). Zastosowane tu „**podejście proteomiczne**” polega na wykorzystaniu maksymalnej ilości technik doświadczalnych: ultrawierwienie analityczne, cryoEM, immunoEM, chromatografia powinowactwa, rentgenografia składników itp., w celu wygenerowania zbioru danych, które są przekładane na więzy przestrzenne wzajemnych relacji między składnikami układu w celu optymalizacji zbioru potencjalnych struktur i wyboru najbardziej prawdopodobnej.

Podsumowując aktualne **trendy badawcze** w biofizyce molekularnej stanowiące ramy ujęcia materiału wykładu [**PLANSZA 6**] można wyróżnić trzy zasadnicze tendencje:

- skoncentrowanie badań na **obiektach o wzrastającej masie cząsteczkowej**, od prostszych modeli do złożonych kompleksów molekularnych i struktur subkomórkowych;
- dążenie do charakterystyki biomolekuł w naturalnych warunkach ich funkcjonowania, czyli rozszerzanie badań *in vitro* na **warunki in vivo**;
- ukierunkowanie metodologii badawczej na kompleksowe działania *high throughput* (w wydaniu nowych dziedzin typu *-omics*).

Oddziaływania elektrostatyczne i hydrofobowe stabilizujące struktury makrocząsteczek i kompleksów molekularnych; rola rozpuszczalnika

Wykład 5

W ścisłym OPISIE KWANTOWYM wszystkie parametry ruchu i własności cząsteczki jako obiektu fizycznego i jej elementów składowych uzyskuje się przez rozwiązanie **równania Schrödingera**, które jest kwantowym odpowiednikiem równania Newtona w klasycznym świecie obiektów makroskopowych. Ma to zasadnicze konsekwencje:

- **energia cząsteczki przyjmuje wartości skwantowane**, tworzące układ poziomów energetycznych charakterystyczny dla tej cząsteczki i zależny od jej budowy chemicznej, struktury przestrzennej a także, w mniejszym stopniu od jej oddziaływań z innymi cząsteczkami i oddziaływań z polami (siłami) zewnętrznymi;
- obowiązuje **zasada nieoznaczoności Heisenberga**: jednoczesne wyznaczenie wartości obserwabli operatorów niekomutujących NIE jest możliwe z dowolną precyzją, np.
 - (a) dla położenia x i pędu p_x ich nieoznaczoności spełniają relację: $\Delta x \Delta p_x \geq \hbar/2$,
 - (b) dla energii E i czasu życia τ poziomu zachodzi: $\delta E \cdot \tau \geq \hbar/2$.

Na skwantowaną **energię pojedynczej cząsteczki składa się** w przybliżeniu zaniedbania wpływu N jąder jako obiektów ciężkich na ruch n elektronów, 1836-razy lżejszych (przybliżenie adiabatyczne i przybliżenie Borna-Oppenheimera):

- energia elektronowa E_{el} ruchu elektronów w polu jąder cząsteczki i energia oddziaływania między elektronami; **potencjał dla drgań jąder**;
- energia drgań E_{osc} jąder: $3N - 6(5)$ współrzędnych normalnych (sztywna cząsteczka);
- energia obrotów E_{rot} cząsteczki: kąty α, β, γ Eulera (liniowa kąty θ, φ);
- energia rotacji wewnętrznej E_{rowe} w przypadku **cząsteczek niesztywnych**, zmieniających konformację w wyniku obrotów wokół wiązań pojedynczych (sprzężona z rotacją cząsteczki jako całości; ([J. Konarski: Teoretyczne podstawy spektroskopii molekularnej, PWN 1991](#)), kosztem energii odpowiedniej ilości stopni swobody drgań);
- energia E_{magn} magnetycznego oddziaływania między orbitalnymi i spinowymi momentami magnetycznymi jąder i elektronów (przykład: oddziaływanie spin - orbita);
- energia translacyjna E_{tr} ruchu postępowego cząsteczek; jedyna o charakterze ciągłym, nieskwantowa dla cząsteczki swobodnej i określona co do wartości przez temperaturę;

- w przypadku przyłożenia pól zewnętrznych, np. magnetycznego o indukcji B lub elektrycznego o natężeniu E energia E_{ex} oddziaływania elektronów i jąder z tym polem. Tak wprowadzony podział energii cząsteczki pozwala wyznaczyć schematycznie układ poziomów energetycznych ważnych np. przy analizie widm spektroskopowych, biorąc pod uwagę relacje pomiędzy zmianami poszczególnych rodzajów energii:

$$\Delta E_{el} > \Delta E_{osc} > \Delta E_{rot} > \Delta E_{ex}.$$

$$\Delta E_{el} \sim 10 \text{ eV}; \Delta E_{osc} \sim 10^{-2} \text{ eV}; \Delta E_{rot}, E_{rota} \sim 10^{-4} \text{ eV}; E_{magn}, E_{ex} \sim 10^{-6} \text{ eV}$$

Każda nierotująca i spoczywająca cząsteczka o danej **strukturze chemicznej**, zadanej przez układ atomów składowych i wiązań oraz **strukturze przestrzennej**, zadanej przez wartości współrzędnych przestrzennych $\mathbf{R}_i = (X_i \ Y_i \ Z_i)$ wszystkich jąder atomów składowych $i = 1, 2, \dots, N$, w dowolnie przyjętym układzie współrzędnych (kartezjańskim), charakteryzuje się określoną energią. Istnieje jedna konformacja o **energii najniższej** (lub kilka o energiach bliskich minimum), zwykle odpowiadająca **strukturze natywnej (funkcjonalnej)**. W przypadku **dużych biomolekuł** wygodnym ujęciem problemu stabilności struktur molekularnych jest analiza wkładów do energii od poszczególnych **kontaktów (oddziaływań) stabilizujących** między atomami (fragmentami molekularnymi), które w sumie przeważają nad destabilizującym strukturę natywną wkładem (głównie) entropii konformacyjnej. Analogicznie analizuje się **stabilizację kompleksów molekularnych**. Dla układu dwóch atomów A i B mających odpowiednio ładunki Z_A i Z_B jąder oraz n_A i n_B elektronów, operator Hamiltona \mathbf{H} można przedstawić w postaci **[PLANSZA 38]** ([W. Kołos: Chemia kwantowa PWN 1978](#)):

$$\mathbf{H} = \mathbf{H}_A + \mathbf{H}_B + \mathbf{H}_{AB} = \mathbf{H}^0 + \mathbf{H}'$$

Pierwsze dwa człony są hamiltonianami izolowanych układów A i B a \mathbf{H}' reprezentuje ich oddziaływanie traktowane jako zaburzenie:

$$\mathbf{H}' = (Z_A Z_B e^2)/R - \sum_{j=1}^{n_B} (Z_A e^2)/r_{Aj} - \sum_{i=1}^{n_A} (Z_B e^2)/r_{Bi} + \sum_{i=1}^{n_A} \sum_{j=1}^{n_B} (e^2)/r_{ij}$$

gdzie i oraz j numerują elektrony, R jest (w przypadku atomów - przyjmujemy dla uproszczenia rozważań) odległością między jądrami oddziałujących atomów, r_{Aj} jest odległością j -tego elektronu od jądra atomu A, r_{Bi} odległością i -tego elektronu od jądra atomu B, r_{ij} jest odległością między elektronami, e jest ładunkiem elementarnym.

Energia oddziaływania:

$$E_{AB} = \int \Psi^* \mathbf{H}' \Psi d\tau$$

uzyskuje wkład w postaci **energii kulombowskiej** E_{cou} po następujących krokach:

1. Przyjęcie funkcji własnej hamiltonianu niezaburzonego \mathbf{H}^0 w **przybliżeniu polaryzacyjnym** (funkcja nie jest antysymetryczna względem permutacji elektronów):

$$\Psi^{(0)} = \Psi_A \Psi_B$$

2. **rozwinięcie multipolowe** drugiego (r_{Aj}), trzeciego (r_{Bi}) i czwartego członu (r_{ij}) hamiltonianu \mathbf{H}' w szeregi potęgowe względem R do rzędu R^{-3} włącznie. Wartości r_{Aj} , r_{Bi} i r_{ij} , wyrażają się przez r_{Ai} , r_{Bj} (rzędu rozmiarów atomów) oraz R (odległość między atomami), co **przy warunku** $r_{Ai}, r_{Bj} \ll R$ pozwala wykorzystać wzór na szereg potęgowy:

$$(1 + x)^{-1/2} = 1 - (1/2)x + (3/8)x^2 - \dots \quad x \ll 1$$

3. Zastosowanie **rachunku zaburzeń** Rayleigha-Schrödingera z ograniczeniem do poprawki pierwszego rzędu $E^{(1)}$.

W energii $E^{(1)}$ występują oddziaływania trwałych (niezaburzonych) multipoli: jon-jon, jon-dipol, dipol-dipol, itd., zależne od orientacji fragmentów molekularnych. Jeśli cząsteczki poruszają się bezładnie w roztworze (w temperaturze rzędu 300K) należy ją **uśrednić** po wszystkich orientacjach przy założeniu rozkładu Boltzmanna:

$$\langle E^{(1)} \rangle = \left\{ \int E^{(1)} \exp(-E^{(1)}/kT) d\Omega \right\} / \left\{ \int \exp(-E^{(1)}/kT) d\Omega \right\}$$

Jeśli energia ta jest rzędu energii ruchów termicznych można się ograniczyć do wyrazów proporcjonalnych do $1/kT$ i otrzymać ją jako szereg członów zależnych od R , jak R^{-1} , R^{-4} , R^{-6} , z których zachowuje się pierwszy, największy człon jako **energię kulombowską**:

$$E_{\text{cou}} = (q_A q_B) / R \quad q_k = e(Z_k - n_k) \quad k = A, B \quad \text{- ładunek k-tego atomu,}$$

W drugim rzędzie rachunku zaburzeń uzyskuje się wkład $E^{(2)}$ do E_{AB} w postaci **oddziaływania indukcyjnego**:

$$E_{\text{ind}} = -(1/R^6)(\alpha_A \mu_B^2 + \alpha_B \mu_A^2) \quad \alpha \text{ - polaryzowalność, } \mu \text{ - moment dipolowy}$$

oraz **dyspersyjnego**:

$$E_{\text{dysp}} = -(3/2R^6) \alpha_A \alpha_B (I_A I_B) / (I_A + I_B) \quad I \text{ - potencjał jonizacji}$$

Oddziaływanie indukcyjne zachodzi między indukowanym momentem dipolowym (i wyższymi momentami multipolowymi), proporcjonalnym do polaryzowalności i trwałym, indukującym momentami dipolowym (wyższym trwałym multipolem). Oddziaływanie dyspersyjne ma klasyczną interpretację przyciągania dwóch indukowanych dipoli (multipoli). W ściślejszym ujęciu kwantowym to oddziaływanie jest wynikiem korelacji ruchu elektronów we fragmentach A i B. Wraz z członem E_{rep} wynikającym z **odpychania**

elektronów walencyjnych w miarę zbliżania się fragmentów A i B (często przyjmowana zależność energii od odległości jak R^{-12}), **przyciąganie** indukcyjne i dyspersyjne zależne jak R^{-6} , dają w sumie **oddziaływanie van der Waalsa**:

$$E_{vdW} = E_{ind} + E_{dysp} + E_{rep} = -CR^{-6} + AR^{-12}$$

W przypadku badania **dużych cząsteczek biopolimerów** pełny opis kwantowy jest obecnie niemożliwy. W przybliżeniu MODELU KLASYCZNEGO biopolimer reprezentuje **zbiór "atomów"** $k = C, O, H$ itd., czyli punktów albo ściślej **sfer materialnych** o ładunkach q_k dodatnich lub ujemnych, na które składają się ładunki jąder wraz z określonym rozkładem ładunku elektronów powłok wewnętrznych i elektronów walencyjnych. Należy podkreślić, że istnieją różne sposoby rozkładu ładunku elektronowego na poszczególne jądra: ładunki Mullikena, ładunki ESP, ładunki Hirshfelda. Potencjał oddziaływania między atomami jest sumą oddziaływań kowalencyjnych, kulombowskich i van der Waalsa **[PLANSZA 39]**:

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

- potencjały **kowalencyjne** w przybliżeniu **harmonicznym** zmian długości wiązań b (stałe siłowe K_b), kątów płaskich θ (K_θ) i torsyjnych ϕ (K_ϕ), i ew. wychyleń δ grupy atomów z płaszczyzny pierścienia do którego są przyłączone (K_δ); b_0, θ_0 są długościami wiązań i wartościami kątów płaskich w stanie równowagi; lepsze przybliżenie uzyskuje się wprowadzając **poprawki anharmoniczne** rzędu wyższego niż 2 ze względu na b, θ, δ ;
- potencjał **kulombowski** oddziaływania punktowych, niespolaryzowanych ładunków q atomów odległych o R w ośrodku o danej stałej dielektrycznej ϵ ;
- potencjał **van der Waalsa** (stałe Lenarda-Jonesa A i C) oddziaływań indukcyjnych i dyspersyjnych ($1/R^6$) oraz odpychania powłok walencyjnych elektronów ($1/R^{12}$).

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 (oddziaływania „**niewiążące**”). Oddziaływanie z otaczającym **roztwornikiem wodnym i jonami** w nim zawartymi można uwzględnić przez wprowadzenie stałej dielektrycznej (ekranowanie ładunków) lub traktując każdą

cząsteczkę otaczającej wody (innego rozpuszczalnika) jako dodatkowe trzy punkty (H₂O) oddziałujące z biomolekułą i między sobą. **To przybliżenie nosi nazwę pola siłowego** (ang. *force field* FF) i jest stosowane przy komputerowych obliczeniach energii biomolekuł MM (*molecular mechanics*) oraz symulacjach ruchu w ramach tzw. **klasycznej dynamiki molekularnej** MD (*molecular dynamics*). Kwantowość jest „ukryta” w wartościach ładunków przypisywanych „atomom”, wartościach stałych siłowych i stałych Lenarda-Jonesa, które trzeba wyznaczyć tylko dla 20 (22) aminokwasów i 5 nukleotydów przez rozwiązanie równania Schrödingera, a następnie przypisać odpowiednim cegiełkom składowym w polimerze. Ze względu na szereg zastosowanych przybliżeń „ukrytych” w końcowej formule potencjału V, tak uzyskane parametry pola siłowego w pakietach MD (AMBER, GROMOS) wymagają przetestowania w procesie **parametryzacji pola siłowego** i zwykle „pracują” najefektywniej dla ograniczonej klasy cząsteczek. Możliwość **klasycznego podejścia** w analizie cząsteczek 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"*.

Siły działające na i-ty atom w punkcie **R_i** wymagane w prowadzeniu symulacji MD czyli komputerowego, numerycznego rozwiązywania równań Newtona układu molekularnego wyznacza się z odpowiednio sparametryzowanego potencjału V:

$$\mathbf{F}_i = -\nabla_{\mathbf{R}_i} V \quad i = 1, 2, \dots, N \text{ „atomów”}$$

Znajomość **rozkładu ładunków** „atomów” biopolimeru co do wartości i położenia umożliwia również wyznaczenie potencjału $\Phi(X,Y,Z)$ i pola elektrostatycznego $\mathbf{E}(X,Y,Z)$ w ośrodku obejmującym rozpuszczalnik i wewnątrz cząsteczki o stałej dielektrycznej $\epsilon(X,Y,Z)$:

$$\mathbf{E} = -\nabla\Phi$$

z komputerowego (numerycznego) rozwiązania **równania Poissona-Boltzmann**:

$$-\nabla \cdot (\epsilon \nabla \Phi) = \rho - 2n_{\infty} \sinh(e\Phi/kT)$$

gdzie ρ jest gęstością ładunku cząsteczki a n_{∞} gęstością jonów, anionów i kationów w rozpuszczalniku. Zwykle dla uproszczenia równanie linearyzuje się do postaci LPB (*linear Poisson-Boltzmann*) zostawiając tylko pierwszy człon rozwinięcia sinusa hiperbolicznego:

$$-\nabla \cdot (\epsilon \nabla \Phi) + eK^2\Phi = \rho \quad K^2 = 2Ie^2/kT\epsilon$$

I – siła jonowa roztworu a 1/K tzw. długość przesłaniania Debye’a - Hückla.

Oddziaływania warunkujące **kształt przestrzenny cząsteczki** (konformacja) i stabilizujące **kompleksy molekularne** mają zasadniczo **charakter elektrostatyczny**. W ramach tego ogólnego typu oddziaływań wprowadza się bardziej szczegółowe typy kontaktów stabilizujących z różnym wkładem pochodzącym od oddziaływań kulombowskich i poszczególnych członów oddziaływania van der Waalsa:

(1) **Wiązania wodorowe** polega na uwspólnieniu protonu między dwoma atomami elektroujemnymi jak **tlen, azot**: donorem D z kowalencyjnie związanym protonem i akceptorem A z wolną parą elektronową. Donorem w słabych oddziaływaniach wodorowych może być węgiel, np. C α w białkach ([review: Panigrahi & Desiraju: Proteins.: Struct. Func. Bioinf. 67, 128, 2007](#)). Jako możliwy akceptor WW przyjmuje się także **pierścień aromatyczny**. Przyciąganie jest **zdominowane ponad 60% przez wkład kulombowski** (40% przez van der Waalsa), co po odjęciu energii odpychania van der Waalsa daje z obliczeń kwantowych szacunkową energię jednego wiązania wodorowego kilka do kilkunastu RT (RT = 2.5 kJ/mol w temperaturze pokojowej 300 K; stała gazowa R = 8.3143 kJ/(mol·K)). Obserwacja sprzężeń skalarnych NMR między jądrami przez wiązania wodorowe dało podstawy do prac teoretycznych, w których postuluje się **kowalencyjny wkład** do energii wiązania w formie delokalizacji elektronów wolnej pary A na orbital $\sigma^*(D-H)$ ([Tuttle et al. J. Am. Chem. Soc. 126, 5093, 2004](#)).

(2) **Wiązania halogenowe**, analogiczne do wiązania wodorowego postuluje się dla oddziaływania atomów chlorowców (Cl, Br, I) pochodnych halogenowych cząsteczek organicznych, np. ligandów z akceptorami w postaci atomów elektroujemnych z wolną parą elektronową: (tlen, azot) i pierścieniem π -elektronowym. Energia wiązania halogenowego jest przedmiotem kontrowersji; postulowane wartości: od 1 kJ/mol do 20 ÷ 30 kJ/mol ([Carter & Ho Cryst. Growth & Design 11, 5087, 2011](#)).

(3) **Mostki solne** tworzą się między grupami atomów silnie naładowanych, o przeciwnych znakach, w wyniku dysocjacji protonu H⁺ z jednej grupy (ładunek ujemny) i przyłączeniu protonu przez drugą (ładunek dodatni), Są stabilizowane przez siłę Coulomba, a energia jest rzędu energii wiązania wodorowego.

(4) **Oddziaływania warstwowe (stacking)** między pierścieniami aromatycznymi $\pi - \pi$ oraz dodatnio naładowanej grupy lub pierścienia aromatycznego z pierścieniem aromatycznym **kation - π** . ([review: Meyer et al. Angew. Chem. Int. Ed. 42, 1210, 2003](#)). Energia stackingu jest rzędu energii wiązań wodorowych i zawiera wkład elektrostatyczny oraz wkład hydrofobowy (patrz dalej). Dla oddziaływania $\pi - \pi$ postuluje się znaczący

udział dyspersyjnego wkładu van der Waalsa, a dla oddziaływań kation - π kulombowskie przyciąganie ładunku dodatniego i momentu kwadrupolowego pierścienia aromatycznego ([Ma & Dougherty Chem Rev. 97, 1303, 1997](#)). Nieco kontrowersyjną sprawą jest wkład w stacking (obu typów) przeniesienia ładunku (*charge transfer force*) między najwyżej obsadzonym orbitalem molekularnym pierścienia donora HOMO (*highest occupied molecular orbital*) i najniższym nieobsadzonym orbitalem pierścienia akceptora LUMO (*lowest unoccupied molecular orbital*). W otoczeniu niewodnym (wnętrze białek) obserwuje się oddziaływanie typu stacking także z prostopadłym ustawieniem pierścieni.

(5) Kontakty van der Waalsa przyciągające, bliskozasięgowe (R^{-6}), „nieukierunkowane“ stanowią pozostałą część energii stabilizacji układów biomolekularnych. W tym sensie pojęcie „van der Waalsa“ używa się z subtelnym rozróżnieniem:

- wkład w energię elektrostatyczną wyrażaną przez potencjał V w przybliżeniu FF
- wkład w stabilizację struktur o charakterze bliskozasięgowym, w przeciwieństwie do dalekozasięgowych oddziaływań kulombowskich: wiązań wodorowych i mostków solnych.

(6) W rozwinięciu multipolowym oddziaływania dwóch atomów pomija się oddziaływania ładunków i trwałych multipoli wyższego rzędu np. uśrednione po orientacjach oddziaływanie dipol - dipol:

$$E = -2/(3kTR^6)\{\mu_A^2\mu_B^2\}$$

Rozpatrując jakościowo oddziaływania stabilizujące fragmentów molekularnych, które zwykle charakteryzują się niejednorodnym rozkładem ładunku elektrycznego można to uwzględnić, np. dla **helis białkowych**. Helisy mają znaczne wartości elektrycznych momentów dipolowych ze względu na sumowanie równoległe ustawionych wektorów μ poszczególnych wiązań peptydowych, 3,5 D (Debye'a) każde.

Wykład 6

Klasyczne metody badania cząsteczek w roztworze zajmują się **makroskopowymi układami molekularnymi**, gdzie obowiązują prawa **termodynamiki** i **mechaniki statystycznej**. Procesy fizyczne i chemiczne w skali makro są regulowane nie tylko relacjami energetycznymi, ale również stopniem uporządkowania (nieuporządkowania) struktur molekularnych opisanym przez pojęcie **entropii**:

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

P - ilość dostępnych mikrostanów układu przy ciśnieniu p, temperaturze T i objętości V oraz innych parametrów makroskopowych

Maksymalną wartość entropii układ izolowany adiabatycznie przyjmuje w warunkach równowagi. Wszelkie procesy zachodzące w makroskopowym, otwartym układzie molekularnym o **energii wewnętrznej** U, który wymienia **ciepło** Q, **pracę** W i **masę** M z **otoczeniem** są opisane przez pierwszą i drugą zasadę termodynamiki, czyli bilans energetyczny na zmianę energii ΔU :

$$\Delta U(S, V, \dots, n_1, n_2 \dots) = W + Q + \Delta M$$

$$W = -p\Delta V + W_{\text{nieobjętościowa}},$$

$$\Delta M = \sum_k \mu_k \Delta n_k \quad \mu_k - \text{potencjał chemiczny } n_k \text{ moli składnika } k$$

$$\mu_k = (\partial U / \partial n_k)_{S, V, \dots, n_j \neq n_k} = \mu_k^0 + RT \ln(n_k) \quad \mu_k^0 - \text{potencjał standardowy 1 mola}$$

$$Q = T\Delta S$$

Energia U makroskopowego układu cząsteczek jest sumą energii E_j (energii wewnątrzcząsteczkowych) i energii oddziaływań między cząsteczkami E_{ij} , w tym oddziaływania z rozpuszczalnikiem. W przybliżeniu słabych oddziaływań międzycząsteczkowych w porównaniu z wewnątrzcząsteczkowymi otrzymuje się **energię układu makroskopowego** N_0 cząsteczek korzystając z ich **rozkładu Boltzmanna** na skwantowanych poziomach energetycznych E_j :

$$U = Z^{-1} N_0 \sum_j E_j \exp(-E_j/kT) \quad Z = \sum_j \exp(-E_j/kT)$$

Każdy układ dąży do stanu **równowagi termodynamicznej**, w której przyjmuje minimum energii wewnętrznej U i maksimum entropii S . W warunkach stałej temperatury i ciśnienia procesy zachodzą w kierunku **osiągania minimum energii swobodnej Gibbsa G** (entalpii swobodnej):

$$G = U + pV - TS$$

(7) **Oddziaływania hydrofobowe** występuje gdy fragmenty cząsteczkowe źle oddziałują z wodą (źle rozpuszczalne w wodzie) i wolą agregować ze sobą, np. grupy alifatyczne CH_3 czy pierścienie benzenowe. W skrótowym, jakościowym ujęciu oddziaływanie wynika z **efektów entropowych**, związanych z reorganizacją wzajemną cząsteczek wody w obecności hydrofobowych powierzchni cząsteczek rozpuszczonych. Związanie reszt hydrofobowych oddziaływaniem przyciągającym zmniejsza ich powierzchnię kontaktu z wodą, a więc pojawia się więcej wody nieuporządkowanej i entropia rośnie. Badania metodą rozpraszania neutronów pokazały (zgodnie z obliczeniami teoretycznymi) zmniejszenie gęstości wody przy powierzchniach hydrofobowych ([Doshi et al. Proc. Natl. Acad. Sci. USA 102, 9458, 2005](#)). Teoria oddziaływań hydrofobowych na gruncie termodynamiki fenomenologicznej i fizyki statystycznej jest ujęta w podręczniku [A. Ben-Naim: Hydrophobic Interactions Plenum Press 1980](#). Dla substancji rozpuszczonej (*solute*) wprowadza się **hydrofobowość** jako **standardową energię swobodną Gibbsa transferu między fazą wodną (aq) i odpowiednio dobraną fazą organiczną (org)**, wyrażoną przez współczynnik podziału substancji $K_{\text{aq} \rightarrow \text{org}}$ między obu fazami:

$$\Delta G^\circ_{\text{h}} = -RT \ln K_{\text{aq} \rightarrow \text{org}}$$

$$K_{\text{aq} \rightarrow \text{org}} = x_{\text{aq}}/x_{\text{org}}$$

x - ułamek molowy (stężenie) substancji w fazach:
wodnej i organicznej

Dysponując **skalą hydrofobowości** dla danego typu cząsteczek, np. aminokwasów w białku można wyznaczać/obliczać energie swobodne Gibbsa stabilizacji struktur związanych oddziaływaniem hydrofobowym, np. przy związaniu białek globularnych (przegląd skal hydrofobowości aminokwasów podaje w *appendix* praca [Cornette et al. J. Mol. Biol. 195, 659, 1987](#)). Najbardziej pożądana byłaby znajomość funkcjonalnej formy **potencjału oddziaływań hydrofobowych**, co pozwoliłoby określić znaczenie tych oddziaływań w stabilizacji struktur makromolekularnych (w porównaniu z oddziaływaniami elektrostatycznymi). Jak dotąd nie skonstruowano uniwersalnego

potencjału hydrofobowego. Przykładem w tym kierunku dla białek jest próba powiązania sił hydrofobowych z tzw. potencjałem energetycznym średniej siły PMF (*potential of mean force*). PMF opisuje oddziaływania cząsteczek w roztworze z uwzględnieniem **uśrednionych efektów solwatacyjnych** cząsteczek rozpuszczalnika ([Casari & Sippl J. Mol. Biol. 224, 725, 1992](#))

Przegląd zaawansowanych TECHNIK badania struktur i dynamiki biopolimerów i ich kompleksów z rozdzielczością molekularną; techniki mikroskopowe; manipulacja pojedynczymi cząsteczkami

Ponieważ duże i skomplikowane biopolimery badane są z punktu widzenia różnych informacji, jakie chcemy uzyskać o ich strukturze i funkcjonowaniu w komórce, rozwinięto znaczną ilość różnych metod. Na wykładzie krótko przedstawię ogólny podział metod ze względu na **zakresy informacji**, jakie pozwalają uzyskać o biopolimerach, oraz problemy do rozwiązywania których są stosowane, a następnie wybrane z metod będą szczegółowo i w ścisły sposób scharakteryzowane przy omawianiu kluczowych tematów na wykładzie.

Pod względem stosowanego warsztatu badawczego **metodologia badań** opiera się na dwóch **komplementarnych i ściśle powiązanych ze sobą podejściach**.

(1) **Metody doświadczalne** które operują na „materialnych” obiektach badań (biomolekułach i ich kompleksach) *in vitro* oraz *in vivo*: wyznaczanie parametrów strukturalnych i dynamicznych, pomiary wielkości fizykochemicznych, analiza procesów z udziałem badanych obiektów, obrazowanie strukturalne i celowa manipulacja obiektami (zasadniczo omawiane na tym wykładzie BIOFIZYKA DOŚWIADCZALNA).

(2) **Metody teoretyczne** które operują na modelach obiektów, matematycznych i komputerowych: obliczenia analityczne i numeryczne, symulacje przebiegu czasowego procesów (zasadniczo omawiane na równoległym wykładzie BIOINFORMATYKA I MODELOWANIE), eksperymenty „*in silico*” (czyli w komputerze).

Pod względem zakresu informacji metody eksperymentalne można zasadniczo podzielić na klasy **[PLANSZA 40]** zilustrowane **charakterystycznymi przykładami**, które będą wspomniane krótko lub omówione szerzej w trakcie analizy zagadnień budowy i funkcjonowania obiektów molekularnych (lista nie obejmuje wszystkich metod).

(A) Metody wyznaczania mas cząsteczkowych i separacji cząsteczek/kompleksów molekularnych na podstawie różnic mas:

ultrawiórowanie analityczne

elektroforeza żelowa GE (*gel electrophoresis*) i kapilarna CE (*capillary electrophoresis*)

spektrometria masowa MS (*mass spectrometry*)

(B) Metody określania struktur chemicznych (wzory strukturalne) monomerów, w tym tautomerii zasad kwasów nukleinowych oraz struktur pierwszorzędowych polimerów liniowych:

dyfrakcja promieniowania rentgenowskiego długości fali $\lambda \sim 10^{-10}$ m [Å] (rentgenografia) na monokryształach molekularnych, spektroskopia molekularna, głównie magnetyczny rezonans jądrowy NMR (*nuclear magnetic resonance*)

sekwencjonowanie białek i kwasów nukleinowych

(C) Metody wyznaczania parametrów fizykochemicznych, energetycznych i termodynamicznych biomolekuł i ich kompleksów; stabilność struktur natywnych oraz równowagi konformacyjne:

stacjonarne metody spektroskopowe, uszeregowane według długości fali stosowanego promieniowania elektromagnetycznego, od spektroskopii Mössbauera (promieniowanie γ : $\lambda \sim 10^{-13}$ m, $\nu \sim 10^{21}$ Hz) do magnetycznego rezonansu jądrowego NMR (fale radiowe $\lambda \sim$ m i $\nu \sim 10^7 \div 10^{10}$ Hz) **[PLANSZA 41]**

metody kalorymetryczne ITC (*isothermal titration calorimetry*) i DSC (*differential scanning calorimetry*) **[PLANSZA 40]**

(D) Metody wyznaczania struktur przestrzennych z rozdzielczością atomową:

rentgenografia, neutronografia, dyfrakcja elektronowa na kryształach 2D

NMR w roztworze i *high resolution* NMR w ciele stałym techniką wirowania pod kątem magicznym MAS (*magic angle spinning*)

(E) Metody wyznaczania struktur przestrzennych ze średnią rozdzielczością

mikroskopia sił atomowych AFM (*atomic force microscopy*)

mikroskopia elektronowa pojedynczej cząsteczki cryoEM (*cryo electron microscopy*)

rozpraszanie niskokątowe promieniowania rentgenowskiego (X) SAXS (*small angle X-ray scattering*) i neutronów SANS w roztworze i DLS (*dynamic light scattering*)

rentgenografia włókien kwasów nukleinowych

(F) Metody badania dynamiki ruchów molekularnych i przekształceń strukturalnych; przejścia konformacyjne, zwijanie białek i RNA, asocjacja kompleksów, w tym kinetyka wymienionych procesów *in vitro*:

fluorescencja czasowo-rozdzielcza
fluorescencja korelacyjna FCS (*fluorescence correlation spectroscopy*)
relaksacja NMR
powierzchniowy rezonans plazmonowy SPR (*surface plasmon resonance*)
termoforeza MST (*microscale thermophoresis*)
wymiana proton-deuter (H/D *exchange*) z detekcją MS lub z detekcją NMR
rentgenografia Lauego reakcji enzymatycznych w kryształach
czasowo-rozdzielcze rozpraszanie synchrotronowego impulsu promieniowania X w roztworze po laserowo wzbudzonej reakcji TRXSS
dwuwymiarowa spektroskopia IR (2D IR *vibrational echo spectroscopy*)
rozpraszanie neutronów QENS (*quasielastic*) i EINS (*elastic incoherent*)

(G) Metody manipulacji i obserwacji pojedynczych makrocząsteczek/kompleksów

szczytce optyczne (*optical tweezers*)
szczytce magnetyczne (*magnetic tweezers*)
AFM (*atomic force microscopy*)
spektroskopia pojedynczej cząsteczki SMS (*single molecule spectroscopy*); smFRET

(H) Metody śledzenia procesów metabolizmu i obrazowania *in vivo*

in vivo NMR oraz *in cell* NMR
obrazowanie MRI i fMRI (*magnetic resonance imaging, functional MRI*)
topical magnetic resonance: połączenie *in vivo* NMR oraz obrazowania MRI
obrazowanie fluorescencyjne.

Fizyka jako „stoi” za wymienionymi metodami obejmuje zasadniczo trzy klasy zagadnień:

- oddziaływanie promieniowania: elektromagnetycznego, neutronów, elektronów, z układami molekularnymi;
- hydrodynamika;
- efekty cieplne.

Fizyka ta może być bardzo prosta, od klasycznego prawa Archimedesesa (ultrawierwienie) i siły Lorentza (spektrometria mas) do kwantowego rezonansu plazmonowego (SPR): wzbudzenie elektronów powierzchniowych (plazmonów) w metalu graniczącym z cieczą celki przepływowej przez promieniowanie „zanikające” (*evanescent field wave*) przy całkowitym odbiciu wewnętrznym. Również aparatura pomiarowa może się bardzo różnić stopniem złożoności, niezależnie od stopnia komplikacji fizyki danej metody.

Profil badań w biofizyce można podzielić na dwa obszary, wymagające postępu prac dla precyzyjnego i wiarygodnego opisu zjawisk biologicznych na poziomie

molekularnym oraz praktycznych zastosowań uzyskanych wyników w biotechnologii i chemioterapii:

- analiza **nowych zjawisk i obiektów** przy zastosowaniu nieraz bardzo skomplikowanej, ale zwykle standardowej metodologii (aparatury badawczej);
- wykorzystywanie sprawdzonych obiektów jako modeli rozwoju i **testowania nowych metod** badawczych; przykłady standardowych modeli tego typu to białko BPTI (*bovine pancreatic trypsin inhibitor*) czy lizozym.

Praktyczna znajomość metod badawczych (lub obliczeniowych w przypadku biofizyka-teoretyka) czyli „kuchni” jest warunkiem koniecznym podjęcia i prowadzenia badań z nadzieją na sukces czyli rozwiązanie postawionego problemu. Jednak podejmując nową tematykę badawczą powinno się rozpoczynać od przestrzegania zasady **prymatu problemu**, do którego trzeba dobrać odpowiednią metodę lub starać się rozwinąć nową metodę. „Naturalna” tendencja w prowadzeniu pracy badawczej jest często odwrotna i wynika zwykle z opanowania określonej metody i umiejętności obsługi przyrządu pomiarowego, co skutkuje „nieodbyłym” podejściem w pracy badawczej, polegającym na selekcji problemów rokujących nadzieję rozwiązania pod kątem określonej metody.

Badania doświadczalne w biofizyce wymagają określonego poziomu umiejętności w zakresie metod teoretycznych (komputerowych). Nieodzownym elementem **analizy danych** jest umiejętność wykorzystywania baz danych i stosowania **narzędzi bioinformatycznych**. Szereg metod doświadczalnych pozwala uzyskać potrzebne parametry przy założeniu określonych modeli teoretycznych, które nie są *implicite* zawarte w stosowanej metodzie. Symulacje komputerowe mogą stanowić nieodzowny element procedury eksperymentalnej, np. wyznaczanie struktur przestrzennych w roztworze metodami NMR zakłada wykonanie symulacji dynamiki molekularnej z więzami wyznaczanymi z widm doświadczalnych rMD (*restrained molecular dynamics; simulated annealing*).

GENOMIKA i sekwencjonowanie DNA; budowa chromosomu, mutageneza, kancerogeneza i systemy naprawcze

Wykład 7

Pierwszym krokiem w ustaleniu pełnej struktury cząsteczki kwasu nukleinowego jest określenie **sekwencji nukleotydów od 5'- do 3'-końca** czyli STRUKTURY PIERWSZORZĘDOWEJ. Naczelnym zadaniem **genomiki** jest ustalenie pełnej sekwencji nukleotydowej DNA człowieka i innych organizmów oraz **określenie wynikających z sekwencji genów**. W 2001 r. badacze w dwóch projektach ogłosili (niezależnie) wykonanie sekwencjonowania DNA genomu ludzkiego: międzynarodowy Human Genom Project ([Lander et al. Nature 409, 860, 2001](#)) i prywatna kompania Celera Genomics ([Venter et al. Science 291, 1304, 2001](#)). Początkowo konkurująca, a pod koniec działań obu projektów, połączona strategia obejmowała:

- pozyskanie i pocięcie DNA przy pomocy endonukleaz restrykcyjnych na zestaw krótkich fragmentów jednoniciowych,
- sklonowanie fragmentów w bakteriach *E. coli*,
- ustalenie sekwencji każdego fragmentu **techniką Sanger**,
- złożenie milionów fragmentów we właściwej kolejności „na zakładki” jednakowych sekwencji, przez wykorzystanie komputerów dużej mocy metodą **whole-genome shotgun cloning** (WGS), opracowaną przez J. Craiga Ventera (Celera Genomics).

Ustalanie sekwencji w krótkich fragmentach do ok. 400 ÷ 600n standardowo wykonywano za pomocą kontrolowanego przerywania replikacji, która 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 lub 2',3'-dideoxy-GTP lub 2',3'-dideoxy-CTP lub 3'-deoxy-TTP, który po włączeniu do nici terminuje syntezę wskutek braku grupy 3'OH. Odpowiednia ilość każdego 2',3'-dideoxy-NTP w każdej z 4 reakcji zapewnia takie włączanie do nowo-syntetyzowanego DNA w stosunku do zwykłych NTP (1:100), że powstają porównywalne ilości fragmentów różniących się o jeden nukleotyd. Primery wymagane do zapoczątkowania działania polimerazy w każdej reakcji posiadają inaczej fluoryzujący znacznik fluorescencyjny. Elektroforeza pod detektorami czuymi na 4 różne długości fali pozwala identyfikować kolejne, coraz dłuższe fragmenty zakończone na 3'-

końcu określonymi zasadami (identyfikacja według fali emisji), dając sekwencję nici komplementarnej, a stąd wyjściowej, po uwzględnieniu sekwencji primera (szczegóły patrz wykład: *Struktura i funkcje makrocząsteczek biologicznych*),

W **automatycznych** sekwenatorach (np. ABI Prism 3730XL) stosuje się do rozdziału elektroforezę **kapilarną CE** (*capillary electrophoresis*) na 96 kapilarach kwarcowych, co daje wydajność ok. **10⁵ nukleotydów/doba**. Zsekwencjonowanie genomu ludzkiego ok. 10¹⁰ n przy pomocy 100 maszyn trwa więc ok. 2,5 roku. W elektroforezie CE **[PLANSZA 42]** (podobnie jak w żelowej GE) rozdzielanie naładowanych cząsteczek o ładunku Z i promieniu a następuje w wyniku różnic ich **ruchliwości elektroforetycznej** v/E, czyli szybkości migracji v w stałym polu elektrycznym o natężeniu E, którego siłę coulombowską: F_e=Z·E równoważy siła tarcia F_t = f·v;

$$F_e = F_t \quad \Rightarrow \quad v = (Z \cdot E) / f \quad f = 6\pi\eta a - \text{współczynnik tarcia (wzór Stokesa)}$$

η - lepkość roztworu

Cząsteczki poruszają się w buforze, w cienkiej kapilarze kwarcowej długości 30 ÷ 100 cm i średnicy 20 ÷ 100 μm, między naczyniami wypełnionymi buforem. Przepływ następuje pod wpływem przyłożenia pola elektrycznego do elektrod zanurzonych w naczyniach przy różnicy potencjałów do 30 kV (pola elektryczne 500 V/cm). Limit natężenia pola wynika z silnego ogrzewania układu przez prąd płynący w kapilarze. Migrująca substancja jest oznaczana w detektorze; w przypadku sekwencjonowania DNA jest to detektor fluorescencyjny, czuły na 4 różne długości fali znaczników fluorescencyjnych. Jonizacji reszt kwasu krzemowego =SiOH → =SiO⁻ na ściankach kapilary (pH buforu > 3) jest (częściowo) zrównoważona sztywno związanymi kationami buforu. Dalej tworzy się warstwa kationów o większej ruchliwości (obszar dyfuzyjny). Powstająca w ten sposób dyfuzyjna warstwa podwójna kationów daje **napięcie ścinające**, które **różnicuje ruchliwość przepływu elektroosmotycznego** cząsteczek, podobnie jak usieciowany żel różnicuje ruchliwość w elektroforezie GE. Ruchliwość v/E cząsteczek jest wypadkową ruchliwości elektroosmotycznej i ruchliwości przepływu („swobodnego”) w buforze. Ogólnie biorąc: cząsteczki naładowane dodatnio poruszają się szybciej niż bufor (przepływ od anody do katody), obojętne tak samo jak bufor, a ujemne wolniej. Małe cząsteczki o dużym ładunku poruszają się szybciej niż duże o małym ładunku. W konsekwencji **ruchliwość elektroforetyczna substancji zależy od masy cząsteczki i jej kształtu** (podobnie jak w elektroforezie żelowej) oraz ładunku i lepkości buforu.

Po ukazaniu się publikacji HGP i Celera Genomics podstawowym celem, w tym praktycznym, genomiki stało się opracowanie **wysoce „zrównoległowanej” techniki sekwencjonowania MPS (*massively parallel sequencing*)**, w której wysoka przepustowość (*high throughput*) jest osiągnięta przez zwiększenie ilości jednocześnie analizowanych fragmentów DNA, jednak (na razie) kosztem długości odczytywanych fragmentów (*read length*). Celem praktycznym takiej metody jest możliwość określenia pełnej sekwencji genomu dowolnej osoby za ok. \$1 000 w przeciągu ok. tygodnia, co stanowi jeden z cyklicznych tematów doniesień prasowych i "reklamowych"*. Nowoczesne metody „*massively parallel*” **[PLANSZA 43]**, zwane też **platformami sekwencyjnymi** dysponują komercyjnie dostępnymi *sekwencjatorami* **[PLANSZA 44]**:

Roche/454 (*GS FLX Titanium*), **Life Technologies** (*SOLID*), **Illumina** (*Hi-Seq 2000*, poprawiona wersja *SOLEXA Genome Analyser*), **Pacific Biosciences** (*PacBio RS*)

Różnią się one od metody Sangera i mają pewne wspólne cechy ogólne **[PLANSZA 43]**.

- (1) Zamiast klonowania w bakteriach i izolacji, tworzenie bibliotek (syntetycznych) DNA;
- (2) Amplifikacja DNA na podłożu stałym (*library amplification method*) lub zastosowanie techniki pojedynczej cząsteczki (Pacific Biosciences);
- (3) Seqwencjonowanie jednocześnie kilkuset tysięcy do kilku milionów fragmentów DNA jako seria powtarzalnych, zautomatyzowanych kroków (zależnie od platformy);
- (4) Możliwość otrzymania informacji o sekwencji z obu końców fragmentów DNA.

Istotnym zagadnieniem i wyzwaniem jest sekwencjonowanie genomów pojedynczych komórek mikrobów (*single-cell sequencing*; [Woyke et al. Nature Methods 14, 1045, 2017](#)).

Zestawienie etapów przygotowania próbek i technologii sekwencjonowania **[PLANSZA 44]** obejmuje pełen przekrój rozwiązań, od pierwszego systemu MPS Roche/454, który ze względu na wysoki poziom błędów i względnie małą przepustowość ma ograniczone zastosowanie, poprzez **konkurujące aktualnie technologie SOLID i Hi-Seq**, aż do techniki „trzeciej generacji” PacBio RS detekcji w czasie realnym fluorescencyjnie znakowanych nukleotydów włączanych podczas syntezy drugiej nici. Przykładowo, w technologii Illumina Hi-Seq „biblioteka” dwuniciowych fragmentów DNA do sekwencjonowania, 300 ÷ 600 bp, przepływa przez przepuszczającą światło celkę z przyczepionymi primerami „*forward*” i „*reverse*”, do których hybrydują i są namnażane techniką PCR (*bridge amplification*). Uzyskuje się **gęstość ~10 mln fragmentów DNA/cm²** w klastrach po ok. 1000 identycznych kopii każdy. Jedna nić jest selektywnie

* Humanizing Genomics **macrogen**: whole genome seq (Human), ...15 days; 599 EUR.

usuwana przed doczepieniem primera „sequencing”. Do każdej nici DNA w klastrze polimeraza inkorporuje ten sam pojedynczy nukleotyd odwracalnie terminujący syntezę podczas jednego cyklu. Nukleotydy identyfikuje się przez rejestrację wzbudzonej laserowo **fluorescencji etykiet** o czterech różnych długościach fali emisji w przypadku każdego rodzaju nukleotydu: A, C, T, G. Fluorofory i grupy blokujące (terminujące) usuwa się chemicznie i następuje kolejny cykl inkorporacji. Składanie całych sekwencji z fragmentów do 150 n (*read length*) odbywa się przez uszeregowanie względem znanych sekwencji (algorytmy BLAST, MAQ i inne), co zapewnia możliwość wyszukiwania różnic w stosunku do wzorca, np. identyfikacja SNP (*single nucleotide polymorphism*).

Kluczowym i trudnym zagadnieniem (szczególnie w eukariotach), dla którego sekwencje stanowią „materiał” wyjściowy jest **annotacja regionów kodujących** czyli „określenie genów” (*gene prediction strategy*; [review: Sleator Gene 461, 1, 2010](#)) stanowiących ok. 3% sekwencji. Wyniki są niezwykle istotne dla badań biomedycznych oraz zrozumienia takich zagadnień jak np. ([review: Lander Nature 470, 187, 2011](#)):

- funkcje biologiczne zakodowane w genomie,
- genetyczne podstawy chorób dziedzicznych i nowotworowych,
- ewolucja ludzkiego gatunku ([Veermach & Hammer Nature Rev. 15, 140, 2014](#)).

Dokładność sekwencjonowani (wykluczenie *sequencing errors*) wydaje się mieć kluczowe znaczenia w wyszukiwaniu rzadkich mutacji w komórkach nowotworowych (*tumor's secret weapon*; [Marx Nature Methods 13, 295, 2016](#)).

Funkcjonowanie maszynarii ekspresji genu jest uwarunkowane przede wszystkim **strukturą DNA** i jej przekształceniami w procesie transkrypcji. Długie cząsteczki DNA, 10^6 bp (bakterie) ÷ 10^9 bp (organizmy wyższe) są **zintegrowane z białkami w nici chromatynowej**. Funkcjonalne DNA przyjmuje STRUKTURĘ DRUGORZĘDOWĄ w formie helisy podwójnej (duplexu), złożonej z dwóch przeciwnie skierowanych i komplementarnych nici polinukleotydowych. Wyróżnia się trzy główne, **klasyczne rodziny helis podwójnych** [\[PLANSZA 45\]](#), w których może występować DNA: **B, A i Z**, przy czym do rodziny B zalicza się **C-DNA** oraz **D-DNA**. Klasyczne struktury DNA zostały wymodelowane na podstawie obrazów dyfrakcyjnych **włókien** [\[PLANSZA 46\]](#), potwierdzone następnie przez wyniki badań krótkich fragmentów w kryształach (dyfrakcja rentgenowska) i w roztworze (NMR). Włókna DNA wytrącane z buforu rozpuszczalnikami organicznymi są regularnymi strukturami tylko w jednym kierunku i stąd **czynnik strukturalny** [\[PLANSZA 47\]](#) we współrzędnych biegunowych jest uśredniony po współrzędnej Φ układu cylindrycznego współrzędnych sieci odwrotnej (R, ζ, Φ):

$$F(R, \zeta) = \langle F(R, \zeta, \Phi) \rangle_{\Phi} = \sum_n \sum_j f_j J_n(2\pi R r_j) \exp\{2\pi i(lz_j - n\phi_j)\}$$

f_j – czynnik rozpraszania j -tego atomu o współrzędnych cylindrycznych r_j, z_j, ϕ_j

J_n – funkcja Bessela rzędu n

l - numer linii warstwy spełniający relację: $\zeta = l/c = n/P + m/h$

gdzie P jest długością skrętu helisy (*pitch*), h jest skokiem helisy na jedną parę zasad (*axial rise*), c jest ilością nukleotydów w motywie powtarzalnym (*repeat unit*), m jest dowolną liczbą całkowitą (dodatnią lub ujemną).

(analiza dyfraktometrów **kryształów** molekularnych - patrz dalej)

Obraz dyfrakcyjny włókna nie daje możliwości odtworzenia położenia poszczególnych atomów. Można jedynie zakładać struktury modelowe helis i badać zgodność ich teoretycznego obrazu dyfrakcyjnego z dyfraktogramem uzyskanym doświadczalnie. Niektóre z helis, np. **helisę D** obserwuje się w syntetycznych polinukleotydach jak poli(dA-T), poli(dG-dC), ale także w fagu T_2 , którego DNA zawiera 5-hydroksymetylocytydynę zamiast cytydyny.

W klasycznych, regularnych helisach wszystkie nukleotydy składowe opisane są przez jeden zestaw wartości kątów dwuściennych łańcucha fosfocukrowego od α do ζ , kąt glikozydowy χ zasady wokół wiązania glikozydowego oraz konformację pierścienia cukru w modelu pseudorotacyjnym; konformacja typu N lub typu S **[PLANSZA 48]**. Podobnie, jeden zestaw parametrów helikalnych dobrze oddaje kształt i charakterystyczne cechy przestrzenne helis.

KĄTY DWUŚCIENNE

| | α | β | γ | δ | ϵ | ζ | χ |
|-------|----------|---------|----------|----------|------------|---------|------------|
| B-DNA | 314 | 214 | 36 (+sc) | 156 (S) | 155 | 264 | 262 (anti) |
| C-DNA | 321 | 200 | 37 (+sc) | 157 (S) | 161 | 254 | 263 (anti) |
| D-DNA | 298 | 208 | 69 (+sc) | 157 (S) | 141 | 259 | 266 (anti) |
| A-DNA | 310 | 172 | 41 (+sc) | 79 (N) | 214 | 282 | 206 (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 $h=P/X$ (Å) | skręcenie/bp $\Omega=360/X$ (°) | nachylenie bp η (°) | odległość bp od osi D(Å) |
|-------|-------------------|---------------|------------------------|------------------------------------|-----------------------------|--------------------------------|
| B-DNA | 10_1 | 33.8 | 3.38 | 36 | -6 | 0.5 |
| C-DNA | 28_3 | 30.9 | 3.31 | 38.6 | -8 | 1.0 |
| D-DNA | 8_1 | 24.2 | 3.03 | 45 | -16 | 1.8 |
| A-DNA | 11_1 | 28.2 | 2.56 | 32.7 | 20 | 4.7 |
| Z-DNA | 6_5 | 43.5 | 7.25 | -60 (-45; -15) | -5 | |

Do parametrów helikalnych zalicza się typ symetrii śrubowej oznaczonej symbolem X_k : obrót helisy o kąt $k \cdot 360^\circ / X$ i jednocześnie przesunięcie wzdłuż osi helisy o $(k \cdot P) / X$ przeprowadza motyw, jeden nukleotyd w helisach A, B (C, D) lub dwa kolejne nukleotydy w Z, w następny motyw. P jest długością jednego skrętu helisy, h jest skokiem helisy na jedną parę (w Z-DNA dwie pary) zasad, kąt Ω jest skręceniem helisy na jedną parę zasad, kąt η jest nachylenie pary zasad do osi, a odstęp geometrycznego środka par zasad od osi helisy określa odległość D. **Dupleksy są stabilizowane** przez trzy rodzaje oddziaływań:

- wiązania wodorowe w parach zasad komplementarnych, A:T i G:C;
- oddziaływania warstwowe (stacking) między sąsiednimi parami zasad;
- ekranowanie ujemnie naładowanych grup fosforanowych przez kationy w roztworze.

Struktura helikalna jest stabilna w określonych warunkach środowiska. Wzrost temperatury, obniżenie stężenia soli, skrajne pH powodują **denaturację helisy** i odwracalne przejście do oddzielnych nici. Profile topnienia helisy wyznacza się:

- (a) spektrofotometrycznie, absorpcja w funkcji temperatury T (hypochromazja);
- (b) metodą DSC (różnicowa kalorymetria skaningowa): c_p w funkcji T.

Analiza profili topnienia pozwala wyznaczyć parametry termodynamiczne ([Marky & Breslauer Biopolymers 26, 1601, 1987](#)). Z fitowania (liniowego lub nieliniowego) spektrofotometrycznie wyznaczanej frakcji formy helikalnej f otrzymuje się z równania van't Hoffa na zależność $K(T)$ **entalpię van't Hoffa** topnienia/tworzenia helisy ΔH_{vH} :

$$\frac{\partial \ln K_a}{\partial T} = \Delta H_{vH} / RT^2 \quad K_a = 2f / (1 - f)^2 c_0 \quad c_0 - \text{stężenie pojedynczej nici}$$

Wyznaczenie temperatury T_m topnienia helisy dla $f = 1/2$ daje również wartość ΔH_{vH} :

$$\Delta H_{vH} = 6RT_m^2(\partial f/\partial T)_{T=T_m}$$

Analiza profilu DSC daje „bezpośrednią” **entalpię kalorymetryczną** ΔH_{cal} i entropię ΔS_{cal} :

$$\Delta H_{cal} = \int c_p(T) dT \quad \Delta S_{cal} = \int [c_p(T)/T] dT$$

Topienie i powstawanie helisy zachodzi w sposób **kooperatywny** (model „*all-or-none*”) $\Delta H_{cal}/\Delta H_{vH} = 1$, lub w poszczególnych fragmentach w warunkach **zmniejszenia kooperatywności** ("spłaszczanie" profilu topienia i jego "pofałdowanie"), $\Delta H_{cal}/\Delta H_{vH} > 1$.

Badania strukturalne syntetycznych helis w **monokryształach** (dyfrakcja rentgenowska) i w roztworze (NMR) pokazały, że struktury helis odbiegają od klasycznych **[PLANSZA 49]** i każdy nukleotyd jest opisany przez indywidualny zestaw kątów dwuściennych i parametrów helikalnych. Dalej jednak nieregularne helisy należą do ściśle określonej rodziny A, B lub Z. Wprowadza się **dotatkowo 11 parametrów helikalnych** opisujących nieregularności we wzajemnym ustawieniu zasad (**przykładowe 4** pokazuje **[PLANSZA 50]**), które przenoszą się na nieregularności łańcucha fosfocukrowego. Parametry te mają zerowe wartości w helisach regularnych.

Polimorfizm strukturalny DNA na poziomie struktury drugorzędowej jest obserwowany w funkcjonalnym DNA *in vivo*, aczkolwiek biologiczna rola struktur typu Z i typu A nie jest w pełni wyjaśniona. Postuluje się m. in.:

- zwiększoną odporność A-DNA na uszkodzenia spowodowane czynnikami fizycznymi;
- powiązanie konformacji Z-DNA z represją genów.

Interesujące są **mechanizmy przejścia między formami**: $B \leftrightarrow A$, a szczególnie $B \leftrightarrow Z$. W tym ostatnim przejściu następuje zmiana skrętności helisy z prawoskrętnej formy B na lewoskrętną formę Z. Problem jest analizowany zarówno poprzez śledzenie procesu przejścia między formami *in vitro* oraz *in vivo*, jak i analizę struktur form przejściowych. Przykładowo, obserwowano przejście $B \leftrightarrow A$ *in vivo* w komórkach pod wpływem odwodnienia metodą spektroskopii ramanowskiej i FT-IR ([Whelan et al. Nucleic Acids Res. 39, 5439, 2011](#)) oraz tworzenie łączy B – Z w syntetycznym fragmencie 13 bp DNA pod wpływem wiązania fragmentu Z_{ADAR1} białka ADAR1 (*double stranded RNA deaminase 1*; [Lee et al. J. Am. Chem. Soc. 134, 5276, 2012](#)).

Z pewnymi wyjątkami (wirusy DNA, plazmidy), funkcjonalne DNA tworzy STRUKTURĘ TRZECIORZĘDOWĄ **chromosomu** w powiązaniu z białkami w wyniku **wieloletapowego upakowania helikalnego DNA**, od 10^3 razy (bakterie) do 10^6 razy (eukariota), jeśli długość nici DNA w chromosomach, ok. 2 m podzielić przez średnicę komórki/jądra ok. 2 μ m, do której musi się zmieścić.

Wykład 8

Pierwszym etapem upakowania DNA jako aktywnego nośnika informacji genetycznej w **chromosomie eukariotycznym** jest struktura **nukleosomu** [PLANSZA 51]. DNA długości 147 bp nawija się lewoskrętnie w formie 1,84 (**super**)skrętu na 8 białek histonowych H2A, H2B, H3 i H4 (102 ÷ 135 AA; po dwa każdego typu, tworząc 11 nm rdzeń (*nucleosome core*). Po „zamknięciu” helisy przez histon H1 (6 podtypów, ~230 AA), kompletne nukleosomy zawierają DNA długości 166 bp i są oddzielone od siebie linkerami zmiennej długości, 10 ÷ 90 bp, tworząc strukturę „paciorków na nici” (*beads-on-a-string*). Jądro komórki człowieka zawiera $3 \cdot 10^7$ nukleosomów. Tendencja podwójnej helisy do tworzenia kolistych struktur bez białek obserwowana np. w plazmidach jest uwarunkowana giętkością helisy i silnie zależy od sekwencji. Miarą giętkości helisy na gruncie **hydrodynamicznego modelu giętkiego pręta** jest tzw. *persistence length* L_p (długość trwała) [PLANSZA 52]. Wielkość tę definiuje się jako funkcję autokorelacji wektorów kierunkowych (stycznych) lokalnych fragmentów helisy:

$$\langle \mathbf{t}(0)\mathbf{t}(s) \rangle = \exp(-s/L_p)$$

Czyli na długości (dużo) większej niż $s = L_p$ lokalne orientacje helisy tracą wzajemną zależność (korelację ustawień):

$$\langle \mathbf{t}(0)\mathbf{t}(L_p) \rangle = (1/L_p) \int_0^{L_p} \mathbf{t}(0)\mathbf{t}(s) ds = \exp(-1) = 1/3$$

Helisa DNA o „przypadkowej” sekwencji nukleotydów ma L_p ok. 150 bp (500 Å) ale wartość ta silnie zależy od temperatury, siły jonowej i obecności cząsteczek wiążących się z DNA, np. spermina lub białka. Możliwość tworzenia struktur kolistych wymaga długości helisy ponad $5L_p$. Tzw. **kinetoplastyczne DNA** zbudowane z traktów $5'-(A_5-T-N_4)_n-3'$ cyklizuje w miniokręgi już przy długości ok. 200 bp. Znaczna krzywizna helisy DNA w nukleozomie jest narzucana przez białka histonowe i realizowana jest poprzez periodyczne zmiany (skoki) wartości kąta roll na odcinku kolejnych par komplementarnych nukleotydów, oraz w mniejszym stopniu kąta tilt [PLANSZA 50]. Zasadniczym oddziaływaniem są **mostki solne** między ujemnie naładowanymi **resztami fosforanowymi DNA** i dodatnio naładowanymi resztami bocznymi zasadowych aminokwasów histonów, głównie **lizyny**.

Istotną cechą struktury trzeciorzędowej DNA jest **superskrętność** (*DNA supercoiling*). Cecha ta przejawia się w DNA kolistym jak w plazmidach lub DNA,

którego fragment jest unieruchomiony na blisko położonych końcach jak w nukleosomie. Rozpatrzmy prosty przykład (*L. Stryer Biochemistry*) **[PLANSZA 53]** 260 bp kolistego B-DNA. Taka zrelaksowana cząsteczka ma $T = +25$ skrętoń (powtórzeń) helikalnych (*twisting number*), gdyż w zrelaksowanym, prawoskrętnym DNA obserwuje się 10,4 bp na jeden skok (skręt), a nie 10 jak w klasycznej formie B we włóknie ($260 : 10,4 = 25$). Podstawową cechą topologiczną helisy jest topologiczna ilość skrętoń L (*linking number*), czyli ile razy dwie nici helisy krzyżują się w jakikolwiek sposób. W naszym przykładzie $L = +25$ (śruba prawoskrętne). Przecięcie jednej nici, odwinięcie o 2 zwoje ($L = +23$) daje po powtórnym połączeniu końców nowy topoizomer w formie helisy $T = +23$ zwoje Watsona-Cricka i pętli wewnętrznej. Taka struktura jest niestabilna i DNA przekształca się spowrotem do poprzedniej formy $T = +25$ skrętoń (260 par zasad Watsona-Cricka), z pojawieniem się $W = -2$ (*writhing number*) ujemnych **superskrętoń**. Trzy parametry: topologiczna liczba skrętoń L , liczba skrętoń narzucana przez strukturę drugorzędową T oraz liczba superskrętoń W związane są zależnością:

$$L = T + W$$

Wprawdzie superhelisa w przykładzie jest prawoskrętne, ale jest topologicznie równoważna lewoskrętnej cewce nawiniętej na cylinder i stąd **ujemny znak liczby superskrętoń**. DNA w nukleosomach jest superskrętne w stopniu

$$\lambda = (L - L_0)/L_0 \sim W/L_0 \text{ od } -0.03 \text{ do } -0.09$$

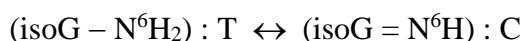
gdzie L_0 jest topologiczną liczbą skrętoń w zrelaksowanym DNA. Odwijanie helisy (bez rozcinania) w jednym miejscu powoduje więc jej "super" zapętlenie w innym, co przez długi czas stwarzało problemy w wyjaśnianiu procesu replikacji DNA. Problem jest rozwiązany przez **topoisomerazy**, enzymy regulujące topologiczną ilość skrętoń. W procesach wymagających rozwijania helisy topoizomerazy przejściowo rozcinają jedną (topoizomerazy typu I) lub dwie (topoizomerazy typu II) nici i trzymają koniec (końce) przez wiązanie kowalencyjne tyrozyna-O-P-O-polinukleotydu. Zwijanie lub odwijanie nici przy udziale ATP z odpowiednią zmianą superskrętności przeciwdziała zapętleniu DNA, a potem następuje zszycie końców. Mikromanipulacje zwijania/rozwijania z możliwością jednoczesnego rozciągania DNA (siły 0,01÷100 pN) wykonuje się za pomocą szczypek magnetycznych dla helisy przymocowanej jednym końcem do stałej powierzchni a na drugim zaopatrzonej w paciorek magnetyczny ([review: Bednar & Dimitrov FEBS J. 278, 2231, 2011](#)). Tak otrzymano m. in. strukturę P-DNA (wykład „Struktura i funkcje makrocząsteczek biologicznych”)

Drugim etapem upakowania DNA w chromosomie jest zwinięcie „włókna” nukleosomowego w solenoidalną **strukturę chromatyny** [PLANSZA 54]. „Klasyczna” struktura o **średnicy 30 nm** zawiera 6 nukleosomów/11 nm (b). Metody mikroskopii elektronowej nie wykazały zbyt dużej ilości tej struktury i postuluje się także inne modele upakowania, np. 12÷15 nukleosomów/11 nm (a). Zarówno same nukleosomy składowa jak i cała nić chromatyny mają struktury wysoce dynamiczne ([Miyagi et al. Biochemistry 50, 7901, 2011](#)), ze względu na konieczność oddziaływania z ponad 500 rodzajami białek zaangażowanych w procesy funkcjonowania DNA: transkrypcja i replikacja ([Ferrari et al. Nature Struct. & Mol. Biol. 21, 118, 2013](#); [752, 45, 2014](#)) oraz systemy naprawcze ([Gospodinov & Herceg Mutation Res. 752, 45, 2013](#)). **Rozwijanie i zwijanie** czyli tzw. „*remodeling*” chromatyny [PLANSZA 17], w tym rozpad i tworzenie nukleosomu oraz ślizganie oktameru histonowego wzdłuż nici DNA bez dysocjacji (*nucleosome sliding*; [review: Planitz et al. Nature Struct. & Mol. Biol. 20, 1026, 2013](#)) zachodzi z udziałem czynników remodelujących, jak np. ATP-azy ([review: Narlikar et al. Cell 154, 490, 2013](#)).

Na tym etapie kończą się oparte na danych doświadczalnych informacje o strukturach trzeciorzędowych funkcjonalnego DNA ([review: Steensel EMBO J. 30, 1885, 2011](#)), ale trwają intensywne prace doświadczalne nad analizą struktur na wyższych etapach zwinięcia nici chromatynowej. **Ujemna superskrętność** chromosomu bakteryjnego (2 ÷ 8 Mb) generuje "*plectonemic loops and branches*" [PLANSZA 55] zorganizowane w odizolowane topologicznie **domeny ~10 kb** (w sumie kilkaset) powiązane z licznymi białkami, w tym kompleksy SMC (*structural maintenance of chromosome*). Domenowa organizacja jest zachowana w chromosomie eukariotycznym [PLANSZA 56] (mysie chromosomy 3 i 18), przy czym domeny są większe, **~100 kb ÷ 1 Mb** [PLANSZA 57]. W reprezentacji macierzowej próbuje się przedstawić oddziaływania poszczególnych fragmentów (*compartments*) [PLANSZA 56D - H]. Doświadczalne techniki wyznaczania oddziaływań pomiędzy odległymi elementami (regulatorowe, promotory) oraz relacje struktura/funkcja, tzw. **techniki 3C** (*chromosome conformation capture*) i ich warianty [PLANSZA 58] ([review: Davies et al. Nature Methods 14, 125, 2017](#)) są oparte na utrwaleniu (*fixation*) komórek, enzymatycznym trawieniu chromatyny i (re)ligacji fragmentów bliskich przestrzennie sekwencji. W połączeniu z sekwencjonowaniem tych fragmentów otrzymuje się **biblioteki 3C** połączeń z profilami rozkładów prawdopodobieństwa tych połączeń. Kluczową sprawą jest, czy wyznaczone kontakty są istotne "*sufficient depth to be reproducible*" czy też są sporadyczne.

Dla obszarów zawierających nici bogate we fragmenty złożone z guanozyn (*G-rich strand*) oraz cytozyn (*C-rich strands*), takich m. in. jak **telomery** („3'-zwisające” końce chromosomów), **centromery** (obszary styku siostrzanych chromatyd gdzie doczepiają się mikrotubule) i **promotory** onkogenów, postuluje się obecność struktur różnych od helisy podwójnej, głównie *in vitro*. Dla powtórzeń (*repeats*) w telomerach dG₂₋₄T₁₋₄A₀₋₁ (ok. 1000 powtórzeń dGGGTTA u człowieka) oraz w promotorach, postuluje się występowanie **struktur G-tetrapleksu** (*G-quadruplex*) różnych typów, od I do IV (**[PLANSZA 59]** [review: Huppert FEBS J. 277, 3452, 2011](#)). Jednak utworzenie G-tetrapleksu *in vivo* stwierdzono na razie tylko w przypadku orzęska *Stilonichia lemnae*. Czteroniciowe fragmenty oligoG w konformacji B-DNA trzymane są przez wiązania wodorowe między guaninami typu **Hoogsteena** i dodatkowo stabilizowane przez **jony K⁺ w środku** planarnego układu czterech guanin. Po każdorazowym skróceniu chromosomów podczas replikacji DNA stabilność genomu utrzymuje **telomeraza** (rybonukleoproteina) katalizująca dołączenie kawałka telomerowego DNA na 3'-końcu korzystając z matrycy RNA. **Struktura C-tetrapleksu. (i-motif)** **[PLANSZA 60]** jest złożona z dwóch równoległych (a nie antyrównoległych) dupleksów. Równoległe duplexy są stabilizowane przez 3 wiązania wodorowe dC⁺:dC i ustawione antyrównoległe do siebie, z naprzemiennym *stackingiem* par cytozynowych. Cała helisa jest prawoskrętna i jej wszystkie cztery nici są konformacyjnie równoważne: należą do rodziny B-DNA. Przejściowymi strukturami w rekombinacji ogólnej (*general recombination*) jest **potrójna helisa (trypleks)**, która zapoczątkowuje proces wymiany części nici w dwóch homologicznych podwójnych helisach DNA (*crossing over*, mejoza). Tworzy się przejściowo struktura złącza Holidaya (*four-way junction*; [Declays & Lilley Curr. Op. Strut. Biol, 18, 86, 2008](#); [wykład Struktura i funkcje makrocząsteczek biologicznych](#)).

Tautomeria zasad azotowych DNA wyznacza **komplementerne parowanie**, na którym opierają się procesy replikacji i ekspresji genu. Rzadkie tautomery **[PLANSZA 61]** enolo (hydroksy) i imino były postulowane przez Watsona i Cricka, a następnie przez Topala i Fresco: ([Nature 263, 285, 1976](#)) jako czynnik inicjujący mutacje punktowe: tranzycje i transwersje **[PLANSZA 9]** wskutek niekomplementarnego parowania zasad (*mismatch*). Wiele nukleozydów z **modyfikowanymi chemicznie zasadami** np. izoguanozyna to promutageny wskutek równowagi tautomerycznej form o porównywalnej zawartości (populacji), z których każda może parować komplementarnie z inną zasadą:



Potwierdzenie hipotezy W.C.T.F. wymaga aby rzadkie tautomery występowały z częstością większą lub równą częstości mutacji punktowych, tranzycji i transwersji, nieskorygowanych przez mechanizmy naprawcze DNA. **W systemach replikacyjnych *in vitro*** prawdopodobieństwo błędnej inkorporacji wynosi $10^{-4} \div 10^{-6}$ **na nowotworzoną parę**, zależnie od typu polimerazy DNA i posiadania przez nią systemu *proofreading* rozpoznania i wycinania błędnie włączonego nukleotydu: polimeraza I *E. coli*, polimeraza bakteriofaga T7 ([Johnson Biochim. Biophys. Acta 1804, 1041, 2010](#)) Wykrycie tak małych populacji rzadkich tautomerów jest bardzo trudne. Przy zastosowaniu pośredniej metody zasadowości **[PLANSZA 62]** uzyskuje się tego rzędu dolne oszacowania wartości stałych równowagi tautomerycznej. W przykładzie z 2'-deoksytydyną:

$$K_T = [\text{rzadki tautomer}]/[\text{dominujący tautomer}] < 10^{-4}$$

Metoda polega na wyznaczeniu stałych protonacji K_P dominującego tautomeru i zafiksowanej (zwykle przez metylację) formy rzadkiego tautomeru. **Założenia metody** to:

- wspólny kation obu tautomerów;

- zbliżone wartości stałych protonacji rzadkiego tautomeru i jego formy zafiksowanej.

Zagadnienie tautomerii zasad zwykłych i modyfikowanych stało się jednym z centralnych problemów doświadczalnych i obliczeń kwantowych 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. Wydawało się, że punktem zwrotnym było wykrycie ok. **50% udziału rzadkiej formy enolowej (hydroksy) w 7-metyloguaninie**, modelowym analogu guanozyny **[PLANSZA 63]** i ok. 10% formy imino w 1-metylocytozynie, analogu cytydyny, techniką spektroskopii IR w niskotemperaturowych matrycach gazowych. Przy rejestracji widma IR 7-Me-Gua, naświetlanie promieniowaniem UV zmienia w sposób zróżnicowany intensywności pasm IR obu tautomerów, co pozwala wyznaczyć widmo tautomeru enolowego (zamiast pasma C = O daje sygnał O-H w zakresie liczb falowych dla obszaru NH₂) i określić jego populację w równowadze tautomerycznej. Wprawdzie nie wykryto mniejszościowych tautomerów adenozyne i tymidyny, ale dokładne dane doświadczalne umożliwiły przetestowanie dokładności teoretycznych obliczeń energii metodami kwantowomechanicznymi. Pionierskie prace prof. Józefa Kwiatkowskiego z UMK w Toruniu pokazały zgodność obliczeń *ab initio* z wynikami IR przy rygorystycznym spełnieniu następujących warunków

- zaawansowana baza funkcji Gaussa,

- optymalizacja geometrii cząsteczek,
- uwzględnienie energii drgań zerowych,
- uwzględnienie energii korelacji elektronów.

Obliczenia kwantowe dla dA i T pokazały że wartości energii (w próżni) dominujących tautomerów są o ok. 60 kJ/mol niższe niż ich rzadkich tautomerów - populacje form rzadkich poniżej 10^{-11} . Ciekawe są również wyniki na energie parowania zasad dla różnych form tautomerycznych, które pokazują przesuwanie równowagi tautomerycznej w stronę tautomerów większościowych przy parowaniu komplementarnym. I tak, energia pary G(oxo):C jest ok. 40 kJ/mol niższa niż G(enol):T, pomimo trzech wiązań wodorowych w obu przypadkach. **Energia pary G(enol):T z trzema wiązaniami wodorowymi jest rzędu energii pary wobble G(oxo):T z dwoma ww.** Ponieważ nie wykryto odpowiednich ilości rzadkich tautomerów dla wszystkich czterech zasad oraz brak jest danych o preferencjach występowania mutacji punktowych dla mismatch'ów z udziałem G i C, hipoteza inicjacji mutacji punktowych przez rzadkie formy tautomeryczne pozostaje nierozstrzygnięta. Ciągłe prowadzone są obliczenia tego typu, i mimo że cząsteczki nie są duże a metody kwantowe stają się coraz bardziej zaawansowane, stale sygnalizowane są trudności z uzyskiwaniem odpowiednich dokładność obliczanych wartości energii ([Piacenza & Grimme J. Comp. Chem. 25, 83, 2004](#)).

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**. Stwierdzono, że adenina naprzeciw cytozyny w syntetycznym, helikalnym 12-merze DNA ([Hunter et al. Nature 320, 552, 1986](#)) **[PLANSZA 64]** w kryształce 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} . Problemem jest **brak stwierdzenia obecności protonu wiązania wodorowego** $A^+N1-H...N3C$. W rentgenografii położenia protonów nie są wyznaczane, a badania odpowiednich fragmentów DNA w roztworze metodami NMR nie wykazały sygnału pochodzącego od tego protonu.

Eksperymenty z włączaniem trifosforanu deoxyrybozydu z **difluorotoluenem zamiast tyminy** **[PLANSZA 65]** naprzeciw adeniny ([Kool et al. Biopolymers 48, 3, 1998](#)) przez polimerazę I DNA 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 [\[PLANSZA 66\]](#), błędne w DNA, jest "normalne" jako parowanie G:U kodon mRNA - antykodon tRNA. Wartości energii par niekomplementarnych w porównaniu z parami komplementarnymi obliczane metodami kwantowomechanicznymi i analogiczne wartości entalpii swobodnych ΔG wyznaczone doświadczalnie wynoszą $\sim 17 \div 28$ kJ/mol, czyli odpowiadają stałej równowagi $K_T \sim 10^{-3}$ do 10^{-5} , zgodnie ze wzorem:

$$\Delta G = RT \ln(K_T) \quad RT = 2,5 \text{ kJ/mol dla } T = 300\text{K}$$

A więc z termodynamicznego punktu widzenia polimeraza, pomimo znacznej dokładności działania, może włączyć niekomplementarny nukleotyd z częstością $10^{-5} \div 10^{-6}$ bez konieczności inicjacji procesu rzadkimi tautomerami lub formami uprotonowanymi/zjonizowanymi.

Podsumowując, żadna z hipotez inicjacji mutacji punktowych nie ma bezwzględnego potwierdzenia. Nie jest wykluczone, że dopuszczalny jest każdy z postulowanych mechanizmów zależnie od typu mutacji punktowej. Aktualne badania są bardziej skoncentrowane na strukturalnej, termodynamicznej i kinetycznej analizie efektywnego i dokładnego (*high fidelity*) powielania informacji genetycznej i utrzymania stabilności genomu przez **polimerazy DNA**, ze szczególnym uwzględnieniem mechanizmu *proofreading*, który o 2 rzędy wielkości podnosi dokładność replikacji ([Freudental et al. Cell 154,157, 2013](#); [Kranc Biochim. Biophys. Acta 1804,1049, 2010](#); [Johnson Biochim. Biophys. Acta 1804,1041, 2010](#)). Jako sondy wykorzystuje się coraz częściej nienaturalne nukleotydy ([Lee & Berdis Biochim. Biophys. Acta 1804,1064, 2010](#)).

Wykład 9

Wielkie cząsteczki **latentnego DNA**, jako złożone układy termodynamiczne, podlegają stałym i niekontrolowanym zmianom oraz **modyfikacjom strukturalnym** (procesy degradacyjne). Jest to o tyle istotne, że modyfikacje struktur DNA mogą prowadzić do zmian niesionej przez nie informacji genetycznej. Nieskorygowane zmiany, utrwalone w procesie replikacji, prowadzą do powstania **mutacji**. Z kolei mutacje wydają się być przyczyną chorób nowotworowych ([review: Jackson & Bartek Nature 461, 1071, 2009](#)). Charakterystycznymi **objawami** (*hallmarks*) nowotworu są **[PLANSZA 67]**: podtrzymywanie sygnałów proliferacji, unikanie supresorów wzrostu, metastazy, nieśmiertelność replikacyjna, angiogeneza, odporność na śmierć komórek (apoptozę). Jedną z hipotez wyjaśniających **proces cancerogenezy** jest skumulowanie mutacji w określonych miejscach DNA, tzw. *cancer-related genes*:

- *tumor suppressors*;
- oncoogeny;
- naprawa DNA ([Kidane et al. Crit. Rev. Biochem. Mol. Biol 49, 116, 2014](#)).

Hipoteza powstawania nowotworów wskutek mutacji nie jest obecnie powszechnie akceptowana i szereg biologów/lekarzy postuluje inne przyczyny transformacji neoplastycznej zdrowej komórki do nowotworowej:

- aberracje chromosomowe (nie wiadomo czy są przyczyną czy skutkiem nowotworu)
- zaburzenia procesów związanych z epigenetyką.

Modyfikacje (uszkodzenia) DNA dzieli się na **spontaniczne**, bez wykrywalnego czynnika inicjującego oraz **indukowane** przez czynniki chemiczne, fizyczne i biologiczne (wirusy).

(A) Chemiczne procesy **spontaniczne**:

- **dezaminacja** cytozyny do uracylu, który paruje z adeniną zamiast z guaniną,
- pękanie wiązania glikozydowego i **oderwanie zasady** (*abasic site*: 10^4 razy/genom i dzień): depurynacja, i ok. 20 razy rzadsza depirymidynacja;

(B) Chemiczne procesy **indukowane**, m. in. przez **czynniki kancero- i mutagenne**, które reagują z zasadami i modyfikują je do form o innym lub niejednoznacznym (tautomeria) typie parowania komplementarnego, ew. zaburzą strukturę helisy:

- alkilacja zasad, O⁶-alkilo-G, O⁶-benzylo-G;
- modyfikacja adeniny do N⁶-metoksyadeniny przez metoksyaminę; równowaga tautomeryczna N⁶OCH₃A - forma amino paruje z T, a forma imino paruje z C;

- związki wielopierścieniowe **interkalują** [PLANSZA 68] między pary zasad DNA zaburzając strukturę helisy: interkalatory „klasyczne”, przebijające (*threading*) z dużymi podstawnikami ulokowanymi w małej i dużej bruzdzie, bisinterkalatory (np. diterkalina) i multiinterkalatory;
- oksydacja przez **rodniki hydroksylowe i nadtlarki** (ROS; *reactive oxygen species*): przecięcie jednej nici, pęknięcie wiązania glikozydowego, cross-linkowanie DNA - białko, tworzenie ok. 50 typów adduktów (*base-oxidized lesions*), najczęściej 8-oxo-G;

(C) Procesy indukowane przez czynniki fizyczne:

- **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;
- **promieniowanie jonizujące**, przecięcie jednej, lub co szczególnie groźne dla integralności DNA, przecięcie obu nici DSB (*double strand break*).

(D) Zmiany w sekwencji wprowadza również **proces replikacji**, włączając do nowej nici nukleotydy niekomplementarne do nukleotydów matrycy. Jeśli *mismatches* nie zostaną skorygowane, zmiany są utrwalane w dalszych procesach replikacji i prowadzą do mutacji punktowych, do których obok tranzycji i transwersji zalicza się również mutacje „*frameshift*”, niewielkie delecje i addycje nukleotydów (zmiany niewykrywalne przez *Southern blotting* *). Bardziej rozległe (odcinkowe) zmiany sekwencyjne obejmują delecje, insercje, duplikacje i translokacje większych fragmentów DNA oraz modyfikacje dużych fragmentów chromosomów.

Zaburzenia struktury DNA, które mogą być przyczyną mutacji, są ustawicznie usuwane przez systemy **naprawy DNA**, które **obniżają częstości mutacji** *in vivo* do 10^{-8} (bakterie) ÷ 10^{-11} (eukariota) na nowo tworzoną parę zasad. **Mechanizmy naprawy DNA** obejmują szereg typów procesów odpowiadających wszystkim wyszczególnionym uszkodzeniom DNA [PLANSZA 69]**:

- **bezpośrednie przywrócenie** nieuszkodzonego DNA (*direct reversal*), np. enzymatyczna fotoreaktywacja dimeru pirymidynowego;

* A ***Southern blotting*** is a method routinely used in molecular biology for detection of a specific DNA sequence in DNA samples. Southern blotting combines transfer of electrophoresis-separated DNA fragments to a filtermembrane and subsequent fragment detection by probe hybridization.

** Czy wszystkim możliwym ???

- naprawa przez wycięcie uszkodzonego fragmentu nici **NER** (*nucleotide excision repair pathway*), w formie wielobiałkowego kompleksu (eukariota) zawierającego enzymy: endonukleaza, polimeraza DNA, ligaza;
- **MMR** (*mismatch repair pathway*); w szczególności rozpoznawane są niekomplementarne pary zasad **po replikacji** i usuwana jest jedna zasada z pary w nowo syntetyzowanej nici, a zachowana "właściwa" zasada na nici starej, dzięki znakowaniu pierwotnej nici poprzez metylację adenin w pozycji N⁶ (N⁶CH₃-dA) w sekwencjach dGATC;
- **BER** (*base excision repair pathway*) z udziałem glikozylaz specyficznych w stosunku do zmodyfikowanej zasady;
- **HR** (*homologous recombination*), usuwanie uszkodzeń DSB przecięcia dwu nici DNA, dominujący w prokariotach;
- **NHEJ** (*non-homologous end joining*), dominujący u kręgowców przy usuwaniu uszkodzeń DSB, rozpoznawanych przez heterodimer białkowy Ku70/Ku80.

Mechanizm BER wymaga lokalizacji, rozpoznania i wycięcia zmodyfikowanej zasady azotowej z bardzo szerokiej klasy produktów reakcji oksydacji, alkilacji i dezaminacji, przez specyficzne glikozylazy ([review: Brooks et al. Biochim. Biophys. Acta 1834, 247, 2013](#)). Wykryształowano i określono struktury dla szeregu kompleksów glikozylaz; przykładowo dla kompleksu ludzkiej glikozylazy uracyl-DNA (UDG) z 10 bp fragmentem DNA, zawierającym w centrum helisy *mismatch* dU:dG **[PLANSZA 70]**. Zakonserwowana leucyna 272 (w kompleksie zastąpiona arginina 272) wchodzi w helisę DNA przez małą bruzdę, ścisła łańcuch fosfocukrowy "flankujący" uracyl i powoduje „wywichnięcie” uracylu na zewnątrz (*flipping out*), gdzie jest stabilizowany przez jedno oddziaływanie stackingowe z Phe158 i 5 wiązań wodorowych (O4, N3H, O2) z łańcuchem głównym (Phe158, Gln144, Asn145) oraz z resztami bocznymi aminokwasów (Asn204, His268). Kompleks jest typu enzym - produkt, ponieważ nastąpiło pęknięcie wiązania glikozydowego i odsunięcie uracylu na odległość 15 Å. Odpowiedź na kluczowe (jak się wydaje) pytanie: **na czym polega mechanizm dyskryminacji zasady uszkodzonej (aberrant) czyli uracylu i zasady prawidłowej, czyli tyminy, które różnią się tylko obecnością jednej grupy metylowej?** została krótko podana w pracy: „*Tyr 147 packs against uracil C5, excluding solvent and selecting against thymine. The mutation Y147A severely reduces k_{cat}.*”

Szerokie możliwości badania mechanizmów naprawczych *in vivo* dają techniki SM (*single molecule*), np. obrazowanie fluorescencyjne znakowanych (PAmCherry) kompleksów białkowych UvrA i UvrB na wstępnym etapie NER w *E. coli* **[PLANSZA 69]**

([Stracy et al. Nature Commun. 7, 1, 2016](#)), w ekspozycjach 15 ms umożliwiającym śledzenie dynamiki kompleksów. Dzięki obrazowaniu PALM (*photoactivated localization microscopy*) cząsteczki lokalizowane są z rozdzielczością przestrzenną kilku nanometrów (patrz dalsza część wykładu: obrazowanie) i pozwalają scharakteryzować dwuetapowy proces napędzany hydrolizą ATP: skanowanie genomu i lokalizacja uszkodzenia przez UvrA, przyłączenie UvrB i dysocjacja UvrA.

Oprócz spełniania funkcji nośnika informacji genetycznej cząsteczki kwasu DNA znajdują zastosowanie w **architekturze wielowymiarowych struktur typu NANO** ([review: Endo & Sudyama ChemBiochem 10, 2420, 2009](#)) oraz niekonwencjonalnych układach elektronicznych. W „naturalny” sposób cząsteczki DNA wchodzi w skład układów magazynowania i przetwarzania informacji (genetycznej), a więc próbuje się je wykorzystać w niekonwencjonalnych systemach komputerowych (temat będzie poruszony podczas ostatniego wykładu). Charakterystycznym przykładem jest struktura **podwójnej helisy M-DNA**, w której zastąpienie jednego protonu wiązań wodorowych w każdej z komplementarnych par zasad przez jon metalu nadało cząsteczce własności drutu molekularnego przewodzącego prąd elektryczny (omówione na wykładzie „Struktura i funkcje makrocząsteczek biologicznych”)

DNA ma unikalne własności organizowania się w możliwe do zaprojektowania nanostruktury:

- (1) system komplementarnego parowania;
- (2) helikalna struktura z periodyczną geometrią ok. 10 bp na skręt;
- (3) możliwość programowania liniowych sekwencji tylko 4 nukleotydów z dalszymi modyfikacjami przez wprowadzanie grup funkcjonalnych, znaczników fluorescencyjnych, dołączania nanocząsteczek złota itp.;
- (4) chemiczna synteza zaprogramowanych sekwencji w zautomatyzowanych syntetyzerach, standardowo do ok. 40 n, ponadstandardowo nawet powyżej 100 n;
- (5) możliwość poddawania biochemicznej manipulacji przez enzymy: kinazy, polimerazy, ligaza, endonukleazy restrykcyjne.

Nanostruktury o rozmiarach 10 ÷ 100nm uzyskuje się techniką „**bottom-up**” przez łączenie (*self-assembly*) odpowiednio zaprojektowanych motywów. Struktury dwuwymiarowe 2D powstają z **bloków budulcowych DX** (*double-crossover tile*) [[PLANSZA 71](#)], zawierających **rozgałęzione złącza Holliday’a** i helisy z czterema jednoniciowymi zwisającymi końcami. Na tej zasadzie konstruuje się także bardziej skomplikowane

motywy: *triple-crossover tile*, *triangular tile*, *4x4 tile*, *four-helix bundled tube* i in. Struktury te można dalej rozbudowywać **[PLANSZA 72]** wprowadzając na DNA organiczne i nieorganiczne łączniki, **uzyskując**:

- **koniugaty** DNA - nanocząstki Au, koniugaty DNA - nanorurki węglowe oraz ich wzajemne połączenia;
- białka (streptawidyna) na macierzach 2D DNA;
- **trójwymiarowe** struktury DNA: nanorurki i wielowarstwowe (wypełnione) wielościany;
- struktury z **akumulacją jonów metali**, do produkcji nanourządzeń elektrycznych, magnetycznych i sensorów;
- struktury **origami**, złożone z rusztowania długich, pojedynczych nici 7,3 Kb (kilozasad) genomu bakteriofaga M13, połączonych z krótkimi, syntetycznymi fragmentami komplementarnymi w antyrównoległą macierz; origami (折り紙, *ori* "zwinięcie", *kami* "papier") jest tradycyjną, od XVII wieku japońską sztuką tworzenia rzeźb ze zwijanych, płaskich kartek papieru bez cięcia i użycia kleju.

Nowa generacja nanostruktur opartych na DNA idzie w kierunku rozwinięcia (wolnych od błędów) technologii „pojedynczej cząsteczki”, tworzenia coraz bardziej skomplikowanych rusztowań DNA i precyzyjnego umieszczania na nich innych cząsteczek i nanostruktur oraz praktycznego wykorzystania uzyskanych nanomateriałów: **nano-obwody elektryczne**, **nośniki magnetyczne** o ultrawysokiej gęstości, **detekcja** pojedynczych cząsteczek DNA i RNA. Niezwykle obiecujące wydaje się połączenie nanomateriałów DNA uzyskanych techniką „*bottom-up*” z materiałami wytworzonymi technikami „*top-down*”, takimi jak techniki nanolitograficzne.

PROTEOMIKA: zastosowania spektrometrii mas (MS); przestrzenne struktury białek: dyfrakcja rentgenowska i magnetyczny rezonans jądrowy NMR, dynamika ruchów molekularnych, interakcje

Wykład 10

Zsekwencjonowanie genomu ludzkiego i ponad 100 genomów innych organizmów zintensyfikowały badania się w kierunku wszechstronnej charakterystyki proteomu, czyli zespołu wszystkich białek kodowanych przez dany genom. PROTEOMIKA, nowa dziedzina, która powstała na gruncie **biologii i biofizyki molekularnej** i szeroko wykorzystująca zdobycze **bioinformatyki**, koncentruje się na **wysoce przepustowym** (*high throughput*) i **kompleksowym charakteryzowaniu proteomu** [PLANSZA 73].

(I) Identyfikacja białek wytwarzanych w określonej komórce, tkance czy całym organizmie, głównie przez wykorzystanie spektrometrii mas. W podejściu *bottom up* (aktualnie standardowym) startuje się z analizy fragmentów peptydowych po proteolizie (szczegóły dalsza część wykładu). W rozwijanym obecnie podejściu *top down* ([review: Ahlf et al. Curr. Op. Chem. Biol. 17, 787, 2013](#)), startuje się z całych białek po separacji technikami elektroforetycznymi (np. *tube gel electrophoresis*). Procedura przy obu podejściach obejmuje:

(a) **określanie struktury pierwszorzędowej**, z możliwie 100% pokryciem sekwencji;

(b) ilościowa charakterystyka białek ulegających ekspresji w komórce np. z wykorzystaniem metod znakowania białek ciężkimi izotopami i analizą MS: SILAC*, iTRAQ** i szereg innych ([Boja & Rodriguez Proteomics 12, 1093, 2012](#)), lub technik mikromacierzy białkowych (*protein microarrays*);

* SILAC *stable isotope labeling by amino acids in cell culture*: C-13/N-15 znakowane lizyny i argininy włączane do białek są odróżnialne w MS przez przesunięcie masy;

** iTRAQ, technika chemicznej modyfikacji białek, których fragmenty różnią się obecnością na widmie MS peptydowych jonów „reporterowych” o niskiej masie; stosunki intensywności sygnałów określają względne ilości peptydów w każdej próbce;

(c) charakterystyka modyfikacji post-translacyjnych (PTM, *post-translational modifications*): fosforylacja, glikozylacja i in.

(II) Określenie struktur przestrzennych cząsteczek z rozdzielczością atomową doświadczalnymi metodami **proteomiki strukturalnej**, niekiedy teoretycznymi (komputerowymi) metodami projektowania molekularnego:

- otrzymanie białek przez nadekspresję w bakteriach/systemach *cell-free*; ważny problem: *difficult-to-express proteins* (DTEPs; <https://www.pegsummit.com/preferences>)

(a) zwykle zawierającego **selenometioninę**, do eksperymentów **dyfrakcji promieni rentgenowskich** w wiązce synchrotronowej; metody: MAD i SAD (*multiplewavelength i singlewavelength anomalous dispersion*) do rozwiązania problemu fazowego, LUB

(b) **podwójnie znakowanego** $^{15}\text{N}/^{13}\text{C}$, a dla większych białek **potrójnie znakowanego** $^2\text{H}/^{15}\text{N}/^{13}\text{C}$, do eksperymentów **nD NMR**: 2D, 3D i 4D NMR (ew. $n > 4$ - wymiarowych);

- znalezienie białek o **znanych strukturach i odpowiednio wysokiej homologii sekwencji** do zwijania komputerowego (*data base search*); ew. zwijanie *de novo*;

- **krystalizacja białka** (dyfrakcja rentgenowska) LUB

uzyskanie **odpowiedniej rozpuszczalności** w wodzie i rozcieńczonych roztworach ciekłokrystalicznych, micelarnych (NMR);

- **rozwiązanie struktury** i budowa modelu cząsteczki;

- **weryfikacja** modelu.

(III) **Określenie współdziałania białek** (PPI, *protein-protein interactions network*) w ramach funkcjonalnych kompleksów molekularnych, czyli **charakterystyka interaktomu**; metody AP-MS (*affinity purification followed by mass spectrometry*), TAP w drożdżach (*tandem affinity purification*), metody bez stosowania MS (głównie białka błonowe): Tango***, F/BRET (*fluorescence/bioluminescence resonance energy transfer*), LUMIER (*luminescence based mammalian interactome mapping*) i szereg innych ([Lam & Stagljar Proteomics 12, 1519, 2012](#)).

*** Tango assay: do badania receptorów błonowych z białkami-ligandami.

W **spektrometrii masowej** lub **spektrometrii mas (MS)** ([review: Burlingame et al. Anal. Chem. 70, 647R, 1998](#)) cząsteczki przeprowadza się w stan gazowy z jednoczesną jonizacją (desorpcja) do ładunku Z i wstępnie przyśpiesza w polu elektrycznym o potencjale V . Po zogniskowaniu, wiązkę jonów poddaje się detekcji w celu uzyskania **wartości stosunku masy jonu do ładunku m/Z** , czyli rejestracji **widma MS** jedną z metod:

- FT-ICR (*Fourier transform ion cyclotron resonance*): pomiar cyklotronowej częstości rezonansowej jonu $\omega = (Z/m)B$ wzbudzonego oscylującym polem elektrycznym w pułapce magnetycznej Penniga (stałe pole magnetyczne o indukcji B z elektrycznymi płytami pułapkującymi);

TOF (*time-of-flight*): pomiar czasu przelotu jonu t na drodze d : $m/Z = t^2(2Ve/d^2)$

- QIT (*quadruple ion trap*); pułapkowanie jonów w (zmiennym) polu elektrycznym; potencjał naczynia o przekroju parabolicznym, którego pochylenie powoduje "wyciekanie" jonów w kolejności wzrastającego m/Z ([March J. Mass Spectrom. 32, 351, 1997](#)).

Technika MS jest **bardzo czuła**: wymaga jedynie nanogramów substancji i osiąga **dokładność** wyznaczania mas biomolekuł kilku Da, praktycznie **bez ograniczeń jeśli chodzi o wielkość badanych biomolekuł**. Podstawowe techniki **desorpcji** polegają na dyssypacji dużej części energii doprowadzonej, w celu odparowywania labilnych cząsteczek, umożliwiając przeprowadzenie ich w stan gazowy bez fragmentacji.

(A) FAB (*fast-atom bombardment*) naładowanych cząsteczek poniżej 1000 Da, w buforze zawieszonym w matrycy (np. glicerol) bombardowanej atomami/jonami (8 keV Xe; 20 keV Cs⁺), co powoduje parowanie jonów [M + H]⁺ z klastrami matrycy.

(B) MALDI (*matrix-assisted laser-desorption ionization*), najczęściej w wersji MALDI-TOF jest podobna do FAB ale bez ograniczeń na masę cząsteczkową. Energia światła laserowego jest pochłaniana przez matrycę małych cząsteczek organicznych, kokryształizowanych z cząsteczkami biopolimeru, które odparowują z fragmentami matrycy.

(C) ES (*electrospray ionization*) stosuje się dla szerokiego zakresu mas cząsteczkowych. Biomolekuły są w kropelkach rozpuszczalnika formowanych na końcu igły i ładowanych przez silne pole elektryczne. Podczas ruchu w kapilarze kropelki parują w przeciwnym strumieniu gazu uzyskując duży ładunek powierzchniowy, który je rozrywa wytwarzając spray drobniutkich i dalej szybko parujących kropelek. Pojedyncze cząsteczki [M + zH]^{z+} w kropelkach spreju przechodzą do fazy gazowej po odparowaniu reszty rozpuszczalnika.

Proteomiczna analiza białek metodą spektrometrii mas w ujęciu "bottom up" **[PLANSZA 74]** polega na wykonaniu szeregu kroków. Ekstrahowane z komórek (tkanek) białka poddaje się trawieniu trypsyną do peptydów, które ze względu na znaczne ilości (setki tysięcy) poddaje się procedurom „upraszczającym”: prefrakcjonacji (*sample prefractionation*) lub wzbogaceniu (*enrichment*). W pierwszej metodzie peptydy są sortowane ze względu na własności fizykochemiczne (ładunek, punkt izoelektryczny, itp.) metodami chromatografii jonowymiennej lub HILIC (*hydrophilic interaction liquid chromatography*). Alternatywnie, specyficzne klasy peptydów modyfikowanych (fosforylowane, dimetylowane, acetylowane) mogą być oddzielane metodą chromatografii powinowactwa (*affinity*) lub przez immunoprecipitację. Dalsza separacja zachodzi w systemie chromatografii cieczowej (UHPLC, *ultra-high performance liquid chromatography*) sprzężonym ze spektrometrem mas (LC-MS). **Identyfikacja peptydów**

odbywa się przez sekwencjonowanie w systemie tandem MS/MS: desorpcja, kontrolowana fragmentacja CID (*collision-induced dissociation*), analiza widm fragmentów różniących się o jeden aminokwas („b” - fragmenty od końca N, „y” - fragmenty od końca C) i określanie **struktury pierwszorzędowej** ([review: Seidler et al. Proteomics 10, 634, 2010](#)), dopasowanie do znanych sekwencji białek czyli **identyfikacja białek** przy pomocy algorytmów przeszukiwania *peptide-search* (Mascot, SEQUEST) **baz danych EST**. Biblioteki EST (*expressed sequence tags*) dzięki sekwencjonowaniu cDNA z mRNA przypadkowo klonowanych zawierają ponad 40 mln ESTs z 1300 organizmów. Celem ostatecznym jest **charakterystyka oddziaływań białko-białko i budowa odpowiednich sieci zależności**.

Ważnym elementem strategii analizy proteomu jest ilościowa charakterystyka czyli określenie **poziomu białek** w analizowanych komórkach. W tym celu rozwija się tzw. metody MS absolutnej kwantyfikacji AQUA (*absolute quantification approach*), np. wspomniane SILAC, iTRAQ: wykorzystanie peptydowych biomarkerów znakowanych izotopami/ chemicznie [\[PLANSZA 73\]](#) ([review: Pan et al. J. Proteome Res. 8, 787, 2009](#))

Pełna charakterystyka proteomu wymaga jednak określenia struktur przestrzennych cząsteczek białkowych metodami PROTEOMIKI STRUKTURALNEJ. **Celem zasadniczym** jest scharakteryzowanie, poprzez strukturę wszystkich białek znanych w wyniku zsekwencjonowania genomu ludzkiego i innych organizmów. Liczbę białek funkcjonujących w samym proteomie człowieka szacuje się na ok. 300 tys., kodowanych przez ok. 30 tys. genów **Celem bardziej realnym** wydaje się uzyskanie **reprezentatywnych struktur** w każdej rodzinie białek o określonym typie foldu. Niemniej, wymaga to rozwiązywania jednej struktury w ok. 2 ÷ 3 tygodnie. Zespoły naukowe biologów molekularnych, krystalografów, spektroskopistów NMR i informatyków (rozwój software'ów komputerowych) powołane do realizacji zadań w zakresie proteomiki strukturalnej 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 i firmami produkującymi aparaturę typu dyfraktometry, spektrometry NMR, spektrometry MS. Finansowanie odbywa się w ramach instytucji takich jak NIH (USA), Programów Ramowych UE. Znaczny udział w finansowaniu mają też firmy prywatne: farmaceutyczne, Celera Genomics (sekwencjująca genomy), firmy proteomiczne GeneProt (Genewa), MDS Proteomics (Toronto). **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 są:

- rentgenografia czyli dyfrakcja promieni X na monokryształach białek: ok. 81%;
- magnetyczny rezonans jądrowy NMR w roztworze: ok. 17%;
- modelowanie teoretyczne na podstawie podobieństwa sekwencji aminokwasowej z homologicznymi białkami o znanej strukturze: poniżej 2%;
- ostatnio również: mikroskopia elektronowa pojedynczej cząsteczki cryoEM.

Rentgenografia pozwala uzyskać struktury przestrzenne cząsteczek w monokryształach przez rozpraszanie promieniowania X i jest najszerzej stosowaną metodą w tym celu. Monokryształ molekularny **[PLANSZA 75]** tworzy regularną, symetryczną strukturę makroskopową wynikającą z symetrii mikroskopowej jednej lub kilku cząsteczek w komórce elementarnej. Wszystkie elementy symetrii można zebrać w jedną z 230 możliwych **przestrzenną grupę symetrii**, i scharakteryzować przez podanie symbolu grupy. Promieniowanie rentgenowskie o **długości fali λ rzędu odległości międzyatomowych** ($\sim 1\text{\AA}$) ulega **dyfrakcji na elektronach atomów** dając w wyniku obraz interferencyjny - **dyfraktogram**. Podstawowe pojęcia z zakresu **krystalografii** umożliwią zrozumienie zjawiska dyfrakcji i interpretacji dyfraktogramów.

(A) **Sieć przestrzenna**: nieskończony w porównaniu z rozmiarami kryształu ($0.1 \div 1 \text{ mm}$) zbiór punktów, tzw. węzłów sieci, z których każdy ma jednakowe otoczenie innych węzłów i powstaje w wyniku **translacji** w trzech kierunkach. W prostych kryształach, np. NaCl **[PLANSZA 76]** jony Na^+ i Cl^- zajmują węzły sieci regularnej (oraz inne położenia). W bardziej skomplikowanych kryształach molekularnych węzły są głównie punktami geometrycznymi. Atomy cząsteczek tworzą **motywy** strukturalne, regularnie powtarzane w trzech kierunkach w wyniku translacji. Każda płaszczyzna przez 3 węzły wyznacza rodzinę równoległych **płaszczyzn sieciowych**, oznaczonych symbolami Millera (hkl) , h , k , l - liczby całkowite, dodatnie lub ujemne $(\bar{h} \bar{k} \bar{l})$, pierwszymi względem siebie. Przykładowo, płaszczyzny ograniczające sześcian NaCl o jednostkowych długościach boków: yz (100), xz (010), xy (001). Proste równoległe przez dwa węzły wyznaczają kierunki krystalograficzne $[uvw]$; kierunek $[hkl]$ jest prostopadły do płaszczyzny (hkl) .

(B) **Komórka elementarna** **[PLANSZA 77]**: równoległościan oparty na 8 węzłach czyli, rozpięty przez trzy wektory **a**, **b**, **c** z kątami między nimi α , β , γ . Wybór jest oparty na dwóch kryteriach: najkrótsze długości a , b i c i symetria komórki, pozwalająca

zaklasyfikować ją jednoznacznie do jednego z 7 **układów krystalograficznych**: trójskośny, jednoskośny, rombowy, tetragonalny, trygonalny, heksagonalny i regularny. W ramach układów krystalograficznych istnieje 14 możliwych typów sieci tzw. **sieci Bravais**, spełniających warunki **jednakowego otoczenia każdego węzła i wypełnianie bez luk przestrzeni przez komórki elementarne**: sieć prosta P, centrowana na dwóch przeciwległych ścianach C, ściennie centrowana F i wewnętrznie centrowana I.

(C) Sieć odwrotna jest zbudowana na wektorach \mathbf{a}^* , \mathbf{b}^* , \mathbf{c}^* zadanych przez iloczynowy wektorowy wektorów \mathbf{a} , \mathbf{b} , \mathbf{c}

$$\mathbf{a}^* = (1/V)\mathbf{b} \times \mathbf{c}, \quad \mathbf{b}^* = (1/V)\mathbf{c} \times \mathbf{a}, \quad \mathbf{c}^* = (1/V)\mathbf{a} \times \mathbf{b} \quad V = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$$

Odległość między płaszczyznami sieciowymi (hkl) wynosi:

$$d = (h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*)^{-1}$$

(D) **Symetria punktowa sieci** jest opisana przez grupę elementów symetrii przecinających się w węźle (C_n , S_n , σ , I), jedną z **32 klas** dopuszczalnych przez warunek zgodności symetrii klasy z symetrią całej sieci, np. brak 5-krotnej osi symetrii i osi o krotności większej niż 6. Układ krystalograficzny wyznacza minimum symetrii sieci. Obok klasy **holoedrycznej** o maksymalnej symetrii należą do niego klasy o niższej symetrii, np. układ tetragonalny obok holoedrii $4/m\bar{m}m$ zawiera klasy $\bar{4}2m$, 422 , $4mm$, $4/m$, $\bar{4}$ i 4 (\bar{n} - oś inwersyjna S_n : złożenie obrotu C_n ($2\pi/n$) z płaszczyzną σ_h prostopadłą do osi obrotu).

(E) **Symetria kryształu** jest opisana przez jedną z 230 **grup przestrzennych** powstałych po złożeniu klas z translacjami. W grupach przestrzennych obok elementów symetrii punktowej i translacji występują ich złożenia: **osie śrubowe** X_k i płaszczyzny poślizgu. Podanie **symbolu grupy przestrzennej** jednoznacznie definiuje symetrię kryształu:

$P4_12_12$ sieć prosta P, układ tetragonalny, klasa 422

Cząsteczka lub jej fragment tworzy w komórce elementarnej **jednostkę asymetryczną**, która po przyłożeniu operacji symetrii klasy daje **motyw**, a translacja motywu w trzech kierunkach wyznacza **kryształ**. **Mikroskopowa struktura kryształu** jest scharakteryzowana przez:

(1) typ grupy przestrzennej

(2) parametry komórki elementarnej a , b , c , α , β , γ

(3) współrzędne N atomów cząsteczki: ułamki wektorów rozpinających komórkę:

$$\mathbf{r}_k = x_k\mathbf{a} + y_k\mathbf{b} + z_k\mathbf{c} \quad k = 1, 2, \dots, N \quad x_k, y_k, z_k \leq 1$$

Wykład 11

Rozpatrzmy prosty model dwóch centrów rozpraszających **[PLANSZA 78]** w położeniach zadanych przez wektor \mathbf{R} , odległych o $R = \|\mathbf{R}\|$. Wiązka padająca w kierunku wyznaczonym przez wektor \mathbf{S}_0 ulega **dyfrakcji** i promieniowanie **ugięte w kierunku \mathbf{S}** lub inaczej **odbite pod kątem 2θ** interferuje wskutek **różnicy faz φ** między obu promieniami:

$$\varphi = (2\pi/\lambda)(OA+OB) = (2\pi/\lambda)\mathbf{R} \cdot (\mathbf{S} - \mathbf{S}_0) = 2\pi \mathbf{R} \cdot \mathbf{H}$$

$$OA+OB - \text{różnica dróg optycznych promieni}, \quad \mathbf{H} = (1/\lambda)(\mathbf{S} - \mathbf{S}_0)$$

Warunek na wektor \mathbf{H} wyznacza maksima natężeń wiązki ugiętej:

$$\mathbf{R} \cdot \mathbf{H} = n \quad n - \text{liczba całkowita} \quad \Rightarrow \quad \varphi = 2\pi n$$

$$\|\mathbf{H}\| = (2\sin\theta)/\lambda$$

Wzmocnienie zachodzi w kierunku θ wyznaczonym przez przecięcie sfery o promieniu $1/\lambda$ i płaszczyzn dwuatomowej „sieci” odległych o $1/R$. Przy **prostej sieciowej** obsadzonej przez N identycznych atomów rozpraszających w położeniach $R_j = ja$ ($j = 0, 1, 2, 3, \dots$) promieniowanie ugięte daje wzmocnienie o amplitudzie A przy warunku $\mathbf{a} \cdot \mathbf{H} = n$:

$$A = f \sum_{j=1}^N \exp(i2\pi j \mathbf{a} \cdot \mathbf{H}) = Nf$$

Atomowy czynnik rozpraszania f jest proporcjonalny do liczby elektronów, a więc do liczby atomowej Z i wyraża się wzorem:

$$f = 2 \iiint_{\text{obj. atomu}} \rho(\mathbf{r}) \cos\{2\pi \mathbf{r} \cdot (\mathbf{S} - \mathbf{S}_0)\} d_3\mathbf{r} \quad \rho(\mathbf{r}) - \text{gęstość elektronowa w atomie}$$

i maleje ze wzrostem długości wektora $\mathbf{S} - \mathbf{S}_0$:

$$\|\mathbf{S} - \mathbf{S}_0\| = (2\sin\theta)/\lambda$$

Promieniowanie ugięte na **trójwymiarowym monokryształe**, czyli ulegające odbiciu rzędu n od płaszczyzn sieciowych (hkl) daje wzmocnienie interferencyjne **[PLANSZA 79]** w kierunkach określonych przez **warunki Lauego**:

$$\mathbf{a} \cdot \mathbf{H} = nh \quad \mathbf{b} \cdot \mathbf{H} = nk \quad \mathbf{c} \cdot \mathbf{H} = nl \quad n = 1, 2, 3 \text{ itd.},$$

równoważne warunkowi Bragga:

$$2d\sin\theta = n\lambda, \quad \theta - \text{kąt między wiązką ugiętą i płaszczyznami sieciowymi}$$

$$d - \text{odległość równoległych płaszczyzn sieciowych}$$

a warunki Lauego są spełnione jeśli wektor \mathbf{H} jest wektorem sieci odwrotnej:

$$\mathbf{H} = h'\mathbf{a}^* + k'\mathbf{b}^* + l'\mathbf{c}^* \quad h' = nh, \quad k' = nk, \quad l' = nl$$

Czyli kierunek wzmocnienia wiązki ugiętej na kryształach (rejestrowany prążek na dyfraktogramie) wyznacza **konstrukcja Ewalda** przecięcia węzła sieci odwrotnej ze sferą o promieniu $1/\lambda$. Względne natężenie prążków dyfrakcyjnych jest dane przez moduł **zespolonego czynnika strukturalnego** (oznaczony **boldem** jako wielkość zespolona, czyli wektor w przestrzeni dwuwymiarowej Re/Im):

$$\begin{aligned} \mathbf{F}(\mathbf{H}) = \mathbf{F}(hkl) &= C \cdot \sum_{j=1}^N f_j \exp\{2\pi i(hx_j + ky_j + zl_j)\} = \\ &= C \cdot V \cdot \iiint \rho(xyz) \exp\{2\pi i(hx + ky + lz)\} dx dy dz \end{aligned}$$

ρ - gęstość elektronowa w komórce elementarnej

$C = \exp(B_{11}h^2 + B_{12}hk + \dots + B_{33}l^2)$ jest czynnikiem temperaturowym, określonym przez średnie kwadraty amplitud drgań atomów, wyrażonym przez **współczynniki**

Debye'a-Wallera B_{ij}

Dyfraktogram **[PLANSZA 75]** można zinterpretować w następujący sposób:

(A) **Położenia prążków dyfrakcyjnych** (kierunki wzmocnienia) są uwarunkowane przez grupę symetrii kryształu i wymiary komórki elementarnej (wartości $a, b, c, \alpha, \beta, \gamma$);

(B) **Natężenia prążków dyfrakcyjnych** I są uwarunkowane, poprzez zespolony czynnik strukturalny F , przez położenia i rodzaje atomów w komórce elementarnej.

Z dyfraktogramu, mierząc intensywności poszczególnych prążków i wykonując odwrotną transformację Fouriera czynnika strukturalnego:

$$\begin{aligned} \rho(xyz) &= V^{-1} C^{-1} \sum \sum \sum_{hkl=-\infty}^{+\infty} \mathbf{F}(hkl) \exp\{-2\pi i(hx + ky + lz)\} = \\ &= 2V^{-1} C^{-1} \sum \sum \sum_{hkl=0}^{+\infty} \|\mathbf{F}(hkl)\| \cos\{2\pi(hx + ky + lz) - \Phi(hkl)\} \\ \|\mathbf{F}(hkl)\| &\sim \{I(hkl)\}^{1/2} \end{aligned}$$

można wyznaczyć **mapę gęstości elektronowej** $\rho(x, y, z)$ w komórce elementarnej **[PLANSZA 80]**, a stąd zlokalizować (wstawić) w odpowiednich miejscach poszczególne atomy i wiązania cząsteczki. Podstawowym problemem w uzyskaniu mapy gęstości elektronowej z dyfraktogramu jest tzw. **problem fazowy**. Wynika on stąd, że czynnik strukturalny, jest liczbą zespoloną:

$$F = a + bi = \sqrt{a^2 + b^2} (\cos\Phi + i \sin\Phi) \quad i = \sqrt{-1}.$$

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 ρ , trzeba niezależnie wyznaczyć **fazę** $\Phi = \arctg(a/b)$ **dla każdego prążka, czyli kilka tysięcy liczb z przedziału** $\Phi \in [0, 360]$. Istnieje szereg metod rozwiązania problemu fazowego.

Kluczowym pojęciem w dyskusji rozwiązań problemu fazowego jest **funkcja Pattersona** reprezentująca mapę wektorów międzyatomowych **[PLANSZA 81]**, która nie zależy od faz.

$$P(uvw) = \iiint \rho(xyz)\rho(x+u, y+v, z+w)dx dy dz = \\ = (1/V)\sum\sum\sum_{hkl} |F(hkl)|^2 \cos\{2\pi(hu + kv + lw)\}$$

W prostych **cząsteczkach z niewielką ilością N atomów** ich pozycje mogą więc być wyznaczone bezpośrednio z map Pattersona co rozwiązuje problem fazowy. W przypadku białek, np. dla $N = 5\ 000$ atomów ilość pików na mapie Pattersona wynosi $N(N - 1) = 25 \cdot 10^6$, co powoduje że przestaje być interpretowalna. **Opracowano szereg metod rozwiązania problemu fazowego w rentgenografii dużych biomolekul.**

(1) Podstawienie molekularne (*molecular replacement*), czyli zastosowanie znanych faz podobnych cząsteczek, które podobnie wykryły, a których struktura przestrzenna w kryształach została już wyznaczona. Ze względu na bardzo duże ilości rozwiązanych struktur zdeponowanych w bazach danych (**ponad 100 000**), zwykle efektywne rozwiązanie problemu fazowego zaczyna się od poszukiwania struktury białka homologicznego, zbliżonej do białka badanego. Procedura polega na właściwym ustawieniu cząsteczki białka badanego w komórce elementarnej czyli **nałożenia jej na znaną strukturę przez rotację z wykorzystaniem map Pattersona**, a następnie **translację**, najczęściej metodą prób i błędów.

(2) Wielokrotne podstawienie izomorficzne MIR (*multiple isomorphous replacement*), historycznie najstarsza metoda o uniwersalnym zastosowaniu (wszystkie rodzaje cząsteczek i dyfraktometrów), polega na: rejestracji i analizie dyfraktogramów od jednakowych (izomorficznych) kryształów zawierających cząsteczki o jednakowej konformacji, z których jeden zawiera dotatkowo ciężki atom (J lub większy) lub jon (Hg) w każdej komórce elementarnej. Ciężkie centrum bardzo silnie rozprasza promienie X i stąd można określić jego położenie w komórce jako punkt odniesienia. Średnie zmiany intensywności prążków $\langle \Delta I \rangle$ spowodowane obecnością ciężkiego centrum są rzędu:

$$\langle \Delta I \rangle / \langle I_P \rangle \approx (\langle I_H \rangle / \langle I_P \rangle)^{1/2}$$

$\langle I_H \rangle$ - średnia intensywność refleksów tylko ciężkiego centrum w komórce elementarnej

$\langle I_P \rangle$ - średnia intensywność refleksów samego białka

Konstrukcja Harkera [\[PLANSZA 82\]](#) oparta na **relacji Arganda** dla czynników strukturalnych: białka z ciężkim centrum F_{PH} , białka natywnego F_P i ciężkiego centrum F_H

$$F_{PH} = F_P + F_H$$

daje jednoznaczny fazę przy dwóch różnych izomorficznych podstawieniach (H_1 , H_2). Pozostaje tylko znaleźć położenie ciężkiego centrum w komórce elementarnej w celu określenia zespolonego czynnika F_H . Jest to względnie proste w grupach z tzw. „projekcjami centrosymetrycznymi”, np. $P2_12_12_1$, dla których F_{PH} i F_P są równoległe (różnica faz 0 lub 180deg) i mapa Pattersona dla **czynnika** $\|F_H\| = \|F_{PH}\| - \|F_P\|$:

$$P(uvw) = (1/V) \sum \sum \sum_{hkl} (\|F_{PH}(hkl)\| - \|F_P(hkl)\|)^2 \cos\{2\pi(hu + kv + lw)\}$$

daje poszukiwane współrzędne ciężkiego centrum. W przypadku grup niecentrosymetrycznych procedura jest bardziej skomplikowana, szczegółowo opisana w podręcznikach: [Blundel & Johnson „Protein Crystallography” Acad. Press 1976](#); [Drenth „Principles of X-ray Crystallography, Springer 1994](#).

(3) Rozpraszanie anomalne przy użyciu promieniowania synchrotronowego, zastosowane do białek zawierających **selenometioninę** (zamiast metioniny) jest metodą „proteomiczną”. Ciężki selen silniej rozprasza promieniowanie X niż siarka i wystarcza do zastosowania procedury. W metodzie z dyfrakcją w kilku długościach fali λ **MAD** (*multiplewavelength anomalous diffraction*) wystarczy wykonanie dyfrakcji dla jednego kryształu białka zawierającego selenometioninę, ale **dla kilku długości fali λ (zwykle od 3 do 5) synchrotronowego promieniowania X w pobliżu krawędzi absorpcji** [\[PLANSZA 83\]](#). Ścisłe ilościowe podejście do problemu fazowego metodą MAD przedstawia [podręcznik Drentha](#) oraz [Hendrickson Quart. Rev. Biophys 47, 49, 2014](#). **Idea** eksperymentu sprowadza się do analizy atomowych czynników rozpraszania: normalnego $f^0(\theta)$ (zwykle ciężkie atomy: C, N, O) i anomalnego $f^\Delta(\lambda)$ (Se), oraz wynikającej z nich całkowitej dyfrakcji ${}^\lambda F_T(hkl)$, złożonej z wkładów: **normalnego wszystkich atomów** (T): ${}^0 F_T(hkl)$, i **anomalnego** ze strony ciężkiego atomu A = Se, wyrażonego przez normalny czynnik rozpraszania tego atomu ${}^0 F_A(hkl)$: $[f^\Delta(\lambda)/f^0(\theta)] {}^0 F_A(hkl)$. Ze zbioru danych całkowitej

dyfrakcji $\|f^{\lambda}F_T(hkl)\|$ dla kilku λ wyznacza się $\|f^0F_T(hkl)\|$, $\|f^0F_A(hkl)\|$ oraz $\Delta\Phi = \Phi_T - \Phi_A$. Z wartości $\|f^0F_A(hkl)\|$ wyznacza się położenie selenu, a stąd Φ_A , co przy znanym (z fitowania) $\Delta\Phi$ daje ostatecznie Φ_T czyli rozwiązanie problemu fazowego. Metoda **SAD** (*singlewavelength anomalous diffraction*) jest elementem metody MAD (podobnie jak SIR jest częścią MIR), ale coraz częściej stosowaną jako "samodzielny" sposób wyznaczania faz ([Bunkoczi et al. Nature Methods 12, 127, 2015](#)), mimo niejednoznaczności trygonometrycznej równania Bijvoeta ([Hendrickson Quart. Rev. Biophys 47, 49, 2014](#)):

$$\|f^{\lambda}F_T(hkl)\| - \|f^{\lambda}F_T(\bar{h} \bar{k} \bar{l})\| = 4(f'/f^0)\|f^0F_A(hkl)\|\sin(\Phi_T - \Phi_A) \quad f' - \text{część urojona } f^{\lambda}(\lambda)$$

(4) Metody **bezpośrednie** (*direct methods*), na razie z ograniczonymi sukcesami, są oparte wyrafinowanych algorytmach, np. maksymalnej entropii (*maximum entropy*)
Rozwijane są zautomatyzowane procedury rozwiązywania problemu fazowego z danych MIR i MAD, np. *software SOLVE* ([Terwilliger & Berendzen Acta Cryst. D55, 849, 1999](#))

Procedura wyznaczania struktury w kryształach jest następująca.

(A) Wyhodowanie kryształu to zwykle nastawienie tysięcy próbek w celkach z „wiszącą (stojącą) kroplą” roztworu białka w buforze, nasączaną rozpuszczalnikiem organicznym, w automatycznych urządzeniach np. Mosquito [\[PLANSZA 84\]](#). Często niezbędne jest wykonanie mutacji białka, o ile nie wpływa to na konformację. Osobnym zagadnieniem jest **kryształizacja białek błonowych** w obecności związków powierzchniowo czynnych (*surfactants*) lub mezofazy lipidowej ([review: Caffrey Annu. Rev. Biophys 38, 29, 2009](#)).

(B) Wstawienie kryształu do dyfraktometru lub na wyjściu wiązki z akceleratora. Zwykle wykorzystuje się szybkie zamrożenie (witryfikacja) kryształu w strumieniu ciekłego azotu, co redukuje rozmycie prążków dyfrakcyjnych spowodowane czynnikiem temperaturowym.

(C) Naświetlanie promieniowaniem X i uzyskanie **dyfraktogramu** odpowiedniej jakości. Kryterium jest **rozdzielczość** w [Å], zdefiniowana jako minimalna odległość w rodzinie płaszczyzn sieciowych (hkl) dających mierzalny prążek. Obecnie uzyskuje się rozdzielczości $\sim 1\text{Å}$; ale $2 \div 3\text{Å}$ zapewnia dokładną lokalizację atomów ciężkich (bez H).

(D) Wyznaczenie **struktury wstępnej** z mapy gęstości elektronowej [\[PLANSZA 80\]](#), wyznaczonej z **intensywności prążków dyfrakcyjnych** i **rozwiązanego problemu fazowego**. Jeśli interpretacja fragmentu(ów) mapy gęstości elektronowej jest wątpliwa wyznacza się tzw. *omit map* tylko dla części cząsteczki o w pełni wiarygodnej strukturze, z pominięciem fragmentów nasuwających wątpliwości. Zdarza się, że dla pewnych fragmentów, np. słabo ustruktrowanych pętli białkowych, nie udaje się w ogóle

wyznaczyć map gęstości elektronowej ze względu na labilność konformacyjną. W **trudniejszych przypadkach** stosuje się mapy gęstości elektronowej dla różnych kombinacji modułów obserwowanych (F_o) i teoretycznych (F_c) czynników strukturalnych, np. $2F_o - F_c$.

(E) Iteracyjne **uzgadnianie dyfraktogramu doświadczalnego i liczonego teoretycznie**, aż do uzyskania wystarczającej zgodności opisanej przez tzw. **czynnik R**, który powinien być mniejszy od 20% ÷ 25%. R określa procentowe odstępstwo intensywności mierzonych z dyfraktogramu i liczonych teoretycznie dla danej struktury:

$$R = \frac{\sum_{hkl} (||\mathbf{F}(hkl)_o|| - ||\mathbf{F}(hkl)_c||)}{\sum_{hkl} (||\mathbf{F}(hkl)_o||)}$$

Czynnik ten zależy oczywiście od jakości kryształu (regularność, brak defektów) warunkującą jego zdolność rozdzielczą.

(F) **Weryfikacja uzyskanego modelu**. Kluczowym zagadnieniem dla rentgenografii jest **relacja między strukturą przestrzenną biomolekuły w kryształach i jej strukturą natywną w roztworze**, gdzie funkcjonuje. Kryształy białek zawierają (na szczęście) znaczne ilości nieustrukturowanej wody, nawet 40% ÷ 60% masy kryształu. Szereg kryteriów (*quality indicators*) omawianych szczegółowo w pracy [Kleywegt Acta Cryst. D56, 249, 2000](#) obejmuje m.in.

- wyznaczanie i analiza *omit maps*;
- zgodność wartości kątów Φ i Ψ z wartościami standardowymi na **wykreście Ramachandrana**;
- analiza sieci **wiązań wodorowych**, pamiętając o udziale ustrukturowanych cząsteczek wody, które wprowadza się zwykle przy konstrukcji mapy gęstości elektronowej;
- pomiary statystyczne **odchyleń r.m.s** (*root-mean square*) dla badanego białka i podobnych konformacyjnie fragmentów białek o bardzo dobrej rozdzielczości z baz danych (*pep-flip values*)

Zdarzają się jednak mimo wszystko spektakularne odstępstwa struktur krystalograficznych od struktur natywnych w roztworze, np. struktura G-kwadrupleksu d[AGGG(TTAGG)₃] jako modelu telomerowego DNA ([Li et al. Nucl. Acids res. 33, 4649, 2005](#))

Duże nadzieje jeśli chodzi o instrumentalny rozwój rentgenografii wiąże się z uzyskiwaniem **promieniowania X z laserów na swobodnych elektronach** (XFEL, *X-ray free-electron lasers*) w postaci niezwykle silnych, femtosekundowych (~50 fm) impulsów o intensywności 10^9 razy większej niż promieniowanie synchrotronowe. Stwarza to duże możliwości niezwykle szybkiej rejestracji dyfraktogramu zanim próbka ulegnie

uszkodzeniu (*diffraction-before-destruction*) [\[PLANSZA 85\]](#). Pierwszy X-ray laser LCLS (*Linac Coherent Light Source*) uruchomiono w 2009 r. w Stanford Linear Accelerator Center w Kalifornii, a potem pojawiły się doniesienia o szeregu następnych. Kilka polskich ośrodków naukowych, m. in. Instytut Badań Jądrowych w Świerku i Instytut Fizyki PAN i są zaangażowane w projekt "*European X-ray Free Electron Laser*" (EuXFEL) przy nadprzewodzącym liniowym akceleratorze DESY w Hamburgu. Urządzenia zostało oddane do dyspozycji użytkowników 2016 r.). Niezwykle cenną zaletą tej techniki jest uzyskiwanie analizowalnych dyfraktogramów dla **bardzo małych kryształów** o wymiarach mikronów (μm). Wymaga to uzyskania obrazu dyfrakcyjnego od dużej liczby mikrokryształów o jednakowych/prównywalnych rozmiarach (nowe techniki krystalizacji) w wiązkę laserową z częstotścią MHz pomiędzy impulsami, np. technika SFX (*serial femtosecond crystallography*) zastosowana w EuXFEL ([Gisriel et. Nature Commun. 10, 1, 2019](#)). Analizę danych i problem fazowy można rozwiązywać standardowymi metodami (podstawienie molekularne, MAD), ale opracowano także nowe metody, które dają się stosować w przypadku bardzo małych kryształów ([Ayyer et al. Nature 530, 202, 2016](#)).

Duże nadzieje wiąże się również z uzyskiwaniem obrazów dyfrakcyjnych pojedynczych dużych kompleksów molekularnych CDI (*coherent diffraction imaging*).

Wykład 12

Do wykonywania obrazów dyfrakcyjnych można oprócz promieniowania rentgenowskiego używać również **wiązki neutronów** (neutronografia) oraz **wiązki elektronów** (krystalografia elektronowa). Neutronografia ([Barkley et al. Curr. Op. Struct. Biol. 18, 593, 2008](#)) jest niezwykle użyteczną metodą w lokalizacji atomów wodoru w strukturach makromolekuł (których nie można wyznaczyć z rentgenografii), ze względu na duże czynniki rozpraszania neutronów przez atomy ^1H , a jeszcze większe przez atomy ^2H (deutery). Neutronografia jest więc metodą komplementarną pod tym względem do dyfrakcji rentgenowskiej, aczkolwiek ilość struktur rozwiązana tą techniką jest mała ze względu na bardzo ostre wymogi odnośnie wielkości i jakości kryształów. Metodę dyfrakcji elektronowej ([Rausner & Walz Annu. Rev. Biophys. 38, 89, 2009](#)) stosuje się do **kryształów dwuwymiarowych**, standardowo do kryształów 2D białek **blonowych**, które powstają w roztworach białek rozpuszczonych w detergencie zmieszany z lipidami, po odparowaniu detergentu. Najpoważniejszym problemem przy zbieraniu obrazu dyfrakcyjnego jest niszczenie próbki przez wiązkę elektronową.

Struktury białek **globularnych** i **blonowych** uzyskiwane głównie na podstawie eksperymentów dyfrakcyjnych (ponad 80% struktur zdeponowanych w bazach danych PDB ([Berman et al. Nucl. Acids Res. 28, 235, 2000](#); <https://www.rcsb.org>) próbuje się **klasyfikować pod względem topologicznym** definiując różne typy trzeciorzędowych foldów domen. W chwili obecnej szacuje się że domeny białkowe są zorganizowane w ponad 700 typów foldów. Klasyfikacja ułatwia zrozumienie struktur 3D białek oraz pozwala bardziej precyzyjnie i ilościowo charakteryzować podobieństwa i różnice strukturalne między cząsteczkami białek na użytek **badania zależności ewolucyjnych** i wyszukiwania **struktur homologicznych** w badaniach **rentgenograficznych** lub **modelowaniu** struktur na podstawie homologii sekwencji. Klasyfikacje opierają się zasadniczo na **podejściu hierarchicznym**:

Baza danych SCOP (*structural classification of protein database*; [Murzin et al. J. Mol. Biol. 247, 536, 1995](#)) organizuje domeny w rodziny (*family*) na podstawie kryteriów homologii sekwencji (ponad 30%) i/lub podobieństwa funkcji przy homologii $> 15\%$, następnie w superrodziny (*superfamily*) prawdopodobnego, wspólnego pochodzenia ewolucyjnego (przy niskiej homologii), oraz wspólne foldy o zbliżonych strukturach drugorzędowych, podobnie powiązanych ze sobą pod względem topologicznym.

Baza CATH z hierarchicznym podziałem białek [**PLANSZA 86**] na klasy (C; *class*) pod względem struktury drugorzędowej: „ α ”, „ β ”, „ α & β ”, na architekturę (A; *architecture*) pod względem ogólnych cech kształtu foldu bez zróżnicowania połączeń struktur drugorzędowych, na rodziny topologiczne (fold families T; *topology*) domen o podobnej ilości, ustawieniu i połączeniach elementów drugorzędowych. Najwyższy stopień podobieństwa strukturalno-funkcjonalnego grupuje białka w superrodziny na poziomie H (*homology level*).

Na drugim „biegunie strukturalnym” występuje znaczna część białek, ponad 30% w eukariotach i ok. 6% w prokariotach, o strukturze nieuporządkowanej (liczne [reviews](#) np.: [Uversky Int. J. Biochem. Cell Bio. 43, 1090, 2011](#); [Babu et al. Curr. Op. Struct. Biol. 21, 432, 2011](#)). Są to białka **całkowicie rozwinięte**, bez struktury drugo- i trzeciorzędowej, w formie kłębka statystycznego czyli zbioru konformacji w równowadze dynamicznej (*random coil*) lub bardziej uporządkowane, w formie stopionej globuły (*molten globule*, *premolten globule*), oraz białka **częściowo rozwinięte** na odcinkach łańcucha dłuższych niż kolejne 30 AA. Z wielu funkcjonujących nazw (*natively unfolded*, *intrinsically unstructured* i innych), ***intrinsically disordered proteins (IDPs)*** jest obecnie najszerzej używana. Białka IDP przyjmują strukturę zwiniętą oddziałując z ligandem - tarczą (*target*) (*folding upon binding*): kwas nukleinowy, jon metalu, błona lipidowa, inne białko. IDPs można badać metodami rentgenografii tylko w formie uporządkowanej, po związaniu w kompleks. Do badań form nieuporządkowanych stosuje się metody spektroskopowe (CD, FT-IR, ROA), które dają informacje o stopniu nieuporządkowania. Natomiast SAXS, DLS (*dynamic light scattering*), ultrawiarowanie (sedymentacja) określają promień hydrodynamiczny cząsteczki (*radius of gyration*). **Badania strukturalne** można prowadzić przy użyciu wysoce wymiarowego NMR (5D i 6D) ze względu na sukcesy w przyporządkowywaniu w cząsteczce rozwiniętej. W **dynamice** znajduje zastosowanie m. in. technika wymiany H/D z detekcją MS ([Balasubramanian & Komives Biochim. Biophys. Acta 1834, 1202, 2013](#)). Brak lub częściowy brak zwiniętej struktury przypisuje się zmniejszonej, w porównaniu z białkami ustrukturowanymi, ilości tzw. aminokwasów sprzyjających uporządkowaniu (*order-promoting*): Ile, Leu, Val, Trp, Phe, Cys, Asn, oraz zwiększonej ilości AA sprzyjających nieuporządkowaniu (*disorder-promoting*): Ala, Arg, Gly, Gln, Ser, Glu, Lys, Pro. Drugim czynnikiem jest silne hydratowanie cząsteczek IDPs i wiązanie jonów. Kluczową cechą biologiczną białek IDP są **procesy sygnałowe i funkcje regulacyjne**, co wiąże się z **koniecznością oddziaływania z kilkoma różnymi partnerami**, np. dwie (z czterech), nieustrukturowane domeny białka p53 (*tumor*

suppressor): N-terminalna (aktywacja translacji) i C-terminalna (regulacyjna) **[PLANSZA 87]**. Środkowy wykres na **PLANSZY** ilustruje oszacowane prawdopodobieństwa nieuporządkowania wzdłuż sekwencji - wartości w zakresie 0,5 ÷ 1,0. Nieuporządkowanie struktury białek IDP daje **szereg korzyści**:

- większą powierzchnię kontaktu niż w przypadku białka ustrukturuwanego;
- „giętkość” konformacyjną ułatwiającą dopasowanie do kilku partnerów;
- względnie wysoką specyficzność wiązania przy relatywnie niskim powinowactwie.

Rentgenografia ma niewątpliwie wiodącą rolę w proteomice strukturalnej, a w szczególności pozwala wyznaczać struktury bardzo dużych układów molekularnych przechodzących w struktury subkomórkowe (rybosom, spliceosom). Jednak podobnie jak w przypadku białek natywnie rozwiniętych, metody rentgenograficzne nie dają możliwości wyznaczania struktur przestrzennych nierozpuszczalnych białek włóknistych ani źle zwiniętych form amyloidalnych białek globularnych, ze względu na brak możliwości uzyskania kryształów. Brak rozpuszczalności stwarza również problemy dla zastosowania spektroskopii magnetycznego rezonansu jądrowego w roztworze, jednak **techniki NMR można również stosować do wyznaczania struktur cząsteczek w fazie stałej, niekryształicznej (amorficznej).** Magnetyczny rezonans jądrowy (NMR - *nuclear magnetic resonance*) **[PLANSZA 88]** jest metodą spektroskopową, w której obserwuje się przejścia pomiędzy poziomami energii jąder o niezerowym spinie $I \neq 0$ cząsteczek w silnym, zewnętrznym, stałym polu magnetycznym o indukcji B_0 od 0.24 do 21 T [tesli], co odpowiada promieniowaniu fal radiowych, 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 (oraz gazach) obserwuje się **widma wysokiej zdolności rozdzielczej (*high resolution*)** w postaci szeregu wąskich linii o szerokościach od dziesiątych części Hz dla małych cząsteczek organicznych do kilkudziesięciu Hz dla biopolimerów. Widma **wysokiej zdolności rozdzielczej w ciele stałym** uzyskuje się **techniką MAS. (*magic angle spinning*)** przez wirowanie próbki stałej pod „kątem magicznym” $\theta = 54,7\text{deg}$ do kierunku indukcji pola B_0 z szybkością, standardowo $\sim 10^4$ obrotów na sekundę a ponadstandardowo $\sim 10^5$, co daje efekt podobny do uśredniania oddziaływań warunkujących szerokie linie rezonansowe wskutek dyfuzji rotacyjnej w roztworze:

$$\langle 3\cos^2\theta - 1 \rangle = 0.$$

Podstawowe relacje w analizie widm NMR prowadzące do ich szerokiego zastosowania w biofizyce molekularnej sprowadzają się do możliwości przyporządkowania pojedynczych sygnałów lub multipletów na widmach poszczególnym jądrom lub grupom jąder równoważnych (chemicznie lub magnetycznie) w cząsteczce.

(A) Częstości rezonansowe poszczególnych jąder w cząsteczce różnią się w zakresie od kilku do kilkudziesięciu kHz wskutek przesłaniania (ekranowania) zewnętrznego pola magnetycznego B_0 przez elektrony. Położenia sygnałów są mierzone jako **przesunięcia chemiczne** $\delta(j)$ w jednostkach ppm (*parts per million*) względem wybranego wzorca.

(B) Sygnały rezonansowe mogą wykazywać strukturę multipletową wskutek rozszczepień wywołanych **sprężeniami skalarnymi** $J(j,k)$ mierzonymi w Hz między jądrami j oraz k (takimi samymi i innego typu) poprzez elektrony wiązań.

(C) Odpowiednio zaprojektowane eksperymenty NMR pozwalają obok wymienionych parametrów $\delta(j)$ i $J(j,k)$ rejestrować szereg dalszych **parametrów związanych ze strukturą i dynamiką** ruchów molekularnych cząsteczek:

- intensywności a_{ij} bezpośredniego oddziaływania dipol-dipol (NOE; *nuclear Overhauser effect*) na widmach NOESY,
- reszkowe sprzężenia dipolowe RDC (*residual dipolar couplings*) oznaczane D_{ij} , w cząsteczkach częściowo zorientowanych w roztworach zawierających micelle lipidowe,
- czasy relaksacji spinowo-sieciowej T_1 i spinowo-spinowej T_2 ,
- transfer protonów między różnymi miejscami w cząsteczce oraz przekształcenia konformacyjne cząsteczek (wymiana chemiczna).

Projektowanie eksperymentów NMR do wyznaczenia potrzebnych parametrów wymaga teoretycznej analizy **ewolucji układu jąder cząsteczki** pod wpływem działania sekwencji impulsów pola magnetycznego B_1 wzbudzającego układ, oddzielonych zmiennymi i stałymi odstępami czasowymi. Wszystkie informacje o jądrowym, **makroskopowym układzie N cząsteczek**, w aktualnie omawianym przypadku cząsteczek w roztworze, są zawarte w **operatorze macierzy gęstości σ układu jąder** [PLANSZA 89] spełniającym równanie Liouville'a-von Neumanna (tu bez uwzględnienia relaksacji). Rozwiązując równanie w kolejnych okresach czasu τ_1, τ_2, \dots : trwania impulsów B_1 (t_p), inkrementowania i rejestracji w eksperymentach wielowymiarowych (t_1, t_2), stałych okresach *mixing* (t_m),

gdy działają hamiltoniany oddziaływań jąder z polem B_0 (przesunięcia chemiczne) i sprzężeń skalarnych, uzyskuje się końcową wartość σ po całym okresie ewolucji $\tau_1 + \tau_2 + \tau_3 + \dots$. Z tej wartości σ wylicza się teoretycznie widmo NMR korzystając z transformacji Fouriera. Interpretacja fizyczna macierzy σ na przykładzie układu złożonego z pojedynczego jądra w cząsteczce jest stosunkowo prosta:

- wyrazy diagonalne oznaczają populacje czyli **boltzmannowskie obsadzenia poziomów** jądra w polu B_0 ;

- niezerowe wyrazy pozadiagonalne pojawiają się w momencie przejść rezonansowych pod wpływem impulsów pola B_1 i oznaczają **koherencje fazowe** jąder próbki, które wykonują przejścia rezonansowe; „pogrubiona kreska” oznacza uśrednienie po wszystkich, w tym przypadku jednakowych jądrach, czyli po wszystkich cząsteczkach próbki.

Niewygodną stroną takiego ujęcia teorii ruchu jąder, w przeciwieństwie do pogładowego modelu wektorowego VM (*vector model*, omawiany na wykładzie [Spektroskopia molekularna](#)) jest brak pogładowej interpretacji wyrazów macierzy σ czyli przypisania im konkretnych efektów w postaci rejestrowanych widm. Wyjściem z sytuacji stał się pewien szczególny **sposób rozwinięcia wyrazów macierzy na operatory bazowe** w postaci iloczynów spinów pojedynczych jąder (*product operator formalism*) I_{kv} , $k = 1, 2, 3, \dots$ jądra, $v = x, y, z$. Dla układu dwóch jąder rozwinięcie jest przedstawione na **PLANSZY**. Operatory tzw. **observable** **[PLANSZA 90]**, na końcu okresu ewolucji w czasie rejestracji, prowadzą do powstawania odpowiedniego **układu linii widmowych**. Ten sposób podejścia zapewnia **ścisłość** kwantowej fizyki statystycznej przy zachowaniu **pogladowości** modelu wektorowego.

Wyznaczanie struktur przestrzennych białek w roztworze metodami NMR przy wykorzystaniu jąder o spinie $I = \frac{1}{2}$ (^1H , ^{13}C , ^{15}N), oraz przy **znanej strukturze pierwszorzędowej** przebiega według następującego schematu **[PLANSZA 91]**:

(1) Wykonanie **przyporządkowania sygnałów NMR** poszczególnym jądrom cząsteczki białka na podstawie unikalnych układów spinowych poszczególnych rodzajów aminokwasów i procedury sekwencyjnej analizy oddziaływań skalarnych i ew. dipol-dipol jąder na sąsiednich monomerach wzdłuż łańcucha:

- widma 2D ^1H NMR **[PLANSZA 92]** dla cząsteczek **do ok. 10 kDa**;

- widma 3D i 4D NMR jąder ^1H , ^{13}C , ^{15}N NMR **[PLANSZA 93]** cząsteczek wzbogaconych podwójnie (^{13}C , ^{15}N) i/lub potrójnie (^{13}C , ^{15}N , ^2H) **do ok. 40 kDa**; wykorzystanie separacji sygnałów wzdłuż osi heterojądra (^{13}C lub ^{15}N) **[PLANSZA 94]**

- 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**.

(2) Wyznaczenie parametrów opisujących strukturę przestrzenną cząsteczki białka na podstawie ich zależności od parametrów mierzonych na widmach NMR **[PLANSZA 91]**:

- **odległości proton-proton** r_{ij} , a ostatnio również odległości między innymi jądrami (proton, węgiel, azot), do ok. 6Å (w przypadku jąder ^{19}F do 8Å), z intensywność oddziaływania dipol-dipol:

$$a_{ij} \sim r_{ij}^{-6}$$

przy założeniu odpowiedniego modelu dynamiki ruchów molekularnych i wyznaczeniu na jego podstawie **czasów korelacji** τ_c z pomiarów czasów relaksacji T_1 i T_2 (patrz dalej);

- **kąty dwuścienne** θ_{ij} wokół wiązań z wartości stałych sprzężenia skalarnego dwóch jąder odległych o trzy wiązania (relacja Karplusa):

$${}^3J(i,j) = A\cos^2\theta_{ij} + B\cos\theta_{ij} + C \quad \begin{array}{l} A, B, C \text{ - stałe zależne od typu fragmentu} \\ \text{chemicznego i rodzaju sprzęgających się jąder;} \end{array}$$

- rejestracja sprzężeń skalarnych przez **wiązania wodorowe** ([*Pervushin et al. Proc. Natl. Acad. Sci. USA 95, 14147, 1998*](#)) $\text{N} - \text{H} \dots \text{O} = \text{C}'$:

$${}^hJ_{\text{NC}'} \sim 0.1 \div 0,9 \text{ Hz}$$

- **przynależność aminokwasu do struktury drugorzędowej** (wartości kątów Φ i Ψ) na podstawie porównania przesunięć chemicznych jąder łańcucha: $\text{C}\alpha$, $\text{H}\alpha$, $\text{C}\beta$, N (amidowy), C' (karbonylowy) trzech aminokwasów: środkowy i 2 sąsiadujące, z wartościami w bazie danych CSI (*chemical shift index*) białek o rozwiązanych strukturach, np. wykorzystując program TALOS ([*Cornilescu et al. J. Biomol. NMR 13, 289, 1999*](#));

- **wzajemna orientacja fragmentów** (struktur drugorzędowych) **cząsteczek** na podstawie resztkowych sprzężeń dipolarnych (RDC, *residual dipolar coupling*) D_{ij} cząsteczek częściowo uporządkowanych przez lipidowe micelle w roztworze,

(3) Przejście od odległości, kątów i innych informacji przestrzennych (długości wiązań, kąty płaskie) **do struktury, a raczej zbioru nałożonych struktur o zbliżonych energiach** [PLANSZA 95]. Współrzędne atomowe wyznacza się w procedurze udokładniania (*refinement*) metodą dynamiki molekularnej z wiązaniami doświadczalnymi NMR **rMD** (*restrained molecular dynamics*) w wersji „*simulated annealing*”. **Pseudopotencjały** [PLANSZA 96] nałożone na parametry geometryczne „ściąają” wyjściową strukturę do struktury zgodnej z danymi NMR. Wcześniej wykorzystywane w tym celu metody czysto geometryczne tzw. geometria odległościowa DG (*distance geometry*) nie są już stosowane. Przedstawienie w formie nałożonych struktur wynika z **niepewności wartości** doświadczalnych parametrów (odległości i kątów), co odpowiada **dynamicznemu charakterowi** cząsteczki.

(4) **Weryfikacja struktury**: metody analogiczne jak w rentgenografii: zgodność wartości kątów Φ i Ψ z wartościami standardowymi na wykresie Ramachandrana; analiza sieci wiązań wodorowych, pomiary statystyczne odchyień r.m.s dla badanego białka i podobnych konformacyjnie fragmentów białek o bardzo dobrej rozdzielczości z baz danych.

Pierwsze dwa zagadnienia (1) i częściowo (2) są ze sobą ściśle powiązane, np. przyporządkowanie oznacza jednocześnie wyznaczenia wartości $\delta(i)$ poszczególnych jąder. Przyporządkowanie jest warunkiem koniecznym ALE niewystarczającym do wyznaczenia wszystkich potrzebnych parametrów NMR.

Wykład 13

Zagadnienia **przyporządkowania sygnałów NMR jądrom cząsteczki** i wyznaczenie **parametrów opisujących strukturę przestrzenną cząsteczki** komplikują się w miarę wzrostu masy cząsteczkowej ze względu na dwa zasadnicze czynniki.

(1) Wzrost ilości jąder i związane z nim **nakrywanie się sygnałów**.

(2) **Poszerzenie linii rezonansowych** wskutek zmniejszania się wartości T_2 czyli spadek ruchliwości molekularnej dużych cząsteczek (długie czasy korelacji τ_c).

O ile w prostych cząsteczkach organicznych można wykonać przyporządkowanie na podstawie widm jednowymiarowych z zakresów przesunięć chemicznych, struktur rozszczepień skalarnych poszczególnych sygnałów i ich intensywności, o tyle nawet w małych białkach długości ~ 50 AA wykonanie takiego przyporządkowania nie jest możliwe. **Widma protonowe** mają tu podstawowe znaczenie ze względu na możliwość uzyskania największej ilości informacji strukturalnych, z wycinalnych stałych sprzężenia 3J proton–proton i efektu Overhausera dla oddziaływań dipolowych (DD) proton–proton, przy jednoczesnej dużej czułości spektroskopii protonowej (100% abundancja 1H , duży moment magnetyczny), co redukuje stężenia substancji i nie wymaga znakowania związków jądrami ^{13}C czy ^{15}N . Dla małych **białek do ok. 100 aminokwasów** (masy cząsteczkowe do ok. 10 kDa) przyporządkowanie widm 1H NMR (powyżej 80%) i potrzebne stąd parametry do określania struktury uzyskuje się na podstawie dwóch podstawowych **eksperymentów dwuwymiarowych (2D NMR): COSY (CORrelated SpectroscopY)** i/lub **TOCSY (TOTAL Correlated SpectroscopY)** oraz **NOESY (Nuclear Overhauser Effect SpectroscopY)** [\[PLANSZA 92\]](#). Widma dwuwymiarowe homojądrowe mają na dwóch osiach przesunięcia chemiczne 1H a na trzeciej osi intensywności sygnałów. Zwykle przedstawia się je w postaci dwuwymiarowych przekrojów „poziomicowych”. W homojądrowych (protonowych) widmach 2D NMR obserwuje się na przekątnej sygnały typu widma jednowymiarowego. **Pozadiagonalne „cross-peaks”, opisane przez dwa przesunięcia chemiczne**, powstają dla takich par protonów które oddziałują ze sobą w określony sposób, uwarunkowany typem eksperymentu. W eksperymencie COSY są to pary protonów oddziałujących skalarnie, czyli wykazujących niezerowe sprzężenie proton–proton; a w eksperymencie NOESY są to pary protonowe oddziałujące dipolarnie, czyli odległe od siebie w przestrzeni do ok. 6 Å. W eksperymencie TOCSY uzyskuje się pozadiagonalne sygnały od całego układu sprzężonych protonów

cząsteczki, jeśli sprzęgają się kolejno parami, np. dla całego łańcucha bocznego aminokwasu z protonami przy kolejnych atomach węgla.

Sekwencyjne przyporządkowanie w łańcuchu peptydowym na podstawie eksperymentów 2D polega na identyfikacji charakterystycznych, unikalnych układów sygnałów protonów poszczególnych aminokwasów, a następnie ich „łączeniu krok po kroku” wzdłuż łańcucha o znanej sekwencji, wykorzystując widma NOESY i COSY zestawione razem nad i pod diagonalą **[PLANSZA 97]**. Zakładając *cross-peak* startowy jako należący do danego AA ("start" w trakcie procedury będzie zweryfikowany), identyfikuje się pary $\text{NH}/\text{C}\alpha\text{H}$ wewnątrz aminokwasu na widmie COSY oraz pary $\text{NH}/\text{C}\alpha\text{H}$ (odległość d_1), NH/NH (odległość d_2) i $\text{NH}/\text{C}\beta\text{H}$ (odległość d_3) danego aminokwasu i aminokwasu poprzedzającego w sekwencji na widmach NOESY, zależnie od struktury drugorzędowej. Prowadząc tę tzw. *spider (snail) game* **przyporządkowuje się kolejno sygnały protonom fragmentu łańcucha z jednoczesnym określeniem struktury drugorzędowej** tego fragmentu (BPTI: Asn24→Ile19). Ponieważ widma białek rejestruje się w H_2O z tłumieniem silnego sygnału wody, a białka zawierają także pętle (inne relacje odległości proton-proton), część potrzebnych sygnałów znika i „*spider game*” trzeba prowadzić szereg razy dla różnych fragmentów. Protony łańcuchów bocznych identyfikuje się przy znanych przesunięciach protonów $\text{C}\alpha\text{H}$ i $\text{C}\beta\text{H}$ na widmach COSY lub TOCSY.

Przy wzroście masy cząsteczkowej białek pojawia się problem nakładania się sygnałów wzrastającej liczby protonów na widmach 2D. **Rozszerzenie korelacji międzyjądrowych przez wiązania na jądra różne od protonów, tzw. heterojądra:** ^{13}C , ^{15}N w trzecim (i ew. czwartym wymiarze) **[PLANSZA 93]** likwiduje chwilowo ten problem **[PLANSZA 94]**. Przykładowo, nakrywające się na widmie 2D NOESY sygnały $\text{NH}/\text{C}\alpha\text{H}$ dwóch aminokwasów ulegają rozseparowaniu w trzecim wymiarze osi azotu. Prawdopodobieństwo jednakowych wartości przesunięć dwóch protonów NH , dwóch protonów $\text{C}\alpha\text{H}$ i dwóch azotów NH jest już dużo mniejsze. Podstawowym mankamentem technik heterojądrowych 3D (4D) NMR jest wymóg białek **podwójnie znakowanych $^{13}\text{C}/^{15}\text{N}$** . Ponadto, zachodzi tzw. **edytowanie**, czyli ograniczenie ilości obserwowanych jąder (w porównaniu z widmami 1D), gdyż „*cross-peaks*” pochodzą tylko od par proton - heterojądro, które są bezpośrednio ze sobą połączone wiązaniem chemicznym (lub w ogóle sprzęgają się ze sobą). Zmianie ulega **procedura sekwencyjnego przyporządkowywania w łańcuchu peptydowym [PLANSZA 98]**. Wykorzystuje się **sprzężenia skalarne przez jedno wiązanie wszystkich jąder wzdłuż łańcucha peptydowego i kolejno identyfikuje**

sprzęgające się skalarnie węgle, azoty i protony. Opracowano szereg metod tzw. "*triple resonance*", których nazwy bezpośrednio określają typy trzech korelowanych jąder, np. w metodzie **HNCA** uzyskuje się trójwymiarowe *cross-peaks* od protonu amidowego **NH**, związanego z nim azotu amidowego **N** i węgla **C α** . **Minimum 5 eksperymentów** potrzeba do przyporządkowania jąder łańcucha, przy czym często stosuje się ich modyfikacje i inne typy korelacji, tak że w literaturze prezentowanych jest **znacznie więcej dróg przyporządkowania**, chociaż są one oparte na jednym schemacie, który dla „przejścia” wzdłuż nici peptydowej 148 AA kalmoduliny (16,7 kDa), **od Lys21 do Asp22** wygląda następująco. Punktem wyjścia jest znajomość dwóch przesunięć chemicznych: Asp20-¹³C(=O) (177,3 ppm) i Lys21-¹⁵NH (117,4 ppm, „płaszczyzna”). Widmo **HNCO** pozwala zidentyfikować proton Lys21-N¹H z *cross-peak* oznaczonego **K21**. Z widma **HNCA** o znanych H i N identyfikuje się skorelowany z nimi Lys21-¹³C α o wartości przesunięcia 58,5 ppm. Widmo **¹⁵N-TOCSY-HMQC** przy tym samym przekroju osi ¹⁵N co poprzednio, daje proton Lys21-C α -¹H przy 4,01 ppm (korelacja protonów NH i H α przy jednym z nich bezpośrednio związanym z N). Widmo **HCACO** przy znanym C α i H α Lys21 wyznacza przesunięcie węgla Lys21-¹³C=O (177,9 ppm), a widmo **HCA(CO)N** identyfikuje azot kolejnego aminokwasu Asp22-¹⁵N (114,1 ppm). Korelacja w tym eksperymencie zachodzi między Lys21-C α /H α aż do Asp22-N z pominięciem węgla karbonylowego, ujętego w nawias. Dalej procedura powtarza się dla Asp22 podobnie jak dla Lys21. Widma analizuje się najczęściej zestawiając dwuwymiarowe przekroje, pozwalające łączyć prostymi skorelowane *cross-peaks* **[PLANSZA 99]**. **Podsumowaniem przyporządkowania** (weryfikacja) jest „odcisk palca” białka w formie widma 2D ¹H,¹⁵N-HMQC (lub ¹H,¹⁵N-HSQC) **[PLANSZA 100]**, na którym każdy aminokwas daje jeden sygnał korelacyjny azot amidowy/proton amidowy, z wyjątkiem proliny, pozbawionej protonu amidowego NH. Dalsze eksperymenty korelacyjne **jąder w łańcuchach bocznych** poszczególnych aminokwasów **[PLANSZA 101]** dodatkowo weryfikują prawidłowość przyporządkowania w łańcuchu peptydowym i kończą pracę dla całej cząsteczki. Zasadniczym problemem jest zapewnienie **stabilności próbki** i parametrów pracy aparatu. Przykładowo, rozpracowanie struktury 269-AA proteazy na podstawie przyporządkowania z wykorzystaniem 10 eksperymentów wymagało 48 dni ciągłej rejestracji (**Remerowski et al. J. Biomol NMR 4, 257, 1994**). Czas ten można próbować skrócić **zwiększając stężenie białka** i stosując techniki **zwiększające czułość rejestracji** sygnału NMR.

Widma 3D i 4D NMR biopolimerów średniej wielkości są w zasadzie do przyporządkowania jeśli chodzi o nakładanie się sygnałów. Niestety, przy masach cząsteczkowych **przekraczających ok. 30 ÷ 40 kDa** czasy relaksacji T_2 stają się na tyle krótkie, że powodują dekoherencje (rozfazowanie) magnetyzacji podczas jej transferu w układzie jądrowym w okresach czasów między impulsami. Powoduje to silne **poszerzenie linii i spadek intensywności rejestrowanych sygnałów**. Podjęto próby eliminacji dominującego w mechanizmie oddziaływania dipol-dipol z protonami używając białek **potrójnie znakowanych $^2\text{H}/^{13}\text{C}/^{15}\text{N}$** , ekspresując białka w ciężkiej wodzie $^2\text{H}_2\text{O}$ z glukozą $^{13}\text{C}/^2\text{H}$ jako źródłem węgla. **Współczynnik żyromagnetyczny jąder ^2H jest 6,5 razy mniejszy niż jąder ^1H** a ponadto deuterium ($I = 1$) ulega relaksacji kwadrupolowej. Oba czynniki znacząco redukują oddziaływania DD, ale brak protonów powoduje dwa negatywne efekty.

(1) utrudnienie przyporządkowanie z wykorzystaniem $\text{H}\alpha$ w łańcuchu **[PLANSZA 98]** i resztach bocznych **[PLANSZA 101]**;

(2) zmniejszenie ilości informacji o strukturze trzeciorzędowej z kontaktów proton-proton. Część protonów próbuje się więc pozostawić (odzyskać), np. protony wymienialne (NH i OH) po rozpuszczeniu białka w zwykłej wodzie $^1\text{H}_2\text{O}$, oraz protony grup alifatycznych w jądrze hydrofobowym hodując białko na wzbogaconych pożywkach wybranych aminokwasów znakowanych $^{15}\text{N}/^{13}\text{C}$, ale zawierających protony ^1H . Zmodyfikowano strategię przyporządkowania sygnałów w łańcuchu peptydowym z pominięciem $\text{H}\alpha$: HN(CA)CB, HN(COCA)CB i 4D HNCACB. Efektem było **niewielkie w sumie przesunięcie granicznych mas cząsteczkowych** układów molekularnych badanych spektroskopią NMR do ok. 60 kDa.

Przełom stanowiło opracowanie przez grupę K. Wütricha z ETH (Zurich) **eksperymentu TROSY** (*T*ransverse *R*elaxation-*O*ptimized *S*pectroscop*Y*; **[PLANSZA 102]***). Eksperyment rejestruje korelację przesunięć chemicznych protonów $I = ^1\text{H}$ i bezpośrednio związanych z protonami azotów lub węgla, $S = ^{13}\text{C}$ lub ^{15}N , i tym różni się od standardowego, heterojądrowego COSY, że następuje **rejestracja tylko jednej, najwęższej linii z dwuwymiarowego kwartetu** rozszczepionego w obu kierunkach przez

W sekwencji TROSY skrót **PFG** oznacza *P*ulse *F*ield *G*radient, standardową obecnie w eksperymentach NMR technikę wyboru drogi transferu koherencji magnetyzacji (*coherence transfer pathway*) w układzie jądrowym, czyli tylko niektórych ze znacznej liczby koherencji, które prowadzą do użytecznego sygnału i supresji pozostałych; odpowiednik tzw. cyklu fazowego impulsów pola B_1 .

sprężenie J_{IS} ($I = {}^1\text{H}$; $S = {}^{13}\text{C}$ lub ${}^{15}\text{N}$). Analiza ewolucji macierzy gęstości [\[PLANSZA 89\]](#) z uwzględnieniem relaksacji (dodatkowy człon $\mathbf{R}\sigma$) pokazuje, że szybkości relaksacji poszczególnych przejść określone przez wyrazy **superoperatora** macierzy relaksacji \mathbf{R} są bardzo zróżnicowane. Dzięki sekwencji TROSY uzyskuje się na widmie **przejście o największym sygnale w obu wymiarach: $1 \leftrightarrow 2$ dla heterojądra S oraz $1 \leftrightarrow 3$ dla protonu $I = {}^1\text{H}$ [\[PLANSZA 103\]](#)**. Dla tej linii następuje wzajemne znoszenie się relaksacji DD (dipol-dipol; człon p) i relaksacji wskutek anizotropii przesunięcia chemicznego CSA (*chemical shift anisotropy*; człon δ_S):

$$p - \delta_S \approx 0 \quad p = (1/2 \cdot 2^{1/2}) \gamma_I \gamma_S \hbar r_{IS}^{-3} \quad \delta_S = (1/3 \cdot 2^{1/2}) \gamma_S B_0 \Delta\sigma_S$$

$$r_{IS} - \text{odległość między jądrami I i S} \quad \Delta\sigma_S \sim 160 \text{ ppm}$$

Residualna relaksacja w eksperymencie TROSY nie zależy praktycznie od masy cząsteczkowej, co przy wysokim stopniu deuteracji cząsteczki zapewnia wąskie linie nawet w bardzo dużych układach molekularnych. Symulacje komputerowe dla hipotetycznego białka 800 kDa dają szerokości na osi protonowej i azotowej odpowiednio: 15 Hz i 50 Hz, w porównaniu z szerokościami dla białka 150 kDa: 5 Hz i 15 Hz, już w polu B_0 odpowiadającym częstości 750 MHz. Obecne aparaty pracują przy częstości 900 MHz zapewniającej większą efektywność metody TROSY, a dla hipotetycznego aparatu 1 100 MHz teoretycznie zachodzi całkowita kompensacja relaksacji DD i CSA i możliwość rejestracji linii „high resolution” niezależnie od masy cząsteczkowej. Sekwencje TROSY można implementować jako część dowolnej sekwencji NMR i uzyskiwać **analizowalne widma do badań strukturalnych niezależnie od wielkości układu**. Dla bardzo dużych układów molekularnych wydajniejszym mechanizmem transferu magnetyzacji między protonami i heterojądrami niż stosowany w TROSY mechanizm oddziaływania skalarnego INEPT (*Insensitve Nuclei Enhancement by Polarization Transfer*) jest **transfer przez cross-relaksację CRINEPT** (*Cross Relaxation-Enhanced Polarization Transfer*). Eksperyment TROSY, limitowany przez transfer INEPT zyskuje nowe możliwości przy transferze CRINEPT. Połączony eksperyment **TROSY-CRINEPT** ([Riek et al. J. Am. Chem. Soc. 124, 12144, 2002](#)) wykorzystuje TROSY w okresie ewolucji i akwizycji, a CRINEPT do transferu polaryzacji. To otwiera **drogę badania NMR struktur supramolekularnych** (mezoskopowych). Przykładowo, uzyskano analizowalne widma NMR dla kompleksu białek opiekuńczych GroEL/GroES o masie cząsteczkowej blisko 900 kDa ([Fiaux et al. Nature 418, 207, 2002](#)). Problemy z ilością sygnałów do analizy w układach o tak dużej ilości jąder:

(8 000 AA) x (6 jąder łańcucha) = 48 tysięcy sygnałów 1D (tylko !)

próbuję się rozwiązywać przez opracowywanie zautomatyzowanych algorytmów komputerowych (wstępnego) przyporządkowania, np. metoda SAGA (*sequential assignment of GS algorithm*; [Crippen et al. J. Biomol NMR 46, 281, 2010](#)). Obiecującym podejściem poprawiającym rozdzielczość i czułość widma NMR supramolekularnych układów molekularnych rzędu 1 MDa z jednoczesną redukcją ilości sygnałów wydaje się **metoda methyl-TROSY** ([Kay J. Magn. Reson. 210, 159, 2011](#)). Widma wykorzystujące efekt TROSY rejestruje się dla cząsteczek „znakowanych” $^{12}\text{C}/^2\text{H}$, w których jądra $^{13}\text{C}/^1\text{H}$ są inkorporowane jedynie w grupach metylowych leucyn, walin, izoleucyn, alanin i metionin, co powoduje minimalizację utraty sygnałów wskutek relaksacji, zmniejszenie ich ilości, a jednocześnie wystarcza do rekonstrukcji struktury przestrzennej (jądra hydrofobowego).

Kluczowe parametry dla określenia globalnego foldu trzeciorzędowego białka to **efekt Overhausera NOE** dla protonów odległych w strukturze pierwszorzędowej i **resztkowe sprzężenie dipolarne RDC**. Efekt Overhausera [\[PLANSZA 104\]](#) jest uwarunkowany **oddziaływaniem dipol-dipol (DD)** i zależy jak r_{ij}^{-6} od odległości między jądrami. Rejestruje się go głównie dla protonów w celu wyznaczenia **odległości proton-proton w cząsteczce**, a ostatnio coraz częściej również **odległości między innymi jądrami** w makromolekułach znakowanych ^{13}C i ^{15}N . Wzory są analogiczne: γ_{H}^4 przechodzi w $\gamma_{\text{N1}}^2\gamma_{\text{N2}}^2$ - współczynniki żyromagnetyczne oddziałujących jąder N1, N2; w mianownikach zamiast częstości rezonansowych protonów ω_{H} są sumy i różnice $\omega_{\text{N1}} \pm \omega_{\text{N2}}$. Efekt Overhausera mierzy się w eksperymencie NOESY w kombinacji z eksperymentami korelacyjnymi 3D i 4D, np. HMQC-NOESY. Intensywności a_{ij} cross-peaks NOESY po czasie mixing t_m pozwalają wyznaczyć odległości proton - proton. Dla izolowanych par protonów i/lub bardzo krótkich czasów mixing (a_{ij} ekstrapolowane do wartości dla $t_m = 0$), wyskalowana intensywność a_{ij} **poprzez wyrazy R_{ij} macierzy relaksacyjnej** daje bezpośrednio **poszukiwaną odległość**. Jednak dla dłuższych t_m , które zapewniają odpowiednio **wysoką intensywność cross-peaks** (szczególne znaczenie dla odległych jąder) trzeba uwzględnić dyfuzję spinową, czyli wpływ wszystkich pozostałych protonów cząsteczki [\[PLANSZA 105\]](#). Jak widać na **PLANSZY** dyfuzja spinowa może silnie zaburzyć intensywności cross-peaks NOESY: intensywność cross-peak między NH Phe-33 a CyH w Tyr-32 w BPTI jest większa od intensywności NH Phe33/NH Arg20 mimo, że odległości wynoszą odpowiednio 4.5 Å i 2.9 Å. Pakiety komercyjne do wyznaczenia

struktur cząsteczek zawierają zwykle programy ortogonalizujące macierz **R** relaksacji DD (np. program MARDIGRAS; [Borgias & James J. Magn. Reson. 87, 475, 1990](#)), co pozwala wyznaczyć jej wartości R_{ij} z intensywności a_{ij} , z uwzględnieniem całego układu protonów cząsteczki.

Wyznaczenie odległości międzyjądrowych z oddziaływania DD wymaga niezależnego określenia czasu korelacji τ_c dla dyfuzyjnego obrotu cząsteczki, który warunkuje relaksację dipol-dipol. Podane wzory dotyczą najprostszego modelu dyfuzji rotacyjnej **cząsteczki jako sztywnej sfery z jednym czasem korelacji**. Najczęściej trzeba je modyfikować tak, aby uwzględnić odbieganie kształtu od sfery (elipsoida obrotowa - dwa czasy korelacji, elipsoida - trzy czasy korelacji) i uwzględniać lokalne ruchy wewnętrzne w cząsteczkach. Różne czasy korelacji można wyznaczać z NMR z **pomiarów czasów relaksacji T_1 i T_2 lub z anizotropii fluorescencji**. Techniki NMR pozwalają wyznaczyć czasy korelacji τ_c :

- pomiar stosunku T_1/T_2 dla par protonów;
- pomiar czasów relaksacji dla jąder ^{13}C związanych przez jedno wiązanie z protonami; dominuje mechanizm DD z bezpośrednio związanym protonem, a znana jest odległość węgiel-proton;
- pomiar T_1 i/lub T_2 jąder $I > 1/2$ (^{14}N , ^2H), gdzie mechanizm relaksacyjny jest zdominowany przez relaksację kwadrupolową.

We wszystkich tych przypadkach są dość proste zależności czasów relaksacji od czasów korelacji τ_c , nie wymagające znajomości struktury cząsteczki.

Wykład 14

Częściowa deuteracja cząsteczek dużych biopolimerów, niezbędna do ograniczenia szerokości linii NMR powoduje redukcję mierzonych parametrów NMR, związanych z bardzo ważnymi w tym przypadku protonami. Podjęto próby znalezienia dodatkowych parametrów NMR, związanych z „globalnymi” parametrami strukturalnymi. Największym powodzeniem cieszy się wykorzystanie **resztkowych sprzężeń dipolarnych RDC** (*residual dipolar couplings*) **[PLANSZA 106]**, zaproponowanych przez **[Tjandra & Bax Science 278, 1111, 1997](#)**. Oddziaływanie DD między jądrami obserwuje się w ciele stałym i jest ono jednym z czynników warunkujących szerokości "szerokich" linii rezonansowych. W widmach wysokiej zdolności rozdzielczej mogą się pojawić resztkowe sprzężenia dipolarne jeśli cząsteczki są **częściowo zorientowane w przestrzeni**. Wartości resztkowych sprzężeń dipolarnych D_{ij} dla par jąder i,j dodają się do, lub odejmują od, wartości sprzężeń skalarnych $J(i,j)$ na widmach korelacyjnych, np. widmo 2D HSQC, dające *cross-peaks* od par amidowych N–H; widmo HSQC niesie podobną informację jak HMQC **[PLANSZA 100]**. **Częściową orientację cząsteczek bez utraty ruchliwości molekularnej** zapewniającej warunki wysokiej zdolności rozdzielczej ("wąskie" linie) można uzyskać na różne sposoby, najczęściej w rozcieńczonych roztworach wodnych fosfolipidów DHPC/DMPC (*dihexanoyl phosphatidylcholine/dimyristoyl phosphatidylcholine*) tworzących fazę ciekłokrystaliczną **bicelli** w formie dysków o długości kilkuset Å i średnicy ok. 40 Å. Bicelle oddziałując z cząsteczkami biomolekuł częściowo je orientują. **Wartości i orientacja osi głównych tensora uszeregowania A nie są znane a priori**, ale mogą być fitowane z porównania mierzonych i liczonych dla różnych struktur wartości D_{ij} . Ponadto, fitowanie dostarcza **zbioru orientacji θ, ϕ wektorów łączących jądra i,j** w stosunku do osi częściowego zorientowania molekuly w układzie laboratoryjnym, czyli **poszukiwaną informację strukturalną**. Najczęściej są to bezpośrednio związane jądra o znanych długościach wiązań r_{ij} ; w białkach N→H, $C\alpha \rightarrow H\alpha$, $C\alpha \rightarrow C'$ (C' - węgiel karbonylowy) $C' \rightarrow N$. Tak uzyskana informacja strukturalna ma charakter **długozasięgowy**, obejmujący całą cząsteczkę. Można przy jej pomocy wyznaczyć np. **wzajemną orientację elementów drugorzędowych w cząsteczce białka** po ich określeniu z „lokalnych” zależności strukturalnych.

Inne eksperymenty NMR pozwalające uzyskać informacje o strukturze przestrzennej białek:

- rejestracja efektów *cross-correlation* mechanizmów relaksacyjnych, które dają wartości kątów ψ w łańcuchu peptydowym bez konieczności znajomości stałych sprzężenia;
- pomiar składowych **tensora dyfuzji rotacyjnej** cząsteczki białka.

Dodatkową możliwość daje obecność centrów paramagnetycznych, od których można wyznaczać odległości do jąder. **Elektronowy rezonans paramagnetyczny EPR** umożliwia pomiary odległości do 80 Å między paramagnetycznymi **etykietami spinowymi** przyczepionymi do białek ([Jeschke ChemPhysChem 3, 927, 2002](#)). Wykorzystuje się również niskokątowe rozpraszanie promieniowania rentgenowskiego i neutronów SAXS/SANS do uzyskiwania dodatkowych więzów strukturalnych cząsteczek w roztworze ([Hennig et al. J. Biomol. NMR. 56, 17, 2013](#)).

Usprawnienia aparaturowe idą w kierunku jak **najwyższych pól magnetycznych**, które zapewniają efektywne wykorzystanie sekwencji TROSY/CRINEPT oraz podnoszą czułość poprzez wzrost różnicy boltzmannowskich obsadzeń poziomów. Największe komercyjnie aparaty NMR o indukcji $B_0 = 21T$, odpowiadają częstości rezonansowej protonów **900 MHz**. Rewolucyjne okazało się również zastosowanie tzw. **kriosond**, gdzie układ detekcji sygnału FID jest chłodzony do temperatury ciekłego helu (próbka pozostaje w temperaturze eksperymentu). Następuje **znacząca redukcja szumów** i wzrost czułości rejestracji - wymagane **stężenia mikromolowe zamiast milimolowych**.

Aktualny postęp jaki dokonał się w ostatnim czasie w wykorzystaniu technik NMR w roztworze do badania różnych zagadnień strukturalnych i dynamicznych oraz zmniejszaniu luki pomiędzy zrozumieniem współzależności struktura - funkcja SAR (*narrow the gap between structure and functional understanding*) jest krótko przedstawiony w [review: Grzesiek & Sass Curr. Op. Struct. Biol. 19, 585, 2009](#). Badanie **białek błonowych metodami spektroskopii NMR w roztworze** wymaga (jak przy wykorzystaniu dyfrakcji rentgenowskiej) rozwiązania **problemu rozpuszczalności**. Obok lipidowych roztworów bicellarnych stosowane są systemy micellarne, „proste” i odwrócone (*reverse*), polimery amfipatyczne (*amphipols*) oraz cząsteczki nanolipoproteinowe NPL (*nanolipoprotein particles*) [[PLANSZA 107](#)].

Unikalną cechą techniki NMR jest możliwość **badania cząsteczek w stanie stałym**, zarówno polikrystalicznym jak i amorficznym. Stąd możliwość badania **białek błonowych, włóknistych** i agregujących w roztworze **form amyloidalnych** niewłaściwie zwiniętych białek. W spektroskopii NMR w ciele stałym **SSNMR** (*solid state NMR*) wyróżnia się dwa podstawowe „obszary badawcze”, oparte na dwóch odmiennych technikach.

(A) Analiza „szerokich linii” **[PLANSZA 108]** ^{13}C NMR ($I = \frac{1}{2}$), ^{31}P NMR ($I = \frac{1}{2}$), ^2H NMR ($I = 1$), ^{14}N NMR ($I = 1$) które powstają w wyniku oddziaływań między jądrami: dipol-dipol (DD), oddziaływań jąder z elektronami (CSA, *chemical shift anizotropy*) i oddziaływań momentów kwadrupolowych jąder o spinie $I = 1$ z gradientami pól elektrycznych. **Techniki 1D i wielowymiarowe** tego typu nie dają zwykle możliwości budowy modelu cząsteczki z rozdzielczością atomową, chociaż znane są przypadki zaproponowania modelu struktury przestrzennej białka, np. jedwab pajaków ([Beek et al. Proc. Natl. Acad. Sci USA99, 10267, 2002](#)), na podstawie analizy widm 2D SSNMR:

- DOQSY (*double-quantum single-quantum correlation spectroscopy*): wyznaczenie orientacji tensorów przesunięć chemicznych węgli $^{13}\text{C}=\text{O}$, które z kolei zależą od rozkładu kątów określających konformację łańcucha peptydowego Φ i Ψ ;
- DECODER (*direction exchange with correlation for orientation-distribution evaluation and reconstruction*): wyznaczenie kątów Eulera (α , β , γ) orientacji łańcucha peptydowego w stosunku do włókna (struktura białka pokazywana na wykładzie „Struktura i funkcja makrocząsteczek biologicznych”).

(B) **Technika MAS** wirowania próbki pod kątem magicznym, która pozwala uzyskiwać **widma wysokiej zdolności rozdzielczej w ciele stałym [PLANSZA 109]**: widmo krystalicznej glicyny przy wzrastających szybkościach wirowania. Szerokości połówkowe linii są gorsze niż dla próbek ciekłych, ale dla dużych biomolekuł nie stanowi to już podstawowego problemu. **Takim problemem są natomiast możliwości techniczne związane z wymaganymi szybkościami wirowania próbki.** Uzyskanie wąskiej linii jest możliwe przy częstościach obrotu przewyższających szerokość połówkową szerokich linii (zakres widmowy danego jądra). **Oddziaływanie DD heterojąder z protonami:** $\sim 5 \cdot 10^4 \text{ Hz} \div 10^5 \text{ Hz}$ jest uśredniane przez szerokopasmowe odsprężanie protonów, co powoduje, że szerokość linii ^{13}C , ^{15}N (^{31}P) jest uwarunkowana jedynie przez CSA \sim kilku kHz (10^3 Hz), a w przypadku **białek znakowanych ^{13}C i/lub ^{15}N , a także przez oddziaływania DD między heterojądrami.** Stosowane obecnie standardowe częstości rotacji próbki $20 \div 35 \text{ kHz}$ oraz ultraszybkie częstości 60 kHz , dają **całkowite uśrednienie oddziaływania CSA** oraz mniejszych niż oddziaływanie heterojądro-proton **oddziaływań DD heterojądro-heterojądro** (^{13}C , ^{15}N). Dodatkowo, wykorzystanie **sekwencji CP** (*cross-polarization*) przy warunku Hartmanna-Hahna na moce impulsów protonu i heterojądra S:

$$\gamma_{\text{H}}B_{1\text{H}} = \gamma_{\text{S}}B_{1\text{S}}$$

umożliwia **podniesienie intensywności** rejestrowanych sygnałów ^{13}C , ^{15}N o czynnik γ_I/γ_S , a eksperyment TOSS (*Total Suppression of Sidebands*) tłumia utrudniające analizę tzw. *sidebands* od wirowania. W przypadku jąder $I = 1$ oddziaływanie kwadrupolowe $\sim 10^6$ Hz nie daje się efektywnie uśrednić, ale możliwość pojawienia się poszerzenia kwadrupolowego występuje również w roztworze. Najpoważniejszym problemem są **jądra ^1H** , kluczowe jeśli chodzi o badania strukturalne białek metodami NMR w roztworze. Uśrednienie oddziaływań DD proton-proton ~ 100 kHz nie daje się łatwo osiągnąć przez wirowanie próbki i jądro to w jest zwykle „**wykluczone**” ze **standardowych eksperymentów „high resolution” SSNMR**. Dopiero ostatnio pojawiły się doniesienia o osiągnięciu superszybkiej częstości MAS 110 kHz ([Ye et al. J. Magn. Reson. 239, 75, 2014](#)) w aparaturze JEOL RESONANCE Inc. Istnieje możliwość zastosowania sekwencji WAHUUH (Waugh-Huber-Hacberlen), dzięki której uśrednianie zachodzi w przestrzeni spinowej (magnetyzacja spędza po 1/3 czasu wzdłuż każdej z osi układu wirującego), ale wymogi techniczne eksperymentu uniemożliwiają rutynowe stosowanie tej techniki.

Eksperymenty 2D i 3D NMR, w tym typu „triple resonance” wykorzystywane do przyporządkowywania sygnałów jąder ^{13}C , ^{15}N w SSNMR zostały zmodyfikowane tak, aby ruch wzdłuż sekwencji od aminokwasu do aminokwasu odbywał się z pominięciem protonów: NCACB (korelacje wewnątrz aminokwasu), NCOCA [[PLANSZA 110](#)] i CACA (korelacje między aminokwasami) ([Baldus J. Biomol NMR 39, 73, 2009](#)). Kosztom rozdzielczości można również, w odpowiednio zaprojektowanych eksperymentach z okresami przerw w szerokopasmowym odsprężaniu protonów, uzyskiwać korelacje z udziałem tych jąder: NHHC (korelacje między aminokwasami). **Likwidacja wszystkich oddziaływań DD** przez wirowanie (i wykluczenie jąder ^1H przez odsprężanie szerokopasmowe) stwarza poważny **problem znalezienia metod określania odległości jądro-jądro do wyznaczania globalnego foldu trzeciorzędowego białek w stanie stałym**. Metody wykorzystujące efekt NOE do pomiaru odległości jądro-jądro są kluczowe dla badań białek w roztworze [[PLANSZA 91](#)]. Dysponując przesunięciami chemicznymi ^{13}C i ^{15}N po wykonaniu przyporządkowania można określić **strukturę drugorzędową** łańcucha peptydowego wykorzystując odpowiednie programy oparte na analizie baz danych przesunięć chemicznych dla rozwiązanych struktur, np. (wspomniany już) program TALOS, CHESHIRE ([Robustelli et al. Structure 16, 1764, 2008](#)).

Dwie podstawowe metody w SSNMR umożliwiają **określenie kontaktów trzeciorzędowych** w strukturze białek:

(A) Homonuklearna ^{13}C - ^{13}C dyfuzja spinowa: pasywna (bez udziału protonów) i PDS (proton driven spin diffusion); szereg sekwencji impulsów jest opisanych w pracy [Lopez et al. J. Biomol. NMR 41, 97, 2008](#);

(B) *Dipolar recoupling* (teoria: [Lin et al. J. Magn. Reson. 197, 145, 2009](#)), np. sekwencje R^2 , REDOR, fpRFDR w których oddziaływania dipolarne są **selektywnie przywracane w celu pomiarów odległości międzyjądrowych**. W przypadku sprzężeń dipol-dipol, technika MAS z częstością rotacji ω przekształca niezależny od czasu hamiltonian (współrzędne polarne R , θ , ϕ) oddziaływania dipol-dipol jąder I oraz S:

$$H_{DD} = (\gamma_I \gamma_S \hbar / 4\pi R_{IS}^3) (1 - 3\cos^2\theta_{IS}) (3I_z S_z - \mathbf{I} \cdot \mathbf{S})$$

w zależny od czasu:

$$H_{DD}(t) = (\gamma_I \gamma_S \hbar / 2\pi R_{IS}^3) (A \sin\omega t + B \cos\omega t + C \sin 2\omega t + D \cos 2\omega t) (3I_z S_z - \mathbf{I} \cdot \mathbf{S})$$

A, B, C, D - funkcje orientacji wektora odległości międzyjądrowej \mathbf{R}_{IS}

w ustalonym układzie współrzędnych względem rotora MAS

Sekwencje impulsów „*dipolar recoupling*” w synchronizacji z rotacją próbki powodują zależność od czasu operatorów spinowych \mathbf{I} oraz \mathbf{S} . Jeśli ich transformaty Fouriera zawierają składowe o częstościach ω i 2ω uśrednianie czasowe nie prowadzi do zera. Odległości $^{13}\text{C} - ^{13}\text{C}$ i $^{13}\text{C} - ^{15}\text{N}$ uzyskiwane z pomiarów oddziaływań dipol-dipol nie przekraczają 6 Å, podobnie jak w roztworze. Większe odległości udało się zmierzyć dla cząsteczek znakowanych trytem ([Yuen et al. J. Am. Chem. Soc. 132, 1734, 2010](#)). Zastosowanie techniki *dipolar recoupling* (sekwencja HORROR) pozwoliła uzyskać pomiar odległości $^3\text{H} - ^3\text{H}$ ok. 14 Å, na razie dla niewielkich związków modelowych.

W ostatnim okresie pojawiło się **alternatywne podejście do rozwiązania problemu dipolarnego poszerzenia linii rezonansowych w eksperymentach MAS białek** wskutek silnego oddziaływania DD z protonami. Rozwiązanie to już poprzednio zastosowano dla bardzo dużych układów molekularnych w roztworze i polega ono na **deuteracji czyli zamianie jąder ^1H na deuterium ^2H** . Deuteracja obejmuje wszystkie niewymienialne pozycje w cząsteczce, czyli protony połączone z węglami (hodowla bakterii na pożywkach deuterowanych), a pozycje wymienialne, czyli połączenia protonów z azotem i tlenem, są powtórnie obsadzane protonami przez rekrytalizację/odparowanie w zwykłej wodzie. Ponadto, odparowanie białka można wykonywać w roztworze $^1\text{H}_2\text{O}/^2\text{H}_2\text{O}$ o zadanej zawartości protonów i w ten sposób regulować zakres sprzężeń DD. Przykładowo, przy 10% zawartości protonów w pozycjach wymienialnych linie rezonansowe na widmach MAS niewielkich białek mają szerokości połówkowe ok. 20 ÷

30 Hz podczas gdy przy rekrytalizacji białek z buforów o 100% zawartości $^1\text{H}_2\text{O}$ są rzędu $80 \div 150$ Hz. Wydaje się, że optymalnym rozwiązaniem zapewniającym akceptowalną szerokość połówkową sygnałów przy zachowaniu odpowiednie czułości ich detekcji jest stężenie 40% ^1H w buforze.

Zastosowanie techniki SSNMR MAS do badań strukturalnych białek przeżywa intensywny rozwój od ok. 2000 r. i ciągle jeszcze pozostaje **w stanie poszukiwania optymalnych rozwiązań i ustalania standardowych procedur postępowania**, z jakimi mają do czynienia badacze struktur białek w roztworze ([review: McDermott Annu. Rev. Biophys. 38, 385, 2009; Reif J. Magn. Reson. 216, 1, 2012](#)). Rutynowo rozwiązuje się struktury raczej mniejszych białek, chociaż spektakularne osiągnięcia dotyczą również białek błonowych [[PLANSZA 111](#)], które w większości są białkami o bardzo dużych masach cząsteczkowych. Niewątpliwym osiągnięciem jest uzyskiwania struktur **form amyloidalnych**, dla których nie ma alternatywnej metody ustalania struktury przestrzennej z rozdzielczością atomową.

Niezależnie od stanu skupienia podstawową cechą struktur białek jest **dynamika ruchów molekularnych** poszczególnych grup i fragmentów, którą obserwuje się nie tylko w roztworze ale również, w bardziej ograniczonym stopniu w stanie stałym. Niekiedy, ze względu na brak dyfuzji rotacyjnej w stanie stałym, badanie ruchów wewnętrznych w tym stanie może mieć przewagę nad badaniami w roztworze ze względu na brak konieczności separacji ruchu na wkład ze strony globalnego ruchu całej cząsteczki i ze strony ruchów wewnętrznych. Ponieważ zastosowanie spektroskopii SSNMR do badań białek znajduje się ciągle w fazie „eksperymentalnej” w znaczeniu poszukiwań optymalnych metod i rozwiązań, **zagadnienie dynamiki ruchów molekularnych białek** będzie omówione dla cząsteczek **w roztworze, zarówno wodnym** jako naturalnym środowisku białek globularnych, jak i **w błonie lipidowej**. Przedstawione będą również wybrane, spektakularne wyniki uzyskane dla białek w stanie krystalicznym.

Wykład 15

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

- (1) globalne obroty i translacje cząsteczki jako całości;
- (2) ruchy wewnętrzne poszczególnych fragmentów cząsteczki;
- (3) dynamika cząsteczek wody w sferze hydratacyjnej białka i ich wymiana oraz wymiana jonów pomiędzy rozpuszczalnikiem i stanem związanym; efekt hydratacyjny i efekt polielektrolityczny dla biopolimerów w roztworze wodnym.

Dynamiczny charakter natywnych struktur białek w roztworze zarówno wodnym (białka globularne), jak i wewnątrz błony lipidowej (białka błonowe) wyraża się wewnętrznymi ruchami molekularnymi w bardzo szerokim zakresie dwóch podstawowych, mierzalnych parametrów:

- szybkości ruchów dyfuzyjnych, określane przez **średnie czasy korelacji** (τ_{int}) od **pikosekund do sekund/minut/godzin**,
- **średnie zmiany położenia atomów** (amplitudy) wyrażane w Å (zmiany odległości) lub stopniach (zmiany kątowe), zwykle skorelowane co do wielkości z szybkościami ruchów; powolne ruchy charakteryzują się większymi amplitudami.

Obserwuje się zmiany strukturalne białek, od ograniczonych ruchów o charakterze lokalnych obrotów wokół pojedynczych (w sensie ilości i rzędowości) wiązań do znacznych przekształceń konformacyjnych pod wpływem oddziaływań międzycząsteczkowych, np. wiązanie ligandów, a także samoistne rozwijanie (częściowe) i powtórne zwijanie do formy natywnej. Z szerokiego spektrum metod badania dynamiki białek **[PLANSZA 40]** stosowanych zależnie od skali czasowej, **techniki:** 2D IR *vibrational echo*, rozpraszanie neutronów QENS (*quasi-elastic neutron scattering*) i fluorescencja czasowo-rozdzielcza (pomiar anizotropii po wzbudzeniu krótkimi impulsami laserowymi) mogą być stosowane do śledzenia bardzo krótkich, **pikosekundowych** procesów dynamicznych, takich jak: wymiana cząsteczek wody białko/solvent i lokalne fluktuacje atomów. Pomimo precyzyjnych pomiarów pikosekundowych i subpikosekundowych szybkości ruchów, można zauważyć trudności w interpretacji charakteru badanych ruchów. „Uniwersalnym” charakterem jeśli chodzi o skalę czasową odznaczają się różne techniki NMR ([reviews: Torchia J. Magn. Reson. 212, 1, 2011;](#)

[Osawa et al. Curr. Op. Struct. Biol. 22, 660, 2012](#)), stosowane do śledzenia dynamiki w szerokim zakresie od sub-nanosekund do sekund (i dłużej).

Pełną informację o ruchach molekularnych, nie tylko w postaci podanych parametrów, ale również zmian położenia poszczególnych atomów uzyskuje się za pomocą komputerowych symulacji **dynamiki molekularnej** MD (*molecular dynamics*). Cząsteczkę przybliża się jako zbiór klasycznych sfer materialnych na które działają siły ze strony innych atomów wyznaczone z pola siłowego FF [\[PLANSZA 39\]](#). Ruch atomów wynika z numerycznego rozwiązania równań Newtona w krótkich skokach czasowych ~femtosekundy (10^{-15} s) przez okres dostępny dla mocy obliczeniowych superkomputerów, standardowo kilkaset nanosekund ($\sim 10^{-7}$ s). W zmodyfikowanych wersjach, np. brownowska MD do zakresu mikrosekundowego (10^{-5} s); granica stale przesuwa się do dłuższych czasów wraz z rozwojem mocy obliczeniowych superkomputerów. Dynamikę molekularną stosuje się do uzyskiwania struktur przestrzennych cząsteczek na podstawie mierzonych parametrów NMR (rMD), wprowadzając pseudosiły "ściąające" cząsteczkę do konformacji opisanej znanymi z NMR parametrami przestrzennymi [\[PLANSZA 96\]](#). To podejście stosuje się także w „czysto” komputerowych symulacjach TMD (*targeted MD*), podobnie jak szereg innych modyfikacji mających na celu wydłużenie dostępnych w symulacji czasów zachodzenia procesów dynamicznych (będą przedstawione na równoległym wykładzie kursowym „Bioinformatyka i modelowanie”).

Lokalne ruchy wewnętrzne w białkach obejmują rotacje grup $-\text{CH}_3$ aminokwasów alifatycznych ($10^{-8} \div 10^{-11}$ s), grupy $-\text{NH}_3^+$ lizyny ($10^{-10} \div 10^{-11}$ s), przeskoki *flip-flop* o 180° pierścieni fenyloalanin i tyrozyn (ale nie tryptofanów), od 10^{-9} do ok. 1s, w zależności od miejsca występowania tych aminokwasów. W skali czasowej, $\sim 10^{-7} \div 10^{-8}$ s zachodzą **wzajemne reorientacje domen** z przemieszczeniami do ~ 10 Å. Ruchy wewnętrzne łańcucha peptydowego w zakresie **nanosekund** ($\sim 10^{-8} \div 10^{-9}$ s) z **rozdzielczością do poszczególnych aminokwasów**, tzn. możliwością określenia indywidualnych czasów τ_{int} dla każdego aminokwasu, analizuje się na gruncie **modelu MFA** (*model-free approach* [\[PLANSZA 112\]](#) ([Lipari & Szabo J. Am. Chem. Soc. 104, 4546, 1982](#))). Ruchy wewnętrzne separuje się od obrotu cząsteczki jako całości wyrażając w przybliżeniu funkcję autokorelacji (patrz wykład „*Spektroskopia molekularna*”) jako **iloczyn funkcji autokorelacji** dla obrotu całej cząsteczki G_0 i ruchu wewnętrznego G_1 :

$$G(\tau) = G_0(\tau) \cdot G_1(\tau)$$

gdzie $G_0(\tau)$ i $G_I(\tau)$ mogą być przedstawione w najogólniejszej formie dla ruchu anizotropowego cząsteczki przez sumę wyrazów eksponencjalnych (wzorów nie podaję). Najistotniejsza dla analizy i interpretacji jest funkcja gęstości spektralnej $J(\omega)$, czyli **transformata Fouriera funkcji autokorelacji**:

$$J(\omega) = \int G(\tau) \exp(-i\omega\tau) d\tau$$

którą można bezpośrednio powiązać z **mierzonymi doświadczalnie czasami relaksacji dla azotów ^{15}N** : $1/T_1 = R_1$, $1/T_2 = R_2$ i **efektem Overhausera** $\{^1\text{H}\}$, ^{15}N - NOE:

$$R_1 = \frac{1}{4} D^2 [J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)] + \frac{1}{3} C^2 J(\omega_N)$$

$$R_2 = \frac{1}{8} D^2 [4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + \frac{1}{18} C^2 [4J(0) + 3J(\omega_N)] + R_{ex}$$

$$NOE = 1 + \left(\frac{\gamma_H}{\gamma_N} \right) \frac{D^2}{4R_1} [6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)]$$

$$D = \frac{\mu_0}{4\pi} \frac{\gamma_N \gamma_H \hbar}{\langle r_{NH}^3 \rangle}, \quad C = \omega_N \Delta\sigma$$

gdzie odległość proton - azot (długość wiązania) $r_{NH} = 1,04 \text{ \AA}$; zakres przesłaniania jąder ^{15}N $\Delta\sigma = -170 \text{ ppm}$; ω_N , ω_H są częstościami rezonansowymi azotu i protonu, γ_N , γ_H są ich współczynnikami żyromagnetycznymi, a człon $R_{ex} = \Phi \omega_N^2$ określa wkład do czasu relaksacji T_2 od wymiany chemicznej procesów o szybkościach mikro- i milisekund oraz efektywności wkładu Φ . W ogólnym przypadku modelu cząsteczki jako dowolnej elipsoidy $J(\omega)$ jest sumą pięciu składników z czasami $\tau_{i,k}$ (trzy czasy korelacji globalnego obrotu wokół trzech osi elipsoidy). Funkcję gęstości spektralnej liczy się zwykle w wystarczającym modelu cząsteczki jako elipsoidy obrotowej z dwiema składowymi tensora dyfuzji rotacyjnej D_{\perp} i D_{\parallel} (dwa czasy korelacji) i trzema składnikami:

$$J(\omega) = \sum_{k=1}^3 A_k(\alpha) J_{MFA}(S, \tau_k, \tau_{i,k}; \omega)$$

α - kąt między wektorem N-H i główną osią symetrii tensora dyfuzji rotacyjnej

$$A_1 = 0.75(\sin^4 \alpha), \quad A_2 = 3(\sin^2 \alpha)(\cos^2 \alpha), \quad A_3 = 0.25(3\cos^2 \alpha - 1)$$

$$\tau_{i,k}^{-1} = \tau_k^{-1} + \tau_{int}^{-1} \quad k = 1, 2, 3$$

$$\tau_1 = 1/(4D_{||} + 2D_{\perp}), \quad \tau_2 = 1/(D_{||} + 5D_{\perp}), \quad \tau_3 = 1/(6D_{\perp})$$

Przyjmując wystarczające dla pojęciowego przedstawienia zagadnienia **przybliżenie izotropowej dyfuzji rotacyjnej** cząsteczki jako sfery:

$$J(\omega) = \frac{2}{5} \left[\frac{S^2 \tau_R}{1 + (\omega \tau_R)^2} + \frac{(1 - S^2) \tau}{1 + (\omega \tau)^2} \right]$$

$$\tau^{-1} = \tau_R^{-1} + \tau_{\text{int}}^{-1}$$

wyraża się „giętkość” aminokwasów w łańcuchu przez **parametr uporządkowania** S^2 :

$S^2 = 1$ dla fragmentów sztywnych,

$S^2 = 0$ dla całkowicie swobodnych ruchów wewnętrznych.

Wyznaczenie wszystkich niezbędnych parametrów: czasy korelacji τ_R , τ_{int} , oraz wartości Φ i S^2 (dla każdego AA) wymaga wyznaczania **czasów relaksacji** T_1 i T_2 i efektu **NOE** $\{^1\text{H}\}$ - ^{15}N poszczególnych aminokwasów dla trzech wartości indukcji pola B_0 (trzy aparaty NMR). Przykładowe wyniki dla białka bakteryjnego Yfia, które powstaje w *E. coli* przy skoku temperatury prezentuje **[PLANSZA 113]**. Widać obraz dynamiki poszczególnych fragmentów łańcucha peptydowego z wyraźnym usztywnieniem obszarów tworzących struktury drugorzędowe (wartości $S^2 \rightarrow 1$) i większą dynamikę pętli (niższe wartości S^2). Przebieg S^2 nie koresponduje całkowicie z przebiegiem czasów korelacji τ_{int} , ponieważ całe fragmenty drugorzędowe mogą wykazywać ruchy w dłuższej skali czasowej, które wpływają na τ_{int} jąder ^{15}N aminokwasów. Są próby wyznaczania S^2 z rentgenografii ([Fenwick et al PNAS www.pnas.org/cgi/doi/10.1073/pnas.1323440111, 2014](http://www.pnas.org/cgi/doi/10.1073/pnas.1323440111))

Dla białek globularnych środowisko buforu wodnego jest zasadniczym elementem warunkującym funkcjonalną strukturę natywną. Zagadnienie struktury wody w obecności powierzchni rozpuszczonych, hydratowanych biomolekuł oraz wpływu środowiska wodnego na oddziaływania międzycząsteczkowe jest przedmiotem intensywnych badań różnymi technikami ([Raschke Curr. Op. Struct. Bio. 16, 152, 2006](http://www.sciencedirect.com/science/article/pii/S0005272806000111)) oraz zmieniających się koncepcji modelowych ([Israelachvili & Wennerström Nature 379, 219, 1996](http://www.sciencedirect.com/science/article/pii/S0028371896000111)). Interesujące wyniki ilościowe dla dynamiki cząsteczek wody w sferze hydratacyjnej białek uzyskuje się przez zastosowanie fluorescencji czasowo-rozdzielczej z rozdzielczością femtosekundową. Fluorymetr z wzbudzeniem impulsowym laserem Ti-szafir w długości fali 800 nm *up-converted* do zakresu bliskiego nadfioletu 210 ÷ 290 nm (techniki mieszania i podwajania częstości) pozwala uzyskać rozdzielczość czasową ~100 fs. Najlepiej „analizowalne” rezultaty uzyskuje się dla białek z pojedynczym tryptofanem, np.

endopeptydaza serynowa, apomioglobina. Ciekawym rozwiązaniem jest więc mapowanie cząsteczki białka pojedynczymi tryptofanami jako lokalną sondą optyczną, wprowadzaną w różne miejsca cząsteczki przez ukierunkowaną metagenезę (*side-directed mutagenesis*). W pracy [Zhang et al. J. Am. Chem. Soc. 131, 10677, 2009](#) wykonano badania dla 16 takich mutantów rejestrując dwufazowe, multieksponecjalne przebiegi czasowe (t) zaniku fluorescencji tryptofanu po wzbudzeniu:

$$I_{\lambda}(t) = \sum_i \alpha_i \exp(-t/\tau_i) + \sum_j \beta_j \exp(-t/\tau_j)$$

Pierwszy człon jest związany z **procesami solwatacyjnymi**, a drugi z zanikiem wewnętrznej (*intrinsic*) emisji, uwarunkowanym przez **otoczenie molekularne tryptofanu**. Parametry α_i oraz τ_i można ekstrahować z przebiegu $I_{\lambda}(t)$ (poprzez tzw. „funkcję korelacji”) jako odpowiedzi lokalnego otoczenia na skokowe zmiany elektrycznego momentu dipolowego Trp pod wpływem wzbudzenia. Dodatkowo, pomiar czasowej zależności **anizotropii fluorescencji** po wzbudzeniu (patrz wykład „Spektroskopia molekularna”):

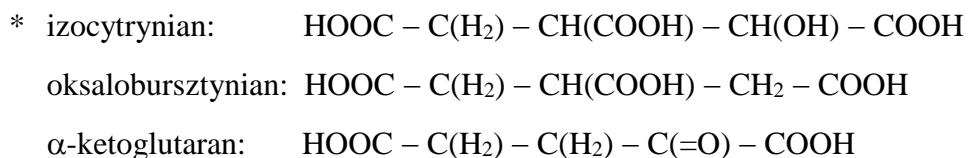
$$r(t) = [I_{\parallel}(t) - I_{\perp}(t)] / [I_{\parallel}(t) + 2I_{\perp}(t)]$$

daje informacje o **dynamice wewnętrznych ruchów molekularnych** cząsteczki (czasy korelacji). Dynamika cząsteczek wody obejmuje dwa procesy o różnej skali czasowej:

- szybkie, pikosekundowe ($\tau_{\text{int}} \sim 0,7 \div 8$ ps), lokalne ruchy relaksacyjne wody w sieci wiązań wodorowych, w sferze hydratujującej białko o grubości ok. 10 Å;
- wolniejsze ($\tau_{\text{int}} \sim 20 \div 200$ ps) ruchy „boczne” (*lateral*) w sferze hydratacyjnej, sprzężone z fluktuacjami strukturalnymi białka;

Dyfrakcja rentgenowska w standardowym wydaniu jest typową techniką dającą statyczne obrazy struktur przestrzennych cząsteczek. Pewną, chociaż ograniczoną informację o lokalnej dynamice atomów w kryształach można uzyskać na podstawie wartości współczynników temperaturowych Debye’a-Wallera B_{ij} **[PLANSZA 79]**. W roku 1912 von Laue i wsp. uzyskali obraz dyfrakcyjny „białego”, tzn. niemonochromatycznego promieniowania rentgenowskiego, ugiętego na kryształach siarczanu miedzi. Metoda ta została zarzucona na rzecz metody z wykorzystaniem wiązki monochromatycznej ze względu na brak odpowiednich źródeł i trudności w analizie znacznie bardziej skomplikowanego układu prążków dyfrakcyjnych. Problem został rozwiązany dzięki zastosowaniu promieniowania synchrotronowego SRS (*synchrotron radiation source*) o dużej intensywności w przedziale 0,2 ÷ 2,5 Å oraz zaawansowanego software’u

komputerowego. Konstrukcja Ewalda dla dyfrakcji Lauego [PLANSZA 114] pokazuje, że warunek przecięcia węzłów sieci odwrotnej z trójwymiarowym obszarem między sferami Ewalda $1/\lambda_1$ i $1/\lambda_2$ zachodzi dla bardzo dużej ilości refleksów. **Skraca to czas rejestracji dyfraktogramu w wiązce SRS do ok. 100 ps.** Metoda Lauego daje możliwość badania dynamiki procesów zachodzących z szybkościami odpowiadającymi czasom dłuższym niż czasy rejestracji obrazów dyfrakcyjnych ([Hajdu et al. Nature 329, 178, 1987](#)). Spektakularne okazało się zastosowanie tej metody do **śledzenia przebiegu reakcji enzymatycznych** ([Hajdu et al. TIBS 13, 104, 1988](#)). Ponieważ wiele enzymów zachowuje aktywność w kryształach szybka rejestracja dyfraktogramów umożliwia uzyskanie obrazów struktur przejściowych na drodze reakcji enzymatycznej. Zmodyfikowany chemicznie substrat, nie ulegający reakcji, jest umieszczony w centrum aktywnym i przekształcany fotochemicznie błyskiem światła do substratu właściwego (fotoliza błyskowa). Od tego momentu startu reakcji enzymatycznej wykonuje się serię dyfraktogramów Lauego. Technika jest trudna eksperymentalnie zarówno pod względem przygotowania układu enzymatycznego jak i analizy obrazów dyfrakcyjnych, ale od czasu do czasu pojawiają się w literaturze doniesienia o jej zastosowaniach do badania enzymów. Jedną z takich przykładowych prac o enzymie w cyklu kwasu cytrynowego (cykl Krebsa), **dehydrogenazie izocytrynianowej** ([Bolduc et al. Science 268, 1312, 1995](#)), przedstawia wykorzystanie zmutowanych białek, Y160F oraz K230M, w celu silnego spowolnienia dwóch etapów reakcji: tworzenie kompleksu potrójnego enzym-izocytrynian*- Mg^{+2} - $NADP^+$ (etap 4) i dehydrogenacja izocytrynianu do przypuszczalnego intermediatu oksalobursztynianu (etap 5), na drodze do produktu α -ketoglutaranu. Rentgenografia Lauego dostarczyła obrazy struktur obu etapów przejściowych [PLANSZA 115].



Wykład 16

Charakterystyka proteomu [PLANSZA 73] pod względem **rodzaju i ilości** funkcjonujących białek oraz określenie ich **struktur przestrzennych** pozwala wyznaczyć **interaktom** czyli obraz współdziałania białek w funkcjonalnych kompleksach molekularnych PPI (*protein-protein interactions*). Kluczową jest sprawa odróżnienia oddziaływań funkcjonalnych od niespecyficznych. Szeroka gama technik detekcji **oddziaływań białko - białko** ([review: Wodak et al. Curr. Op. Struct. Biol. 23, 941, 2013](#)), sprzężonych z proteomicznym podejściem „*high throughput*”, najczęściej opiera się na **spektrometrii mas** (*proteomic pipeline* [PLANSZA 116] i obejmuje [PLANSZA 117]:

(a) techniki elektroforetyczne CNE (*clear-native electrophoresis*; [Vittig & Schrägger Proteomics 9, 5214, 2009](#)) do izolacji, szczególnie hydrofobowych (białka błonowe), ale także rozpuszczalnych w wodzie, kompleksów i superkompleksów;

(b) chromatografia powinowactwa ([Walther & Mann J. Cell. Biol. 190, 491, 2012](#)) ze znacznikami TAP (*tandem affinity purification*; immunoprecipitacja) lub ostatnio coraz częściej używanym GFP (*green fluorescence protein*), także do **lokalizacji białek w komórce** za pomocą mikroskopii fluorescencyjnej;

(c) oczyszczanie przez powinowactwo do przeciwciał IP (*immunoaffinity purification*) w połączeniu z analizą MS ([Miteva et al. anal. Chem. 85, 749, 2013](#));

(d) metody optyczne ([Lam & Stagljar Proteomics 12, 1519, 2012](#)): FRET (*fluorescence resonance energy transfer*) między białkami *bait* w fuzji z donorem (np. CFP, *cyan fluorescent protein*) oraz *prey* w fuzji z akceptorem (YFP, *yellow fluorescent protein*) oddziałującymi w żywej komórce; BRET (*bioluminescence resonance energy transfer*), modyfikacja metody FRET z wykorzystaniem świecenia lucyferazy jako donora podczas oksydacji substratu (minimalizacja efektu fotobłaknięcia, *photobleaching*); LUMIER (*luminescence-based mammalian interactome mapping*);

(e) metody NMR ([O'Connell et al. Proteomics 9, 5224, 2009](#)): zaburzenia przesunięć chemicznych na widmach ^1H , ^{15}N -HSQC, efekt Overhausera NOE, resztkowe sprzężenie dipolarne RDC do określenia orientacji oddziałujących białek, oddziaływanie z sondami paramagnetycznymi (PRE, *paramagnetic relaxation enhancement* oraz PCS (*pseudocontact shift*) na nieznakowanych promotorach, z detekcją efektów na białkach znakowanych ^{13}C i/lub ^{15}N w odległości do 30Å; NMR z wiązami SAXS: intensywność

promieniowania X rozproszonego pod małymi kątami na roztworach molekuł w funkcji kąta rozproszenia; uzyskuje się globalne kształty i rozmiary kompleksu białko-białko;

(f) **mikromacierze** białkowe ([Tomizaki et al. FEBS J. 277, 1796, 2010](#)): detekcja białek przez wychwyt różnymi czynnikami, umocowanymi na podłożu stałym o dużej gęstości (ilości cząsteczek chwytających na jednostkę powierzchni), takimi jak: przeciwciała, białka fuzyjne, aptamery DNA/RNA, małe ligandy; detekcja fluorescencyjna znaczników lub tzw. *label-free*: MS, SPR;

(g) **projektowanie** białek wiążących z analiz struktur trzeciorzędowych białek badanych (*peptide grafting and library approach*; [Umetu et al. FEBS J. 277, 2006, 2010](#)). W przypadku białek błonowych ich natywna struktura jest stabilizowana w kompleksach z ligandami niskocząsteczkowymi, w tym lipidami, co ma szczególne znaczenie w desorpcji MS natywnych kompleksów (nMS; *native mass spectrometry*); publikacja [Gault et al. Nature Methods 13, 333, 2016](#) prezentuje opracowaną w tym celu *Orbitrap platform*, dzięki której można precyzyjnie identyfikować i rozróżniać związane ligandy.

W funkcjonującej komórce, proteomiczna, wysokoprzepustowa analiza odkrywa setki wielobiałkowych kompleksów, których funkcje są regulowane przez tysiące oddziaływań białko-białko PPI. **Podstawowym problemem**, podobnym do problemu ograniczonej przepustowości metod strukturalnych, jest możliwość uzyskania bardziej szczegółowych informacji warunkujących funkcjonowanie oddziaływań białko-białko jedynie dla stosunkowo niewielkiego ich ułamka. Z pomocą w modelowaniu oddziaływań (kompleksów) białko-białko przychodzi **bioinformatyka strukturalna**, która integruje i uzupełnia metody doświadczalne [[PLANSZA 118](#)], np. *search strategy* w połączeniu z analizą *top-down* w pracy [Skinner et al. Nature Methods 13, 337, 2016](#). W konstruowaniu interaktomu bierze się pod uwagę **różne typy kompleksów**:

- duże, wielobiałkowe asocjaty o powierzchniach kontaktu $>2\ 000\ \text{\AA}^2$, często z elementami co-wiązującymi (*interlocking*), uczestniczącymi w kooperatywności procesu wiązania;
- oddziaływania domena-domena (DDIs, *domain-domain interaction*) o powierzchniach $1\ 000\ \div\ 2\ 000\ \text{\AA}^2$;
- oddziaływania za pośrednictwem peptydów z interfejsami $\sim 350\ \text{\AA}^2$, często w kompleksach o charakterze przejściowym.

Obok wymienionych metod doświadczalnych rutynowo stosuje się **dokowanie** białko-białko i **modelowanie homologiczne**. Techniki bioinformatyczne mają również za zadanie wypracowanie **efektywnych i czytelnych wizualizacji interaktomów** (*PPI network*)

[PLANSZA 119], z reguły bardzo złożonych. Efektywny interaktom powinien „udźwignąć” zdolność badacza do zrozumienia współpracy różnych białek w funkcjonowaniu komórki. Tworzenie sieci PPI jest z matematycznego punktu widzenia rysowaniem **grafów** $G(V,F)$ z V węzłami i F krawędziami, w których każdy węzeł reprezentuje pewne białko; przegląd [Fung et al. Proteomics 12, 1669, 2012](#) prezentuje sposoby wizualizacji interaktomów. Kluczowym zadaniem jest konstrukcja sieci PPI w formie powszechnie dostępnych i wygodnych w użyciu **baz danych interaktomu**, pozwalających zunifikować ponad 130 istniejących „niewyspecjalizowanych” repozytoriów (MINT, DIP, BioGRID) i „wąskich” grup białek (MPIDB, MatrixDB).

Celem praktycznym proteomiki jest wybór "tarcz" czyli receptorów białkowych dla różnego typu inhibitorów czy aktywatorów jako środków farmakologicznych: antywirusowych, antynowotworowych, cytostatycznych czy sprzyjających proliferacji określonych typów komórek. **Podstawowym warunkiem racjonalnego wyboru „tarcz”** jest **analiza interaktomu**: [cytat z Taylor & Wrana Proteomics 12, 1706, 2012](#):

*„several studies have shown that the structure and dynamics of protein networks are disturbed in complex diseases such as cancer progression. These relationships suggest a novel paradigm for treatment of complex multigenic disease where the **protein interaction network is the target of therapy more so than individual molecules within the network**”*

Analiza stabilizujących kontaktów modyfikowanych chemicznie ligandów w centrum aktywnym wybranego receptora (enzymu) jest kolejnym krokiem w **projektowaniu środków farmakologicznych** (*rational drug design*; patrz dalsza część wykładu).

ZWIJANIE BIAŁEK (PROTEIN FOLDING): mechanizmy zwijania *in vitro* i *in vivo*:
badania doświadczalne i teoretyczne; „błędne” zwijanie i agregacja (formy amyloidalne)

Stabilizacja struktur natywnych cząsteczek białkowych ΔG° w stosunku do struktury rozwiniętej kłębka statystycznego nie jest wysoka i wynosi od kilkunastu do kilkadziesiąt kilojouli na mol w temperaturze 25°C. Dla białek globularnych w przeglądzie o termodynamice białek ([Makhatadze & Privalov Adv. Protein Chem. 47, 307, 1995](#)) podane są doświadczalne wartości ΔG° dla szeregu białek globularnych: od niewielkiej domeny SH3 (57 AA), $\Delta G^\circ = 14,5$ kJ/mol, do lizozymu T4 (163 AA) $\Delta G^\circ = 68,4$ kJ/mol. Porównanie entalpii swobodnych stabilizacji struktur natywnych w białkach

organizmów **mezofilnych** (zwykłych) i ich homologów z **bakterii termofilnych** pokazuje **ok. dwukrotny wzrost wartości ΔG°** dla tych ostatnich, mierzonych w temperaturach maksymalnej stabilności białek ([Kumar et al. Biochemistry 40, 14152, 2001](#)). Przykładowo, dla mezofilnej domeny Btk-SH3 $\Delta G^\circ = 16$ kJ/mol, podczas gdy dla ośmiu homologicznych białek termofilnych średnie $\Delta G^\circ = 30$ kJ/mol. Dla średniej wielkości białek termofilnych wartości $\Delta G^\circ \sim 65 \div 72$ kJ/mol. Stabilizacja struktur natywnych **białek globularnych** w danych warunkach roztworu i temperatury wynika z oddziaływań stabilizujących (o których mówiłem): hydrofobowe, mostki solne, wiązania wodorowe, oddziaływania vdW, które przeważają nad wkładem destabilizującym, głównie entropii konformacyjnej łańcucha i reszt bocznych, przy czym udział mają również efekty entalpowe i entropowe oddziaływania z rozpuszczalnikiem: zrywanie/powstawanie wiązań wodorowych z cząsteczkami wody, efekt hydratacyjny i efekt polielektrolityczny. W **białkach błonowych** poszczególne wkłady stabilizujące i destabilizujące są nieco inne ze względu na hydrofobowe otoczenie cząsteczek białka, ale zasadniczo można wyróżnić takie same typy oddziaływań. Dla białek globularnych i błonowych zmiany warunków środowiska mogą spowodować **odwrócenie relacji stabilizacja/destabilizacja struktury natywnej** i przejście białka w formę zdenaturowaną kłęбка statystycznego, w wyniku działania czynników denaturujących: podwyższona temperatura lub dla niektórych białek niska temperatura (zimna denaturacja), chemiczne denaturanty (chlorek guanidyny) i skrajne pH roztworu (z wyjątkiem np. enzymów trawiennych). Dla wielu białek usunięcie czynnika denaturującego powoduje powtórne zwinięcie się białka zdenaturowanego do formy natywnej czyli **refolding** podobnie jak zwijanie do struktury natywnej ko- i posttranslacyjnie czyli **folding**, z wyjątkiem białek natywnie rozwiniętych. **Informacja o strukturze przestrzennej formy natywnej białka, zazwyczaj o minimum energii swobodnej Gibbsa, jest określona przez sekwencją aminokwasów** (hipoteza Anfinsena). Temat ZWIJANIE BIAŁEK jest centralnym problemem biofizyki molekularnej i obejmuje:

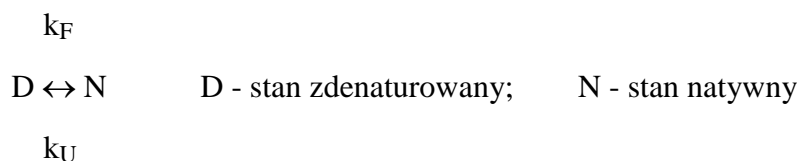
- (a) poszukiwanie algorytmów komputerowych wyznaczania struktury natywnej na podstawie sekwencji aminokwasów;
- (b) poszukiwanie molekularnych mechanizmów (dróg) zwijania różnymi metodami doświadczalnymi i teoretycznymi;
- (c) zwijanie *in vivo* i komórkowe mechanizmy wspomagające.

Krótkim, ale kompletnym przeglądem aktualnych osiągnięć jest [review: Dill & MacCalum Science 338, 1042, 2012](#), natomiast wyczerpujące przeglądy, skoncentrowane głównie na problematyce mechanizmu zwijania to: [Dill et al. Annu. Rev. Biophys 37, 289, 2008](#); [Chen et al. Arch. Biochem. Biophys. 469, 4, 2008](#). Zagadnienia zwijania białek (niektórych) błonowych są omówione w [review: Otzen & Andersen Arch. Biochem. Biophys. 531, 34, 2013](#), a jedynym (znalezionym) przeglądem o zwijaniu białka włóknistego jest praca o kolagenie: [Boudko et al. Int. J. Biochem. Cell Biol. 44, 21, 2012](#). Na wykładzie przedstawione zostaną najnowsze wyniki w zakresie analizy doświadczalnej i konstruowania **modeli mechanizmów zwijania** (punkt b), obrazowanych przez **lejki zwijania** [[PLANSZA 120](#)], które usuwają słynny **paradoks Levinthala**. W zagadnieniach zwijania *in vivo* (punkt c) szczególna uwaga będzie zwrócona na nieprawidłowe zwijanie, czyli *misfolding* i agregację do form amyloidalnych, ze względu na ich **działanie patogenne**. Kilka zasadniczych elementów przechodzenia białka od formy rozwiniętej na obrzeżu lejka zwijania (wysoka wartość entropii S) do formy natywnej o minimum entalpii swobodnej G na dnie lejka (niska wartość entropii S):

(A) Niezależny **ruch dyfuzyjny** poszczególnych cząsteczek białka „po ścianach lejka” można ująć w ramach dyfuzyjno-kolizyjnego modelu reakcji zwijania ([Veawer'a & Karplusa Nature, 260, 1976](#)), w którym fragmenty struktur drugorzędowych, tzw. mikrodomeny, dyfundują w roztworze pod wpływem ruchów termicznych i zderzając się tworzą stopniowo większe fragmenty uporządkowane, aż do zakończenia reakcji zwijania. Model Veawer'a-Karplusa dobrze koresponduje z **hierarchicznym** ujęciem procesu zwijania ale trudności występują w przypadku koncepcji zwijania nie-hierarchicznego jako **kolapsu hydrofobowego**: kontrakcji łańcucha peptydowego pod wpływem efektu hydrofobowego ([Dill et al. Protein Sci. 4, 561, 1995](#)). Problem hierarchiczności zwijania białek nie jest jak dotąd rozstrzygnięty.

(B) Proces zwijania jest **kooperatywny**: tworzenie wcześniejszych kontaktów sprzyjających zwiniętej strukturze natywnej sprzyja tworzeniu kolejnych kontaktów tego typu [[PLANSZA 121](#)], co nadaje hiperpowierzchni konformacyjnej białka kształt lejka,

(C) **Kinetyka** procesu zwijania zależy od kształtu lejka. W szczególności, prosty dwustanowy model procesu refoldingu [[PLANSZA 122](#)]:



$$K_{D \leftrightarrow N} = k_F/k_U = [native]/[unfolded] \quad - \text{stała równowagi}$$

odpowiada „gładkim” (niepofałdowanym) ścianom lejka. Wykres zależności stężenia formy zdenaturowanej [D] od czasu po usunięciu denaturanta ma charakter **monoeksponencjalny**, z obserwowaną stałą szybkości k_{obs} :

$$[D](t) = \{D_0/(k_F + k_U)\} \{k_F \exp(-(k_F + k_U)t) + k_U\}$$

$$k_{obs} = k_F + k_U$$

W warunkach równowagi stacjonarnej zależność k_{obs} od stężenia czynnika denaturującego

$$k_{obs} = k_F^0 \exp(-m_F[\text{denaturant}]/RT) + k_U^0 \exp(m_U[\text{denaturant}]/RT)$$

k_F^0 i k_U^0 – szybkości zwijania i rozwijania w nieobecności denaturanta

[denaturant] - stężenie denaturanta

m_F i m_U współczynniki liniowej zależności entalpii swobodnej zwijania (F) i

rozwijania (U) od stężenia denaturanta:

$$\Delta G^\circ_F = \Delta G^\circ_F(\text{H}_2\text{O}) + m_F[\text{denaturant}]$$

Zależność $\ln(k_{obs})$ od [denaturant] tworzy charakterystyczny „*chevron plot*”. Zakres **czasów zwijania** τ_F dla takich białek mimo prostoty modelu kinetycznego może być bardzo szeroki, od mikrosekund (10^{-5} s) do sekund, czyli pokrywa cały obserwowany dla wszystkich białek **zakres czasów zwijania** $10^{-5} \div 10^2$ s. Szereg białek zwija się zgodnie z **kinetyką wieloeksponencjalną**, z utworzeniem wyraźnie populowanych i strukturalnie zdefiniowanych **stanów pośrednich**, lub niekoniecznie określonych strukturalnie **pułapek kinetycznych**, co odpowiada lejkom pofałdowanym **[PLANSZA 120]**.

(D) [Plaxco et al. J. Mol. Biol. 277, 985, 1998](#) wprowadzili pojęcie **rzędu kontaktu** CO (*contact order*) jako miary porównawczej liczby natywnych kontaktów bliskozasięgowych ($< 6 \text{ \AA}$) między aminokwasami blisko położonymi w sekwencji i liczby takich kontaktów między aminokwasami oddalonymi w łańcuchu peptydowym:

$$CO = (1/LN) \sum_{i < j} \Delta S_{ij}$$

ΔS_{ij} - ilość aminokwasów w sekwencji dzielących dwa oddziałujące, i-ty oraz j-ty

N - całkowita liczba kontaktów, L - ilość aminokwasów w białku.

CO reprezentuje **topologiczną złożoność** białka jako względną miarę udziału (znaczenia) w strukturze natywnej lokalnych kontaktów stabilizujących w stosunku do nielokalnych. Pokazano, że **CO koreluje ujemnie z szybkością zwijania** k_F . Im niższa wartość CO (mniej złożone białko) tym większa szybkość zwijania **[PLANSZA 123]**. Trzy zaznaczone

na **PLANSZY** białka tworzą 70% natywnej topologii w stanie przejściowym TS (*transition state*), jak wyznaczono na podstawie wyników symulacji molekularnej dynamiki Langevina z eksperymentalnymi wiązaniami, które uzyskano z doświadczalnych wartości Φ .

(E) Wygodną miarą ułamka struktury natywnej w zbiorze mikrostanów przejściowych na drodze zwijania wokół miejsca mutacji jest „ Φ -value” wprowadzona przez Fershta ([Itazaki et al. Journal of Molecular Biology 254, 260, 1995](#)):

$$\Phi = (\Delta\Delta G^\ddagger)/(\Delta\Delta G^\circ)$$

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

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

gdzie k_{mut} , k_{wt} są szybkościami zwijania mutantu i białka „dzikiego”, $\Delta\Delta G^\ddagger$ są odpowiednimi różnicami entalpii swobodnych barier, a $\Delta\Delta G^\circ$ różnicami entalpii swobodnych stanów natywnych. Wartość $\Phi \sim 1$ oznacza, że zmiana entalpii swobodnej w wyniku mutacji jest podobna w stanie natywnym i przejściowym, czyli natywne kontakty z udziałem zmutowanego aminokwasu są już utworzone w stanie przejściowym. Natomiast wartość $\Phi \sim 0$ oznacza, że lokalne otoczenie miejsca mutacji ma charakter rozwinęty.

(F) Analiza procesu **rozwijania białek** może dostarczyć danych na temat zwijania jako procesu odwrotnego, szczególnie jeśli symulacje MD procesu rozwijania ([Nölting et al. J. Theor. Biol. 251, 331, 2008](#)) są weryfikowane przez eksperymentalne dane z manipulacji pojedynczą cząsteczką białka przy wykorzystaniu mikroskopii AFM. Rozciąganie wielodomenowego białka daje charakterystyczny obraz (*sawtooth pattern*) na wykresie zależności siły rozciągającej (pN) od rozciągnięcia (nm) [**PLANSZA 124**]. Pojedynczy pik odpowiada rozwinięciu jednej domeny. W ten sposób można dokładniej charakteryzować **stany przejściowe**.

Wykład 17

Znaczna ilość pojawiających się stale publikacji poświęconych doświadczalnemu badaniu foldingu *in vitro* z wykorzystaniem najróżniejszych metod dotyczy kilku aspektów, kluczowych dla opracowywania modelowych mechanizmów procesu zwijania, głównie dla białek globularnych **[PLANSZA 120]**:

- (a) **kinetyka** refoldingu i wyznaczenie stałych szybkości procesu (czasów zwijania);
- (b) **stabilność** białek w różnych warunkach środowiska i **parametry termodynamiczne** przejścia między stanami, natywnym i zdenaturowanym;
- (c) **intermediaty** na drodze zwijania, struktura i **dynamika** form przejściowych;
- (d) inicjacja foldingu i **załączki** zwijania;
- (e) charakterystyka form **rozwiniętych** białek;
- (f) struktura i mechanizm tworzenia form **amyloidalnych**.

Kinetykę procesu refoldingu białka zdenaturowanego po usunięciu czynnika denaturującego oraz rozwijania, czyli **unfoldingu** białek pod wpływem denaturanta można śledzić stosując różnorodne metody spektroskopowe, w których **rejestrowany sygnał zależy od stopnia zwinięcia/rozwinięcia cząsteczki**:

- fluorescencja (głównie) tryptofanów;
- absorpcja UV reszt bocznych aminokwasów aromatycznych;
- techniki NMR: poszerzenie linii, zmiany intensywności i położenia sygnałów 1D, 2D i 3D;
- dichroizm kołowy CD w zakresie absorpcji wiązań peptydowych (180÷240 nm);
- absorpcja IR, ATR-FTIR (*Attenuated Total Reflexion-FTIR*), rezonansowy Raman (RR), w zakresie rozciągania C = O i zginania N – H, 1600 ÷ 1700 cm⁻¹ **[PLANSZA 125]**.

Stopień zwinięcia cząsteczki można określić poprzez **ilościowe wyznaczenie zawartości struktur drugorzędowych** (helisy, β-kartki, pętle) na widmach CD oraz ATR-FTIR; dopasowanie widm mierzonych do modelowych. W ATR-FTIR ([Shivu et al. Biochemistry 52, 5176, 2013](#)) wyznacza się drugą pochodną absorpcji i dopasowuje pola pod krzywymi dla wybranych pasm. Z analizy kinetyki procesu zwijania/rozwijania **[PLANSZA 122]** **stałe szybkości** można uzyskać w dwóch typach eksperymentów.

(A) Pomiar **stacjonarne** zmian parametrów spektroskopowych w funkcji stężenia denaturanta w warunkach równowagi formy natywnej i rozwiniętej **[PLANSZA 126]**. Dodatkowo, z analizy zależności ilości frakcji rozwiniętej/zwiniętej α w funkcji stężenia denaturanta można uzyskać wartość **stałej równowagi** $K_{D \leftrightarrow N} = [unfolded]/[native] = (1-\alpha)/\alpha$

(B) **Rejestracja widm w czasie** po usunięciu/dodaniu denaturanta w stężeniu zapewniającym warunki „skrajne” tylko dla jednej z form: natywnej (refolding) lub zdenaturowanej (unfolding).

Pomiary stacjonarne wymagają jednoczesnego wyznaczenia większej ilości parametrów [PLANSZA 122], a więc preferowane są techniki rejestracji dochodzenia białka w czasie do stanu natywnego/zdenaturowanego. **Dofitowanie** odpowiednich krzywych mono- lub wieloeksponencjalnych do punktów eksperymentalnych, zależnie od złożoności modelu kinetycznego, pozwala uzyskać poszukiwane **stałe szybkości i czasy zwijania/rozwijania** jako odwrotności tych stałych. W przypadku długich czasów zwijania można, po zapoczątkowaniu reakcji, rejestrować widmo lub mierzyć intensywność sygnału w jednej długości fali w ustalonych odstępach czasu. **Kinetyki w skali milisekund** ($> 10^{-3}$ s) i większej wyznacza się z wykorzystaniem techniki zatrzymanego przepływu (*stopped-flow*) [PLANSZA 127] z rejestracją sygnału fluorescencji [PLANSZA 126], absorpcji UV, CD, lub NMR. Przebieg procesów zwijania *de novo*, powtórnego zwijania po usunięciu działania czynnika denaturującego i rozwijania białka zielonej fluorescencji GFP (*green fluorescent protein*) wykonane w Zakładzie Biofizyki prezentuje [PLANSZA 128]. Przepływ reagentów ze strzykawek do kuwety pomiarowej jest zatrzymywany [PLANSZA 127], wyznaczając punkt startowy pomiaru zmiany sygnału w czasie. Dolną granicę rozdzielczości czasowej wyznacza „czas martwy” aparatu rzędu kilku ms. Czas martwy jest uwarunkowany charakterystycznym czasem dyfuzji cząsteczek t do osiągnięcia właściwego wymieszania reagentów, który zależy od kwadratu drogi dyfuzji r :

$$t = r^2/D \sim 1 \text{ ms}$$

przy stałej dyfuzji $D = 10^{-5} \text{ cm}^2/\text{s}$ oraz $r = 1 \text{ }\mu\text{m}$

Szybkie kinetyki w skali mikrosekund ($10^{-6} \div 10^{-3}$) uzyskuje się przez zastosowanie turbulentnego mieszania ([Roder et al. Methods 34, 15, 2004](#)), które zapewnia „skrócenie” odległości r . Turbulentne mieszanie zachodzi przy przepływie scharakteryzowanym przez parametry osiągające wartość liczby Reynolds $Re > 2000$:

$$Re = \rho v d / \eta$$

ρ - gęstość płynu [g/cm^3]; v - prędkość przepływu [cm/s]; d - wymiar kanału [cm];
lepkość płynu $\eta = 0.01 \text{ P}$ [poise] dla wody w temperaturze 20°C .

Turbulentne mieszanie implementuje się na aparatach z **mikserami kapilarnymi** i dowolnego typu detektorem (absorpcja UV, fluorescencja, CD, NMR):

- omówione już aparaty „*stopped-flow*” ;

- aparaty „*continuous-flow*”; technika poprzedzająca technikę stopped-flow a następnie zarzucona, przeżywa obecnie „renesans”; wykorzystuje miksery kapilarne do turbulencji a także do ograniczenia zużycia substancji badanych (likwidacja głównej wady metody);
- aparaty „*quenched-flow*”; technika wykorzystuje skokowe zmiany warunków roztworu, np. pH, temperatura, fotoliza błyskowa, które inicjują proces refoldingu/unfoldingu

Termodynamiczna analiza pozwala uzyskać informacje o **stabilności** białek ΔG° z podziałem na **wkłady: entalpowy** ΔH° i **entropowy** $T\Delta S^\circ$, oraz ciepło właściwe (pojemność cieplna) procesu zwijania przy stałym ciśnieniu Δc_p° . Na podstawie zmierzonych parametrów termodynamicznych można wnioskować o **mechanizmie foldingu**, tworzeniu **struktur czwartorzędowych** i procesach **agregacji** przy niewłaściwym zwijaniu. Kluczową techniką w tym obszarze badań jest **mikrokalorymetria**, przede wszystkim **DSC** (*differential scanning calorimetry*), pozwalająca wyznaczyć wartości ciepła właściwego przy stałym ciśnieniu c_p procesu przejścia (fazowego) $D \leftrightarrow N$ w funkcji temperatury T **[PLANSZA 129]**. Niektóre zagadnienia, np. procesy agregacji można badać techniką izotermicznej kalorymetrii miareczkującej **ITC** (*isothermal titration calorimetry*). Ostatnio pojawiła się nowa technika **PPC** (*pressure perturbation calorimetry*; [Doyle et al. Arch. Biochem. Biophys. 531, 44, 2013](#)), polegająca na pomiarze efektów cieplnych przy przykładaniu małych impulsów ciśnienia do roztworu białka. PPC daje temperaturową zależność objętościowego **współczynnika rozszerzania białka** (*coefficient of the partial volume*), uwarunkowanego przez hydratację (oddziaływanie białko-rozpuszczalnik) powierzchni cząsteczki eksponowanej do roztworu. **Technika DSC** dostarcza wartości **trzech podstawowych parametrów** termodynamicznych.

- (1) Temperatura przejścia T_m , która jest bezpośrednią miarą stabilności białka, w maximum temperaturowej zależności ciepła właściwego $c_p(T) = (\partial H/\partial T)_p$.
- (2) Różnica ciepła właściwych w stałym ciśnieniu formy zdenaturowanej (D) i formy natywnej (N): $\Delta c_p^\circ = c_p^\circ(D) - c_p^\circ(N)$. Wartość Δc_p° jest dodatnia i można ją powiązać z „zagrzebywaniem” powierzchni hydrofobowej (*non-polar*) ΔA_{npol} oraz powierzchni polarnej (hydrofilowej) ΔA_{pol} w wyniku zwijania:

$$\Delta c_p^\circ = -0,32\Delta A_{npol} + 0,14\Delta A_{pol} \quad \text{wartości powierzchni } \Delta A \text{ w } [\text{\AA}^2].$$

- (3) Entalpia kalorymetryczna: $\Delta H_{cal} = \int c_p^\circ dT$ jako całki po paśmie $c_p^\circ(T)$

$$\text{i entropia } \Delta S_{cal} = \int [c_p(T)/T]dT$$

Dla białek **jednodomenowych**, wykazujących **całkowicie kooperatywne** zwijanie/rozwijanie (model „wszystko-albo-nic”) zachodzi relacja równości entalpii kalorymetrycznej ΔH_{cal} oraz entalpii van't Hoffa ΔH_{vH} , uzyskiwanej z pośrednich pomiarów zależności temperaturowej stałej równowagi $K_{D \leftrightarrow N}$:

$$\Delta H_{cal}/\Delta H_{vH} = 1$$

$$\partial \ln K_{D \leftrightarrow N} / \partial T = \Delta H_{vH} / RT^2$$

Dla białek **wielodomenowych** z (względnie) **niezależnym** zwijaniem/rozwijaniem poszczególnych domen:

$$\Delta H_{cal}/\Delta H_{vH} > 1$$

Dla białek ze zwijaniem/rozwijaniem poprzez **formy pośrednie**:

$$\Delta H_{cal}/\Delta H_{vH} < 1$$

Analiza DSC białka **BPTI** (6 alanin) i jego mutantów o różnej zawartości alanin (BPTI-21, BPTI-22 i BPTI-27) pokazuje spadek stabilności białka w miarę wzrostu ich ilości. Zależność $c_p(T)$ wskazuje na jednodomenową budowę białka. Złożony z 430 AA, trójdomenowy chaperon *E. coli* **TF** (*trigger factor*) wykazuje na niezależne rozwijanie trzech domen poprzez dwudomenowe intermediaty, co uzyskano przez fitowanie takiego modelu procesu do zależności $c_p(T)$. W przeciwciałach **IgG** zróżnicowane stabilności domen (Fab, C_{H2} i C_{H3}) prowadzą do trzech oddzielnych pików DSC. Stabilności poszczególnych domen można modyfikować bez wpływu na pozostałe domeny. Pojemność cieplna w funkcji temperatury form amyloidalnych zmutowanej domeny **SH3** białka *alfa-spectrin* (N47A) o różnym czasie tworzenia (linia ciągła i linia przerywana) wykazuje prostą i typową zależność, z wartościami T_m wyższymi niż dla białka natywnego. Jest to zgodne ze strukturą włókna amyloidального, zbudowanego z rozciągniętych i częściowo ustrukturuowanych monomerów o wysokim stopniu hydratacji w porównaniu z natywną domeną SH3.

Intermediaty czyli formy przejściowe na drodze zwijania próbuje się identyfikować strukturalnie z wykorzystaniem typowych dla tego celu (już omawianych) metod oraz metod specjalnie rozwiniętych w sposób ukierunkowany na analizę tego zagadnienia zwijania:

- wymiana proton-deuter (*H/D-exchange*) z detekcją NMR lub MS;
- wykorzystywanie mostków dwusiarczkowych;
- znakowanie oksydacyjne z detekcją MS; $\Delta\Delta$

- analiza pojedynczych cząsteczek: mikroskopia AFM i pomiar FRET (*fluorescence resonance energy transfer*), z ew. wykorzystaniem szczypiec optycznych.

Badanie form przejściowych stanowi z jednej strony wyzwanie dla metodologii doświadczalnej, z drugiej strony rośnie liczba dowodów na znaczenie tych form w funkcjonowaniu białek, m. in. **modyfikacjach post-translacyjnych i wiązaniu ligandów**.

Formy przejściowe mogą być bardzo różnorodne. Obok **wyraźnie populowanych i**

(a) **strukturalnie dobrze zdefiniowanych intermediatów** - minima lokalne na leжку zwijania [**PLANSZA 120**], wyróżnia się:

(b) niewykrywalne w eksperymentach kinetycznych „**ukryte**” **intermediaty** (*hidden intermediats*) nawet w bardzo małych, „dwustanowych” białkach (~100 AA lub mniej);

(c) krótko żyjące **intermediaty kinetyczne** (*kinetic intermediats*);

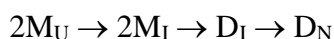
(d) występujące w małych zawartościach (populacjach), **intermediaty termodynamiczne** (*thermodynamic intermediats*);

(e) powszechnie przyjmowane w zwijaniu białek globularnych formy przejściowe **stopionej globuly** (*molten globule, premolten globule*; [Baldwin & Rose Curr. Op. Struct. Biol. 23, 4, 2013](#)).

NMR można wykorzystywać w badaniach **struktur intermediatów** w standardowy sposób **stabilizując formy przejściowe**: inżynieria genetyczna (mutacje i delecje), dodawanie chemicznych denaturantów, zmiana pH roztworu, domieszanie ko-solwentu. Zwykle sama obecność intermediatów jest już zweryfikowana kinetycznie. W przykładowej analizie **rybonukleazy H** [**PLANSZA 130**] zaprojektowano delecyjny mutant z usuniętą ostatnią α -helisą oraz dwiema niemi b2 i b3 β -kartki. Mutant imituje strukturę intermediatu na wczesnym etapie zwijania (*early-folding intermediate*), który tworzy się (jak wynika z analizy) w skali milisekundowej lub krótszej. Struktura intermediatu została zaproponowana m. in. na podstawie eksperymentu **wymiany proton-deuter z detekcją NMR** [**PLANSZA 131**]. Rozwinięte białko w D₂O (wszystkie protony przyłączone do tlenów i azotów wymienione na deuter) poddaje się refoldingowi inicjowanemu przez rozpuszczenie w buforze H₂O, w warunkach fizjologicznych, czyli sprzyjających formie zwiniętej. W kolejnych odstępach czasu rejestruje się wymianę deuterów nie związanych wiązaniami wodorowymi na protony. Chodzi tu głównie o detekcję protonów amidowych N–H w wiązaniach peptydowych. Po przerwaniu wymiany np. przez obniżenie pH białko poddaje się całkowitemu refoldingowi. Można w ten sposób na widmach ¹H NMR **zidentyfikować miejsca wymiany** w strukturze natywnej i wyciągać wnioski odnośnie struktur intermediatów.

W białkach tworzących mostki dwusiarczkowe można identyfikować struktury przejściowe, stabilizowane przez mniejszą niż wymagana, ilość takich mostków, wychwytywanych przez zatrzymywanie procesu w niskich pH (zakwaszanie roztworu). W przeciwieństwie do poprzednio omówionej metody, analiza strukturalna dotyczy bezpośrednio form częściowo zwiniętych. Jak pokazały klasyczne i powszechnie cytowane prace dla BPTI (wykład „*Struktura i funkcja makrocząsteczek biologicznych*”), technika daje możliwość określania zarówno struktur na drodze prawidłowego zwijania jak i struktur w „ślepych uliczkach”. Analogiczny do BPTI przykład dla rybonukleazy H obrazuje **[PLANSZA 132]** native vs. nonnative –S–S–.

Komplementarną techniką do wymiany H/D jest **impulsowe znakowanie oksydacyjne z detekcją MS**. Poddanie roztworu białka z niewielką zawartością nadtlenu wodoru (H₂O₂) mikrosekundowemu impulsowi lasera UV generuje **rodniki •OH** i oksydację dostępną dla solwentu reszt bocznych aminokwasów (Met, Cys, Trp, Phe, Tyr, His, Pro, Leu, Ile). Produkty łatwo się rozpoznaje po wzroście masy cząsteczkowej o 16 Da, rejestrowanym jako przesunięcie na widmie masowym. Monitorowany metodą MS refolding S100A11, homodimeru wiążącego wapń **[PLANSZA 133]**, ze stanu zdenaturowanego pH 2 po skoku do pH 7 daje **mapowanie NOL** (*normalized oxidation levels*) cząsteczki. Stąd wyznacza się różne obszary dostępności solwentu; w tym przypadku zaznaczone na czerwono, jak w formie rozwiniętej oraz na niebiesko, jak w formie natywnej. Uzyskane dane pozwalają zaproponować drogę zwijania S100A11:



gdzie M oznacza monomer, D dimer, I intermediat, U *unfolded*, N *native*.

Niewątpliwą zaletą technik badania pojedynczej cząsteczki (*single-molecule methods*) jest „uwolnienie się” od obrazu uśrednionego po zespole statystycznym cząsteczek. Techniki manipulacji i SMS wykorzystuje się w badaniach strukturalno-dynamicznych foldingu białek (*review: Schuler & Hoffman Curr. Op. Struct. Biol. 23, 36, 2013*), foldingu RNA (*Keller et al. J. Am. Chem. Soc. 136, 4534, 2014*) i innych procesów, np. rekombinacji DNA (*Ma et al. J. Mol. Biol. 426, 793, 2014*). Najszerze zastosowanie w problemie „*protein folding*” uzyskały dwie podejścia:

- **fluorescencja pojedynczej cząsteczki** z pomiarem rezonansowego transferu energii wzbudzenia **smFRET** (*single-molecule resonance energy transfer*), uzupełniony w pomiarach dynamicznych szybkości o technikę PET (*photo-induced electron transfer*; *review: Schuler & Hoffman, 2013*);

- eksploracja lejków zwijania czyli **pejzażu energetycznego** (*energy landscape*) przez rozciąganie (*unfolding*) z wykorzystaniem **mikroskopu sił atomowych AFM**, **szczypiec optycznych** i **szczypiec magnetycznych**.

Wykorzystanie efektu FRET [**PLANSZA 134**] umożliwia **pomiar odległości r** pomiędzy wzbudzonym donorem i akceptorem energii wzbudzenia w zakresie zbliżonym do promienia Förstera R_0 czyli dla $r \sim 30 \div 70 \text{ \AA}$, na podstawie wydajności transferu energii wzbudzenia E (zadanie na wykładzie „*Spektroskopia molekularna*”):

$$E = \frac{k}{\tau_D^{-1} + k} = 1 - F_{DA}/F_D = 1 - \tau_{DA}/\tau_D$$

$$k = \frac{1}{\tau_D} \left(\frac{R_0}{r}\right)^6 - \text{szybkość procesu transferu;}$$

τ_D - czas życia stanu wzbudzonego donora przy braku akceptora

τ_{DA} - czas życia stanu wzbudzonego donora w obecności akceptora

F_{DA} - intensywność fluorescencji donora w obecności akceptora

F_D - intensywność fluorescencji donora przy braku akceptora

W wersji smFRET, zliczanie pojedynczych fotonów w układzie konfokalnym (**prawa część PLANSZY**) w trakcie procesu zwijania (**a**) umożliwia wyznaczenie i **śledzenie dynamiki zmian odległości donor-akceptor w czasie (b) ÷ (d)**. Technika smFRET wydaje się szczególnie sprzyjająca śledzeniu szybkiego (skala czasowa submikrosekundowa) procesu kolapsu do struktury stopionej globuły.

Technika fluorescencji pojedynczej cząsteczki białka α_3D (73 AA) unieruchomionego na podłożu (biotyna - streptawidyna - szkło pokryte PEG) pozwoliła wyznaczyć oprócz czasów zwijania także **szybkości przechodzenia przez barierę folded/unfolded** (*Chung & Eaton Nature 502, 685, 2013*)

Pomiary sił molekularnych techniką AFM (*review: Zoldak & Rief Curr. Op. Struct. Biol. 23, 48, 2013*) pozwalają wyznaczać pejzaż energetyczny foldingu [**PLANSZA 135**].

Wykład 18

Szczególna rolę w mechanizmie procesu zwijania pełnią struktury stanowiące **zależki formy natywnej** białka. Zagadnienie to bezpośrednio wiąże się z problemem hierarchiczności zwijania białek: kolejne tworzenie struktur o wzrastającej rzędowości: zależki zwijania → struktury drugorzędowe → globalny fold trzeciorzędowy vs. kolaps hydrofobowy do globalnego foldu z następującym tworzeniem struktur drugorzędowych i optymalizacją położeń reszt bocznych aminokwasów. Niezależnie od przyjętego modelu struktury zależkowe wydają się być wstępnym etapem procesu. Badania doświadczalne koncentrowały się na tworzeniu struktur α -helikalnych i dwuniciowych, antyrównoległych struktur β -kartkowych (β -hairpin) w odpowiednio zaprojektowanych, krótkich fragmentach peptydowych. Dobrym modelem do badania propagacji helisy okazały się np. syntetyczne peptydy długości 12 ÷ 16 AA, modyfikowane analogi III-pętli wiążącej wapń kalmoduliny ([Goch et al. Biochemistry 42, 6840, 2003](#)), a modelem w tworzeniu β -hairpin C-końcowy fragment (41-56) białka G ze streptokoków ([Munos et al. Nature 390, 196, 1997](#)). Szybkości zwijania tych form wynosiły:

- ok. 200 ns dla α -helisy ($\sim 10^{-7}$ s)

- 6 μ s dla β -hairpin ($\sim 10^{-6}$ s)

Badania wczesnych etapów zwijania spełniają również bardzo ważną rolę przy porównywaniu wyników doświadczalnych z komputerowymi symulacjami dynamiki molekularnej. **Metody dynamiki brownowskiej BMD** (*Brownian molecular dynamics*) uproszczonych modeli białek w ciągłym ośrodku (potencjał średniej siły, równanie Langevina) umożliwiały osiągnięcie mikrosekundowych czasów symulacji, dostatecznie długich do utworzenia zależków. Znacznie bardziej wiarygodne wyniki **symulacji pełnoatomowych** małych białek z uwzględnieniem *explicite* cząsteczek wody pojawiły się stosunkowo niedawno, np. szybkozwijających się mutantów domeny Pin1 WW ($\sim 30\,000$ atomów) w czasie 10- μ s ([Freddolino et al. Biophys. J. 94, L75, 2008](#)). Zgodność czasów zwijania z symulacji i równolegle wyznaczanych czasów z pomiarów kinetycznych stanowiło zasadniczy warunek **weryfikacji** drogi zwijania uzyskanej dzięki symulacjom: kolaps z formy rozwiniętej do struktury stopionej globuły z jednoczesnym utworzeniem dwóch obszarów helikalnych i trójniciowej β -kartki połączonej dwoma skrętami.

Badania strukturalno-dynamiczne form **rozwinętych białek** (na obrzeżu lejków zwijania) ma kluczowe znaczenie, szczególnie dla zrozumienia początkowych etapów

foldingu oraz procesów błędnego zwijania (*misfolding*). Zgodnie z **teorią kłębka statystycznego** Flory'ego dowolny homopolimer w stanie rozwiniętym jest scharakteryzowany przez promień hydrodynamiczny R_g (*radius of gyration*):

$$R_g = R_0 n^\nu \quad R_0 - \text{czynnik skalujący (funkcja persistence length);}$$

n - ilość monomerów; $\nu \cong 0,6$

W przypadku białek czyli heteropolimerów szereg danych doświadczalnych wskazuje na szczątkowe, natywne elementy strukturalne nawet w warunkach silnie denaturujących (8 M mocznik). Prace teoretycznych pokazały brak sprzeczności między tymi wynikami: zdenaturowane białko spełnia relacje geometryczne Flory'ego i jednocześnie **zachowuje szczątkowe struktury natywne**. To uzgodnienie wymagało tzw. **renormalizacji strukturalnej**. Białko dzieli się na grupy lokalnie oddziałujących aminokwasów (*native-like structures*), które zachowują się jak homopolimer:

$$R_g = R_0 (n/L)^\nu \quad L \approx 10 - \text{średnia ilość AA segmentu typu native-like}$$

Drugorzędowe elementy typu natywnego ułatwiają pokonanie wysokich barier entropowych w zwijaniu (scenariusz „*guided folding*”). **Badania strukturalne form rozwiniętych** uzyskały znaczące przyspieszenie od czasu udanych prób wykonywania przyporządkowania sygnałów wielowymiarowych widm NMR o małej dyspersji **[PLANSZA 136]**, czyli rozrzucie na skali przesunięć chemicznych. W tego typu analizach z powodzeniem stosuje się widma wielowymiarowe 5D i 6D NMR z niejednorodnym próbkowaniem NUS (*non-uniform sampling*) dla skrócenia czasu rejestracji ([Hoch et al. Accounts Chem. Res. 47, 708, 2014](#)); a także w **ogólnym przeglądzie z osiągnięciami NMR w proteomice strukturalnej** [Frueh et al. Curr. Op. Struct Biol. 23, 734, 2013](#)).

Obok decydującej roli sekwencji aminokwasowej szereg innych elementów ma udział w zwijaniu białek *in vivo*. Przykładowo, w „zatłoczonym” otoczeniu molekularnym w komórce cząsteczki mogą być schwyte w struktury metastabilne i albo rozpoznane i poddane degradacji/refoldingowi (chaperony) albo asocjują w nieprawidłowo zwinięte agregaty **amyloidalne**. W obrazie lejków zwijania odpowiada to wzrostowi stopnia **pofaldowania lejka** (*landscape ruggedness*) oraz **drugiemu dnu**. Niewłaściwie zwinięte białka **agregują** i mogą być przyczyną chorób, np. neurodegeneracyjnych (omówione dalej przy foldingu *in vivo*). Podobnie jak badanie białek rozwiniętych, badania amyloidów jest domeną NMR, ALE w ciele stałym ([review: Tycko Quart. Rev. Biophys. 39, 1, 2006](#)), szczególnie omawiane techniki strukturalne w połączeniu z wymianą H/D. Znaczące rezultaty uzyskuje się również przy zastosowaniu szeregu technik spektroskopii

fluorescencyjnej, takich jak FCS, FRET, mikroskopia fluorescencyjna *in vivo* ([review: Munishkina et al. Biochim. Biophys. Acta 1768, 1862, 2007](#)) oraz mikroskopii elektronowej STEM (*scanning transmission electron microscope*) i AFM. STEM jest odmianą transmisyjnego mikroskopu elektronowego (TEM), która różni się od konwencjonalnego zogniskowaniem wiązki elektronów w bardzo wąską „plamkę” skanującą cienką próbkę linią po linii (*raster*; jak w *scanningu* SEM).

Intensywnie badanym technikami ^{13}C MAS NMR (7 typów eksperymentów) oraz EM „modelem” złożeń amyloidalnych są **włókna amyloidalne $\text{A}\beta_{1-40}$ związane z chorobą Alzheimera** [\[PLANSZA 137\]](#). Para peptydów, każdy w **konformacji β -karkowej** tworzy protofilament o symetrii C_2 , Hydrofobowe reszty boczne aminokwasów (zielone) są głównie w jego środku, a polarne (purpurowe), naładowane dodatnio (czerwone) i ujemnie (niebieskie) na zewnątrz, w obrazie NMR nałożonych 10 struktur. Protofilament rozciąga się w **strukturę czterowarstwą włókna** (reszty 12÷21 zaznaczone na czerwono, reszty 30÷40 na niebiesko, reszty 22÷29 na zielono, a niezaznaczone reszty 1÷8 stanowią fragmenty nieuporządkowane). Kolorowe strzałki oznaczają kontakty łańcuch boczny - łańcuch boczny oraz łańcuch boczny - łańcuch główny (peptydowy) na podstawie pomiarów sprzężeń dipolarnych ^{13}C - ^{13}C (purpurowe) i ^{13}C - ^{15}N (niebieskie) metodą **dipolar recoupling** (REDOR). Wartość **krytycznej koncentracji** C_r , w procesie elongacji/dysocjacji włókna odpowiada stałej dysocjacji K_d (w równowadze nie ma praktycznie zmiany molarnego stężenia włókna):

$$\text{K}_d = [\text{fibryl}_n][\text{monomer}] / [\text{fibryl}_{n+1}] = [\text{monomer}] = \text{C}_r$$

Wyznaczona stąd **energia swobodna Gibbsa wydłużania włókna** $\Delta G = -38 \text{ kJ/mol}$, czyli rzędu entalpii swobodnej foldingu niewielkich białek globularnych.

Badania innych złożeń amyloidalnych dały w wyniku szereg modeli tworzenia włókien [\[PLANSZA 138 & 139\]](#) (nie będę ich omawiał szczegółowo). Tworzenie amyloidu $\text{A}\beta$ zachodzi zgodnie z **modelem NP** (*nucleated polymerization*) i/lub NDP (*nucleation-dependent polymerization*; rys. C na [PLANSZY 139](#)). W ogólnym ujęciu, pokazane mechanizmy można w zasadzie sprowadzić do jednego schematu ([Jahn & Radford Arch. Biochem. Biophys. 469, 100, 2008](#)):

- **nukleacja**: częściowo zwinięte struktury pośrednie, które powstały z częściowo rozwiniętych form natywnych lub częściowo ustruktrowanych form całkowicie rozwiniętych;

- tworzenie **uporządkowanych agregatów** stosownie do różnych mechanizmów: stacking β -nici, asocjacja nici (sa, *strand association*), asocjacja krawędź-krawędź (ee, *edge-edge association*), wymiana domen (*domain swapping*);
- **self-asocjacja** agregatów połączona ze zmianami konformacyjnymi do włókien amyloidalnych.

Szczególne znaczenie przypisuje się mechanizmowi agregacji ***domain-swapping*** [PLANSZA 140], w którym wewnętrzne fragmenty nici peptydowej, pozbawione ochrony przed oddziaływaniem z innymi nićmi (tak jak fragmenty zewnętrzne) mogą wiązać się z innymi fragmentami wewnętrznymi sąsiadujących domen.

Pełne poznanie mechanizmów zwijania w przypadku konkretnych białek zakłada wyznaczenie ich lejków zwijania, czyli dokładnego obrazu **zbioru stanów (konformacji) przejściowych TSE** (*transition state ensemble*), od jego obrzeża ze strukturami rozwiniętymi do zakończenia lub ściślej zakończeń obrazujących prawidłowo zwiniętą, funkcjonalną formę natywną i/lub nieprawidłowo zwiniętą formę amyloidalną (lejek z kilkoma końcami; *multiple folding funnel*). Ta dziedzina badań procesu foldingu/refoldingu/misfoldingu zakłada ściśle **współdziałanie metod doświadczalnych**, dostarczających kluczowe informacje dla **mechanizmu zwijania**:

- struktury przestrzenne form natywnych, amyloidalnych i rozwiniętych;
- parametry kinetyczne (stałe szybkości/czasy zwijania);
- parametry termodynamiczne (ΔG° , ΔH° , ΔS° , ΔC_P°),
- Φ -values,
- wartości rzędów kontaktów CO,

oraz **symulacji komputerowych** procesu zwijania/rozwijania, w oparciu o znane struktury przestrzenne. Problem symulacji przebiegu foldingu do znanej struktury i przewidywanie *de novo* struktury na podstawie symulacji łączy wspólny problem złożoności struktury białka w roztworze, scharakteryzowanej dużą ilością stopni swobody. Szczegółowa problematyka modelowania struktur i symulacji komputerowych będzie omawiana na równoległym wykładzie „Bioinformatyka i modelowanie”. Podczas wykładu „Biofizyka doświadczalna” nacisk będzie położony na **współdziałanie i uzgodnienia eksperymentu z teorią**, co wydaje się warunkiem koniecznym uzyskania prawidłowego obrazu TSE.

Symulacje MD białek dla **modeli pełnoatomowych z uwzględnieniem rozpuszczalnika *explicite***, z czasami symulacji do **kilkuset μ s**, udaje się (jak na razie) tylko dla małych białek i domen [PLANSZA 141]. Rozwój *hardware* (superkomputer

ANTON), programów symulacyjnych (kody: GROMACS, NAMD) i prowadzenie szeregu niezależnych symulacji na wielu komputerach daje aktualnie możliwości osiągnięcia **czasu rzędu milisekundy dla białek 50 ÷ 100 AA w roztworze wodnym. Standardowym** jak na razie podejście polega na zastosowaniu::

- uproszczonych modeli białek, tzw. **modeli gruboziarnistych** (*coarse-grained models*)
- oszczędniejszych czasowo **próbekowań Monte Carlo** przestrzeni konformacyjnej TSE (*discrete molecular dynamics*) zamiast „ciągłych” symulacji MD.

W 1975 r. został zaproponowany, a następnie szczegółowo przedstawiony w 1981 r. w 2-częściowej publikacji [Go & Abe Biopolymes 20, 991-1011; 1013-1031](#) uproszczony model gruboziarnisty białka zwijanego metodą próbkowania statystycznego Monte Carlo na dyskretnej sieci z potencjałem oddziaływania średniej siły (*mean force*) **uwzględniającym jedynie kontakty natywne**, tzn. obserwowane w strukturze natywnej białka. Ujęcie to dało początek szeregowi modeli tzw. *Go-type lattice models*. Poszczególne fragmenty składowe („ziarna”) białka mają zwykle formę sfer (elipsoid) **centrowanych na C α i wiązaniu peptydowym reprezentujących łańcuch peptydowy oraz na węglu C β jako reprezentacja łańcucha bocznego. Potencjał oddziaływania** w powszechnych implementacjach modeli „Go-type” przyjmuje się w formie ([Paci et al. Biophys. J. 83, 3032, 2002](#)):

$$V^{\text{Go}} = \gamma \sum_{IJ} N_{IJ}^{\text{nat}}$$

N_{IJ}^{nat} - ilość kontaktów między atomami ciężkimi fragmentów białkowych I oraz J

γ - współczynnik proporcjonalności określany w procedurze fitowania.

Pomimo „mocnego” założenia ignorowania kontaktów nie-natycznych „modele Go” sprawdzają się w konstrukcji TSE, co potwierdzają testy porównawcze z zastosowaniem bardziej „realistycznych” potencjałów oddziaływań i dane doświadczalne. Stosowalność modeli Go przypisuje się dość powszechnie akceptowanej hipotezie, że powstałe w procesie ewolucji sekwencje białkowe wykształciły zdolność do zwijania na zasadzie „*minimal frustration*”, z redukcją udziału kontaktów wykluczających się (*in conflict*) w stanach przejściowych. Modele Go dobrze oddają proces zwijania obrazowany przez „gładkie” lejki, natomiast odstępstwa zachodzą gdy lejki są pofałdowane. W obrazie TSE pofałdowanie przypisuje się właśnie kontaktom typu nie-natycznego. W kilku przypadkach zaobserwowano wartości Φ większe od 1 lub mniejsze od 0, co przypisuje się słabszemu udziałowi tych mutowanych aminokwasów w strukturze natywnej przy ich dużym znaczeniu w strukturach przejściowych (*frustration*). Z oczywistych powodów modeli Go

nie można stosować w analizie procesów błędnego zwijania. Rozwijane są modele zwijania wychodzące poza proste ujęcie Go ([Clementi Curr. Op. Struct. Biol. 18, 10, 2008](#)):

- bardziej „realistyczne” **potencjały Lenarda-Jonesa 6-12** z uwzględnieniem członów odpowiedzialnych za wiązania wodorowe i efekty hydrofobowe;
- zwijanie bez użycia dyskretnych sieci „*off-lattice models*”
- **podejście wieloskalowe** (*multi-scale modeling approach*), zakładające łączenie stosowanie dokładniejszych symulacji MD ze zwijaniem w oparciu o modele gruboziarniste.

Ciekawą klasą stanowią białka zawierające **węzły w łańcuchu peptydowym**. Jak pokazują dane eksperymentalne i wyniki symulacji białka zawierające węzły zwijają się o ok. rząd wielkości wolniej niż podobne białka (podobna mapa kontaktów) bez zapętleń. Symulacje odwracalnego zwijania i zapętlania oraz rozwijania konstruktu białkowego 2ouf (*domain fusion*) z wykorzystaniem modelu gruboziarnistego i potencjałów oddziaływań 6-12 Lenarda-Jonesa [[PLANSZA 142](#)] pokazały, że „siłą napędową” zapętlania są natywne kontakty we wstępnie uporządkowanych intermediatach. Równoległe symulacje z odpowiednikiem niezapętlonym pokazały wyższe bariery energetyczne (dłuższe czasy zwijania) na drodze zwijania dla białka z węzłem.

Środowisko wodne odgrywa kluczową rolę w zwijaniu białek globularnych ([Mattos & Clark Arch. Biochem. Biophys. 469, 118, 2008](#)) nie tylko jako ośrodek umożliwiający utworzenie docelowej struktury natywnej, ale także poprzez **wpływ na „pejzaż energetyczny”, który określa się jako jego wygładzanie** (*minimizing frustration*). Białka nierozpuszczalne w wodzie, błonowe i włókniste mogą więc charakteryzować się innymi własnościami TSE i przebiegu procesu zwijania. Uzyskanie dokładnych struktur białek błonowych, kanałów i transporterów umożliwiło prowadzenie symulacji komputerowych foldingu i analizy TSE z zastosowaniem metod wieloskalowych (omówionych poprzednio) oraz metod mezoskopowych ([Khalili-Araghi et al. Curr. Op. Struct. Biol. 19, 128, 2009](#)). Symulacje wydają się być w istotny sposób zależne od przyjętego **modelu warstwy lipidowej** (*coarse-grained*) otaczającej białko. Drugim czynnikiem, aktualnie jeszcze niemożliwym do uwzględnienia wydaje się być skład błony; większość symulacji zakłada homogeniczną dwuwarstwę (pojedynczy lipid). Równoległe wzrasta ilość prac doświadczalnych dotyczących zwijania *in vitro* czyli refoldingu ([Stanley & Fleming Arch. Biochem. Biophys. 469, 46, 2008](#)) w roztworach micelarnych detergentów i dwuwarstwych pęcherzykach lipidowych, SUV (*small unilamellar vesicles*; 30 ÷ 50 nm)

lub LUV (*large unilamellar vesicles*; > 100 nm). Punktem startowym jest właściwy wybór czynników denaturujących, które różnią się w przypadku dwóch zasadniczych typów białek błonowych:

- białka typu β -baryłki: chemiczne denaturanty, mocznik, chlorek guanidyny
- białka typu wiązki helis: związki powierzchniowo czynne (*surfactants*), np. SDS (*sodium dodecyl sulphate*).

Folding białek błonowych **dzieli się zarówno *in vitro* jak i *in vivo* na dwa zasadnicze etapy** ([Bowie Nature 438, 581, 2005](#)) [[PLANSZA 143](#)]:

(1) **insercja w błonę**, ukierunkowana i katalizowana *in vivo* przez kompleks translokacyjny (*translocation complex*);

(2) **powstawanie struktur drugo- trzecio- i czwartorzędowych**.

Kinetyka foldingu *in vitro* jest bardzo silnie zależna od **grubości membrany i składu lipidowego**. Kluczowym zagadnieniem jest stworzenie warunków eksperymentalnych dla uniknięcia histerezy, czyli uzyskania pełnej odwracalności w procesach refoldingu i denaturacji: pokrycie się punktów doświadczalnych na drodze obu procesów. W rezultacie otrzymuje się charakterystyczny „*chevron plot*” [[PLANSZA 122](#)], podobnie jak w przypadku kinetyki zwijania białek globularnych. Podobnie jest także z możliwością tworzenia źle zwiniętych agregatów (*misfolding*), np. w warunkach ograniczonej ilości micelli przy zwijaniu białka OmpA (*outer membrane protein A*) pojawiają się dimery, trimery i oligomery wyższego rzędu zgodnie z mechanizmem wymiany domen (*domain swapping*) [[PLANSZA 139, mechanizm G](#)].

Najwięcej uwagi w **zwijaniu białek włóknistych** poświęcono różnym rodzajom kolagenu [[PLANSZA 20](#)], od I do XIX, głównemu składnikowi kości, skóry, ścięgien, błon np. w naczyniach krwionośnych. Kluczowym elementem w zwijaniu kolagenów są **domeny trimeryzacji** (*trimerization domains*), segmenty białkowe od 35 AA do 250 AA na N-końcu i/lub C-końcu i nie należące do łańcucha peptydowego [[PLANSZA 144](#)]. Domeny trimeryzacji **definiują punkt startowy tworzenia potrójnej helisy**, a po zwinięciu kolagenu są usuwane przez specyficzne enzymy. Brak domen trimeryzacyjnych powoduje możliwość nukleacji potrójnej helisy w różnych miejscach na zasadzie statystycznego spotkania łańcuchów, co prowadzi do misfoldingu kolagenu. Domeny trimeryzacji są przykładem mechanizmów wspomagających zwijanie *in vivo* (omawiane dalej).

Kontrolowane przez mechanizmy komórkowe *in vivo* **rozwijanie białek** pełni istotną rolę w trzech (przynajmniej) procesach ([Matouschek Curr. Op. Struct. Biol. 13, 98, 2003](#)):

- **degradacja** białek przez ATP-zależne proteazy;
- **translokacja** białek przez niektóre błony;
- **pasywne elastyczność** mięśni prążkowanych.

We wszystkich tych przypadkach mechanizm jest podobny, zależy od struktury i stabilności białka i w istotny sposób różni się od niekontrolowanego unfoldingu w komórce. Badanie rozwijania białek *in vitro* towarzyszy wielokrotnie badaniom procesu zwijania (omówione poprzednio), ale podejście skoncentrowane na samym procesie rozwijania jest oparte na kilku współpracujących ze sobą technikach ([Galera-Prat et al, M. Cieplak. Curr. Op. Struct. Biol. 20, 63, 2010](#)) i prowadzi do pojawienia się nowej dziedziny badań: **nanomechaniki białkowej** (*protein nanomechanics*):

- manipulacja pojedynczymi cząsteczkami: mikroskopia sił atomowych, szczypce optyczne, szczypce magnetyczne;
- symulacje dynamiki molekularnej MD;
- inżynieria białkowa

Przykładem może być połączona z mikroskopią sił atomowych spektroskopia sił molekularnych AFM-SMFS (*atomic force microscopy - single molecule force spectroscopy*) [\[PLANSZA 145\]](#). Symulacje dynamiki molekularnej wyjaśniają **molekularne podstawy mechanicznych sił, 150 ÷ 350 pN, stabilizujących struktury natywne** białek i charakterystyczny „zębaty” przebieg (*saw-tooth pattern*) zależności między siłą rozciągania (pN) i rozciągnięciem (nm) [\[PLANSZA 124\]](#).

Wykład 19

Zwijanie białek globularnych *in vivo* w komórce zachodzi przy dwóch elementach zasadniczo różnych od warunków zwijania *in vitro*.

(A) **Wysokie stężenie molekuł** (*macromolecular crowding*) wyznaczające ostrą granicę dla objętości dostępnej dla polipeptydowego łańcucha zwijającego się białka.

(B) **Udział mechanizmów komórkowych** podnoszących efektywność procesu zwijania. Zwijanie *in vivo* dotyczy białek dysponujących odpowiednią informacją w sekwencji do zachodzenia foldingu *in vitro* - pomoc, ale nie „wymuszanie” zwijania ze strony komórki.

„Zatłoczenie” makromolekularne [PLANSZA 146] wydaje się działać w kierunku **zwiększenia „upakowania” białek**, ale brak jest danych jednoznacznie wskazujących na jego dodatni/ujemny wpływ na folding. Wydaje się, że zależy to od **typu i własności** zwijającego się białka, **wielkości otaczających** makromolekuł oraz **lepkości otoczenia** molekularnego, która jest ważnym czynnikiem modyfikującym szybkości i mechanizm foldingu. Można wyróżnić szereg pomocniczych mechanizmów komórkowych.

(a) Ko-translacyjny **wpływ mRNA i rybosomu** w przypadku białek wielodomenowych (*for example: Netzer & Hartl Nature 388, 343, 1997*); przy szybkości translacji 50 ÷ 300 AA/minuta wystarcza czasu na (częściowy) folding;

(b) **Foldazy**, sekwencje PRO w niektórych protobiałkach działające jak wewnętrzne katalizatory foldingu przez stabilizację stanów przejściowych, np. *α-lytic protease* (*Sauter et al. Nature Struct. Biol. 5, 945, 1998*), czy omówione poprzednio domeny trimeryzacji w zwijaniu kolagenu [PLANSZA 144].

(c) **Reorganizacja mostków dwusiarczkowych** –S–S– (*shuffling*) między właściwymi parami cystein przez izomerazę dwusiarczkową (*protein disulfide isomerase*; *Wilkinson & Gilbert Biochem. Biophys. Acta 1699, 36, 2004*);

(d) **Izomeryzacja trans→cis** wiązań peptydowych poprzedzających od strony N-końca prolinę, przez izomerazę peptydylo-prolylową (*peptidyl-prolyl cis/trans isomerase*); *Fischer Angew. Chem. Int. Ed. Engl. 33, 1415, 1994*

(e) **Białka opiekuńcze** (*chaperons*) są białkami, które oddziałują z innymi białkami, stabilizują je lub pomagają w uzyskaniu funkcjonalnych konformacji natywnych. W cytoplazmie, reticulum endoplazmatycznym, mitochondriach i chloroplastach występuje szereg klas chaperonów tworzących kooperatywnie sieci złożone z tzw. **białek stresu albo szoku termicznego HSP** (*stress proteins, heat-shock proteins*). Ich klasyfikacja opiera się

na masie cząsteczkowej: HSP40, HSP60, HSP70, HSP90, HSP100 i małe HSPs. **Chaperony są zaangażowane w funkcje utrzymania proteomu, w tym:** folding *de novo*, refolding zdenaturowanych białek, asocjacja oligomerów, degradacja proteolityczna. Centralną rolę w foldingu i zapobieganiu agregacji odgrywa HSP70, który „sekwencyjnie współpracuje” z **systemami chaperoninowymi** (*chaperonins*): HSP wspomaga zwijanie nowo-syntetyzowanych (*nascent*) polipeptydów, a chaperoniny działają w końcowych etapach zwijania [PLANSZA 147]. Najlepiej poznane to systemy ATP-zależne: chaperonin klasy I **GroEL/GroES** z *E. coli* o znanej strukturze krystalograficznej ([Coyle et al. Current Biol. 2, R93, 1997](#)), omawiany na wykładzie „Struktura i funkcje makrocząsteczek biologicznych” i chaperonin klasy II **TriC** (archaea, eukaryota) ([Gutsche et al. J. Mol. Biol. 293, 295, 1999](#)). Dla białek opiekuńczych GroEL/GroES postuluje się dwa typy mechanizmów działania, ale problem nie jest do końca rozstrzygnięty:

- (1) model iteratywnego „zgrzewania” **IAM** (*iterative annealing model*) w periodycznych cyklach rozwijania niewłaściwych (*off-pathway*) intermedatów przez związanie i uwalniania do roztworu dla prawidłowego zwinięcia;
- (2) model klatki Anfinsena **ACM** (*Anfinsen cage model*), w której białko zwija się pod osłoną przed agregacją.

W aktualnych przeglądach zagadnienia preferowany jest obraz, w którym chaperony działają według schematu (1) uwalniając rozwinięte polipeptydy do roztworu dla prawidłowego zwinięcia, a chaperoniny do których środka wchodzi **białka o masie do 60 kDa**, działają według mechanizmu (2).

Dwuetapowy proces **zwijania *in vivo*** (oraz *in vitro*) **białek błonowych** był już sformułowany poprzednio [PLANSZA 143]. Składa się z częściowo nakładających się kroków zakotwiczenia segmentów TMS (*trans-membrane sequence*) sprzężonego z translacją (omawiana już topogeneza) z udziałem kompleksu RTC (*Sec61 translocon complex*) [PLANSZA 148]; różne warianty na przykładzie foldingu akwaporyn AQP.

Ciekawym przykładem udziału chaperonów w zwijaniu **białek włóknistych o strukturze β -helis** [PLANSZA 149] są tzw. **chaperony wewnątrzcząsteczkowe IMC** [PLANSZA 150] (*intramolecular chaperones*). Białka β -helikalne w formie równoległej β -spirali z jednej nici peptydowej, potrójnej β -spirali oraz potrójnej β -helisy z trzech owiniętych łańcuchów, występują w *tail fibers* bakteriofagów (T4), adenowirusie i reowirusie. Niektóre z nich zwijają się przy pomocy **domen trimeryzacji**, jak omówiony kolagen [PLANSZA 144], inne wymagają wyspecjalizowanych chaperonów. W strukturze

białka Phi29-TSP [\[PLANSZA 150\]](#), po N-końcowej wiązce α -helis i dwóch lewoskrętnych, równoległych β -helikalnych domenach, występuje C-końcowa domena chaperonowa IMC, dwie β -baryłki z pętlami ok. 50 AA. Folding Phi29-TSP kończy się **odcięciem IMC**, zgodnie z mechanizmem proteazy asparaginowej: nukleofilowy atak z utworzeniem tetrahedralnego intermediatu oksyanionowego kwasu glutaminowego.

Niewłaściwie zwinięte i agregujące białka tworzą złogi zewnątrzkomórkowe, które mogą być przyczyna **chorób amyloidozowych** (*amyloidoses*), z których najbardziej znane to: choroba Alzheimera, choroby wywołane prionami (Creutzfelda-Jacoba, wściekłych krów) oraz prawdopodobnie choroba Parkinsona [\[PLANSZA 151\]](#). Mimo dużej ilości nagromadzonych danych, przyczyny i mechanizm patogenicznego charakteru złogów amyloidalnych nie są jednoznaczne i do końca poznane. Raczej klasyfikuje się i postuluje różne ich własności, np. dla **prionów** [\[PLANSZA 152\]](#):

- efektywna replikacja i (*self*)propagacja;
- unikanie mechanizmów „oczyszczających” organizm z patogennych agregatów (degradacja, wydalanie, autofagia, oporność na działanie proteaz);
- zdolność do osiągnięcia docelowej tkanki;
- przenoszenie zmian fenotypowych.

Jedynie w oparciu o dobrze scharakteryzowane mechanizmy patogenicznego działania form amyloidalnych można poszukiwać odpowiednich środków farmakologicznych. Trwają ustawiczne wysiłki w tym kierunku. Przykładem mogą być próby skonstruowania cząsteczek małych, **farmakologicznych chaperonów** ([Arakawa et al. Biochem. Biophys. Acta 1764, 1677 2006](#)). Mają one charakter **osmolitów**, które podnoszą wewnątrzkomórkowe ciśnienie osmotyczne i wywołują wzrost stabilności białek bez wpływu na ich funkcjonowanie.

Struktury przestrzenne i zwijanie RNA; TRANSKRYPTOMIKA: mikromacierze i profile ekspresji, biosensory, sekwencjonowanie

Struktura trzeciorzędowa RNA jest bogata [\[PLANSZA 153\]](#) i przypomina struktury przestrzenne białek globularnych. Bogactwo struktur pokazały wyniki badań najpierw podjednostek a następnie całych rybosomów (odnośniki podane poprzednio):

- prokariotycznego z bakterii *Thermus thermophilus*,
- eukariotycznego z drożdży *Saccharomyces cerevisiae*.

Metoda dyfrakcji promieniowania X okazała się tu niezwykle przydatna, aczkolwiek **kluczowym problemem** była krystalizacja złożonego i **dynamicznego** układu z wykorzystaniem odpowiednio dobranych substancji stabilizujących. Przykładowo, w przypadku rybosomu eukariotycznego zastosowano na krótko pożywkę pozbawioną glukozy do wzrostu drożdży, co dało w wyniku zahamowanie inicjacji translacji i nagromadzenie **homogenicznej** populacji rybosomów. **Problem fazowy** rozwiązywano metodą podstawienia molekularnego (*molecular replacement*) wykorzystując fazy uzyskane z rozwiązania struktur podjednostek. Fazy dla podjednostek uzyskiwano z kolei z kombinacji metod izomorficznego podstawienia MIR (*multiple isomorphous replacement*) i rozpraszania anomalnego MAD (*multiwave anomalous diffraction*), np. wprowadzając na kompleks rybosomu drożdżowego pentaminę **osmu*** i heksaminy **irydu****.

Zastosowanie wielowymiarowych pomiarów NMR do badań struktur RNA w roztworze napotyka trudności w uzyskaniu **cząsteczek znakowanych** $^{13}\text{C}/^{15}\text{N}$. Standardowa, ale kosztowna procedura przygotowywania **znakowanego RNA** (e.g. Nikonowicz et al. Nucl. Acids Res. 25, 1390, 1997) polega na transkrypcji *in vitro* syntetycznej matrycy DNA z wykorzystaniem polimerazy T7 RNA oraz odpowiednio znakowanych nukleozydo-5'-trifosforanów NTPs. Znakowane NTPs uzyskuje się z hodowli bakteryjnej na wzbogaconej pożywce (^{15}N -siarczan amonu, ^{13}C -glukoza lub $^{13}\text{C}, ^2\text{H}$ -glukoza i 88% $^2\text{H}_2\text{O}$), po izolacji, hydrolizie RNA i chromatografii HPLC.

Dla krótkich fragmentów helikalnych RNA do ok. 20 bp, sekwencyjne przyporządkowanie metodami 2D ^1H NMR (van de Ven & Hilbers Eur. J. Biochem. 178, 1, 1988) opiera się na odległościach **poniżej 5 Å** pomiędzy protonami zasady C8–H (puryny) i C6–H (pirymidyny) a protonami cukru, w tym samym nukleotydzie i w nukleotydzie poprzedzającym w sekwencji, oraz H5'H5''/H1' dla układu nukleozyd/nukleozyd poprzedzający. Cross-peaks NOESY mają relatywne intensywności zależne od typu helisy. Trudniejsza sytuacja występuje w pętlach, gdzie dodatkowo wykorzystuje się **rezonans ^{31}P NMR** i metody heterojądrowe $^1\text{H}, ^{31}\text{P}$ NMR. **Przyporządkowanie fosforów** uzyskuje się poprzez sprzężenia z przyporządkowanymi protonami H5', H5'', H3' i H4' (2D heteroTOCSY, inaczej ^{31}P - ^1H J-cross polarization) oraz poprzez **selektywne wprowadzanie ^{17}O i/lub ^{18}O do reszt fosforanowych**. Izotop ^{17}O ma moment kwadrupolowy i wywołuje poszerzenie sygnału sąsiadującego jądra P, a

* pentamina: $\text{H}_3\text{C} - \text{CH}_2 - \text{CH}_2 - \text{CH}_2 - \text{CH}_2 - \text{NH}_2$

** heksamina (urotropina): 3 skondensowane "cykloheksany" z 4 azotami.

^{18}O daje jego przesunięcie *upfield* (niższe wartości ppm).. W miarę wzrostu masy cząsteczkowej fragmentów RNA metody dwuwymiarowe, homojądrowe ^1H NMR i heterojądrowe $^1\text{H}/^{31}\text{P}$ NMR okazywały się nieefektywne, chociaż starano się je rozszerzyć do homojądrowych 3D NMR, np. 3D TOCSY/NOESY oraz wykorzystując jako trzecią oś korelacje dwukwantowe 2Q dla skalarnie sprzężonych protonów w eksperymencie 3D NOESY/2QC ([Borer et al. Adv. Biophys. Chem. 6, 173, 1987](#)).

Dla większych, podwójnie $^{13}\text{C}/^{15}\text{N}$ - i $^2\text{H}/^{13}\text{C}/^{15}\text{N}$ -potrójnie znakowanych RNA starano się wykonać przyporządkowanie podobnie stosując metody 3D i 4D *triple resonance* (jak w białkach). Wykorzystuje się **metodologię TROSY** i/lub **koherencji wielokwantowych MQ** (*multiple-quantum*) w celu zwężenia szerokich linii wskutek relaksacji ([Fiala et al. J. Biomolec. NMR 16, 291, 2000](#); [Farmer II et al. J. Biomolec. NMR 4, 129, 1994](#)). Na podstawie eksperymentów *triple resonance* HCN, HCNCH i HC(N)C (z dodatkowymi subscriptami b = base, s = sugar) oraz $^1\text{H}, ^1\text{H}$ -NOESY, z rozciągnięciem wzdłuż osi ^{13}C (3D NOESY-HMQC), dokonuje się przyporządkowania, odpowiednio **wewnątrz-** i **międzynukleotydowego**. Niemagnetyczne jądro ^{16}O w łańcuchu **uniemożliwia sekwencyjne przyporządkowanie na podstawie korelacji skalarnych** jak w białkach. Dodatkowa trudność to mała dyspersja sygnałów na widmach. **Największą strukturą RNA** wyznaczoną techniką NMR jest jak dotąd **155n** (ok. 50 kDa) **region Ψ^{CES} "5'-leader" genomu HIV-1** [[PLANSZA 154](#)]. Do przyporządkowania sygnałów i wyznaczania odległości ^1H - ^1H zastosowano widma 2D NOESY wysoce deuterowanych próbek (^2H -*edited* NMR), selektywnie w stosunku do nukleotydów i/lub ich fragmentów (zasada, ryboza). Ponadto, rozwinięto "*fragmentation/annealing strategy*" niezależnej rejestracji widm fragmentów cząsteczki i ich niekowalencyjnych połączeń.

Łańcuch RNA tworzy elementy **struktury drugorzędowej**, jak łańcuch DNA:

- **helisy**: podwójne, potrójne (brak doniesień w literaturze) i poczwórne G-kwadrupeksy na 3'UTR mRNA ([Beaudoin & Perreault Nucl. Acids. Res. 38, 7022, 2010](#)) i w telomerycznym RNA (TERRA) z UUAGGG *repeats* (np. [Deng et al. Mol. Cell 35, 403, 2008](#)), gdzie dwa kwadrupeksy prawdopodobnie tworzą **strukturę ośminiciową** z naprzemiennym stackingiem kolejnych czwórek zasad G, A i U (podobnie jak i-motif DNA; [[PLANSZA 60](#)]) ([Xiao et al. I. Am. Chem. Soc. 2016 submitted/w recenzji](#));
- **szpilki do włosów**: siedmionukleotydowe, np. pętla antykodonu tRNA [[PLANSZA 155](#)], czteropętla (*tetra-loop*) [[PLANSZA 156](#)] 5'-GNRA, UNCG, GANC, AGNN (N - dowolny nukleotyd, R = G lub A) w intronach grupy I, rRNA, retrowirusach oraz trzyopętla UNR;
- **3'-zwisające końce** (*dangling ends*) [[PLANSZA 157](#)], nie wymagające terminacji nici;

- **pętle wewnętrzne**, w tym wybrzuszenia (*buldge*), z często występującą kanoniczną 3-nukleotydową formą **k-turn** [PLANSZA 158]

Oddziaływania elementów drugorzędowych prowadzą do powstania zestawu charakterystycznych **motywów strukturalnych**, które warunkują globalne zwinięcie nici polirybonukleotydowej **na poziomie trzeciorzędowym** (omówione częściowo na wykładzie „*Struktura i funkcje makrocząsteczek biologicznych*):

- **pseudowęzły** różnych typów zależnie od pętli, z którą paruje (ww W.C.) pojedyncza nić; typ H (szpilka do włosów, *hairpin*), B (wybrzuszenie; *buldge*), I-type (pętla wewnętrzna; *inner loop*), M-type (pętla wielogałęziowa; *multibranched*), H-H (dwie pętle *hairpin*; *kissing loops*) [PLANSZA 159], oraz pseudowęzły podwójne jeden w drugim (*nested*);

- **tRNA** [PLANSZA 153] jako odrębny typ struktury trzeciorzędowej, aczkolwiek topologicznie odpowiadający podwójnemu pseudowęzłowi [PLANSZA 16];

- **złącza helis**, od dwóch w tRNA do dziesięciu [PLANSZA 160];

- **motywy A-minor** typu I i II [PLANSZA 161];

- **zamek rybozowy** [PLANSZA 162];

- **Z-kotwice**, jakby „reminiscencja” struktury P-DNA [PLANSZA 163; I(ii) oraz IC];

- motyw **RA** (*right angle motif*) [PLANSZA 164]: dwa motywy G:A minor (*blue*) stabilizowane przez oddziaływania *along-groove* (*pink*), w różnych fragmentach 23S i 16S rRNA oraz złącza P2.1-P3-P8 rybozomu *Tetrahymena*;

- **platforma adenozynowa** [PLANSZA 165];

Motywy trzeciorzędowe są stabilizowane przez oddziaływania:

- **wiązania wodorowe**, typowe W.C. i nietypowe: *wobble*, np. G:U, Hoogsteen (dwójki i trójki zasad), odwrócone W.C., puryna-puryna, A⁺:C (tRNA^{Lys}) często z udziałem **nukleotydów modyfikowanych**, oraz wiązania wodorowe z udziałem grup 2'OH rybozy,

- **stacking**: „typowy” W.C., interkalacyjny oraz koaksjalny na złączach dwóch helis. Nukleozydy modyfikowane w tRNA ([Duechler et al. Cell Mol. Life Sci. 73, 3073, 2016](#)) pełnią nie tylko rolę stabilizacyjną, ale także regulacyjną w syntezie białek. Przykładowo, modyfikacje zasad w obszarze antykodonu wpływają na powinowactwo parowania kodon-antykodon, a stąd na szybkość translacji (*fine-tuning mechanism*).

Strukturalna zależność funkcjonowania cząsteczek RNA jest w znacznym stopniu uwarunkowana przez **dynamikę** w szerokim zakresie czasowym (stałych szybkości), od pikosekund (10^{-12} s) do ponad 10^3 s (minuty) [PLANSZA 166]:

- **lokalne ruchy** molekularne: drgania wiązań i kątów oraz fluktuacje kątowe o większej amplitudzie łańcucha fosfodiesterowego i przejść (*repuckering*) N ↔ S pierścieni cukrowych w skali piko- i nanosekund;

- **przekształcenia konformacyjne** (*conformational dynamics*): otwieranie par zasad w strukturach helikalnych i ruchy wzajemne poszczególnych fragmentów (domen) od nanosekund do milisekund (10^{-8} s ÷ 10^{-2} s), oraz „*riboswitching*” aptamerów RNA przy wiązaniu ligandów i globalne przekształcenia konformacyjne całych cząsteczek w skali ponad ~1s;

- **zwijanie** RNA w skali 10^{-5} s ÷ 10^2 s (podobnie jak w przypadku białek).

Struktura i dynamika RNA są (podobnie jak w przypadku innych biomakromolekuł) kluczowe dla regulacji i funkcjonowania genów. Początkowe oparcie **analizy transkryptomów** na określaniu struktur pierwszorzędowych i profili ekspresji uległo znacznemu poszerzeniu na skutek rozwoju metod doświadczalnych określania struktur przestrzennych i metod projektowania molekularnego - strukturalna organizacja transkryptomu eukariotycznego ([*Wan et al. Nature Rev. 12, 641, 2011*](#))

Wykład 20

Metabolizm RNA powstającego w procesie transkrypcji zapoczątkowuje zwijanie (*RNA folding*), które *in vivo* jest bezpośrednio i w różnym stopniu związane z **szeregiem procesów**, zależnie od typu RNA ([Fürtig et al. Biopolymes 86, 360, 2007](#)):

- **transkrypcja** z szybkością $20 \div 80$ n/s;
- „**dojrzewanie**”, czyli ostateczne ukształtowanie cząsteczek RNA w procesach splicingu (self-splicing) oraz wpływ exonów jako self-chaperonów (*exon context*), np. w foldingu rybozymów grupy pierwszej *Tetrahymena*;
- **translacja** i synchronizacja jej szybkości (~ 10 wiązań peptydowych /s) z transkrypcją w przypadku mRNA, ponieważ powstający (*nascent*) łańcuch mRNA jest bezpośrednio procesowany przez rybosom;
- **udział białek** w zwijaniu ([Doetsch et al. FEBS J. 278, 1634, 2011](#)): **chaperony** (*annealer proteins*) wiążące RNA w kompleksy destabilizują przejściowe struktury RNA, przykładowo **helikazy** RNA (*DEAD-box proteins*) **[PLANSZA 167]** przyspieszają rozwijanie wielu RNA przy przekształceniach konformacyjnych, które zachodzą w RNA i RNP.

Wymienione czynniki mają na celu **przyspieszenie foldingu RNA, które *in vivo* jest ok. 10-krotne**. Silne pofałdowanie lejków zwijania RNA prowadzi *in vitro* do **pulapek kinetycznych** i wydłużania czasów zwijania **do sekund i minut**. Badania w zakresie szerokiej problematyki „RNA folding” obejmują podobnie jak w przypadku zwijania białek **trzy zasadnicze obszary**:

- (1) zarysowany właśnie obszar badań **foldingu *in vivo***,
- (2) teoretyczne **przewidywanie struktur** przestrzennych RNA na podstawie sekwencji nukleotydowej przy użyciu metod komputerowych **[PLANSZA 168]**,
- (3) badania **doświadczalne dróg zwijania** cząsteczek RNA o znanej strukturze przestrzennej i próby konstrukcji odpowiednich **modeli teoretycznych** „pejzaży energetycznych” **[PLANSZA 169]**.

Modelowanie struktur przestrzennych RNA **[PLANSZA 168]** pozostaje w tyle za modelowaniem białek globularnych, co wynika z wpływu kilku (zasadniczych) czynników:

- (a) większa ilość parametrów opisu konformacji łańcucha polinukleotydowego w porównaniu z peptydowym: 9 kątów dwuściennych: $\alpha \div \zeta$ i χ , kąt pseudorotacji P i

pofałdowanie ν_{\max} w porównaniu z dwoma kątami Φ i Ψ , plus kąty w łańcuchach bocznych (mniejsze znaczenie dla konformacji łańcucha głównego);

(b) czułość struktur RNA na obecność jonów, przede wszystkim Mg^{+2} , wpływ składu solwentu i obecność innych biomolekuł;

(c) zdefiniowanie struktury drugorzędowej przez wiązania wodorowe W.C., relatywnie silniejsze niż wiązania specyficzne dla struktur 3D i pozostałe kontakty van der Waalsa.

Hierarchiczny charakter zwijania RNA (punkt c) znajduje również odzwierciedlenie w podejściu teoretycznym, gdzie przewidywanie struktur drugorzędowych ([methodological review: Mathews J. Mol. Bio. 359, 526, 2006](#)) osiągnęło znaczny (zadawalający ?) postęp i nie wymaga znajomości wpływu struktur trzeciorzędowych. Szczególne znaczenie miało rozwinięcie **algorytmów dynamicznego programowania** (*dynamic programming*), które umożliwiają sprawdzenie możliwych struktur drugorzędowych (10^{25} dla sekwencji 100 n) bez ich generowania i zastosowanie tych algorytmów do **efektywnego próbkowania** zespołu (suboptymalnego) Boltzmannna struktur drugorzędowych. Zamiast wyboru par zasad generuje się je probabilistycznie zgodnie z zespołem statystycznym dla wszystkich możliwych fragmentów sekwencji.

Przewidywanie **struktur trzeciorzędowych**, chociaż intensywnie rozwijane przez wiele grup badawczych przy pomocy różnych metod, ciągle **nie jest satysfakcjonujące**. Opracowane algorytmy różnią się danymi wejściowymi (*input data*), dokładnością przewidywań i wykorzystują szeroka gamę reprezentacji cząsteczki RNA, od jednego pseudoatomu reprezentującego cały nukleotyd lub grupę atomów (*coarse-grained models*; iFoldRNA) do reprezentacji pełnoatomowej (*all-atom representation*; FARFAR, NAST, MC-SYM) **[PLANSZA 168]**. Różne wersje algorytmów to przewidywanie *ab initio*: z wykorzystaniem struktur drugorzędowych, z wykorzystaniem kontaktów trzeciorzędowych i/lub homologii sekwencji nukleotydowej. Opracowany w zespole kierowanym przez prof. Ryszarda Adamiaka (Laboratorium Chemii Strukturalnej Kwasów Nukleinowych Instytutu Chemii Bioorganicznej PAN w Poznaniu) program **RNAComposer**, dostępny na dwóch serwerach:

<http://rnacomposer.cs.put.poznan.pl/>

<http://rnacomposer.ibch.poznan.pl/>

pozwala przewidywać struktury 3D RNA długości do 500n na podstawie zdefiniowanej przez *usera* struktury drugorzędowej (RNA FRABASE *database*). Szerzej rozwijana grupa metod opiera się na założeniu, że oddziaływania stabilizujące strukturę trzeciorzędową nie

definiują tej struktury, ale raczej selekcjonują specyficzne konformery z topologicznie predefiniowanego rozkładu konformerów (*topological constraints*), opartego na strukturze drugorzędowej ([Bailor et al. Curr. Op. Struct. Biol. 21, 296, 2011](#)).

Badania kinetyki, termodynamiki i dróg zwijania [PLANSZA 169] wykorzystują omawiane już podejścia analogicznie jak w białkach ([reviews: Chen Annu. Rev. Biophys. 37, 197, 2008; Tinoco & Bustamante J. Mol. Biol. 293, 271, 1999](#)).

(A) **Metody doświadczalne:** wielowymiarowe techniki NMR, SAXS, analiza mutacyjna i wyznaczanie Φ -values, fluorescencja z wykorzystaniem sond takich jak piren, badania *single-molecule*, mapowanie dostępności powierzchni RNA dla rodników $\cdot\text{OH}$ z radiolizy wody synchrotronowym promieniowaniem X (*synchrotron hydroxyl radical footprinting*) i szereg innych;

(B) **Teoretyczne analizy „energetic landscapes”:** symulacje dynamiki molekularnej i statystyczne próbkowanie zbioru mikrostanów przejściowych (TSE)

Ogólnie biorąc, skala czasowa hierarchicznego foldingu/refoldingu RNA obejmuje najpierw tworzenie struktur drugorzędowych, od ok. 10 do 100 μs , a następnie trzeciorzędowych, z czasami dochodzącymi do minut, przy refoldingu, wolniejszym niż folding. Czasy wydają się w niewielkim stopniu zależeć od sekwencji nukleotydowej. Modelową cząsteczką RNA badaną zarówno doświadczalnie jak i teoretycznie jest **rybozym *Tetrahymena***. **Cechy charakterystyczne procesu zwijania RNA:**

- silna zależność od jonów Mg^{+2} ;
- **brak addytywności** entalpii swobodnej w strukturach trzeciorzędowych (szczególnie pseudowęzły);
- **pofaldowany lejek zwijania** z licznymi stanami pośrednimi (pułapki kinetyczne) nawet dla małych cząsteczek;
- względnie szybkie tworzenie zwartej, ale niefunkcjonalnej struktury przypominającej formę natywną, **analogicznej do molten globule** w białkach [PLANSZA 170; Ic].

Cząsteczki RNA ulegają w komórce **złożonym procesom metabolicznym**, począwszy od „dojrzewania”, a zakończonych procesem degradacji, głównie przez egzorybonukleazy. Ważnym elementem funkcjonowania kwasu rybonukleinowego jest **lokalizacja cząsteczek RNA w komórce**. Przykładowo, wewnątrzkomórkowa lokalizacja mRNA (lokalizacja transkryptów), m. in. ich aktywny transport jest ważnym mechanizmem **kierowania białek**, obok omówionych peptydów sygnałowych, do miejsc ich działania i odgrywa kluczową rolę od wyznaczania osi ciała do wyznaczania

plastyczności synaps ([Meignin & Davis Curr. Op. Struct. Biol. 22, 112, 2010](#)). Kontrolowanie szybkości **degradacji RNA** jest jednym z kluczowych procesów regulacji ekspresji genów w komórce eukariotycznej. Trawienie RNA może zachodzić od obu końców:

- degradacja 5' → 3' przez enzymy z rodziny Xrn;
- degradacja 3' → 5' przez kompleks **eksosomu** (*exosome*), który bierze udział w **licznych drogach metabolizmu RNA** [[PLANSZA 171](#)].

Jednym z podstawowych zadań transkryptomiki jest pełna **charakterystyka puli funkcjonujących cząsteczek RNA, ulegających ekspresji** w wyspecjalizowanych komórkach organizmów. Początkowo badania w tym zakresie oparte były na technologii **mikromacierzy** (*microarrays*) wykorzystującej hybrydyzację. W ciągu kilku ostatnich lat mikromacierze są wypierane przez tzw. **technologie RNA-seq** ([Ozsolak & Milos Nature Rev. 12, 87, 2011](#)) sekwencjonowania komplementarnego **cDNA** (*copied DNA*) z **wykorzystaniem przede wszystkim platform nowej generacji NGS** (*next-generation sequencing*), które były omówione poprzednio (Roche 454, Illumina Hi-Seq, Life Technologies, Pacific BioSciences).

Mimo wszystko, **mikromacierze** w postaci chipów DNA i chipów białkowych pozostają niezwykle interesującym przykładem biologicznego i medycznego wykorzystania **wysokiej specyficzności rozpoznawania wzajemnego biomolekuł**. W szczególności mikromacierze białkowe są stosowane do ilościowej charakterystyki białek w proteomice [[PLANSZA 73](#)]. Przytaczam w skrócie podane na wykładzie „*Struktura i funkcja makrocząsteczek biologicznych*” szczegóły techniczne dotyczące mikromacierzy. W przypadkach obu typów mikromacierzy łańcuchy polinukleotydowe jednoniciowego DNA lub łańcuchy peptydowe są unieruchamiane jednym końcem na powierzchni (szklanej, krzemowej, złotej, polimeru) przez odpowiednio dobraną cząsteczkę linkera. Każde miejsce próbkujące, tzw. *probe site* lub *spot* zawiera jednakowe cząsteczki biopolimeru (ta sama sekwencja) 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. Potoczna nazwa „chip” bierze się stąd, że pierwsze tego typu obiekty były wykonywane techniką fotolitograficzną jak w produkcji mikroprocesorów komputerowych (*computer chips*). **Chipy DNA** ([review: Pirrung Angewandte Chemie International Edition 41, 1276, 2002](#)) są złożone z **syntetycznych oligomerów** długości do 60n lub **cDNA**

produktów PCR do 300 nukleotydów, które mogą **wiązać komplementarnie** pojedyncze nici DNA lub RNA. 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 w transkryptomice** (*gene-expression profiling*): które geny w tkance uległy ekspresji.

Przykład drugiego zastosowania pokazuje testowanie zmian ekspresji genów np. jako reakcja na podanie leku. Chip zawiera DNA fragmentów **tysiący genów**, po jednym na każdy spot. Po pobraniu dwóch próbek komórek, 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" **[PLANSZA 172]**:

- silny wzrost aktywności genu po podaniu leku - czerwony,
- silny spadek aktywności po podaniu leku - zielony,
- geny jednakowo aktywne („złożenie” kolorów) - żółty,
- brak aktywności danego genu w obu przypadkach - bez koloru.

Dysponując dodatkowo różnymi **profilami reakcji dla znanych toksyn** można porównać z nimi profil testowanego leku i szybko wykluczyć go na podstawie podobieństwa profili (kolorowego schematu) i/lub wytypować „kandydatów” nietoksycznych. Innym przykładem zastosowania *gene-expression profiling* jest "prognozowanie" rozwoju nowotworów (na podobnej zasadzie). Wydajność hybrydyzacji jest na tyle duża, że pozwala na charakteryzowanie indywidualnych (osobniczych) różnic w sekwencjach genomów **SNP** (*single nucleotide polymorphism*).

Chipy białkowe ([review: Zhu & Snyder Current Opinion in Chemical Biology 7, 55, 2003](#)) 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 w próbce krwi. Dodanie **fluorescencyjnie znakowanych przeciwciał** w stosunku do innych fragmentów tych białek ujawnia zwiążanie się antygenu i przeciwciał na chipie przez kompleks typu "sandwich".

Zabarwienie określonego miejsca na chipie sugeruje wystąpienie we krwi antygenu pochodzącego od jakiegoś czynnika patogennego. W wersji tzw. **funkcjonalnych chipów białkowych** wprowadza się na chip same białka a nie przeciwciała.

Struktury typu mikromacierzy (ale nie tylko) mogą być również stosowane w rozszerzonej/uzupełniającej wersji do „konwencjonalnych” mikromacierzy DNA i mikromacierzy białkowych jako **biosensory**, używane do różnego typu detekcji. „*Biosensing*” obejmuje wykorzystanie różnych technik dla wygenerowania **odczytu (read-out) oddziaływań międzycząsteczkowych między (nano)układem molekularnym biosensora MRE (molecular recognition element) a interesującą substancją (analyte), lub odczytu aktywności enzymatycznej**. Biosensory są więc jakby „nosami chemicznymi” wyczuwającymi określone substancje. Jako MRE stosuje się:

- **sondy DNA** oparte na strukturze trójwymiarowej nano-DNA do detekcji (przykładowo) miRNA, białek (trombina), kokainy, oraz jako immunosensory elektrochemiczne i sensory do badań energetycznych i procesów przeniesienia ładunku (CT) **[PLANSZA 173]**;
- **RNA aptamery-przełączniki (riboswitch)**, np. w allosterycznie sterowanych przez liganady rybozymach (*aptazymes*; **[PLANSZA 174]**);
- **białka**, np. z wykorzystaniem efektu FRET (*fluorescence resonance energy transfer*) między donorem i akceptorem, np. mutantami GFP **[PLANSZA 175]**), do detekcji aktywności enzymatycznej, zmian konformacyjnych MRE po związaniu *analyte* oraz oddziaływań białko-białko;
- **inne**: związki **niskocząsteczkowe** i krótkie peptydy ([Fong et al. Curr. Op. Struct. Biol. 16, 234, 2012](#)), układy **nano**, np. fluoryzujące polimery, nanocząsteczki złota ([Miranda et al. Curr. Op. Struct. Biol. 14, 728, 2010](#)).

Technologie RNA-seq w transkryptomice oparte są na dwóch sposobach sekwencjonowania RNA.

(1) Sekwencjonowanie **klonowanego, podwójnie-niciowego cDNA** z wykorzystaniem platform NGS. Wymagana jest tu strategia uzyskiwania informacji o sekwencji właściwej nici DNA (*strand-specific information*), odpowiadającej pojedynczej nici RNA (co najmniej trzy metody), np. przez selektywne znakowanie markerami drugiej nici cDNA w syntezie.

(2) **Bezpośrednie sekwencjonowanie RNA** (DRS, *direct RNA sequencing*) w przypadkach trudności z zastosowaniem techniki sekwencjonowania cDNA.

Aktualnie, zdecydowana większość prac badawczych w zakresie transkryptomiki jest prowadzona z wykorzystaniem **sekwencjonowania cDNA [PLANSZA 176]**, z szerokim

zastosowaniem **algorytmów bioinformatycznych** do rozwiązania kluczowego problemu uszeregowania (*alignment*) uzyskanych sekwencji (*reads*) do genomu:

- już zsekwencjonowanego i stanowiącego punkt odniesienia;
- o nieznannej sekwencji DNA: charakterystyka *de novo* transkryptomu.

Rozwój metod RNA-seq jest związany z pokonaniem trudności specyficznych dla RNA. Mało wiadomo o heterogeniczności ekspresji małych RNA (*small-RNA profiling*). [Faridani et al. Nature Biotechnology 34, 1264, 2016](#) prezentuje metodę sekwencjonowania *single-cell* transkryptomu małych RNA i jej zastosowania do separowania różnych typów komórek, a stąd do analizy heterogeniczności tkanek nowotworowych. W **neurobiologii** charakterystyka aktywności różnych typów komórek nerwowych (*what happens where*) wymaga metod profilowania ekspresji genu w miejscach powstawania, czyli uzyskanie *spatial transcriptomic information* ([Marx Nature Methods 14, 1037, 2017](#)): *in situ* fluorescencja, hybrydyzacja (FISH) i sekwencjonowanie RNA; w perspektywie: atlasy komórek różnych tkanek i ich funkcji.

Modyfikacje chemiczne rybonukleotydów są aktualnie oceniane na ok. 100 i występują w tRNA, rRNA, snRNA (*noncoding RNAs*). Ich wpływ na metabolizm i funkcjonowanie RNA określa się jako **epitranskryptom** przez analogię do epigenetycznych regulacji poprzez modyfikacje DNA i histonów. Osiągnięcia technologii sekwencjonowania transkryptomu pod kątem modyfikowanych nukleotydów (*epitranscriptomic marks*): [Li et al. Nature Methods 14, 23, 2016](#). Równoległe sekwencjonowanie genomu (DNA) i transkryptomu (RNA) pojedynczych komórek jest bardzo użytecznym narzędziem do ustalania zależności między zróżnicowaniem genetycznym, epigenetycznym (metylacja DNA) i transkrypcyjnym. Przykładowo, protokół **scM&T-seq** ([Angermueller et al. Nature Methods 13, 229, 2016](#)) umożliwił, poprzez równoległe profile metylomu i transkryptomu (średnio 2,7 mln *reads* na komórkę), wyznaczenie zależności między metylacją DNA (516 genów) i transkrypcyjną heterogenicznością w embrionalnych komórkach macierzystych (*embryonic stem cells*), w których zachodzi stochastyczne przełączanie między stanami transkrypcyjnymi. Sekwencjonowanie RNA umożliwia **próbkiowanie *in vivo* struktur drugorzędowych RNA**. Opracowana w tym celu metoda **DMS-MaPseq** (*dimethyl sulphate mutational profiling with sequencing*; [Zubradt et al. Nature Methods 13, 75, 2016](#)) wykorzystuje chemiczne modyfikacje niesparowanych nukleotydów DMS-em w pociętym RNA jako *mismatches* dla odwrotnych transkryptaz, które syntetyzują cDNA do sekwencjonowania.

Kompleksy biomolekularne: kinetyka i dynamika asocjacji

Wykład 21

Pierwszym i zasadniczym etapem w wyznaczaniu mechanizmów specyficznej asocjacji cząsteczek jest charakterystyka **równowagi kompleks-składniki i kinetyki procesu wiązania**, czyli wyznaczenie **parametrów** charakteryzujących te procesy:

- (a) **stechiometria** składników w kompleksie;
- (b) **równowagowa stała asocjacji** K_{as} , która wiąże się bezpośrednio z **entalpia swobodną tworzenia** kompleksu **[PLANSZA 177]**:

$$\Delta G^\circ = -RT \ln K_{as},$$

- (c) **stałe kinetyczne** wiązania i rozpadu kompleksu, k_{on} i k_{off} (patrz kinetyka zwijania białek), które wiążą się z **entalpiami swobodnymi bariery** przejścia (w obie strony) zgodnie z równaniem Eyring'a:

$$k = (kT/h) \exp(-\Delta G^\ddagger/RT)$$

przy czym zachodzi zależność

$$K_{as} = k_{on}/k_{off}$$

Temperaturowe zależności K_{as} oraz k_{on} i k_{off} pozwalają na wyznaczenie **parametrów termodynamicznych**: entalpii ΔH° i entropii ΔS° asocjacji oraz entalpii ΔH^\ddagger i entropii ΔS^\ddagger barier. Z nieliniowego w ogólnym przypadku równania van't Hoffa otrzymuje się dla pomiarów równowagowych zależne od temperatury ΔH° , ΔS° i ew. ciepło właściwe asocjacji w stałym ciśnieniu p (*Ha et al. J. Mol. Biol. 209, 801, 1989; Sturtevant Proc. Natl. Acad. Sci. USA 74, 2236, 1977*):

$$[\partial \ln(K_{as})/\partial(1/T)]_p = \Delta H^\circ/R$$

$$\Delta G^\circ = -RT \ln K_{as} = \Delta c_p^\circ [(T - T_H) - T \ln(T/T_S)]$$

$$\Delta H^\circ = \Delta c_p^\circ (T - T_H) \quad \text{dla } T = T_H \quad \Delta H^\circ = 0$$

$$\Delta S^\circ = \Delta c_p^\circ \ln(T/T_S) \quad \text{dla } T = T_S \quad \Delta S^\circ = 0$$

Temperatury T_H i T_S wyznaczają **zakresy "napędu" procesu asocjacji**, $\Delta G^\circ < 0$ przez entropię (*entropy-driven, enthalpy-opposed*), przez entalpię (*enthalpy-driven, entropy-opposed*) i przez oba czynniki (*enthalpy and entropy-driven*). Tworzeniu dużych kompleksów białkowych (i ogólnie makromolekularnych) towarzyszy multum oddziaływań dziesiątek aminokwasów i **tysiące Å² powierzchni kontaktu** (Δc_p°).

Analiza kompleksów na poziomie molekularnym (w tym projektowanie farmaceutyków) napotyka na trudności związane z określeniem **oddziaływań** zapewniających wysokie powinowactwo. Z termodynamicznego punktu widzenia kluczową rolę odgrywa czynnik entropowy, a szczególnie znaczny wkład **entropii konformacyjnej**. Rozkład całkowitej energii swobodnej Gibbsa wiązania na składniki:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ = \Delta H^\circ - T\Delta\{S^\circ_{\text{protein}} + S^\circ_{\text{ligand}} + S^\circ_{\text{solvent}}\}$$

zawiera w czynniku entropowym białka S°_{protein} entropię konformacyjną S°_{conf} oraz entropię translacyjno-rotacyjną S°_{rt} . **Problem z wkładami entropowymi, w przeciwieństwie do entalpowych**, wynika z nieuprawnionego dzielenia entropii na addytywne wkłady; ściśłość relacji symbolizuje zapis z Δ przed całkowitym czynnikiem entropowym. Próbuje się w przybliżeniu określać i sumować wkłady ΔS°_α , ($\alpha = \text{protein, ligand, solvent}$), ale ich obliczanie i doświadczalne wyznaczenie jest trudne. Interesujące publikacje grupy z Univ. Pensylwania ([Marlow et al. Nature Chem. Biol. 6, 352, 2010](#); [Wand Curr. Op. Struct. Biol. 23, 75, 2013](#)) postulują **liniową korelację** entropii konformacyjnej białek S°_{conf} i całkowitej entropii asocjacji S° . Z kolei entropię konformacyjną wyznacza się zastępczo (*proxy*) na podstawie parametrów dynamiki ruchów molekularnych. Wymaga to kalibracji tzw. **miernikiem entropii** (*entropy meter*). Autorzy wykorzystali tu **liniową relację** między całkowitą entropią asocjacji z pomiarów metodą ITC dla serii kompleksów tego samego białka oraz parametrem uporządkowania S^2 (*order parameter in model-free approach*) rotacji grup C^2H_3 białka z czasów relaksacji 2H NMR.

Dla **asocjacji wieloskładnikowych** zachodzących **po różnych drogach** "reakcji" określa się mikroskopowe stałe asocjacji oraz stałe szybkości dla każdego procesu składowego. **Obserwowane** stałe **makroskopowe** wiążą się ze stałymi **mikroskopowymi** relacjami zależnymi od mikroskopowego **modelu asocjacji**. W weryfikacji modeli stosuje się metody **analizy statystycznej**, np. testowanie hipotez dopasowania modelu do danych **testem F-Snedecor'a** ([S. Brandt Analiza danych, PWN 1999](#)) lub na podstawie informacyjnego kryterium **Akaike** ([IEEE Transactions on Automatic Control AC-19, 716, 1974](#)). **PROBLEM** jaki pojawia się na samym wstępie wynika z konieczności **dofitowania** (nieliniowego) zależności funkcyjnej do zarejestrowanych punktów doświadczalnych [**PLANSZA 178**] dla odpowiedniego **modelu stechiometrycznego** (**równowaga**) lub **modelu sekwencji etapów asocjacji** (**kinetyka**), a z drugiej strony dane doświadczalne stanowią główną **przesłankę do sformułowania** takiego modelu. Wyjściem

z tej sytuacji jest **numeryczne całkowanie** odpowiednich równań na stężenia równowagowe czy przebiegi czasowe procesów asocjacyjnych np. z użyciem programu **DYNAFIT**: symboliczna charakterystyka stechiometrii/reakcji (na wejściu) połączone z nieliniową regresją i weryfikacją statystyczną (*Kuzmič Analytical Biochem. 237, 260, 1996*).

Wyznaczanie zarówno **stałych asocjacji** K_{as} jak i **stałych kinetycznych** k_j wymaga zastosowania metody opartej na zmianach mierzalnego parametru pod wpływem tworzenia kompleksu. Do doświadczalnego wyznaczania wartości równowagowych, **stężeniowych stałych asocjacji** K_{as} stosuje się **miareczkowanie** jednego składnika kompleksu drugim składnikiem o wzrastającym stężeniu w obecności ustalonych stężeń lub braku pozostałych składników, jeśli kompleks jest więcej niż dwuskładnikowy. **Termodynamiczna, równowagowa stała asocjacji** K_{as} (T,p), zależna tylko od temperatury i ciśnienia, wymaga określenia aktywności kompleksu, jego składników i wszystkich pozostałych składników reakcji wiązania. Wyznaczane eksperymentalnie równowagowe stałe stężeniowe K_{as} mogą zależeć od dodatkowych parametrów, takich jak pH, siła jonowa, itp. Standardowe metody miareczkowania, z których wymieniam najpopularniejsze, są oparte na rejestracji zmian mierzalnych parametrów pod wpływem **wzrostu stężenia całkowitego** [L] jednego ze składników (titranta) przy ustalonych stężeniach pozostałych.

(A) **Metody spektroskopowe**: absorpcja, CD, NMR, EPR (przegląd technik analizy wiązania ligandów w review: Meyer & Peters Angew. Chem. Int. Ed. 42, 846, 2003) i fluorescencja. W eksperymencie miareczkowania należy uwzględnić wszystkie czynniki wpływające na pomiar. Pomiar **wygaszanie fluorescencji** makromolekuły pod wpływem przyłączania ligandu, np. kompleks o najprostszej stechiometrii 1:1 białka eIF4E (*eukaryotic initiation factor 4E*) i analogu mRNA 5' kap, wymaga stosowania poprawek na efekt filtru wewnętrznego, rozcieńczenia przy dodawaniu titranta, niestabilność sygnału emisji. Spektroskopię emisyjną można stosować w miareczkowaniu z rejestracją **depolaryzacji fluorescencji** (*depolarization assay*) oraz wykorzystując przyczepioną do jednego ze składników **sondę fluorescencyjną**; pomiar wygaszania i/lub anizotropii.

(B) **Kalorymetria miareczkująca** ITC (*isothermal titration calorimetry*) [**PLANSZA 179**] w wersji mikro lub ostatnio nano, co zapewnia redukcję ilości używanych substancji.

(C) **Ultrawirowanie analityczne**, które przeżywa renesans [**PLANSZA 180**] (*Schuck Biophys. Rev. 5, 159, 2013*), stosuje się w dwóch typach eksperymentów: *sedimentation velocity* (SV) i *sedimentation equilibrium* (SE) [**PLANSZA 181**]. Składniki kompleksu

mogą mieć **wysoką masę**, czyli metoda jest bardzo użyteczna dla wyznaczania stałych asocjacji dla kompleksów złożonych z dużych składników, np. białko–białko, białko-RNA. Detekcja absorpcji, fluorescencji lub rozproszenia Rayleigha substancji rozpuszczonej daje jej odpowiednie **profile stężenia** χ ([Cole et al. Meth. Cell. Biol. 84, 143, 2008](#)).

(1) W niestacjonarnym eksperymencie SV, z detekcją absorpcji $A(r, t)$ makrocząsteczek w chwilach t w punktach r próbki wirowanej z dużą prędkością kątową $\omega \sim 50\ 000$ rpm [**PLANSZA 180**], uzyskuje się **rozkład współczynnika sedymentacji** $s = (dr/dt)/(\omega^2 r)$. Metoda pochodnej czasowej $d\chi/dt$ daje $g(s^*)_t = (\chi_0)^{-1}(d\chi/dt)[\omega^2 t/\ln(r/r_m)(r/r_m)^2]$, r_m - położenie menisku. Programu SEDPHAT daje $c(s)$ z: $A(r, t) = \int_{s_{\min}}^{s_{\max}} c(s)\chi(s, D(s), r, t) ds$, D to współczynnik dyfuzji. Maksimum $c(s)$ lub $g(s^*)$ określa współczynnik sedymentacji cząsteczki w sevdbergach, $1S = 10^{-13}$ s, a stąd jej **masę cząsteczkową**: $M = N_{Av}om = RTS/D(1 - \rho v)$, gdzie m - masa cząsteczki, $v = 1/\rho_{cz}$ - jej objętość właściwa (odwrotność gęstości ρ_{cz}), ρ - gęstość rozpuszczalnika. Czasową ewolucję radialnego rozkładu stężenia χ podczas sedymentacji określa **równanie Lamma** [**PLANSZA 181**]:

$$\partial\chi/\partial t = (1/r)\partial/\partial r\{rD\partial\chi/\partial r - s\omega^2 r^2\chi\}$$

Integracja $c(s)$ po s daje wagowo-uśredniony współczynnik sedymentacji w funkcji stężenia titranta (4E-BP1), czyli **izotermę** s_w (**Fig 1, inset**), przy założeniu określonego modelu asocjacji, w tym przypadku $A + B \leftrightarrow AB$:

$$s_w = [\varepsilon_{ASACA} + \varepsilon_{BSBCB} + (\varepsilon_A + \varepsilon_B)s_{AB}K_{as}c_{ACB}]/(\varepsilon_{ACA}^{tot} + \varepsilon_{BCB}^{tot})$$

ε - współczynnik ekstynkcji; c_α , $\alpha = A, B, AB$ - stężenia molowe,

$$\text{całkowite stężenia molowe: } c_A^{tot} = c_A + K_{as}c_{ACB} \quad c_B^{tot} = c_B + K_{as}c_{ACB}$$

(2) W stacjonarnym eksperymencie SE przy małych szybkościach $\omega \sim 5\ 000$ rpm ustala się niezależny od czasu rozkład stężeń χ . Z warunku równowagi termodynamicznej $\Delta G_{siły\ odśrodkowej} = \Delta G_{dyfuzji}$, gdzie $\Delta G_{siły\ odśrodkowej} = -\int_{r_1}^{r_2} m(1 - v\rho)\omega^2 r dr$, $\Delta G_{dyfuzji} = (RT/N_{Av})\ln(\chi_1/\chi_2)$ i pomiaru stężeń w r_1 i r_2 otrzymuje się bezpośrednio masę M . Wzór na absorpcję $A(r)$ w celu dopasowania pełnego rozkładu (nie tylko dwu) stężeń można znaleźć np. w [Vistica et al. Analytical Biochem. 326, 234, 2004](#).

Analogicznie, w pomiarach kinetycznych wyznacza się tzw. pseudo-pierwszorzędowe stałe kinetyczne wykorzystując dwa zasadnicze typy metod.

(A) **Metody relaksacyjne**, w których rejestruje się powrót do stanu równowagi pod wpływem zaburzenia (skok temperatury, skok pH).

(B) Metody przepływowe, daleko od położenia równowagi, do których należą omawiana już technika zatrzymanego przepływu (*stopped-flow*) i powierzchniowy rezonans plazmonowy SPR (*surface plasmon resonance* [PLANSZA 182]. W metodzie SPR jeden ze składników kompleksu jest unieruchomiony na cienkiej powierzchni ze złota (metal), która przedziela dwa nieabsorbujące ośrodki (*chip*) i bufor celki przepływowej (*flow channel*) o współczynnikach załamania odpowiednio n_1 i n_2 , a drugi składnik przepływa w celce i wiąże się z unieruchomionym w kompleks. Przy **całkowitym odbiciu wewnętrznym** część światła padającego na granicę ośrodków wnika jako *evanescent field wave* i zanika eksponencjalnie. Przy określonym kącie padania, spolaryzowane w płaszczyźnie pole elektryczne światła wzbudza zdelokalizowane elektrony powierzchniowe czyli **plazmony**, co powoduje silny spadek natężenia wiązki odbitej, rejestrowany jako sygnał rezonansowy R. **Rezonans plazmonowy** zachodzi przy zrównaniu wektora falowego światła k_x w kierunku rozchodzenia fali i wektora falowego oscylujących plazmonów k_{sp} . Tworzeniu kompleksu na chipie towarzyszy wzrost współczynnika załamania, a w efekcie zmiana rejestrowanego kąta SPR (I)→(II). Czasowa zależność sygnału rezonansu R(t) czyli **sensogram**, ma dla modelu wiązania „1:1” składnika o stężeniu c(t) następujący przebieg:

$$dR(t)/dt = k_{on}c(t)[R_{max} - R(t)] - k_{off}R(t)$$

k_{on} - obserwowana stała szybkości wiązania; zakres $10^3 \div 10^8 \text{ M}^{-1}\text{s}^{-1}$

k_{off} - obserwowana stała szybkości dysocjacji; zakres $10^0 \div 10^{-6}\text{s}^{-1}$

gdzie sygnał równowagowy dla $dR/dt = 0$ ma postać tzw. izotermy Langmuira:

$$R_{eq} = [1/(k_{off}/k_{on} + c)]cR_{max}$$

a stała równowagowa $K_{as} = k_{on}/k_{off}$

W bardziej skomplikowanych przypadkach jak np. **wiązanie ligandów przez makromolekułę z kilkoma centrami wiążącymi** (niezależnymi lub kooperatywnymi) zasadniczym problemem jest powiązanie obserwowanych parametrów: natężenie emisji, rozkład mas, kąt odbicia SPR, ze stałymi mikroskopowymi:

- asocjacji, poprzez odpowiednie równania równowagowe na stężenia składników,
- kinetycznymi, poprzez równania na czasowe zależności stężeń składników,

w ramach określonych modeli (schematów „reakcji”), odpowiednio **równowagi asocjacji i kinetyki asocjacji**. Już przypadek eksperymentu wiązania przez białko dwóch

różnych ligandów prowadzi do równania trzeciego stopnia na stężenia ([Wang FEBS Lett. 360, 111, 1995](#)), którego rozwiązanie (w tym przypadku ściśle) pozwala dofitować krzywą zależności mierzalnego parametru (np. fluorescencji) do rejestrowanych punktów miareczkowania. **Problem wynika z konieczności wyznaczenia wartości mierzonego parametru w funkcji całkowitego, znanego stężenia [L] roztworu liganda, a nie liganda swobodnego** [\[PLANSZA 178\]](#). Wymaga to skomplikowanych relacji stężeń i stałych asocjacji w ramach założonego modelu wiązania lub schematu kinetyki „reakcji”. Eftink ([Methods in Enzymology 278, 221, 1997](#)) klasyfikuje w sposób ogólny różne przypadki równowag makromolekuła–ligandy i ich związki z natężeniem (zmianami natężenia) fluorescencji i stosuje do różnych przykładów. **Program DYNAFIT** redukuje problem do "schematycznego" przedstawienia modelu reakcji przez użytkownika.

W przeciwieństwie do skomplikowanych wzorów na mierzone parametry w funkcji całkowitego stężenia wprowadzonego ligandu [L] procedura otrzymywania takich zależności w funkcji stężenia wolnego ligandu [X] jest w miarę prosta i standardowa [\[PLANSZA 177\]](#). Na podstawie: [Hill „An Introduction to Statistical Thermodynamics” Addison-Wesley, London 1960](#) podali ją [Johnson & Straume Methods in Enzymology 323, 155, 2000](#), jako średnią liczbę ligandów $\langle N \rangle$ związanych przez makromolekułę lub inaczej średni ułamek nasycenia Y , $\langle N \rangle = nY$, n - całkowita ilość miejsc wiążących:

$$\langle N \rangle = ([X]/\Xi) \partial \Xi / \partial [X]$$

gdzie Ξ jest sumą statystyczną dużego zespołu kanonicznego układu otwartego:

$$\Xi(V, T, \mu) = \sum_{j, N} \exp[-E_j(N, V)/kT] \exp(N\mu/kT) \quad \mu - \text{potencjał chemiczny}$$

Dla makromolekuły wiążącej kilka jednakowych ligandów, Ξ uzyskuje formę wielomianu wiązania, zależnego od mikroskopowych stałych asocjacji $K_\alpha = \exp(-\Delta G_\alpha/RT)$ i potęg $[X]$. Przykładowo, dla dwóch miejsc wiążących, czyli czterech konfiguracji: bez ligandu ($\Delta G_\alpha = 0$; poziom odniesienia), jeden ligand w miejscu $\alpha = 1$, jeden ligand w miejscu $\alpha = 2$ i ligandy w obu miejscach $\alpha = 12$: można wyliczyć:

$$\Xi = 1 + K_1[X] + K_2[X] + K_{12}[X]^2$$

$$\langle N \rangle = \{(K_1 + K_2) [X] + 2 K_{12}[X]^2\} / \{1 + [(K_1 + K_2) [X] + 2 K_{12}[X]^2]\}$$

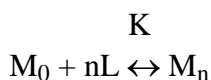
W konsekwencji, stosunkowo łatwo i ściśle można wyrazić mierzalne parametry przez stałe wiązania i stężenia w **reakcjach enzymatycznych**, gdzie kinetykę wyznacza się przy dużym nadmiarze substratu w stosunku wiążącej substrat makromolekuły, i stężenie wolnego ligandu jest w przybliżeniu równe całkowitemu stężeniu, $[L] \cong [X]$.

Wygodnym, chociaż przybliżonym ujęciem **różnych aspektów wiązania w ligandów przez makromolekulę z n identycznymi miejscami wiążącymi** jest liniowa **zależność Scatcharda** na ilość moli ligandu na jeden mol makromolekuły, czyli $\langle N \rangle$, w funkcji stężenia wolnego ligandu $[X] \equiv [L]$:

$$\langle N \rangle / [L] = n / K_d - \langle N \rangle / K_d \quad \text{stała dysocjacji } K_d = 1 / K_{as}$$

Zakrzywienie wykresu Scatcharda może być spowodowane przez dwa czynniki:

- nieidentyczność miejsc wiążących, np. tzw. przebieg dwufazowy (*biphasic*) dla wiązania dwóch ligandów;
- kooperatywność miejsc wiążących, czyli zmiana powinowactwa makromolekuły do ligandu („pustego” miejsca wiążącego) po jego związaniu w innym centrum wiążącym, co prowadzi do modyfikacji stałych dysocjacji/asocjacji przy częściowym zapełnieniu centrów. W warunkach tzw. „nieskończonej kooperatywności” (reakcja „*all-or-none*”):



$$K^n = [M_0][L]^n / [M_n]$$

średni ułamek wysycenia:

$$Y = \langle N \rangle / n = \{ [L]^n / K^n \} / \{ 1 + [L]^n / K^n \}$$

Nieskończona kooperatywność nie jest nigdy obserwowana i wprowadza się pół-empiryczną relację w formie **równania Hilla**:

$$Y = \{ [L]^{\alpha_H} / K^{\alpha_H} \} / \{ 1 + [L]^{\alpha_H} / K^{\alpha_H} \} \quad 1 \leq \alpha_H \leq n$$

w której parametr α_H , **stała Hilla**, jest **wskaźnikiem kooperatywności**. System jest doskonale kooperatywny dla $\alpha_H = n$ oraz całkowicie niekooperatywny dla $\alpha_H = 1$. Ponieważ równanie (Hilla) nie jest w ogólności spełnione w całym zakresie zmienności Y , α_H może być funkcją stopnia wysycenia. Stąd definicja współczynnika Hilla:

$$\alpha_H = d \{ \ln Y (1 - Y) \} / d (\ln [L])$$

Równanie Hilla jest parametryczną zależnością, która może być użyta do **analizy wykresów Scatcharda kooperatywnych asocjacji**. W przypadku wiązania kilku różnych ligandów sytuacja się komplikuje. Podręcznik: [Cantor & Schimmel Biophysical Chemistry III \(Freeman & Co. 1980\)](#) dostępny on-line podaje (za pracą [Wyman Adv. Prot. Chem. 19, 223, 1964](#)) użyteczne relacje na równowagi asocjacji (Y_1 i Y_2) dla dwóch rodzajów ligandów.

Wykład 22

Asocjacja cząsteczek w kompleksy molekularne jest **procesem dynamicznym**, w którym mniej lub bardziej specyficzne rozpoznanie i wiązanie jest sprzężone ze zmianami konformacyjnymi składników. **Dopasowanie sztywnych struktur** molekuł na zasadzie „klucz do zamka”, chociaż jest bardzo uproszczonym, najwcześniej opracowanym modelem asocjacji, w różnym zakresie znajduje wykorzystanie w efektywnych (szybkich) algorytmach **dokowania komputerowego**. W bardzo różnych klasach takich algorytmów, dostosowanych do różnych typów molekuł stosuje się różne sposoby przeszukiwania przestrzeni konformacyjnej układu (*search algorithms*) i różne metody ilościowej weryfikacji wyników dopasowania (*scoring function*). Algorytmy „*search*” dla (dużych) układów, np. białkowych, opierają się na dopasowaniu sztywnych cząsteczek (*rigid body docking*; [Kohlbacher et al. J. Biomol. NMR 20, 15, 2001](#)), natomiast przy dokowaniach ligandu do receptora makromolekularnego (białka) powszechnie stosowane protokoły zakładają konformacyjnie „giętki” ligand i sztywny receptor makromolekularny ([Taylor et al. J. Computer-Aided Mol. Design 16, 151, 2002](#)).

Dynamiczny charakter asocjacji zapewnia bardziej optymalne dopasowanie powierzchni molekularnych cząsteczek, a stąd efektywniejsze oddziaływanie (efekt entalpowy). W modelach asocjacji **z uwzględnieniem dynamiki** postuluje się dwa alternatywne podejścia, na przykładzie wiązania makromolekuła-ligand. **[PLANSZA 183]:**

- **wzajemne konformacyjne dopasowanie** asocjujących składników (*induced fit*), które optymalizuje wiązanie po utworzeniu kompleksu spotkaniowego;
- **selekcja konformacyjna** (*conformational selection*) składników kompleksu o właściwej (optymalnej) konformacji z ich równowagowego rozkładu, która poprzedza wiązanie.

Bezpośredni wgląd daje analiza pojedynczych cząsteczek ([Hatzakis Biophys. Chem. 186, 46, 2014](#)). W przybliżeniu „*rapid equilibrium*”, gdy wiązanie i dysocjacja są szybkie w porównaniu ze zmianami konformacyjnymi, obserwowana stała szybkości k_{obs} albo hiperbolicznie rośnie (*induced fit*) albo hiperbolicznie maleje (*conformational selection*) w funkcji stężenia ligandu. Ten efekt może stanowić przesłankę określenia właściwego modelu w konkretnym przypadku asocjacji. Konsekwencje termodynamiczne modelu „*induced fit*”, postulowanego z termodynamicznej analizy kompleksów białek z kwasami nukleinowymi ([Spolar & Record Science 263, 777, 1994](#)) są następujące:

- duża, zwykle ujemna wartość ciepła właściwego asocjacji przy stałym ciśnieniu $\Delta_c p^\circ$;

- znaczny koszt entropowy asocjacji, który może być zredukowany przez usztywnienie (*fixation*) niektórych reszt w niezwiązanym białku (**koncepcja:** [de Poupulana et al. Biochemistry 35, 8095, 1996](#)).

Odkryciu białek natywnie rozwiniętych (IDP) towarzyszyły liczne badania **sprężonego, kooperatywnego mechanizmu zwijania i wiązania** (*folding upon binding*) ([Kiefhaber et al. Curr. Op. Struct. Biol. 22, 21, 2012](#)). Ujęcie problemu na gruncie molekularnym wymagało opracowania nowych modeli tego procesu, ponieważ dotychczasowe modele dynamiczne asocjacji opierały się na oddziaływaniu cząsteczek w strukturach (zbliżonych do) natywnych. Wydaje się, że dobrze ugruntowanym i wyczerpującym podejściem do zagadnienia jest **model „fly-casting”** („wędkowanie na muchę”). Nieuporządkowane białko wiąże się **względnie słabo i niespecyficznie z cząsteczką „tarczą”** i **zwija na niej całkowicie lub częściowo w miarę specyficznego rozpoznania** centrów wiążących [**PLANSZA 184**]. Większy promień (*radius of gyration*) nieustrukturuwanej cząsteczki białka przyspiesza znalezienie właściwego centrum wiązania. Ilościowe sformułowanie modelu „fly-casting” zostało przedstawione w dwóch obszernych pracach: [Shoemaker & Wolynes. J. Mol. Biol. 287, 657, 1999](#); [Shoemaker et al. J. Mol. Biol. 287, 675, 1999](#); zarysowuję je w dużym skrócie. Definiuje się **funkcjonał energii swobodnej Gibbsa**:

$$F = F_{\text{protein}}[\{q_{ij}^p\}] + F_{\text{binding}}[\mathbf{R}_{\text{cm}}, \mathbf{\Omega}, \{q_{ij}^p\}, \{q_{ij}^s\}]$$

w którym pierwszy człon (wzór podany w [Shoemaker et al. J. Mol. Biol. 287, 675, 1999](#).) reprezentuje kontakty **wewnątrz białka**, zależne od ułamka (prawdopodobieństwa) **poszukiwanych obsadzeń** $\{q_{ij}^p\}$ **kontaktów wewnętrznych** i,j , a drugi reprezentuje **wiązanie i sprzęga tworzenie struktur lokalnych i oddziaływań wiążących kosztem entropii** kontaktów powierzchniowych białko-tarcza, o ułamkowych, **poszukiwanych obsadzeniach** $\{q_{ij}^s\}$, przy ustalonym wektorze odległości środków masy obu cząsteczek \mathbf{R}_{cm} i orientacji białka i tarczy zadanej przez kąty Eulera $\mathbf{\Omega}$:

$$F_{\text{binding}} = E_{\text{binding}}[\{q_{ij}^s\}] - TS_{\text{binding}}[\mathbf{R}_{\text{cm}}, \mathbf{\Omega}, \{q_{ij}^p\}, \{q_{ij}^s\}]$$

Dysponując funkcyjnymi zależnościami na F , w tym na E_{binding} i na S_{binding} , **minimalizację się funkcjonal** względem $\{q_{ij}^p\}$ i $\{q_{ij}^s\}$, przy zachowaniu stałych wartości wszystkich kontaktów Q_p i Q_s , czyli wyznacza się ekstremum związane z mnożnikami Lagrange’a:

$$Q_p = \sum_{ij} \{q_{ij}^p\} = \text{const}$$

$$Q_s = \sum_{ij} \{q_{ij}^s\} = \text{const}$$

Minimalizacja daje **poszukiwane prawdopodobieństwa kontaktów** $\{q_{ij}^P\}$ i $\{q_{ij}^S\}$ oraz odpowiadające im **stałe szybkości wiązania**, **wyższe** w przypadku (całkowicie) rozwiniętego białka ($k_{QP=0}$) **niż** zwiniętego białka ($k_{QP=1}$).

Dynamikę ruchów molekularnych cząsteczek w trakcie asocjacji można badać przy pomocy szeregu metod doświadczalnych, z których dwie uzyskały (lub uzyskują aktualnie) szeroką popularność.

(1) **Wymiana proton/deuter** z detekcją **MS** (*H/D exchange MS*) lub z detekcją **NMR** (*H/D exchange NMR* **[PLANSZA 131]**) dostarcza informacji o strukturalnej lokalizacji fragmentów molekularnych tworzących **powierzchnię kontaktu w trakcie asocjacji**.

(2) **Niskokątowe rozpraszanie promieniowania rentgena SAXS** (*small-angle X-ray scattering*) w roztworze asocjujących cząsteczek dostarcza informacji o dynamice zmian i globalnej giętkości konformacji (kształtu) kompleksu.

Wykorzystanie wymiany proton/deuter przy **detekcji MS** **[PLANSZA 185]** do badania asocjacji, tu białka z błoną lipidową, polega na inkubacji badanego białka w roztworze ciężkiej wody zawierającym drugi składnik kompleksu. Wymiana protonów białka na deuterium rozpuszczalnika jest przerywana w kolejnych odstępach czasu przez obniżenie pH i temperatury, dla utrwalenia miejsc deuteracji. Zdenaturowane białko jest ładowane na kolumnę HPLC z proteazą, która hydrolizuje białko na fragmenty. Odseparowane na kolumnie peptydy są kierowane bezpośrednio do spektrometru mas (LC-MS) gdzie peptydy są identyfikowane (patrz wykład o proteomice). **Obszary kontaktu białka z błoną są osłaniane**, a więc **mniej** skłonne do wymiany, co pozwala je zidentyfikować na podstawie różnicy mas (deuter/proton). Podobnie można charakteryzować **obszary zmian konformacyjnych** w trakcie asocjacji, jeśli prowadzą do ekspozycji tych obszarów do rozpuszczalnika lub likwidacji kontaktów z rozpuszczalnikiem. Porównawczo przeprowadza się wymianę w samym białku. Widać to na przykładzie zmian konformacyjnych związanych z fosforylacją metyloesterazy CheB (demetylacja specyficznych reszt metyloglutaminowych) **[PLANSZA 186]**. Ponieważ fosforylacja na Asp56 powoduje 10-krotny wzrost aktywności enzymu, przypuszczalnie pod jej wpływem zachodzi zmiana pozycji domen. Zagadnienie badano przez porównanie wymiany H/D w białku fosforylowanym i nefosforylowanym. Wprawdzie nie udało się określić stałych szybkości wymiany (zbyt szybka), co umożliwiłoby ilościową analizę problemu, jednak określono **liczbę wymienianych protonów w funkcji czasu** dla fragmentów białka

trawionego pepsyną z obwoluty masowej (*inset*: widmo MALDI-TOF fragmentu 56-66), przerywając proces wymiany w kolejnych momentach co 1 min. Po 5 min.:

(i) CheB bez deuteracji;

(ii) Asp56Asn mutant CheB po 5 min. deuteracji;

(iii) CheB w obecności czynnika fosforylującego po 5 min. deuteracji.

Zmiany konformacyjne odzwierciedlone poprzez zmiany dostępności solwentu można zobrazować kolorami na strukturze krystalograficznej białka (Asp56 w reprezentacji *ball-and-stick*): obszary nie zanalizowane na szaro, obszary bez zmian deuteracji pod wpływem ufosforylowania na niebiesko, obszary o zwiększonej deuteracji na żółto. Znakowanie deuterowanych obszarów **metodą NMR** jest podobne i wynika z zamiany protonów wymienialnych na deuterium i znikania odpowiadających tym protonom sygnałów w wielowymiarowych widmach korelacyjnych proton-heterojądrowo ([review: Hoshino et al. Biochi. Biophys. Acta 1768, 1886, 2007](#)). Analizowanymi sygnałami są z reguły sygnały protonów amidowych, które w regularnych strukturach drugorzędowych (helisy, β -kartki) są zaangażowane w wiązania wodorowe. Protony na powierzchni wymieniają się zbyt szybko. Analizie poddaje się nie ilość wymienianych protonów jak w MS, ale tzw. **czynniki protekcji**:

$$P = k_{\text{int}}/k_{\text{ex}}$$

k_{int} - stała szybkości wymiany (*intrinsic*) przy pełnej ekspozycji protonu do rozpuszczalnika;

k_{ex} - obserwowana stała szybkości wymiany w białku.

Ze względu na utrudnioną wymianę protonów wiązań wodorowych można łatwiej śledzić proces wymiany w czasie, a więc lepiej uzyskiwać stałe szybkości wymiany, Jednak warunkiem określania miejsc wymiany jest wykonanie przyporządkowania sygnałów protonowych, które dla większych białek jest trudniejsze niż analiza MS.

Metoda SAXS rozpraszania promieniowania X o długości fali λ na cząsteczkach w roztworze [[PLANSZA 187](#)], czyli we wszystkich orientacjach, nie daje bezpośredniej informacji o położeniach atomów uginających promieniowanie w regularnej sieci krystalicznej, a jedynie **kształt, tworzenie kompleksów, agregację i przekształcenia konformacyjne** makromolekuł. Bezpośrednio, SAXS dostarcza ilościowych wartości R_g (*radius of gyration*), który opisuje rozkład masy wokół środka masy molekuly. Wzrost R_g z reguły oznacza rozluźnienie/rozciąganie struktury (*opening*) a jego spadek

przechodzenie do struktury zwartej (*compaction*). **Profil rozpraszania** $I(q)$ w w funkcji długości wektora q w przestrzeni odwrotnej

$$|q| = (4\pi/\lambda)\sin\theta \quad 2\theta - \text{kąt rozproszenia}$$

przekształca się w **rozkład odległości międzyatomowych** $P(r)$. Analiza $I(q)$ może również stanowić bezpośredni test:

- struktur krystalograficznych w metodzie dyfrakcji X na monokryształach;
- symulacji komputerowych dynamiki molekularnej MD.

Zmiany konformacyjne, które towarzyszą tworzeniu kompleksów można szczególnie efektywnie śledzić w sposób półilościowy za pomocą **rezonansowego transferu energii wzbudzenia FRET** między odpowiednimi sondami fluorescencyjnymi, wykorzystując zależność efektywności rezonansowego sprzężenia sond od odległości między nimi [**PLANSZA 134**]. Dużym krokiem na drodze do badania **asocjacji przez optyczne obrazowanie *in vivo*** wydaje się zastosowanie **sond o widmach absorpcji i fluorescencji w zakresie bliskiej podczerwieni NIR (*near infra red*)**, ze względu na niższe rozpraszanie i absorpcję fotonów NIR w tkance. Umieszczenie takich sond: donora R = Cys5.5 oraz akceptora X = Cy7 oraz 800CW [**PLANSZA 188**] na helisach oligodeoxyrybonukleotydowych o odpowiedniej sekwencji pokazało możliwość **detekcji specyficznego wiązania** transkrypcyjnego czynnika NF- κ Bp50. Widma emisji donora i wzbudzenia obu akceptorów pokrywają się w zakresie 700 ÷ 800 nm. Transfer energii widać poprzez wzrost intensywności fluorescencji akceptora (widma emisji: krzywa 3), zależnie od ilości nukleotydów (**panele A ÷ D**) dzielących donor (widmo emisji, krzywa 1, wzbudzenie 675 nm) oraz akceptor (widmo emisji, krzywa 2). Wprowadzeniu białka NF- κ Bp50 towarzyszy wzrost stosunku emisji w długościach fali 700 nm do 800 nm.

Dynamiczny charakter tworzenia kompleksów molekularnych, który jest z kolei uwarunkowana „giętkością” konformacyjną makromolekuł, jest podstawą omawianej wcześniej kooperatywności miejsc wiążących makromolekuły (enzymu), którą można zaobserwować zarówno w przypadku jednakowych miejsc wiążących jak i w przypadku inhibicji/aktywacji katalizy enzymatycznej w centrum katalitycznym przez przyłączenie ligandu w innym centrum, tzw. allosterycznym. Pojęcie **dynamicznej allosterii** (Monod, Wyman, Chageux) w zasadzie obejmuje większość przypadków kooperatywności w procesach asocjacji [**PLANSZA 189: B, C**]. Więcej niż jedno miejsce wiążące w białku (**A**) nie jest warunkiem koniecznym efektu allosterycznego, np. γ -immunoglobuliny (IgG) dysponując dwoma identycznymi centrami wiążącymi do rozpoznania antygeny nie

kooperują ze sobą. Z drugiej strony **efekty allosteryczne regulują** szeroka gamę różnorodnych procesów, takich jak: kontrola metabolizmu, funkcje receptorowe makromolekuł, transport ligandów, ruchliwość komórki. **Efekt allosteryczny** może mieć charakter:

- dodatni, wzmacniania wiązania w innym centrum,
- ujemny, osłabiania wiązania innego ligandu.

Ogólne ujęcie efektu allosterycznego w terminologii „pejzażu energetycznego” zakłada przynajmniej **trzy możliwości przy wiązaniu efektora allosterycznego**:

- zmiana energetycznych relacji między konformerami (minimów energii swobodnej);
- zmiana relacji energetycznych minimów energii swobodnej jak w poprzednim przypadku wraz z redystrybucją konformacji, „pre-egzystujących” w stosunku do wiązania efektora;
- zwiększenie fleksybilności białka bez zmian relacji energii minimów: poszerzenie przestrzeni konformacyjnej.

Mimo ogólnego charakteru tego sformułowania, proponowane przez grupy badawcze **molekularne mechanizmy allosterii** różnią się w przypadkach poszczególnych makromolekuł/zagadnień:

- „prototypowy” **model MWC** (Monod, Wyman, Chageux; [Monod et al. J. Mol. Biol. 12, 88, 1965](#)) tetramerycznej hemoglobiny wiążącej tlen (omawiany na wykładzie „[Struktura i funkcja makrocząsteczek biologicznych](#)”);
 - mechanizm allosterii w **dimerycznych hemoglobinach**, związany ze zmianami struktury trzeciorzędowej (w przeciwieństwie do poprzedniego modelu) ([Royer et al. Proc. Natl. Acad. Sci. USA 93, 14526, 1996](#));
 - zależna od PLP (*pyridoxal 5'-phosphate*) **syntaza tryptofanu** $\alpha_2\beta_2$ ([Barends et al. Curr. Op. Chem. Biol. 12, 593, 2008](#));
 - **wiązanie tetracykliny** (Tc) do homodimerycznego Tc-represora, kontrolujące oddziaływanie białka represora z podwójną helisą (operatora) DNA ([Reicheld et al. Proc. Natl. Acad. Sci. USA 106, 22263, 2009](#));
 - **reduktaza rybonukleotydoma**: allosteryczne definiowanie specyficzności substratowej obok kontroli aktywności enzymatycznej ([Reicheld Proc. Biochem. Biophys. Res. Commun. 396, 19, 2010](#)),
- i inne ([review: Peracchi & Mozzarelli Biochim. Biophys. Acta 1814, 9212, 2011](#)).

W pracy [Ma et al. Structure 19, 907, 2011](#) postuluje się szczególną rolę **giętkiego linkera** między kooperującymi domenami białkowymi w propagacji efektu allosterycznego, "zakodowaną" w sekwencji linkera.

Długozasięgowe perturbacje allosteryczne propagują się nie tylko przez zmiany strukturalne miejsca aktywnego **ale także przez modulacje w samej jego dynamice**, poprzez wzrost ruchów w różnej skali czasowej, ps ÷ ms, bez zauważalnych efektów strukturalnych w wyniku przyłączenia regulatora allosterycznego ([Selvaratnam et al. Proc. Natl. Acad. Sci. USA 108, 6133, 2011](#)). Na tej podstawie autorzy pracy opracowali eksperymentalną metodę CHESCA (*chemical shifts covariance analysis*) mapowania sieci zależności allosterycznych w białkach, łączących miejsce regulatorowe i katalityczne, oparte na analizy kowariancji* przesunięć chemicznych NMR. Po przyporządkowaniu sygnałów np. w widmie $^1\text{H},^{15}\text{N}$ -HSQC w białku *apo* i *holo*, konstruuje się **macierz korelacji przesunięć chemicznych** wszystkich aminokwasów i **identyfikuje sprzężone aminokwasy na drodze działania efektora allosterycznego** na podstawie liniowej korelacji między zmiennością tych przesunięć chemicznych. Odległe aminokwasy „czują” jak gdyby tę samą równowagę konformacyjną „stan aktywny/stan nieaktywny”.

* korelację między wielkościami X_j wyraża macierz kowariancji $\mathbf{C} = [C_{ij}]$

$$C_{ij} = \langle (X_i - \langle X_i \rangle) (X_j - \langle X_j \rangle) \rangle$$

gdzie $\langle \rangle$ oznacza uśrednienie według funkcji rozkładu prawdopodobieństwa zmiennych losowych (wektora losowego) \mathbf{X} .

**SPECYFICZNE ODDZIAŁYWANIA biopolimerów w kompleksach: białko - ligand
białko-DNA i białko - RNA, projektowanie leków (*DRUG DESIGN*); struktury
subkomórkowe - motory molekularne**

Wykład 23

Centralnym punktem zainteresowania biofizyków i biologów molekularnych są wysoce specyficzne i złożone kompleksy **białko-DNA** i **białko-RNA**, które stanowią podstawę zachodzenia i regulacji procesów **replikacji i ekspresji genu** na poziomie transkrypcji, translacji i splicingu, w tym **reakcji enzymatycznych z udziałem m. in. polimeraz i endonukleaz**. W dziedzinie kompleksów białko - RNA osiągnięto olbrzymi postęp w wyznaczaniu dużych struktur metodami rentgenograficznymi, gdzie wychodzi się poza układy modelowe w kierunku **pełnych struktur subkomórkowych**, czyli jak to się określa układów **mezoskopowych** (pomiędzy mikroskopowymi i makroskopowymi): 70S rybosom prokariotyczny (2.7 mln Da) oraz 80S rybosom eukariotyczny (3.3 mln Da). Niewątpliwie najwięcej prac poświęcono badaniu oddziaływań **białek enzymatycznych** z różnego typu ligandami. Badania te miały szczególnie cenny wymiar praktyczny przy poszukiwaniu **inhibitorów** określonych reakcji enzymatycznych jako **środków farmakologicznych** o charakterze cytostatycznym: przeciwbakteryjnym i przeciwnowotworowym, oraz przeciwwirusowym. Miarą efektywnego blokowania enzymu jest niska wartość stałej inhibicji K_i . Enzymy są jednymi z możliwych **receptorów, których zablokowanie daje określony efekt farmakologiczny**, np. selektywne zablokowanie enzymu kodowanego przez wirusa przy słabszym wiązaniu inhibitora przez enzym gospodarza może uniemożliwić namnażanie i spowodować zniszczenie wirusa przez system immunologiczny. **Praktyczne wykorzystanie zdobyczy proteomiki** i biofizyki molekularnej opiera się na analizie struktury centrum wiążącego ligand w wybranym receptorze białkowym czyli tarczy (analiza *PPI network*). Do niedawna poszukiwano inhibitorów **metodą „prób i błędów”** modyfikując chemicznie związek wyjściowy o określonym działaniu aby zwiększyć efekt. **Racjonalne projektowanie leków (*rational DRUG DESIGN*)** w oparciu o znaną strukturę receptora-tarczy, np. białka enzymatycznego, jest zdecydowanie **bardziej wydajne**. W połączeniu z testowaniem doświadczalnym tylko dobrze rokujących związków modelowanych, podejście racjonalne **skraca ok. trzykrotnie czas** od badań podstawowych do klinicznych

zastosowań leku. Mimo to, ilość całkowicie nowych leków dopuszczanych do użycia farmakologicznego nie jest zbyt wysoka. W USA, Food and Drug Administration (FDA) dopuszcza ok. 5 do 6 NMEs (*new molecular entities*) rocznie (dane za 2004 - 2009).

Przypomnijmy (wykład „*Struktura i funkcje makrocząsteczek biologicznych*”) w skrócie schemat procedury racjonalnego projektowania leków **[PLANSZA 190]**. Jeśli w zainfekowanej komórce pojawia się **nowe białko, kluczowe** dla rozwoju wirusa lub bakterii (analiza PPI), albo białko tego typu w komórce nowotworu jest **zmodyfikowane** w stosunku do analogu w komórce zdrowej, procedury proteomiczne umożliwiają jego wyłowienie, uzyskanie w odpowiedniej ilości i określenie struktury przestrzennej. Analiza **kontaktów stabilizujących ligand w centrum aktywnym umożliwia zaprojektowanie cząsteczki** ligandu-inhibitora, która powinna charakteryzować się:

- **lepszym powinowactwem** do receptora (enzymu) niż naturalny substrat;
- **maksymalnie selektywnym** blokowaniem niepożądanego białka, bez efektów ubocznych.

Przy słabszym wiązaniu przez enzym gospodarza działanie inhibitora może uniemożliwić szybkie namnażanie wirusa lub nowotworu i zyskać czas na jego zniszczenie przez system immunologiczny.

Projektowanie leków jako jedna z dziedzin **projektowania molekularnego** zakłada ściśle współdziałanie **technik doświadczalnych** w ustalaniu **struktur białek, screeningu ligandów** pod kątem potencjalnych własności farmakologicznych i dalszych modyfikacji strukturalnych: MS, NMR, dyfrakcja rentgenowska, techniki biochemii i biologii molekularnej, **technik komputerowych analizy kompleksów** receptor-ligand: mechanika molekularna, dynamika molekularna, dokowanie, techniki „wirtualnej rzeczywistości” (*virtual reality*), grafika komputerowa oraz **teoretycznych poszukiwań** mniej lub bardziej ogólnych relacji, dotyczących **katalizy** enzymatycznej ([Babtie et al. Curr. Op. Chem. Bio. 14, 200, 2010](#)) i **czynników fizykochemicznych**, warunkujących powinowactwo ligandów do makromolekularnych „tarcz” ([Kuntz et al. Proc. Natl. Acad. Sci. USA 96, 9997, 1990; Stockwell & Thornton J. Mol. Biol. 356, 928, 2006](#)). To ostatnie zagadnienie prowadzi do ilościowych **relacji QSAR** (*quantitative structural activity relationship*) między wartościami stałych asocjacji a wybranymi **parametrami fizykochemicznymi** dla serii cząsteczek danej klasy: stała Hammetta σ , hydrofobowość wyrażana przez współczynnik podziału K i in., wyczerpująco opisane w [reviews: Kubinyi Drug Discovery Today 2, 457 \(part I: methodology\); 538 \(part II: applications and problems\), 1997](#).

Analiza kontaktów stabilizujących sprowadza się zasadniczo do komputerowych obliczeń energii swobodnej wiązania kolejnych **związków, modyfikowanych** w stosunku do wyjściowego wzorca (*alchemical binding free energy calculations*: [Chodera et al. Curr. Op. Struct. Biol. 21, 150, 2011](#)). Analizy statystyczne pokazują, że nawet umiarkowana dokładność obliczeń $\sim \pm 8$ kJ/mol jest wystarczająca dla bardziej efektywnego projektowania ligandów niż docking. Jest to optymistyczne spostrzeżenie w świetle stosowanych przybliżeń klasycznej mechaniki i dynamiki molekularnej. Aktualnie stosowane (*state-of-the-art*) podejścia zakładają albo obliczenia wartości absolutnych energii albo względnych powinowactw, czyli różnic energii pomiędzy ligandami analizowanej klasy. Ścisłe przestrzeganie **zasadniczych aspektów procedury** pozwala uniknąć rozbieżności między obliczeniami i doświadczeniem:

- (1) wymodelowanie chemicznej (wyjściowej) struktury ligandu w oparciu o *screening*;
- (2) właściwa dla danej klasy związków parametryzacja pola siłowego;
- (3) odpowiednia metoda próbkowania (*sampling*) „przestrzeni **chemicznej**” (alchemia komputerowa): wymienianie atomów na inne, wprowadzanie podstawników, oraz przestrzeni **konformacyjnej**, biorąc pod uwagę dostępne możliwości komputerowe (oprogramowanie, czas CPU) i czas „realny”;
- (4) analiza wyników, szczególnie pod kątem próbkowania.

Potencjalny środek farmakologiczny wymaga doświadczalnego sprawdzenia efektywności wiązania przez białko oraz przetestowania pod kątem trzech kluczowych czynników, niezbędnych do jego **działania *in vivo***:

- (a) **odpowiednia rozpuszczalność i efektywne wnikanie/wprowadzenie do komórki**;
- (b) **brak efektów ubocznych** (nietoksyczność);
- (c) **stabilność, m. in. oporność na degradację** (enzymy, system immunologiczny).

Problem z wprowadzeniem leku do komórki wynika przede wszystkim z konieczności **przenikania dostarczanego z zewnątrz związku przez błony komórkowe**. Wymienione czynniki (a-c) próbuje się osiągnąć stosując systemy dostarczania środków farmakologicznych (*drug-delivery systems*) [\[PLANSZA 191\]](#), najczęściej (temat jest bardzo obszerny) w formie $5 \div 50$ nm **samoorganizujących się**: (1) **nano-micelli (liposomów)**, (2) **ko-polimerów** (blokowych) zbudowanych z PEO (*polyethylene oxide*), PCL (*polycaprolactone*), PAsp (*polyaspartate*) lub (3) **dendrytycznych** (rozgałęzionych) **polimerów** na bazie poliglicerolu, *polyethylene imine*, i in. Wnikanie do komórki wraz z niesionym w środku lekiem odbywa się na zasadzie **endocytozy**. W przypadku komórek nowotworowych wykorzystuje się np. **efekt EPR** (*enhanced permeability and retention*)

[PLANSZA 192] do wnikania nanocząsteczek złota, które absorbując światło lasera podgrzewają tkankę i zwiększają przenikalność oraz wiązanie nośników (lipozomów) z lekiem. Używa się również **nośników peptydowych i technologii LbL** (*layer-by-layer*; ([Sakr & Borchard Biomacromolecules 14, 2117, 2013](#))). Szereg nośników stosuje się także do dostarczania czynników kontrastowych **[PLANSZA 191]** w obrazowaniu.

Oprócz **wprowadzenia** farmaceutyku do komórki i jego **stabilności *in vivo***, pozostaje jeszcze „otwarte” pytanie, czy oddziaływanie z „tarczą” ***in vivo* nie różni się od oddziaływania *in vitro*** ? Jedną z intensywnie rozwijanych metod badania asocjacji makromolekuła-ligand, ogólnie tworzenia kompleksów w cieczach biologicznych (*biological liquids*) takich jak osocze krwi czy lizaty komórkowe, jest **termoforeza MST** (*microscale thermophoresis*), znana od 1856 r. jako efekt Soreta **[PLANSZA 193]**. Metoda jest komplementarna do SPR oraz do kalorymetrii ITC i szeregu innych metod (biosensory firmy Fortebio). W metodzie MST mierzy się fluorescencyjnie **zmianę stężenia makrocząsteczek** c_i w kapilarze wskutek ich **termodyfuzji** z prędkością v , wywołanej skokiem temperatury T za pomocą lasera IR. Strumień masy cząsteczek spełnia relację:

$$\mathbf{J}_i = \mathbf{v}c_i = -c_i D_{Ti} \nabla T$$

D_{Ti} - współczynnik dyfuzji termicznej (ruchliwość termoforetyczna)

Sygnał emisji w MST ma kilka faz: *T-jump*, ruch termoforetyczny, stan równowagi termodyfuzji i dyfuzji:

$$\mathbf{J}_i = -D_i \nabla c_i$$

o ustalonym stężeniu zależnym od **stałej Soreta** $S_{Ti} = D_{Ti}/D_i$, odwrotny skok temperatury po wyłączeniu lasera i dyfuzja zwrotna. W badaniach oddziaływań profil zmiany stężenia frakcji związanej (f) i niezwiązanej:

$$\Delta c_{\text{measured}} = f \Delta c_{\text{bound}} + (1 - f) \Delta c_{\text{unbound}}$$

w funkcji stężenia jednego ze składników kompleksu (liganda) daje **wartość stałej asocjacji**, która **może się różnić w warunkach *in vitro* i w lizacie komórkowym**.

W badaniu wiązania ligandów przez białka w ich naturalnym środowisku nienaruszonej komórki znajdują zastosowanie **profile termostabilności** ([Huber et al. Nature Methods 12, 1055, 2015](#)). Wykorzystuje się w tym celu efekt wzrostu termostabilności kompleksu białko-ligand w porównaniu z niezwiązanym białkiem, monitorowany przy pomocy różnych metod, np. strategię **znakowania w spektrometrii mas** TMT (*isobaric tandem mass tags*) lub podobną do kalorymetrii DSC **różnicową**

fluorymetrię skanującą DSF (*differential scanning fluorimetry*). Metodologia jest bardziej nastawiona na **identyfikację białek komórkowych** wiążących ligandy: metabolity i środki farmakologiczne (*drugs*), w porównaniu z np. MST.

Analiza strukturalna kompleksów białek z kwasami nukleinowymi, kwasów nukleinowych z ligandami niskocząsteczkowymi, oraz kompleksów enzymatycznych (białka, RNA) na poziomie molekularnym była omówiona na wykładzie „Struktura i funkcje makrocząsteczek biologicznych”. W ramach obecnego wykładu uwaga będzie skoncentrowana na wykorzystaniu metod **obrazowania, manipulacji i spektroskopii pojedynczej cząsteczki** (*single-molecule level*) do analizy procesów z udziałem dużych kompleksów molekularnych, warunkujących zachodzenie ekspresji genów (*gene expression*) oraz replikacji i transportu ([review: Dulin et al. Nature Rev. 14, 9, 2013](#)).

Bezpośrednie określanie struktur z rozdzielczością znacznie poniżej 1Å (rentgenografia, NMR), nie jest na razie możliwe przy zastosowaniu **mikroskopów**. Niemniej, informacja strukturalna o **kształcie** kompleksów i **oddziaływaniach** składników z rozdzielczością do kilku Å oraz **dynamice ruchów molekularnych** stanowi źródło **bezpośredniej informacji o funkcjonowaniu kompleksów w czasie realnym**: detekcja intermedatów procesu i rzadkich efektów, maskowanych przez uśrednianie po makroskopowej ilości cząsteczek. Udział **technik przetwarzania danych** w obrazowaniu (nie do uniknięcia), jest znacznie mniejszy niż w przypadku metod pośrednich.

Konwencjonalna mikroskopia optyczna jest w wyniku falowej natury światła obarczona fundamentalnym ograniczeniem rozdzielczości w formie tzw. **granicy dyfrakcyjnej Rayleigha-Abbe** (dysk Airy'ego po przejściu przez kołową aperturę), $0.61\lambda/NA \sim 200 \text{ nm}$, dla maksymalnej wartości $NA \sim 1,5$ ($\theta = 70^\circ$) i promieniowania widzialnego $\lambda \sim 500 \text{ nm}$. Przykładowo, dla soczewki apertura numeryczna $NA = n\sin\theta$, n - współczynnik załamania ciekłego ośrodka imersyjnego, θ - połowa maksymalnego kąta (ro)zbieżności wiązki.

Wykorzystuje się szereg **sposobów zwiększenia rozdzielczości przesuując granicę dyfrakcyjną** (zmniejszyć λ i/lub zwiększyć NA). W mikroskopii **bliskiego pola NSOM** (*near-field scanning optical microscopy*) **[PLANSZA 194]** uzyskuje się obraz o rozdzielczości rzędu 50 nm. **Granica rozdzielczości NSOM** jest uwarunkowana aperturą sondy skanującej 20÷100 nm (pokryte na końcu metalem włókno optyczne), która oświetla „punktowo” (skanuje) kolejne fragmenty powierzchni z bliskiej odległości $\sim 10 \text{ nm}$ falą o zanikającym eksponencjalnie (*evanescent*) polu elektrycznym bez absorpcji (patrz SPR).

W mikroskopii fluorescencyjnej **dalekiego pola STED** (*stimulated emission depletion microscopy*) wzrost apertury uzyskuje się bardzo silną wiązką wygaszającą (w kształcie obwarzanka) poprzez emisję wymuszoną części wzbudzanych cząsteczek poza centralnym minimum w płaszczyźnie ogniskowej. Zdolność rozdzielcza dla próbek biologicznych osiąga $20 \div 100$ nm. Podobnie, wycięcie sygnału emisji z płaszczyzn pozaogniskowych przy detekcji jest podstawą fluorescencyjnego **mikroskopu konfokalnego**.

W **mikroskopii elektronowej EM** rozdzielczość wynika ze skrócenia długości fali elektronów w porównaniu ze światłem widzialnym. Wyróżnia się mikroskopy transmisyjne (TEM) i skaningowe (SEM). W przypadku najlepszych **transmisyjnych mikroskopów elektronowych cryoEM** pojedynczej cząsteczki (300 kV mikroskop firmy Thermo Fisher Scientific) obrazowanie dużej liczby **witryfikowanych cząsteczek w roztworze**, osadzonych w próżni na metalowych siatkach (*grid*) w niskich temperaturach daje **strukturę 3D po komputerowej obróbce obrazu**. Standardowo zyskuje się **rozdzielczości** $\sim 3 \div 5 \text{ \AA}$ dla dużych kompleksów **[PLANSZA 195]**. Przekroczenie bariery **rozdzielczości** 2 \AA przy rozmiarach obrazowanych obiektów **100 kDa** (i poniżej) ([Merk et al. Cell 165, 1, 2016](#)) uzyskuje się dzięki coraz nowszym rozwiązaniom technicznym, np. nowy typ płytki fazowej dla wysokiego kontrastu obrazu (*Volta phase plate*; [Li et al. J. Struct. Biol, 2019 in press](#)). „Rekordową” rozdzielczości 1.54 \AA uzyskano dla 500 kDa apoferrytyny o ([Kato et al. Microscopy and Microanalysis 25, 998, 2019](#)). Przełomowy charakter w „rewolucji rozdzielczości” cryoEM (nagroda Nobla z chemii w 2017 r.) miały trzy czynniki: wynalezienie bezpośredniego detektora elektronów, opracowania sposobu przygotowania próbek na siatce oraz metody obliczeniowe do uśredniania setek tysięcy obrazów pojedynczej cząsteczki, czyli rekonstrukcja obrazu 3D w formie mapy gęstości elektronowej z projekcji 2D, przez zastosowanie transformacji Fouriera). Nową techniką wykorzystania wiązek elektronów do obrazowania 3D jest **tomografia elektronowa cryoET** ([Yahav et al. Curr. Op. Struct. Bio. 21, 670, 2011](#)), chociaż o rozdzielczości ~ 5 nm lepiej niż cryoEM pozwala uchwycić **natywne asocjaty molekularne in situ**.

Rozdzielczości $\sim 5 \div 10 \text{ \AA}$ obrazu makromolekuł uzyskuje się za pomocą **mikroskopów siły atomowej AFM** (*atomic force microscopy*) **[PLANSZA 196]**. Pojedynczą makrocząsteczkę w kropli roztworu na płaskiej powierzchni skanuje się bardzo ostrym rylcem. **Sily molekularne van der Waalsa** między ostrzem i podłożem **rzędu pikonewtonów** ($1 \text{ pN} = 10^{-12} \text{ N}$) 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 poprzez odpowiednie przesunięcie rylca, tak aby odległość między ostrzem a powierzchnią (siła van der Waalsa) skanowania pozostawała stała. Służy do tego celu najczęściej **serwomechanizm** w postaci wiązki laserowej odbitej od wspornika. W ten sposób rejestrowany jest **obraz powierzchni, a więc kształt umieszczonej na niej cząsteczki lub kompleksu** molekularnego. Rejestracja obrazów **w czasie realnym** pozwala śledzić zmiany zachodzące w cząsteczkach (**A ÷ D**). Aktualne wersje HS-AFM (*high-speed AFM*; [Preiner & Hinterdorfer G.I.T. Imaging & Microscopy 1, 18, 2017](#)) umożliwiają rejestrację **dynamiki zmian strukturalnych z rozdzielczością czasową kilku milisekund** (i rozdzielczością przestrzenną rzędu nanometra): wędrówka myozyny V po aktynie, zmiany konformacyjne statora syntazy ATP, dyfuzja pierścieni syntazy ATP w błonie, usztywnienie powierzchni *E. coli* pod wpływem peptydu antybakteryjnego. Zaczepienie rylcem cząsteczki przytwierdzonej do podłoża pozwala poruszać nią 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** [**PLANSZA 124**]. Mikroskopię AFM sprzęga się z innymi technikami mikroskopowymi (NSOM) i technikami spektroskopowymi, np. TERS (*tip-enhanced Raman scattering*; [Rasmussen & Deckert J. Raman Spectr. 37, 311, 2006](#)).

Techniki manipulacyjne z wykorzystaniem sił pikonewtonowych (10^{-12} N) stosuje się w **szczypcach optycznych** [**PLANSZA 197**]. Do makrocząsteczki przyczepia się paciorek dielektryczny (np. polistyrenowy) ok. $1\div 3$ μm . Wiązka laserowa utrzymuje paciorek w środku zogniskowanej wiązki na zasadzie **przekazu pędu fotonu**: $p = (h\nu)/v$, (gdzie $v = cn$, prędkość światła w ośrodku o współczynniku załamania n) i umożliwia jego przesuwanie. W **szczypcach magnetycznych** [**PLANSZA 198**] pole magnetyczne wywołuje rotację i przeciąganie paciorka ferromagnetycznego wraz z przyczepioną cząsteczką, np. DNA, i umożliwia pomiary siły i momentu siły (**b**). Rozwój **spektroskopii pojedynczej cząsteczki SMS** (*single molecular spectroscopy*) umożliwia **wizualizację** manipulowanych cząsteczek. *in vitro* i ich **fluorescencyjne obrazowanie** (nagroda Nobla z chemii w 2014 r.) *in vivo* wykorzystując barwniki organiczne i kropki kwantowe jako sondy fluorescencyjne ([Yang Curr. Op. Chem. Biol. 14, 3, 2010](#)). Standardowe techniki różnią się obszarem wzbudzenia/detekcji (na zielono) [**PLANSZA 199**]: **szerokopolowa** (*wide-field*) ze wzbudzeniem blisko skolimowaną wiązką, **konfokalna** (*confocal*) ze wzbudzeniem skolimowaną wiązką laserową, fluorescencja **całkowitego odbicia**

wewnętrznego **TIRF** (*total internal reflexion fluorescence*) oraz **ZMW** (*zero-mode waveguides*) z użyciem fali *evanescent*. Techniki **ISM** (*image scanning microscopy*; [Gregor et al. Nature Methods 14, 1087, 2017](#)) i **SOFI** ([Dertinger et al. Proceedings National Academy of Sciences USA 106, 22287, 2009](#)) podwajają rozdzielczość przez zastosowanie procedur "smart" wzbudzenia i detekcji: wzbudzenia dwufotonowe, generacja drugiej harmonicznej lasera wzbudzającego, a także przez zastosowanie **efektów kwantowych korelacji fotonów Q-ISM** (*quantum image scanning microscopy*; [Tenne et al. Nature Photonics 13, 116, 2019](#)). Hybrydowa metoda **SOFIZM** łącząca ISM i SOFI i podnosząca rozdzielczość 2,5 raza została opracowana w IFD Wydziału Fizyki UW przy współpracy z Weizman Institute of Sciences (Izrael) ([Sroda et al. Optica 7, 1308, 2020](#)). **Lokalizację** cząsteczek czyli ich znaczników fluorescencyjnych kamerami CCD lub fotodiodami uzyskuje się z rekonstrukcji obrazów emisyjnych **plamek dyfrakcyjnych PSF** (*point spread function*) przez wsparcie rozwiązań doświadczalnych zaawansowanymi, komputerowymi technikami obliczeniowymi (klasteryzacji) ([Baddley Nature Methods 12, 1019, 2015](#)). Technika **PALM** (*photoactivated localization microscopy*; [Betzig et al. Science 313, 1642, 2006](#)) z obrazowaniem **TIRF** była omawiane poprzednio: śledzenie kompleksów białkowych w procesie **NER** naprawy DNA. Redukcja szumów (*background noise*) poprzez detekcję odpowiedniej liczby fotonów ($N \sim 10^4$) z jednego chromoforu zanim zajdzie *photobleaching* pozwala zredukować błąd statystyczny $N^{1/2}$ razy i uzyskać położenia centrum emisji z dokładności **kilku nanometrów**. Technika **COLD** (*cryogenic optical localization in 3D*; [Weisenburger et al. Nature Methods 14, 141, 2017](#)) umożliwia wzajemną lokalizację chromoforów przyczepionych do (małego) białka z rozdzielczością $\sim 6 \text{ \AA}$. a efekt **smFRET** [**PLANSZA 134**] do **kolokalizacji** fluoroforów przy śledzeniu zmian konformacyjnych i ilościowego wyznaczania sił $\sim 3 \div 5 \text{ pN}$ wewnątrz komórki (FL-TSM, *ferredoxin linker sensor module*; [Ringer et al. Nature Methods 14, 1090, 2017](#)). **Dalsze rozwiązania to łączenie technik** obrazowanie fluorescencyjne/mikroskopia elektronowa **CLEM** (*correlative light and electron microscopy*; [Zhang Curr. Op. Struct. Biol. 23, 763, 2013](#)), wykorzystanie **spektroskopii FTIR** z zastosowaniem promieniowania synchrotronowego dużej intensywności (mikroskopia w podczerwieni) do korelacji struktur cząsteczek z ich biologicznymi funkcjami ([Miller & Dumas Curr. Op. Struct. Biol. 20, 649, 2010](#)), **mikroskopia fototermiczna PTIM** (*photothermal imaging microscopy*; [Gruszecki et al. Analytical Chem. 2015, in press](#))

Wykład 24

Działanie dużych, wieloskładnikowych kompleksów, MOTORÓW MOLEKULARNYCH zostanie przedstawione na przykładach typowych procesów, charakterystycznych dla funkcjonowania komórki. Punktem wyjścia są struktury przestrzenne z rozdzielczością atomową (dyfrakcja rentgenowska, NMR), natomiast dynamikę funkcjonowania w czasie realnym uzyskuje się na podstawie **eksperymentów SM** (*single-molecule*). Analiza pojedynczych, wieloskładnikowych kompleksów molekularnych pokazuje, że **procesy ekspresji genu** wywodzące się z pojedynczej cząsteczki DNA w komórce mają **charakter stochastyczny** [PLANSZA 200]. Poziom ekspresji RNA i białek wykazuje **fluktuacje czasowe**, co ma istotny wpływ na fenotyp komórki.

Replikacja DNA [PLANSZA 201] jest dynamicznym procesem sprzężenia molekularnych motorów. Heksameryczna **helikaza** poruszając się wzdłuż DNA i działając **mechanicznie**, rozdziela nici podwójnej helisy dsDNA (*double stranded DNA*) poprzez termodynamiczne sprzężenie z reakcją **hydrolizy nukleozydotrifosforanów** (NTP). W ten sposób **koordynuje sprzężoną** (w zwykłej replikacji) **syntezę na niciach *leading* oraz *lagging***, oddziałując z enzymami tworzącymi primery (*primases*) oraz z wykorzystującymi te primery polimerazami (*leading and lagging strand polymerases*). Poszczególne białka charakteryzują się określonymi szybkościami działania [PLANSZA 202], a najwolniejsze z nich wyznacza globalną szybkość całego procesu (*rate limiting step*). Badania z wykorzystaniem pojedynczych molekuł dostarczają bezpośrednich danych odnośnie **szybkości zachodzenia i sił mechanicznych** dla poszczególnych etapów replikacji, np. **rozwijanie** dsDNA przez helikazę (szczypce optyczne i magnetyczne) i **synteza** nici potomnej przez polimerazę (szczypce optyczne) [PLANSZA 203]. W analogiczny sposób można badać proces **transkrypcji**: polimeraza RNA (RNAP) poruszająca się wzdłuż nici DNA [PLANSZA 204] ([review: Larson et al. Mol. Cell 41, 249, 2011](#)).

Dwie **pompy protonowe** F-ATPaza (*ATP-synthase*) i V-ATPaza (*endomembrane proton pump*) [PLANSZA 205] mają różne funkcje fizjologiczne, ale są bardzo podobne pod względem struktury i mechanizmu działania. Składają się z fragmentów zewnątrz-błonowych (*membrane extrinsic F1 or V1*) i wewnątrz-błonowych (*membrane intrinsic Fo or Vo*), **łącząc syntezę/hydrolizę ATP z transportem protonów przez błonę za pomocą**

mechanizmu rotacyjnego. Obserwacja odbicia światła laserowego od złotego paciorka o optymalnych rozmiarach $40 \div 60$ nm [PLANSZA 206a] przyczepionego do rotora (γ) [PLANSZA 205] oraz równoległe pomiary FRET między sondami fluorescencyjnymi przyczepionymi do rotora i statora ($\alpha_3\beta_3\text{ab}_2$) umożliwiły wyznaczenie **czasowego przebiegu obrotu** ze średnią szybkością ok. 400 rps ([206b]) przedzielonego pauzami ok. 2s ([206c]); rozzrzut statystyczny czasu pojedynczego obrotu $2 \pm 1,5$ ms ([206d]). Wyniki pozwoliły skonstruować mechanizm działania ATPazy [PLANSZA 207], w którym **reakcja enzymatyczna i przebieg czasowy oddziaływań między podjednostkami** (obrot γ w funkcji czasu) są **skoordynowane** dla obniżenia bariery energii aktywacji.

Miozyny (35 klas białek), wędrujące po filamentach aktynowych [PLANSZA 208] i **kinezyny** poruszające się po mikrotubulach [PLANSZA 209], są przykładami motorów aktywnego ruchu i transportu molekularnego, napędzanego energią hydrolizy ATP. Ruch miozyn [PLANSZA 208] opisuje **hipoteza „kołyszącego ramienia dźwigni”** (*swinging lever arm*): przed wykonaniem suwu (*powerstroke; red*) przez ramię dźwigni i po jego wykonaniu (*rigor state; purple*). Klasa VI miozyn różni się od pozostałych przeciwnym kierunkiem ruchu dzięki dodatkowym domenom i innym mechanizmem sprzężenia ruchu z hydrolizą ATP. Szczegółowych danych do **konstrukcji modeli** dostarcza **analiza czasowa przemieszczeń białka** z wykorzystaniem pułapek (szczypiec) optycznych, np. 8 nm „kroki” kinezyny, przedzielone zmiennymi okresami spoczynku [PLANSZA 209].

Motory molekularne, jak omówione wyżej miozyny, mogą aktywnie transportować inne cząsteczki. **Transport DNA** ma zasadnicze znaczenie dla życia komórki eukariotycznej, od separacji chromosomów w procesie podziału komórki do tworzenia otoczki wirusowego DNA (*encapsidation*). Wskutek okrągłej, zamkniętej formy wielu chromosomów bakteryjnych rekombinacja homologiczna (*homologous recombination*) wytwarza dimery chromosomowe, które muszą się rozdzielić przed podziałem komórkowym. **Podstawowym białkiem katalizującym proces jest FtsK, pompujący schwytny dimer do specyficznych rekombinaz XerC/D** [PLANSZA 210c]. N-terminalna domena FtsK wiąże się z błoną a jego C-terminalna część jest heksamerem z centralnym otworem, w który wchodzi dsDNA [210a,b] Badania *in vitro* metodą szczypiec optycznych i magnetycznych [210d] aktywnej formy FtsK dostarczają danych o zmianach rozciągnięcia DNA, a stąd o procesywności FtsK oraz prędkości pompowania DNA.

Wspominany już (kilkakrotnie) **kompleks błony jądrowej NPC** (*nuclear pore complex*) jest przypuszczalnie **największą nanomaszyną** w komórce (50 mln Da), w formie wydrążonego cylindra w kopercie jądrowej o średnicy 100 nm **[PLANSZA 211]**. NPC jest „bramą” między jądrem a cytoplazmą o ośmiokrotnej symetrii kompleksu 30 typów różnych białek, tzw. nukleoporyn (Nups), w sumie ponad 450 białek. Woda i metabolity przechodzą swobodnie. Inne cząsteczki poniżej 60 kDa mogą przejść, jeśli niosą specyficzny sygnał importu i eksportu, np. cząsteczki mRNA: *β-actin mRNA transport*; pomiary pozycji względem POM121 - *fluorescence marker fusion protein*. Szczegółowy mechanizm eksportu jest słabo poznany **[PLANSZA 212]**. Szczególnie istotne jest wyjaśnienie sposobu(ów) transportu biopolimerów o bardzo zróżnicowanych rozmiarach i kształcie, co stanowi „prawdziwe wyzwanie” dla kompleksu NPC. W szerokim zestawie metod badania NPC stosowane są również badania „*single - molecule*” z wykorzystaniem głównie mikroskopii fluorescencyjnej: SMT (*single-molecule tracking*) i SPEED (*single point edge excitation subdiffraction*) do wyznaczania położenia cząsteczek, oraz *super-registration microscopy*, 4-Pi microscopy (rozszerzenie mikroskopii konfokalnej) do badania oddziaływań molekularnych, FCS (*fluorescence correlation spectroscopy*) do śledzenia dynamiki ruchu dyfuzyjnego.

BŁONY BIOLOGICZNE oraz

PROCESY BIOENERGETYCZNE (transfer elektronów i protonów)

Błona lipidowa (wykład „*Struktura i funkcje makrocząsteczek biologicznych*” **[PLANSZA 35]**) jest molekularną, quasi-dwuwymiarową strukturą **heterogeniczną** zawierającą zróżnicowane elementy o wymiarach w skali nano, 2 ÷ 200 nm. Szereg stwierdzeń i założeń w przypadku pierwszego modelu błony *fluid mosaic model* (jednolita płynna struktura błony) sformułowanego ponad 40 lat temu, zostało zrewidowanych w świetle późniejszych dowodów, wskazujących na **strukturalnie asymetryczny charakter organizacji matrycy lipidowej w domeny** (*lateral domains, rafts*) Wyróżnia się dwa typy domen błonowych **[PLANSZA 213]**.

(1) *Raft membrane (micro)domains* (*membrane/lipid rafts*, MLR; [Head et al. Biochem. Biophys. Acta 1838, 532, 2014](#)), dynamiczne agregaty warstwy lipidowej wzbogacone cholesterolem, białkami (*raft-associated proteins*) i nasyconymi sfingolipidami (*sphingomielin*), w granicach od klastrów niewielkich białek o krótkim czasie życia do czasowo ustabilizowanych domen indukowanych przez receptory sygnałowe, np. T- i B-

receptory. Organizacja i *clustering* MLR w aktywne **platformy sygnałowe** zależy od oddziaływań z cytoskeletonem i dynamiki jego przekształceń.

(2) **Apposed domains** wynikające z kompartmentalizacji błony przez usieciowanie (*membrane skeleton*), połączone białkami z siecią filamentów aktynowych - ograniczają ruchy dyfuzyjne cząsteczek w błonach: szybkie przemieszczenia w ramach domen i długie w skali czasowej przeskoki między domenami.

Do badań błon stosuje się szereg metod podobnie jak w badaniach białek i kwasów nukleinowych. Szczególnie dobre rezultaty w **obrazowaniu** struktur błon uzyskano przy zastosowaniu mikroskopii AFM ([Muller Biochemistry 47, 7986, 2008](#)) oraz **obrazowania fluorescencyjnego FLIM** (*fluorescence life-time imaging*; [Korczyński & Włodarczyk Postępy Biochemii 55, 434, 2009](#); aparatura pomiarowa jest dostępna w Zakładzie Biofizyki). Oprócz **obrazów mikroskopowych** technika FLIM dostarcza informację o **oddziaływaniach międzycząsteczkowych** w błonie poprzez wykorzystanie **efektu FRET**, którego wydajność wyraża się przez czasy życia:

$$E = 1 - (\tau_{DA}/\tau_D)$$

Pomiar nanosekundowych czasów życia wzbudzonych fluoroforów (czasu życia fluorescencji) można przeprowadzić zasadniczo na dwa sposoby (wykład „Spektroskopia molekularna”):

- w **domenie czasowej** po wzbudzeniu krótkim impulsem lasera z rejestracją zaniku emisji przy użyciu licznika pojedynczych fotonów TCSPC (*time correlated single photon counting*);

- w **domenie częstości** poprzez ciągłe wzbudzenie, modulowane z częstością ω rzędu odwrotności czasu życia τ ($\omega \sim 1/\tau$).

Najczęściej spotykane **układy FLIM** to mikroskopy szerokiego pola w domenie częstości oraz mikroskopy skaningowe w domenie czasowej. Jako rezultat pomiarów otrzymuje się **mapę krzywych zaniku fluorescencji**, np. różnych domen w błonie [[PLANSZA 214](#)], a stąd po dopasowaniu odpowiednich modeli zaniku (wieloeksponencjalnego), rozkład czasów życia w danym fragmencie ROI (*region of interest*). Rozkład przedstawia się w formie barwnego obrazu mikroskopowego jak w przypadku znakowanych sondą fluorescencyjną NBD-cholesterol (A) następujących typów błon na [PLANSZY](#) (biały odcinek odpowiada długości 10 μm):

I - pęcherzyk GUV (*giant unilamellar vesicle*) „sztucznej” błony;

II - GPMV (*giant plasma membrane vesicle*) z komórek HeLa;

III - komórki HepG2.

W części hydrofilowej błony na granicy błona-cytoplazma, w tzw. interfazie (*interphase*; **[PLANSZA 215]**) znajdują się związane cząsteczki strukturalnej wody, która modyfikuje strukturę/funkcję składników błony oraz jej parametry termodynamiczne, np. napięcie powierzchniowe. **Ultraszybkie eksperymenty** echa wibracyjnego (2D IR *vibrational echo*; wykład, „*Spektroskopia molekularna*”) są używane do badania dynamiki cząsteczek wody w interfazie lipidowej w skali czasowej pikosekund.

Błony charakteryzują się **silną integralnością w płaszczyźnie** (*lateral integrity*), wywołaną głównie hydrofobowym oddziaływaniem składnika lipidowego. Mała grubość dwuwarstwy stwarza znaczne możliwości **wyginania powierzchni** błony i tworzenia pęcherzyków. Przykładowo, energia E na jeden lipid dla wygięcia błony w pęcherzyk o średnicy 50 nm:

$$E = 8\pi K/N_{\text{lipids}} \leq k_B T \quad \text{moduł na zginanie } K \sim 10 \div 40 \text{ } k_B T$$

k_B - stała Boltzmanna; T - temperatura

Komórki wykształciły mechanizmy remodelowania błon (*membrane remodeling*; **[PLANSZA 216]**), polegające na tworzeniu nowych, zamkniętych przedziałów błonowych (*fission*) oraz łączenia poszczególnych fragmentów błon (*fusion*). Mechanizm jest uwarunkowany zarówno przez oddziaływania lipid - lipid jak i przez udział cząsteczek białkowych (*Sens et al. Curr. Op. Cell Biol. 20, 476, 2008*).

Jednym z fundamentalnych procesów w żywej komórce jest **transport jonów i cząsteczek przez błony** za pośrednictwem białek błonowych, które tworzą **kanały i transportery błonowe** **[PLANSZA 217]**.

(A) **Kanały błonowe** zapewniają pasywny przepływ przez błonę bez zużycia energii (*downhill*), o optymalnej, ale zróżnicowanej selektywności i przepuszczalności, oraz przepustowości (włączenie/wyłączenie, *gating*) zależnej od czynników otoczenia, np. akwaporyny do przepływu wody, alkoholi liniowych, gazów, selektywne kanały potasowe, nieselektywne kanały kationowe (Na^+ , K^+) czułe na siły mechaniczne Ms (*mechanosensitive*), zależne od pH, *signal peptide-gated* i in. **Przepływ jonów przez błony** jest jednym z fundamentalnych procesów w komórce sprzężonym z wytworzeniem **elektrycznego napięcia transmembranowego** (*review: Cooper et al. Progr. Biophys. Molec. Biol. 46, 51, 1985*). Procesy uwarunkowane tym napięciem to m. in.:

- generowanie potencjałów działania (*action potentials*) w *excitable cells*: nerwowe, mięśniowe, wydzielające hormony, niektóre roślinne (kanały *voltage gated*);
- regulacja objętości komórki;
- *epithelial transport* ECF (*extracellular fluid*)/lumen, czyli transport przez nabłonek;
- przewodzenie sygnałów nerwowych (potencjały postsynaptyczne);
- aktywacja limfocytów;
- ruchy komórkowe.

Teoria transportu jonów przez kanały błonowe może być ujęta w różny sposób:

(a) analitycznie, na podstawie modeli stochastycznych, które prowadzą do równań przepływu teorii elektrodyfuzyjnej Nernsta-Plancka (patrz [review Cooper et al.](#)) oraz **teorii Eyringa**, przy założeniu szeregu (i) kolejnych minimumów energetycznych w błonie na drodze przepływu, oddzielonych o λ ($i \rightarrow i + 1$), oraz oddzielonych barierami energetycznymi $\Delta G^{\#}_{i,i+1}$:

$$\text{przepływ: } J = k' \lambda c_i - k' \lambda c_{i+1} \text{ [ilość substancji/(cm}^2 \cdot \text{s)]}$$

$$k' = (1/2\pi k_B) D k_i k_{i+1} \exp(\Delta G^{\#}_{i,i+1}/RT)$$

c_i - stężenie jonów w i-tym minimum; k_i - stała siłowa i-tego minimum energii w przybliżeniu harmonicznym.

(b) zastosowanie brownowskiej dynamiki molekularnej BMD (*Brownian molecular dynamics*).

(B) Transportery i nośniki (*carriers*) do aktywnego transportu cząsteczek przez błony dzielą się na dwa zasadnicze typy **[PLANSZA 217]**:

- **transportery ABC** (ATP - *binding cassettes*), wykorzystują hydrolizę ATP jako źródło energii; przykłady: transporter maltozy (MalEFGK), transporter witaminy B₁₂ (BtuB);

- **secondary transporters**, sprzęgają aktywny transport substratów z potencjałem elektrochemicznym jako źródłem energii, w formie ko-transportu jonów (Na⁺, H⁺) zgodnie z ich gradientem stężenia ([Forrest et al. Biochim. Biophys. Acta 1807, 167, 2011](#)).

Wymienione aspekty strukturalne i dynamika błon oraz procesy transportu przez błonę są uwarunkowane przez oddziaływania lipid - lipid LLI (*lipid - lipid interactions*; [Quinn Progr. Lipid Res. 51, 179, 2012](#)) i oddziaływania białko - lipid PLI (*protein - lipid interactions*; [Cho et al. Proteomics 12, 3273, 2012](#)). Jednym z kluczowych wyzwań stało się wyjaśnienie strukturalnych cech kompleksów LL, które są odpowiedzialne za asocjację

i organizację białek błonowych w funkcjonowaniu *rafts*. Metodologia badań doświadczalnych PLI oparta jest na **podejściu proteomicznym**, (omówionym wcześniej podczas wykładu), wzbogaconym o symulacje dynamiki molekularnej zawiązania się (*self-assembly*) struktury lipidowo-białkowej.

Wykład 25

Specyficzną klasę motorów molekularnych stanowią **związane z błoną kompleksy białkowe** dostarczające **energię komórce**. W biofizyce procesów pozyskiwania energii dwa zagadnienia wydają się najistotniejsze z punktu widzenia wykładu:

- (a) mechanizmy molekularne bioenergetyki, które bazują na tworzeniu **dużych kompleksów białkowych w błonach lipidowych**;
- (b) sprzężone procesy **transferu elektronów i jonów H⁺**.

Pozyskiwanie energii jest niezbędne do zachodzenia wszystkich procesów w komórce jako **układzie otwartym w sensie termodynamicznym**, daleko od stanu równowagi. Pierwotnym źródłem energii w **biosferze** jest promieniowanie elektromagnetyczne Słońca przekształcane zależnie od typu organizmu. **Energia słoneczna** jest wykorzystywana w wymiarze 0,1% z całkowitej ilości $\frac{1}{2}$ razy $5 \cdot 10^{21}$ kJ/(mol·rok), czyli połowy całkowitego rocznego wydatku energii słonecznej, która dociera do powierzchni Ziemi, po odbiciu i rozproszeniu w atmosferze.

Energia słoneczna jest przetwarzana bezpośrednio w procesie **FOTOSYNTETY** w sinicach (*cyanobacteria*), algach i roślinach (chloroplasty). **[PLANSZA 218]**. Proces ma charakter dwuetapowy.

- (1) *Light reaction*: **wzbudzenia elektronowe** w barwnikach, które prowadzą do syntezy adenozyntrifosforanu ATP **[PLANSZA 219]** i ufosforylowanego dinukleotydu nikotynoamidoadeninowego NADPH **[PLANSZA 218]**;
- (2) *Dark reaction* **synteza cukrów** z CO₂ i H₂O z wykorzystaniem ATP i NADPH (cykl Calvina-Bensona)

Wypadkowy efekt fotosyntezy: $h\nu + 6\text{CO}_2 + 6\text{H}_2\text{O} \rightarrow \text{C}_6\text{H}_{12}\text{O}_6$ (glukoza) + 6O₂

Dwa typy fotosystemów działają kolejno w celu wytworzenia odpowiedniego **potencjału oksydo-redukcyjnego związanego protonu**.

- (1) **Fotosystem II** ([Dau et al. Curr. Op. Chem. Biol. 16, 3, 2012](#)) indukuje **rozszczerzenie wody przez światło na molekularny tlen i wodór** (H⁺ + e⁻) oraz zwiążanie wodoru przez plastochinon (PQ) i powstanie formy zredukowanej PQH₂, o potencjale redukcyjnym +80 mV i zysku energetycznym $\Delta G(h\nu) \approx 217$ kJ/mol.
- (2) **Fotosystem I** ([Busch & Hippler Biochim. Biophys. Acta 1807, 864, 2011](#)) daje podniesienie potencjału redukcyjnego do poziomu NADPH, -320 mV. PQH₂ działa jako

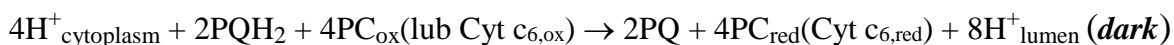
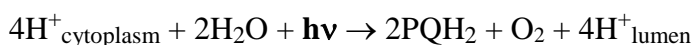
reduktant w kompleksie cytochromu b6f, który jest donorem elektronów do PSI: redukcja **plastocjaniny** (PC) w roślinach lub cytochromu c₆ (Cyt c₆) w sinicach. Efekt to **transport elektronów z H₂O na NADP⁺** (czerwone strzałki) i **sprzężony z nim transfer jonów H⁺**.

W przypadku organizmów beztlenowych działa tylko jeden z dwóch systemów. Do rozwiązania struktur fotosystemów w formie kryształów o rozmiarach mikrometrowych (μm) użyto lasery na swobodnych elektronach, LCLS do PSII (3,5Å; [Ayver et 34 co-authors..., Chapman, Nature 530, 202, 2016](#)) i EuXFEL do PSI (2,9Å; [Gisriel et 78 co-authors..., Zatsepin, Nature Commun. 10, 1, 2019](#)) **[PLANSZA 220]**. W pracy o PSII zastosowano **nowe metody analizy dyfraktogramu** (m. in. *iterative phasing*), a w pracy o PSI **technikę SFX** (*serial femtosecond crystallography*) dostarczania kolejnych mikrokryształów w wiązce z częstotliwością ~1,1 MHz, w przerwach między impulsami lasera.

Przestrzenne rozdzielenie utleniania i redukcji (transfer elektronów) z utworzeniem **transmembranowej różnicy potencjału elektrochemicznego protonów ΔpH** **[PLANSZA 219]** jest „siłą napędową” **syntazy ATP** **[PLANSZA 205]**, ΔG(hν) > ΔG(ATP). Promieniowanie słoneczne jest absorbowane przez chromofory trzech typów barwników **[PLANSZA 221]**: (bakterio)chlorofile, phycobiliny (*phycocyanobilin*) i β-karotenoidy. Badania doświadczalne pokazały, że tylko jedna cząsteczka tlenu powstaje na ok. 2 500 cząsteczek barwników przy oświetlaniu komórek krótkimi impulsami światła 10 μs, co prowadzi do modelu organizacji barwników w dwa układy **[PLANSZA 222]**, dla każdego z systemów fotosyntetycznych PSI i PSII:

- **system antenowy** zbierania światła (*light harvesting antenna*), który zawiera przeważającą ilość barwników i o dwa rzędy wielkości **podwyższa przekrój czynny** na absorpcję kwantów światła;

- **centrum reakcji RC** liniowego transferu elektronów z H₂O na NADP⁺, z pobieraniem protonów z cytoplazmy i ich uwalnianiem w lumenie (błona tylakoidalna):



Szybkość transferu elektronów jest limitowana przez obrót PQ/PQH₂ systemu oksydacyjno-redukcyjnego. **Sprawność systemu fotosyntezy** próbowano porównywać ze sprawnością urządzeń fotowoltaicznych (PV) do elektrolizy wody ([Blankenship et al. Science 332, 805, 2011](#)). Energia swobodna Gibbsa **rozszczenia** cząsteczki wody na

tlen i wodór jest w obu przypadkach podobna, odpowiednio 1,24 i 1,23 eV. System fotowoltaiczny (sprawność 10% ÷ 11%) okazuje się wydajniejszy od fotosyntezy (1% ÷ 7%) w długiej skali czasowej, ale fotosynteza jest unikalnym systemem produkcji szeregu złożonych związków organicznych.

Fotosynteza ma niezwykle ważną konsekwencję dla biosfery przez utworzenie atmosfery tlenowej z powstającego w reakcji fotosyntetycznej „waste’ product” jakim jest O₂, co otwiera drogę dla znacznie wydajniejszego procesu pozyskiwania energii drogą oddychania tlenowego w organizmach heterotroficznych. **Utlenianie cząsteczek biologicznych** z pożywienia roślinnego i zwierzęcego, z ostatecznymi produktami w postaci CO₂ i H₂O i pozyskiwaniem „wysokoenergetycznego” ATP oraz NADH, zachodzi w trzyetapowym katabolizmie.

- (1) **Degradacja makromolekuł**, białek, polocukrów, tłuszczu (głównie pozakomórkowa);
- (2) **Degradacja monomerów** makromolekuł do **acetylokoenzymu A*** **[PLANSZA 221]** (adenozyno-3'-fosforan-PO₃⁻-PO₃⁻-kwas pantoinowy-β-alanina-cysteamina), poprzez beztlenową **glikolizę** z wytwarzaniem niewielkich ilości ATP i NADH;
- (3) Całkowita **oksydacja** acetyloCoA do CO₂ i H₂O w **mitochondriach**, od cyklu Krebsa z wytwarzaniem dużych ilości NADH do **oksydacyjnej fosforylacji z transferem elektronów (redukcja) na O₂**, czyli **ODDYCHANIE KOMÓRKOWE** z wytwarzaniem dużych ilości ATP:



Sprzężenia procesów oksydo-redukcyjnych oddychania komórkowego z fosforylacją ADP w błonie mitochondrialnej prowadzi do powstania wysokoenergetycznego ATP w organizmach heterotroficznych analogiczne do procesu fotosyntetycznego. **Chemiosmotyczna teoria Mitchella** (*oryginalna publikacja: Mitchel Biol. Rev. 41, 445, 1966; powtórnie opublikowana: Biochim. Biophys. Acta 1807, 1507, 2011; nagroda Nobla w 1978 r.*) **[PLANSZA 223]** ujmuje zarówno sprzężone procesy oksydacyjnej fosforylacji oraz omówione procesy fotosyntetycznej fosforylacji. Podstawowe **postulaty** teorii chemiosmotycznego sprzężenia:

1. The membrane-located ATPase systems of mitochondria and chloroplasts are hydro-dehydration systems with terminal specificities for water and ATP; and their normal function is to couple reversibly the translocation of protons across the membrane to the

flow of anhydro-bond equivalents between water and the couple ATP/(ADP+Pi)
[PLANSZA 224].

2. *The membrane-located **oxido-reduction chain systems** of mitochondria and chloroplasts catalyse the flow of reducing equivalents, such as hydrogen groups and electron pairs, between substrates of different oxido-reduction potential; and their normal function is to **couple reversibly the translocation of protons across the membrane to the flow of reducing equivalents during oxido-reduction** **[PLANSZA 225].***

3. *There are present in the membrane of mitochondria and chloroplasts **substrate-specific exchange-diffusion carrier systems** that permit the effective reversible trans-membrane exchange of anions against OH^- ions and of cations against H^+ ions; and the normal function of these systems is to **regulate the pH and osmotic differential across the membrane, and to permit entry and exit of essential metabolites** (e.g. substrates and phosphate acceptor) without collapse of the membrane potential.*

4. *The systems of postulates 1, 2, and 3 are located in a specialised **coupling membrane** which has a **low permeability to protons and to anions and cations** generally.*

Potencjał błonowy ΔV i potencjał elektrochemiczny ΔpH niezbędne do zapewnienia fosforylacji ATP/ADP powstają w czterech kompleksach **[PLANSZA 226]** systemu oddychania komórkowego:

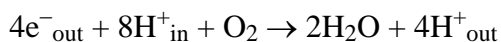
$$\Delta V - 60\Delta pH \cong 210 \text{ mV}$$

Szczegółowe badania doświadczalne i teoretyczne w zakresie biofizycznych podstaw procesów bioenergetycznych są skoncentrowane na dwóch typach zagadnień:

- **analiza strukturalna kompleksów błonowych**, zaangażowanych w funkcjonowanie procesów przetwarzania energii;
- **fizyczne podstawy procesów transferu elektronów i protonów**.

Przykładem pierwszej grupy zagadnień jest strukturalna analiza funkcjonowania **ATP-az** łączących syntezę/hydrolizę ATP z transportem protonów przez błonę **[PLANSZE 205 ÷ 207]**. **Oksydaza cytochromu c** jest drugim przykładem pompy protonowej, jako końcowy enzym w łańcuchu transportu elektronowego systemu oddechowego w wewnętrznej błonie mitochondrialnej (lub bakteryjnej) **[PLANSZA 226]**. Katalizuje redukcję tlenu do wody

generując elektrochemiczny gradient protonowy przez błonę [\[PLANSZA 227\]](#) do produkcji ATP [\[PLANSZA 205\]](#). Podczas reakcji [\[PLANSZA 227\]](#) pobiera 4 elektrony z dodatnio naładowanej (zewnątrznej) P-części odgraniczonej błoną i 4 jony H^+ z ujemnie naładowanej N-części (wewnętrznej):



Cztery dodatkowe protony są przetrzucane przez błonę ze strony N do P przeciwnie do kierunku potencjału błony zwiększając dwukrotnie całkowitą separację ładunków. Pomimo wszechstronnych badań biofizycznych i biochemicznych, w tym rentgenografii, **kluczowe aspekty mechanizmu pompowania protonów** pozostają niejasne . Separacja ładunku przez błonę zachodzi w dwóch oddzielnych etapach. Najpierw, redukcja O_2 do H_2O przez elektrony i protony wzięte z obu stron błony prowadzi do przeniesienia jednego elementarnego ładunku przez błonę na jeden „zużyty” elektron, zgodnie z teorią Mitchella. Potem, dodatkowy proton jest przeniesiony przez błonę podwajając separację ładunku. **Sprężone z reakcją redox pompowanie protonów jest centralnym zagadnieniem bioenergetyki** i stąd propozycje wielu modeli do wytłumaczenia mechanizmu procesu, z uwzględnieniem **stanów przejściowych białka i przejściowych etapów reakcji**. Mimo to, modele nie dają odpowiedzi na podstawowe pytanie:

jak energia uwalniana przy redukcji tlenu jest używana do przenoszenia protonów w jednym kierunku, przy możliwym przecieku protonów w kierunku odwrotnym, napędzanym przez przeciwny do pompowania potencjał błony?

Szereg prac dotyczy roli wody w fotosystemie PSII jako substratu w reakcji rozszczepienia pod wpływem energii słonecznej i jako nośnika protonu, oraz jej lokalizacji w kompleksie PSII ([review: Linke & Ho Biochim. Biophys. Acta 1837, 14, 2014](#)). Jedną z metod mapowania powierzchni kontaktu białek PSII z wodą jest jej radioliza przy użyciu promieniowania synchrotronowego, co prowadzi do oksydacji aminokwasów rodnikami OH^\bullet i identyfikacji odpowiednich fragmentów białek w techniką MS ([Frenkel et al. J. Mol. Biol. 288, 23565, 2013](#)).

Sprężenie pomiędzy transferem protonów i transferem elektronów zachodzi nie tylko w procesach fotosyntezy i oddychania komórkowego, ale w bardzo dużej liczbie procesów technologii chemicznej, zarówno naturalnych jak i „sztucznych”. Badania w ostatnim okresie koncentrują się na **procesach zachodzących w sposób zsynchronizowany** (CPET, *concerted proton-electron transfer*) i równoczesny, w przeciwieństwie do mechanizmu złożonego z kolejno następujących etapów (*stepwise*).

Zaletą zsynchronizowanego sprzężenia wydaje się omijanie wysokoenergetycznych intermediatów, ale może to stwarzać koszty kinetyczne. Wymagane w tym przypadku **tunelowanie protonów** narzuca niewielką odległość transferu między grupą generującą proton przy oksydacji i akceptorem protonu, lub odwrotnie w procesie redukcji. Zwykle jest to odległość rzędu wiązania wodorowego tworzonego między tymi grupami. Idea mechanizmu pozwalającego wydłużyć odległość przekazu protonu między źródłem protonów uwalnianych przez transfer elektronów a akceptorem protonów zakłada pośrednictwem grupy tzw. „*hydrogen-bond relay*” [PLANSZA 228]. Grupa ta zawiera wodór i jest zdolna zarówno do utworzenia wiązania wodorowego jako akceptor z resztą utlenianą (punkt startowy transferu protonu) jak i jako donor wiązania wodorowego dla reszty przejmującej proton (punkt docelowy transferu protonu), bez przechodzenia przez stan uprotonowany w trakcie reakcji. Może to być grupa –OH w cząsteczce zawierającej fenol i pirydynę lub woda w rejonie Tyr_Z–His190–OEC (*oxygen-evolving complex*) fotosystemu PSII. Cykliczna woltametria (prąd *i* w funkcji potencjału elektrody E) pokazuje, że oksydacja fenolu zachodzi podobnie jak w cząsteczce aminofenolu (AP), w której oksydacja nie wymaga mechanizmu „H-bond relay”.

Badanie **tunelowania protonów** wymaga uwzględnienia **efektów kwantowych**. Prowadzi się analizę transferu protonów w relatywnie niewielkich, modelowych układach molekularnych. Przykładem może być porficyna (izomer porfiryny) [PLANSZA 229], w której można obserwować **podwójny transfer protonów**. Każdy z trzech tautomerów (trans, cis-1, cis-2) ma identycznego chemicznie partnera, co prowadzi do powstania symetrycznego potencjału o dwóch minimach. Tautomer *trans* ma najniższą energię. Badania spektroskopowe pokazują charakterystyczne rozszczepienia pasm (*tunneling doublets*), wynikające z kwantowej delokalizacji wewnętrznych protonów*. Klasyczo-kwantowa **dynamika molekularna** Car-Parrinello pokazuje lokalne ruchy protonów wokół położeń równowagi oraz przeskoki do drugiego minimum energii potencjalnej. Z 20 przejść w czasie symulacji 60 ps, 4 przejścia miały charakter przeskoków pojedynczych protonów z przejściowym tworzeniem tautomeru *cis-1* a pozostałe 16 to kooperatywne, synchroniczne przeskoki dwóch protonów.

* patrz wykład *Spektroskopia molekularna*: podwojenie inwersyjne pasma IR w cząsteczce amoniaku

Biopolimery w „żywej komórce” (*in vivo* NMR, *in cell* NMR)

Wykład 26

Analiza biologicznie aktywnych cząsteczek, zarówno metabolitów niskocząsteczkowych jak i biopolimerów, w naturalnym środowisku żywej komórki i tkanki stanowi znaczący krok naprzód w charakterystyce ich funkcjonowania. Lokalizacja i określanie zawartości biocząsteczek w płynach ustrojowych jest rutynowo wykonywane w **diagnostyce medycznej**. W badaniach o charakterze podstawowym w zakresie **METABOLOMIKI** ([Fien et al. Nature Biotechnol. 18, 1157, 2000](#)) poddaje się całościowej **analizie „high throughput” fenotyp**, czyli: zawartość metabolitów (*metabolite profiling, footprinting*), **drogi przepływu** metabolitów, oraz **czasowe zmiany metabolizmu** (*metabonomics*), pod wpływem różnych czynników środowiska: stres, dieta, konsumpcja leków, a także procesy chorobowe. „*Metabolite footprinting*” w zakresie analizy „*high throughput*” metabolitów wydzielanych pozakomórkowo ma bezpośredni związek z diagnostyką medyczną i odgrywa coraz bardziej znaczącą rolę w diagnozie nowotworów, identyfikacji nowych biomarkerów nowotworowych i rozwijaniu terapii antynowotworowej [**PLANSZA 230**]. Wymaga to niekiedy „wyznakowania” (*derivatization*) cząsteczek dla ich identyfikacji (sondy, izotopy), lokalizacji i kwantyfikacji metodami chromatografii (GS, *gas chromatography*, UPLC, *ultrahigh performance liquid chromatography*) spektrometrii mas (MS) i NMR.

Kluczową sprawą jest możliwość **wykonywania badań *in vivo*** w sposób możliwie **nieinwazyjny**, tzn. w nienaruszonych układach. NMR stwarza unikalne możliwości śledzenia metabolizmu ze względu na **bezpośrednią identyfikację biomolekuł na widmach**. Widmo NMR można rejestrować nie tylko dla wyizolowanego związku chemicznego w odpowiednio przygotowanej próbce, ale także ***in vivo***, dla nienaruszonych zawiesin komórek bakteryjnych, roślinnych i drożdżowych, wypreparowanych tkanek w specjalnych układach perfuzyjnych, zapewniających ciągłość procesów metabolicznych, lub **całkowicie nieinwazyjnie** dla wybranego obszaru całego żywego organizmu, umieszczonego wewnątrz magnesu aparatu NMR. Te ostatnie eksperymenty wymagają dodatkowego zastosowania **technik (mikro)obrazowania** (omówione w dalszej części wykładu). W warunkach eksperymentu ***in vivo* NMR** (*in vivo* MRS; *magnetic resonance spectroscopy*) [**PLANSZA 231**] **wysokiej zdolności rozdzielczej** rejestruje się sygnały pochodzące od jąder cząsteczek metabolitów wykazujących:

- (a) stosunkowo **niewielką masę** cząsteczkową;
- (b) dostateczną ruchliwość tzn. **nie związanych** w dużych kompleksach molekularnych, organellach komórkowych lub w błonach;
- (c) dostatecznie wysokie **stężenie**, powyżej $\sim 10 \mu\text{M}$.

Nie rejestruje się zwykle sygnałów pochodzących od dużych kompleksów biopolimerów i struktur komórkowych w stałym stanie skupienia. Przykładowe widma ^{31}P NMR, ^{13}C NMR, ^1H NMR, ^{14}N NMR (standardowo wykorzystywane jądra) mięśnia wołowego pozwalają **zidentyfikować szereg różnych metabolitów**, m. in. zawierającą fosfor fosfokreatynę (PCr) i ATP, kwas mlekowy (L), dwupeptyd karnozynę ($\beta\text{-Ala-His}$; patrz acetylo CoA), glikogen (Gl), tłuszcze (F). Wykonując widma NMR w określonych odstępach czasu można śledzić **zmiany zawartości metabolitów i ich przekształcanie w inne metabolity**, np. wnikanie związku do komórek i reakcje biochemiczne zachodzące w jej wnętrzu, z możliwością identyfikacji **rozkładu** związków w organellach subkomórkowych. Można w ten sposób również śledzić wnikanie do komórek i metabolizm **środków farmakologicznych**, np. zawierających **fluor** ^{19}F , np. 2'-fluoro-5-methyl- β -L-arabinofuranosyluracil do erytrocytów ([Xu et al. Biochem. Pharmacol. 55, 1611, 1998](#)), lub **znakowanych** ^{13}C , ^{15}N , ze względu na jednoznaczność i prostotę identyfikacji tak oznakowanych metabolitów na widmach: ^{19}F , ^{13}C i ^{15}N NMR.

Szczególnie intensywnie badanym zagadnieniem od początku stosowania metody w końcu lat 70-tych XX w. był proces **glikolizy** [**\[PLANSZA 232\]**](#), rozkładu sześciowęglowej glukozy do trójwęglowego pirogronianu z powstawaniem wysokoenergetycznego ATP, oraz odwrotny do glikolizy proces **glukoneogenezy**. Zgodnie z (poprzedni wykład) **teorią chemiosmotyczną Mitchella** [**\[PLANSZA 223\]**](#) sprzężenia oksydacji (oddychania) i fosforylacji, w komórkach bakteryjnych w warunkach beztlenowych powstaje **transmembranowy gradient potencjału elektrycznego** (gradient elektrochemiczny) wskutek przeniesienia na zewnątrz jonów H^+ [**\[PLANSZA 233\]**](#). Proces zachodzi kosztem energii hydrolizy ATP powstającego w glikolizie i jest wykonywany przez ATPazę związaną z błoną komórkową. Powstały gradient umożliwia aktywny transport molekularny przez błonę. W warunkach tlenowych gradient powstaje poprzez przeniesienie elektronu na O_2 . [**\[PLANSZA 225\]**](#). ATPaza pracuje wtedy w odwrotną stronę produkując ATP kosztem energii gradientu elektrochemicznego. Potwierdzenie teorii Mitchella bezpośrednio pokazują **widma ^{31}P NMR zawiesiny komórek *E. coli***

($5 \cdot 10^{11}$ komórek w mililitrze). W nieobecności glukozy obserwuje się szereg ufosforylowanych metabolitów: difosforany nukleozydów, NADP^+ , fosfoenolopirogronian [\[PLANSZA 232\]](#), (PEP) i fosforan nieorganiczny P_i . Po dodaniu glukozy do zawiesiny [\[PLANSZA 233\]](#), w wyniku glikolizy powstają (kosztem difosforanów) trifosforany nukleozydów, w tym ATP. Widać powstawanie kluczowego intermediatu glikolizy, **fruktozo-1,6-difosforan** (βFBP) [\[PLANSZA 232\]](#). Dwa sygnały od fosforu P_i [\[PLANSZA 233\]](#) odpowiadają fosforanowi wewnątrz (*in*) i na zewnątrz (*ex*) komórki. Różnica ich przesunięć chemicznych $\Delta\delta$ wynika z **różnicy wartości pH** po obu stronach błony bakteryjnej wskutek gradientu pH przez błonę, a pomiar $\delta(\textit{in})$ pozwala wyznaczyć **wewnątrzkomórkowe pH**, po odpowiedniej kalibracji δ fosforu w funkcji pH. Rozdzielenie sygnału nie następuje w mutantach ATPase^- lub po dodaniu inhibitora ATPazy, np. DCC (*dicyclohexylcarbodiimid*).

Wykonanie kolejnych widm ^{31}P NMR i ^{13}C NMR naturalnej abundancji [\[PLANSZA 234\]](#) zawiesiny komórek drożdży (500 mg suchych komórek w 5 ml) od momentu podania glukozy pozwala śledzić przebieg **fermentacji alkoholowej**. Zmiany stężeń metabolitów procesu z integracji sygnałów obrazują zanik glukozy i fosforanu nieorganicznego, wzrost a następnie zanik fruktozo-1,6-difosforanu [\[PLANSZA 232\]](#) i ATP, narastanie produkowanego alkoholu etylowego do stanu nasycenia [\[PLANSZA 234\]](#).

Znakowanie wyjściowych metabolitów pozwala na śledzenie **dróg metabolizmu**. Ekstrakt komórek zboża (a więc nie całkiem ściśle eksperyment *in vivo*) po podaniu octanu znakowanego ^{13}C kolejno na C1 i C2 [\[PLANSZA 235\]](#) pokazuje na widmach ^{13}C NMR:

- węgiel C1 jest tracony jako CO_2 ; brak różnic **widma B** w stosunku do widma naturalnej abundancji **A**.

- węgiel C2 (pik AcO^- ; **widmo C**) wchodzi poprzez cykl Krebsa do aminokwasów: Glu, Gln, Asn, Ala, Ser.

W analogicznym metodologicznie eksperymencie ([review of in vivo plant metabolism: Schneider Planta 203, 1, 1997](#)) śledzono biosyntezę glutaminy znakowanej $\alpha, \gamma\text{-}^{15}\text{N}$ i kwasu glutaminowego $\alpha\text{-}^{15}\text{N}$ w *Bacillus azotofixans* po podaniu znakowanego azotanu $^{15}\text{NO}_3^-$. Z azotanu powstaje $^{15}\text{NH}_4^+$, włączany do tych aminokwasów po czasie $3 \div 30$ min. Problemem było określenie dominującego mechanizmu włączania amoniaku do aminokwasów z dwóch zakładanych modeli:

- (a) synteza glutaminy katalizowana przez syntetazę glutaminową (GS), a następnie synteza kwasu glutaminowego z glutaminy i α -ketoglutaranu przez syntetazę kwasu glutaminowego GOGAT;
- (b) bezpośrednia synteza kwasu glutaminowego z α -ketoglutaranu przez dehydrogenazę kwasu glutaminowego (GDH).

Szybkie narastanie glutaminy (60% całkowitej intensywności γ - ^{15}N po 3 min) i stopniowe narastanie kwasu glutaminowego (6% po 3 min) sugerują pierwszy mechanizm (a). Podanie inhibitorów GOGAT, które jednocześnie nie są inhibitorami GDH blokuje tworzenie kwasu glutaminowego w ponad 90%

Przy rejestracji **widm NMR tkanek i narządów *in vivo*** początkowo stosowano metodę inwazyjną preparowania i umieszczenia narządu (tkanki) w układzie perfuzyjnym, który zapewniał zachodzenie metabolizmu. Obecnie używa się metody nieinwazyjnej przez **stworzenie warunków rezonansu jądrowego tylko w określonym miejscu** organizmu umieszczonego w polu magnetycznym aparatu NMR. Wybór **obszaru VOI** (*volume of interest*) **[PLANSZA 236]** uzyskuje się **technikami obrazowania** (mikroobrazowania) **MRI** (*magnetic resonance imaging*; wykład „**Spektroskopia molekularna**”; szczegóły następny wykład). Połączenie *in vivo* NMR i MRI określa się niekiedy jako **TMR** (*topical magnetic resonance*) lub 3D MRSI (*magnetic resonance spectroscopy imaging*). Widmo ^{31}P NMR fragmentu 5 x 5 x 5 cm pokazuje ufosforylowane **metabolity tkanki nerwowej mózgu**. Szereg prac poświęcono metabolizmowi mięśni, w tym mięśnia sercowego, w których **zasadnicze reakcje [PLANSZA 237]** to **transfer reszty fosforanowej** między wysokoenergetycznymi związkami: ATP i fosfokreatyną, katalizowany przez kinazę kreatyny, oraz **hydroliza ATP do ADP**. W procesie niedotlenienia (*ischemia*) mięśnia sercowego **[PLANSZA 238]** na widmie ^{31}P NMR widać początkowo spadek zawartości fosfokreatyny, a następnie wyczerpywanie się ATP i powstawanie fosforanu nieorganicznego. Powtórne doprowadzenie tlenu przywraca w ok. 70% poziom ATP i fosfokreatyny. Intensywna praca mięśni prowadzi do powstawania kwasu mlekowego [PLANSZA 239], identyfikowanego na podstawie grupy CH_3 na widmach ^1H NMR. Słumienie wody uzyskano sekwencją CPMG (Carr-Purcell-Meiboom-Gill), która polega na wykorzystaniu różnicy czasów relaksacji T_2 wody i innych związków (sekwencja używana do pomiaru T_2). **Wydzielony sygnał CH_3 mleczanu** uzyskano z widm różnicowych dla selektywnych soft-pulsów (odsprzęgających) po obu stronach tego sygnału.

Szeroki wachlarz **zastosowań *in vivo* NMR** obejmuje również kinetykę reakcji enzymatycznych: przepływy, stałe szybkości, w warunkach żywej komórki lub tkanki ([Brindle *Progrwss NMR Spectro.* 20, 257, 1988](#)). Coraz intensywniej rozwijają się badania z zastosowaniem technik wielowymiarowych, począwszy od jednej z pierwszych prac dotyczącej metabolizmu azotu ^{15}N ($(^{15}\text{NH}_4)_2\text{SO}_4$) drożdży metodą dwuwymiarowej spektroskopii korelacyjnej *double quantum* (DQ) ^1H NMR/ ^{15}N NMR ([Juretschke *FEBS Lett.* 178, 123, 1984](#)), do prac z ostatniego okresu: metabolizm azotu w wątrobie mysiej przy użyciu ^1H , ^{14}N HSQC [\[PLANSZA 240\]](#).

Znaczny postęp w **rozwoju metabolomiki NMR** (*magnetic resonance metabolomics*) został osiągnięty dzięki zastosowaniu techniki MAS (*magic angle spinning*) do uzyskiwania widm wysokiej zdolności rozdzielczej w ciele stałym HR MAS (*high resolution MAS*). HR MAS umożliwia **detekcję i śledzenie metabolizmu związków o większych masach cząsteczkowych i/lub zintegrowanych w kompleksach molekularnych lub strukturach komórkowych** (*rigid cell components*). Przykładowo, bezpośrednie badania wykonano dla O-antygeny (A), lipopolisacharydów (LPS) długości $n > 100$ krótkich fragmentów oligocukrowych na zewnętrznej błonie bakteryjnej [\[PLANSZA 241\]](#), w różnych fazach rozwoju bakterii. Zwraca uwagę wykorzystanie jąder ^1H przy MAS z relatywnie niską częstotliwością wirowaniem 6 kHz (!?). Badania HR MAS okazują się niezwykle istotne jako biotechnologiczne, nieinwazyjne narzędzie w analizie metabolomiki nowotworów [\[PLANSZA 242\]](#), w szczególności następujących zagadnień:

- **biomarkery detekcji** nowotworów: różnice metabolomicznego profilu tkanki nowotworowej w porównaniu z tkanką zdrową: metabolity zawierające cholinę, poziom mleczanu i glukozy (mózg, jelito grube), poliaminy i cytrynian (prostata);
- **charakterystyka** nowotworu: typ, lokalizacja, ewolucja (*histopatologic grade*);
- **monitorowanie** działania środków leczniczych w trakcie **terapii** (*treatment*) i wybór strategii leczenia oraz ocena ryzyka (*risc stratification*).

Wykorzystanie HR MAS NMR bezpośrednio wiąże się z **techniką „*in cell* NMR”**, skierowanej na uzyskanie danych na temat biopolimerów, głównie białek, w **naturalnym środowisku komórkowym** ([review: *Seber & Dötsch Biochemistry* 40, 14317, 2001](#)). Badania tego typu, jak na razie, dotyczą głównie obserwacji **zmian strukturalnych, dynamiki i oddziaływań biopolimerów w funkcjonującej komórce**. Przykładowo, badanie fosforylacji białek *in vitro* może być utrudnione ze względu na brak informacji o

właściwych kinazach, miejscu wprowadzania grupy fosforanowej, czy izolowaniu odpowiednich ilości materiału do badań. Te problemy dają się usunąć przy zastosowaniu metod „*in cell* NMR” ([Selenko et al. Nature Struct. Mol. Biol. 15, 321, 2008](#)). W ostatnim okresie zaczęły się pojawiać prace prezentujące wielowymiarowe widma *in cell* NMR o odpowiedniej jakości (szerokości połówkowej linii), które umożliwiają wyznaczenie struktur białek *in vivo* ([Hamatsu et al. J. Am. Chem. Soc. 135, 1688, 2013](#)): 2D ^1H - ^{15}N HSQC oraz 3D triple resonance; system ekspresji: komórki owadzie sf9/wektor baculowirusa.

Stosowanie techniki *in cell* NMR wymaga rozwiązania dwóch zasadniczych problemów ([Xu et al. Biochemistry 53, 1971, 2014](#)):

(1) odróżnienie sygnałów NMR interesującego białka od sygnałów „tła” innych związków, szczególnie niskocząsteczkowych (*in vivo* MRS);

(2) zwiększenie szerokości połówkowych sygnałów rezonansowych w „zatłoczonym” środowisku o wysokiej lepkości i przy oddziaływaniu białek ze strukturami komórkowymi.

Pierwszy problem rozwiązuje się przez znakowanie izotopowe $^{15}\text{N}/^{13}\text{C}$. Białka mogą być również obserwowane przez wyznaczenie fluorem ^{19}F : inkorporacja 3-fluorotyrozyny, 3-fluorofenyloalaniny. Do wytwarzania w komórkach białka w odpowiednim stężeniu stosuje się sterowaną nadekspresję, np. pod kontrolą promotora polimerazy RNA T7 bakteriofaga. Zastosowanie inhibitorów polimeraz komórkowych (ryfampicyna), które nie blokują polimerazy T7, hamuje wtedy ekspresję innych białek komórkowych.

Drugi problem jest rozwiązywany w sposób ogólny poprzez stosowanie

- sekwencji TROSY, najczęściej dla częściowo deuterowanych cząsteczek białka;
- techniki HR MAS wirowania pod kątem magicznym.

Odpowiednikiem „*in cell* NMR” w obszarze **badań chemicznych** jest metodologia „*in situ* NMR”, np. analiza procesów spalania metanu mieszanego z ksenonem, z wykorzystaniem rezonansu ^{129}Xe NMR ([Anala et al. J. Am. Chem. Soc. 125, 13298, 2003](#)).

Obrazowanie NMR w DIAGNOSTYCE MEDYCZNEJ (*MRI* i *fMRI*); porównanie z tomografią komputerową (*X-ray CT*) i pozytonową (*PET*)

Wykład 27

Podstawowym sposobem wykorzystania NMR w diagnostyce i terapii medycznej jest nieinwazyjne obrazowanie MRI (*magnetic resonance imaging*) [PLANSZA 243], jako jedna z szeregu technik obrazowania molekularnego (*review: Massoud & Gambhir Genes & Development 17, 545, 2003*). W zakresie obrazowania struktur narządów wewnętrznych i tkanek MRI jest techniką analogiczną jak tomografia komputerowa CT (*computer tomography*) z wykorzystaniem promieniowania rentgenowskiego. W zakresie obrazowania procesów biologicznych na podstawie przepływów płynów ustrojowych, MRA (*magnetic resonance angiography*) i fMRI (*functional MRI*), jest techniką komplementarną do PET (*positron emission tomography*). W MRI stosuje się niewielkie indukcje statycznego pola B_0 , do ok. 7 T, dla którego można zachować homogeniczność w dużej komorze (*wide-bore-magnet*) mieszczącej całego pacjenta. Na tej samej zasadzie obrazowania oparta jest technika nieinwazyjnej mikroskopii NMR (MRM) niewielkich obiektów. W nowoczesnych aparatach MRS/MRI łączy się eksperymenty *in vivo* NMR i obrazowanie spinowe przez rejestrację obrazu, wydzielenie pożądanego fragmentu VOI i rejestrację widma określonego typu dla tego małego obszaru (patrz poprzedni wykład).

W MRI wykorzystuje się głównie, ale nie jedynie, rezonans ^1H NMR wody i tłuszczów zawartych w tkankach. Różnice intensywności sygnałów NMR w różnych punktach obrazowanego obiektu, które stanowią podstawę uzyskiwania obrazu, wynikają zasadniczo z różnic:

- zawartości wody [PLANSZA 244]: dużo w płynach ustrojowych i tkankach miękkich zależnie od typu (mięśnie, mózg, wątroba, skóra), mało w kościach (*femur cortex*), praktycznie brak w “pustym” obszarze płuc;
- czasów relaksacji T_1 i T_2 wody.

Idea eksperymentu jest prosta [PLANSZA 245]. Przykładowo, gdybyśmy umieścili obiekt, np. dwa próbki z wodą w stałym polu magnetycznym B_0 otrzymalibyśmy jeden zbiorczy sygnał NMR. Jeżeli wzdłuż osi x ustawienia próbników przyłożymy liniowy gradient g_x statycznego pola magnetycznego, zjawisko rezonansu będzie zachodziło przy innych częstościach dla każdego z tych próbników:

$$\omega = \gamma(B_0 + g_x x).$$

Otrzymamy dwa sygnały osobno od każdego z próbników, a **mierzonym różnym częstościom rezonansowych ω odpowiadają różne położenia x próbników:**

$$x = \omega / (\gamma g_x).$$

Eksperyment MRI składa się z dwóch etapów **[PLANSZA 246]**:

(1) Okres przygotowawczy (*preconditioning period*) decyduje o **kontraście** obrazu, dzięki zastosowaniu sekwencji *inversion recovery* oraz *spin-echo* (sekwencje do pomiaru czasów relaksacji) i wykorzystaniu różnic czasów relaksacji T_1 i/lub T_2 .

(2) Tworzenie obrazu, najczęściej przedstawianego w formie przekroju 2D obiektu 3D, lub coraz częściej w formie obrazu przestrzennego 3D. Technik tworzenia obrazu jest bardzo dużo, zależnie od wymaganego przeznaczenia i szybkości zbierania obrazu. Techniki *point measurement* są oparte na powolnym skanowaniu obiektu przez selektywne wzbudzenie rezonansu punkt po punkcie; stosowane m. in. w wyborze VOI dla TMR. Szybkie techniki *plane measurements* stosuje się w obrazowaniu medycznym, np. wielowymiarowy eksperyment ***Fourier imaging*** **[PLANSZA 247]**. Po ostatnim impulsie wzbudzającym B_1 (90°) sekwencji *preconditioning* przykłada się kolejno gradienty wzdłuż osi x oraz y przez okresy czasu t_1 oraz t_2 (okres ewolucji), inkrementowanej długości, od zera do wartości końcowej. W ten sposób informacja o współrzędnych punktu (x, y) dającego sygnał NMR jest fazowo zakodowana w rejestrowanym w czasie t_3 sygnale FID, dla którego trzeci, stały gradient wzdłuż osi z koduje tę współrzędną. **Po trójwymiarowej transformacji Fouriera sygnału $s(t_1, t_2, t_3)$ powstaje obraz intensywności $S(x, y, z)$ poszczególnych punktów obiektu.** Widać to dokładniej dla przypadku dwuwymiarowego **[PLANSZA 248]**. Czas t_1 trwania gradientu g_x kode współrzędną x na sygnale $s(t_1, t_2)$, a częstość precesji w gradiencie g_y **[PLANSZA 245]** daje drugą współrzędną y . **Modulację fazową** rejestrowanego sygnału pokazuje dokładnie dolna część **[PLANSZA 248]** z widmem $S(\omega_1, \omega_2) \equiv S(x, y)$ dwóch próbników z wodą, rozseparowanych wzdłuż y : wartość ω_1 odpowiada współrzędnej x , a wartość ω_2 odpowiada współrzędnej y w postaci **dubletu** $S(t_1, \omega_2)$.

Jakość i rozdzielczość obrazu są uwarunkowane przez następujące czynniki:

(a) jednorodność statycznego pola B_0 i liniowość gradientów g_i decydują o dokładności odtworzenia, czyli **braku zniekształceń** obrazu;

(b) ilość eksperymentów w czasach ewolucji t_1 , t_2 , ilość punktów próbkowania w czasie rejestracji t_3 oraz nachylenie (wartości) gradientów warunkują w sumie zdolność rozdzielczą, czyli **minimalną objętość, z której jest zbierany osobny sygnał rezonansowy**.

W dużej komorze magnesu MRI stosuje się niższe niż w MRS indukcje pola B_0 , co przy gradientach rzędu 10 mT/metr daje rozdzielczość rzędu $1 \div 2$ mm w każdym z trzech wymiarów. Stąd możliwość detekcji zmian chorobowych tkanek, a w szczególności nawet mało zaawansowanych nowotworów, przede wszystkim na skutek **różnic w czasach relaksacji między tkanką normalną i nowotworową** [PLANSZA 244]. Można dodatkowo zróżnicować obraz obu tkanek przez wprowadzenie w odpowiednie obszary **czynników kontrastujących**, które modyfikują czasy relaksacji. Przykładem jest zawierający Gd^{+3} *dimeglumine gadapentate* z grupy kompleksów organicznych **jonów lantanowców** (*review: Saini et al. Adv. Contrast-Enhanc. MR Imag. 156, 235, 1991*). Obiecujące są szczególnie wprowadzane do tkanki nowotworowej *smart contrasting agents*, które uaktywniają się dopiero po dotarciu i wnikięciu do komórek nowotworowych. Pod względem diagnostyki zmian chorobowych (nowotworów) tkanek miękkich MRI przewyższa pod względem jakości i rozdzielczości obrazowanie CT przy pomocy promieniowania rentgenowskiego [PLANSZA 243], w którym obrazy otrzymuje się w wyniku zróżnicowanej absorpcji promieniowania X przez tkanki. Jednak CT wydaje się z kolei lepszą techniką przy obrazowaniu kości.

Modyfikacje eksperymentu MRI przez stosowanie:

- różnych sekwencji impulsów i gradientów;
 - specjalistycznego oprzyrządowania spektrometru NMR;
 - różnych metod obróbki sygnałów rezonansowych,
- można uzyskać obrazy do określonych celów.

(A) **Obszary „puste”** czyli wypełnione protonami o małym stężeniu jak np. płuca są trudne do konwencjonalnego obrazowania NMR ze względu na słaby sygnał i zniekształcenia na granicy gaz/tkanka wskutek różnic podatności magnetycznej, a w konsekwencji niejednorodności pola magnetycznego prowadzące do powstawania „szkodliwych” mikrogradientów. Dobre efekty obrazowania MRI płuc uzyskuje się przy zastosowaniu rezonansu ^3He NMR z **hiperpolaryzowanym helem** (*review: Viale & Aime Cur. Op. Chem. Biol. 14, 90, 2010*). Metastabilny stan hiperpolaryzacji uzyskuje się przez wstępne umieszczenie helu w niskiej temperaturze 4,2 K i polu magnetycznym o wysokiej

indukcji ponad 4 T lub przez oddziaływanie nadsztywne z rubidem o spolaryzowanych spinach elektronowych. Po wpuszczeniu do płuca (edema) stan hiperpolaryzacji ^3He utrzymuje się dostatecznie długo dla wykonania eksperymentu MRI i daje zwiększoną $10^2 \div 10^4$ razy różnicę boltzmannowskich obsadzeń poziomów spinowych, czyli wzrost intensywności sygnału.

(B) Jeśli ograniczy się rozmiary obrazowanego obiektu (poniżej 1 cm) można w polach magnetycznych ~ 14 T zastosować bardzo strome gradienty, do 1 T/m i uzyskać rozdzielczość przestrzenną obrazu Δx do 10^{-5} mm³:

$$\Delta x = 2/(\gamma g_x T_2)$$

W ten sposób powstaje **obraz mikroskopowy NMR** ([review: Kuhn Angew. Chem. Int. Ed. 29, 1, 1990](#)). Po pierwszym, niewyraźnym obrazie komórki jaja ropuchy amerykańskiej [[PLANSZA 243](#)] ([Aguayo et al. Nature 322, 190, 1986](#)) nastąpił szybki rozwój techniki. Przykładowo, obrazowanie z akwizycją *multiecho* (polepszenie stosunku sygnał/szum) struktur wewnętrznych muchy unieruchomionej w wosku bez narkozy [[PLANSZA 249](#)] pokazuje szczegóły anatomiczne mózgu z **dokładnością do 20 μm^3** w dwóch przekrojach ([panele A i C](#)) wraz z odpowiednimi schematami anatomicznymi: LC – *lobula complex*, La – *lamina*, M – *medulla*, ON – *ocular nerve*, Es – *esophagus*. MMR rokuje perspektywę obrazów w skali molekularnej ([Hemmer Science 339, 529, 2013](#))

(C) Technika **flow imaging** daje obraz MRA (*magnetic resonance angiography*) [[PLANSZA 243](#)] tylko cieczy o ciągłym przepływie: naczynia krwionośne, naczynia limfatyczne ([Nordon et al. Analyst 126, 260, 2001](#)). Efekt uzyskuje się albo po nasyceniu sygnałem rezonansowym rf (*radiofrequency*) i przepływie jąder o zmienionej magnetyzacji M_z do obszaru detekcji (*time of flight MRA*), albo zmieniając fazę magnetyzacji poprzecznej (M_x , M_y) w technice *phase contrast MRA*. Technika jest komplementarna do **obrazowania PET**, w której wykorzystuje się detekcję dwóch kwantów γ rozchodzących się z punktu anihilacji pozytronu e^+ po rozpadzie β radioaktywnego izotopu o krótkim okresie półrozpadu, wprowadzanego do naczyń.

(D) **Lokalizację aktywnych neuronów** z rozdzielczością do milimetra i w czasie rzędu sekund od zadziałania bodźca [[PLANSZA 250](#)] wykonuje się techniką funkcjonalnego rezonansu fMRI. Metoda przewyższa PET i może dać informacje uzyskiwane dotychczas przez wprowadzanie elektrod w wybrane obszary mózgu. Aktywna praca komórek nerwowych wiąże się ze **zwiększonym zużyciem** tlenu transportowanym

przez krew. Najprawdopodobniej (dokładny mechanizm nie jest do końca poznany), **redukcja żelaza hemoglobiny** w aktywnym neuronie zmniejsza podatność magnetyczną i w ten sposób powstaje skontrastowany obraz (T_2 - *weighted*) aktywnego obszaru w stosunku do otaczającej tkanki nerwowej [PLANSZA 251]. Hitem stały się badania tzw. *working memory* [PLANSZA 250], zlokalizowanej w płatach czołowych obu półkul (*prefrontal cortex*), czyli odpowiednika RAM w komputerze, dokąd kopiowana jest informacja z pamięci trwałej (ROM) do porównań, wyborów i innych reakcji na bodźce (*Cohen et al. Nature 386, 604, 1997*). Typowe **testy psychologiczne *delayed choice***, polegające na sygnalizowaniu określonej odpowiedzi na bodziec zależnie od tego czy jest znany czy nie, pozwalają lokalizować obszary kory zaangażowane w dokonywanie wyboru. Pokazano w ten sposób, że *prefrontal cortex* jest zorganizowana w regiony, które czasowo przechowują informację o różnych aspektach bodźca: aspekty przestrzenne, ogólne cechy wyglądu, itp. Spektakularna praca w *Science* z 2001 roku [PLANSZA 251] relacjonuje wyniki badań stopienia emocjonalnego zaangażowania w rozwiązywaniu dylematów moralnych („eksperymentalna” filozofia).

(E) **Selektywne obrazowanie CSI (*chemical shift imaging*)** z wykorzystaniem jednego sygnału dowolnego jądra o określonym przesunięciu chemicznym (*Brown et al. Proc. Natl. Acad. Sci USA 79, 3523, 1982*) pozwala ograniczyć mapowanie do wybranej substancji i w ten sposób określić jej **rozkład w tkance**. „Dodatkowy wymiar” przesunięcia chemicznego w CSI-MRI uzyskuje się np. przez wprowadzenie dodatkowego pulsującego gradientu. Wykonanie obrazowania CSI ^1H NMR na podstawie tylko sygnału wody i analogiczne obrazowania tylko na podstawie odległego o 3,5 ppm sygnału tłuszczu (*review: Jingfei J. Magn. Reson. 28, 543, 2008*) umożliwia wgląd w odkładanie się tkanki tłuszczowej między tkanką mięśniową [PLANSZA 252] w dystrofii miotonicznej (choroba Steinerta). **CSI jest podobnie jak TMR połączeniem technik MRI i MRS**, przy czym nacisk położony jest na obrazowanie, a parametr spektroskopowy (przesunięcie chemiczne) spełnia rolę dodatkowego czynnika różnicującego obraz ze względu na uzyskanie odpowiednich informacji (odwrotnie niż w TMR).

(F) Przy wspomnianym poprzednio połączeniu technik MRI i MRS w **TMR** [PLANSZA 236] obraz uzyskiwany techniką STEAM (*stimulated echo*) ma charakter pomocniczy w wyznaczeniu VOI na (gradienty 3 mT/m wraz z gaussowskimi impulsami rf). Poszukiwaną informację dostarcza widmo NMR (MRS). **Połączenie obu technik**

stwarza teoretycznie nieograniczone możliwości jednoczesnego śledzenia **morfologii i anatomii** w stanach zdrowych i patologicznych w sposób nieinwazyjny.

(G) Równolegle prowadzone są intensywne badania nad usprawnieniem i implementacją nowych technik obrazowania MRI, które poprawiają **jakość** oraz **rozdzielczość przestrzenno-czasowej obrazu** ([Yu et al. Nature Methods 13, 337, 2016: single vessel, 100 ms fMRI](#)). Uwalniają również pacjenta od „niekomfortowej”, głośnej i zamkniętej przestrzeni elektromagnesu, Metoda *travelling-wave* MRI ([Brunner et al. Nature 457, 994, 2009](#)) idzie w tym kierunku. Zaprojektowano już funkcjonujące aparaty MRI z **otwartymi komorami** magnesu MRI. Prowadzone są również próby całkowitego **wyeliminowania magnesów nadprzewodzących** i rejestrację sygnału NMR techniką *optical atomic magnetometry* ([Xu et al. Proc. Natl. Acad. Sci USA 103, 12668, 2006](#)).

Technika obrazowania CSI jest przykładem jednego z szeregu możliwych zastosowań **obrazowania molekularnego** z detekcją molekularnych mechanizmów metabolizmu *in vivo*, stanów patologicznych (chorób) oraz przebiegu terapii. Obrazowanie wykonuje się z zastosowaniem m. in. **fluorescencji, spektrometrii mas SIMS** (*secondary ion mass spectrometry*), **stymulowanego rozpraszania Ramana SRS** (*stimulated Raman scattering*) [**PLANSZA 253**] oraz wspomniany już EPR. Oprócz samego obrazu można próbować uzyskać **wartości parametrów fizykochemicznych komórek in vivo** wykorzystując ich wpływ na rejestrowany obraz. Charakterystycznym przykładem jest pomiar lepkości w komórkach na podstawie obrazowania FLIM rotorów molekularnych wprowadzanych do środka komórki w wyniku endocytozy ([Kuimova et al. J. Amer. Chem. Soc. 130, 6672, 2008](#)). W badaniach *in vitro* pokazano wzrost fluorescencyjnego czasu życia (i wydajności kwantowej fluorescencji) fluoroforu ze wzrostem lepkości roztworu; lepkość warunkuje szybkość obrotu wewnętrznej cząsteczki. Po skalibrowaniu zależności czas życia – lepkość wykonano obrazowanie FLIM komórek inkubowanych z cząsteczkami rotora. Uzyskano obraz rozkładu fluorescencyjnego czasu życia rotora, a stąd stosunkowo wysoką wartość lepkości (140 ± 40 cP). Równoległe pomiary anizotropii fluorescencji wykluczyły ograniczenie rotacji wewnętrznej wskutek wiązania cząsteczki rotora przez „tarczę” makromolekularną wewnątrz komórki.

Dygresja. Nagroda Nobla w 2003 r. z "Fizjologii i medycyny" została przyznana za MRI dla P.C. Lauterbura i P. Mansfielda, a pominięto Raymonda Damadiana. R. Damadian zamieścił na 1. stronie New York Times "ogłoszenie": The shameful wrong that must be righted. Albo "honory" naukowe albo "pieniądze". R. Damadian uzyskał tylko z firmy Siemens \$155 mln za naruszenie praw patentowych przy produkcji aparatów MRI.

„Niekonwencjonalne” wykorzystanie polimerów biologicznych

Wykład 28

Omówionym poprzednio przykładem „niekonwencjonalnego” wykorzystania kwasów deoxyrybonukleinowych była **architektura wielowymiarowych struktur typu NANO [PLANSZA 72]**. Coraz częściej próbuje się wykorzystywać cząsteczki prostych związków organicznych oraz biopolimerów do **konstrukcji niekonwencjonalnych elementów komputerowych**. Próby te ilustrują możliwości budowy komputerów „nowej generacji” z pojedynczymi molekulami jako „cegiełkami” budowy obwodów logicznych o dużym stopniu miniaturyzacji. Badanie **własności elektrycznych** pojedynczych biomolekuł stwarza możliwość stosowania **technologii hybrydowych**, w których cząsteczki biologiczne i półprzewodniki użyte są łącznie. Przykładem może być wykazanie efektu *proximity-induced superconductivity* w podwójnej helisie DNA długości 16 μm umieszczonej między dwiema nadprzewodzącymi elektrodami ren/węgiel (Re/C) w temperaturze poniżej 1 K (*laboratorium CNRS w Orsay: Science 291, 280, 2001*).

Ciekawą próbę stanowi budowa układów **pamięci trójwymiarowej** (w przeciwieństwie do aktualnych dwuwymiarowych) z szybkim dostępem, na bazie cząsteczek białkowych. Duże zainteresowanie wzbudza pod tym względem **rodopsyna bakteryjna [PLANSZA 254]** zawierająca retinal jako absorbujący chromofor. Pod wpływem światła widzialnego o różnych długościach fali zachodzi zmiana konformacji chromoforu, która z kolei indukuje przejścia konformacyjne w samym białku między różnymi formami: bR, K, M, O, P, Q. Zmiany te mają charakter trwałe lub przejściowe. Bakteriorodopsyna wykorzystuje energię kwantów światła do pompowania protonów przez błonę w celu wygenerowania gradientu elektrochemicznego, a podobne białko rodopsyna jest podstawowym elementem w procesie widzenia. Oba białka można zastosować do **przechowywania informacji** przez wykorzystanie dwu dowolnych, dostatecznie **długożyciowych stanów jako reprezentantów jednostek binarnych 0 oraz 1**. Zapis i odczyt informacji odbywa się za pomocą układu wiązek laserowych i detektorów otaczających sześciąt z cząsteczkami białka. Przykładowo, impuls światła zielonego pobudza wszystkie cząsteczki wybranej płaszczyzny sześciąta dając początek fotocyklu ze stanu konformacyjnego bR, a światło czerwone z laserów prostopadłych do zielonych na wybrane obszary płaszczyzny przeprowadza zlokalizowane tam molekuly białka w konformacyjny **stan Q - zapis binarnego stanu 1**. Pozostałe cząsteczki relaksują do **stanu**

podstawowego bR, reprezentujący zapis stanu binarny 0. Odczyt jest możliwy wiązką zieloną oraz niskoenergetyczną wiązką czerwoną: różnice długości fali absorbowanej przez chromofor białka w konformacjach bR i Q, z wykorzystaniem detektora. W ten sposób uzyskuje się odczyt w postaci układu **ciemnych i jasnych prążków** na powierzchni detektora odpowiadającej **plaszczyźnie zapisu**. **Zaletą** takiej pamięci jest:

- trójwymiarowość, czyli duża pojemność;
- duża szybkość adresowania poprzez równoległy dostęp (*parallel processing*).

Przewiduje się, że pierwszy parametr może dać ok. 300-krotną poprawę w stosunku do urządzeń 2-wymiarowych; w granicy ok. 10^{12} bitów/cm³. Drugi parametr daje 10 mln bitów na sekundę.

Przykładem wykorzystania cząsteczki **DNA jako procesora komputerowego** do rozwiązania określonego problemu jest poszukiwanie tzw. **drogi Hamiltona** [**PLANSZA 255**]. Chodzi o wyznaczenie drogi z, przykładowo, miasta startowego Atlanta do miasta końcowego Detroit, przechodząc tylko raz przez wszystkie miasta pośrednie i przestrzegając zadanych kierunków możliwych połączeń. **Nie istnieją** wydajne algorytmy rozwiązywania tego typu problemów i przy 100 punktach najlepsze komputery potrzebowałyby kilku lat pracy CPU (problem NP-complete). **Algorytm** działania jest następujący.

- (1) Wygenerować układ przypadkowych połączeń n punktów.
- (2) Sprawdzić kolejno dla każdego połączenia:
 - (a) punkt startowy i punkt końcowy i odrzucić niewłaściwe połączenia ze zbioru,
 - (b) przejść przez dokładnie n punktów i odrzucić niewłaściwe,
 - (c) dla każdego punktu sprawdzić, czy idzie przez niego połączenie.
- (3) Niepusty zbiór oznacza istnienie rozwiązania problemu a pusty brak rozwiązania.

Problem przedstawiony na **PLANSZY** wymagał 7 dni pracy laboratoryjnej. Poszczególne punkty algorytmu (1 ÷ 3) są realizowane przez **syntezy** odpowiednich fragmentów DNA. Enzymatyczne **łączenie** i **separacja** właściwych produktów czyli cząsteczek DNA o właściwej długości i sekwencji jest **rozwiązaniem problemu**. Dokładna procedura w każdym z kolejnych kroków od (1) do (3) jest następująca.

- (1) Przyporządkowanie **punktom-miastom** ośmioznakowych nazw z nukleotydów: cztery pierwsze znaki określa się jako „imię”, cztery dalsze znaki jako „nazwisko”. Przyporządkowanie **każdemu połączeniu** ośmioznakowej nazwy: cztery pierwsze znaki to „nazwisko” punktu wyjścia, cztery pozostałe to „imię” punktu dojścia. **Synteza** chemiczna fragmentów DNA:

- nazwy miast (*DNA names*)
- komplementarne nazwy miast (A odpowiada T, G odpowiada C, itd.; *complements*);
- nazwy połączeń (*DNA flight numbers*);

Następnie zmieszanie fragmentów DNA w równych stężeniach i dodanie ligazy łączącej fragmenty. **Zbiór produktów reakcji zawiera rozwiązanie.**

(2a) Powielenie właściwych sekwencji metodą PCR z primerami: „nazwisko” punktu wyjściowego GCAG i komplementarne „imię” punktu docelowego GGCT;

(2b) Migracja elektroforetyczna na żelu wydziela DNA poszukiwanej długości 24 nukleotydy.

(2c) Rozdział przez powinowactwo (*affinity separation*) za pomocą krótkich fragmentów próbnych DNA.

(3) Sekwencja DNA długości 24n: **GCAGTCGGACTGGGCTATGTCCGA** jest rozwiązaniem problemu: **Atlanta – Boston – Chicago – Detroit.**

Dynamiczny rozwój zaprezentowanych na wykładzie obszarów badawczych biofizyki molekularnej jest oparty na **interdyscyplinarnym** (multidyscyplinarnym) wykorzystaniu zdobyczy doświadczalnych **nauk przyrodniczych**: fizyki, chemii, biologii (molekularnej) oraz nowatorskiego podejścia w analizach teoretycznych, szczególnie w zakresie **projektowania molekularnego i bioinformatyki**. Podsumowując wykład, warto jeszcze raz scharakteryzować aktualne tendencje rozwoju biofizyki molekularnej w wybranym ujęciu kluczowych tematów.

1. Podejście „high throughput” w całościowej analizie złożonych polimerów biologicznych.
2. Wzrastająca masa cząsteczkowa układów molekularnych, analizowanych pod względem strukturalnym i dynamicznym; przechodzenie od układów modelowych (uproszczonych) do "rzeczywistych", funkcjonalnych kompleksów.
3. Przechodzenie od analiz własności i oddziaływań biomolekuł w modelowych warunkach *in vitro* do charakterystyki biomolekuł w ich naturalnym środowisku funkcjonowania *in vivo*.

Prezentowane podczas wykładu informacje są w sposób ciągły **aktualizowane w świetle nowych danych**, zawartych w stale publikowanych pracach.